NBER WORKING PAPER SERIES

BUSINESS GROUP SPILLOVERS

S. Lakshmi Naaraayanan
Daniel Wolfenzon

Working Paper 31107
http://www.nber.org/papers/w31107

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023

For helpful comments, we thank Sumit Agarwal, Morten Bennedsen, Nittai Bergman, Janis Berzins, Kirill Borusyak, Claudia Custodio, Francesco D'Acunto, Sudipto Abhiroop Mukherjee, Erik Loualiche, Indrajit Mitra, Kasper Nielsen, Gordon Philips, Shivaram Rajgopal, Rik Sen, Giorgo Sertsios, Matthew Spiegel, Philip Strahan, Parth Venkat, Wei Wang, Tracy Wang, Alminas Zaldokas and seminar participants at the 2020 Minnesota Corporate Finance Meetings, 2019 EFA Meetings (Carcavelos), 2019 FIRS Meetings (Savannah), 2019 China International Conference in Finance (Guangzhou), 2018 FMA Annual Meetings (San Diego), 2017 SFS Cavalcade Asia Pacific (Beijing), 2017 Emerging Markets Conference (Mumbai), Yale School of Management, Georgia Tech (Scheller), HKUST, Northeastern University (D'Amore-McKim) and University of Oklahoma (Price). We thank Manpreet Singh for sharing the list of product dereservation dates. We gratefully acknowledge support from the Chazen Institute for Global Business at Columbia Business School. We are grateful to the London Business School's Research and Materials Development Grant and Wheeler Institute for Business and Development for supporting this research. Naaraayanan thanks Columbia University for hosting him for a part of the time during which this research was conducted. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Business Group Spillovers
S. Lakshmi Naaraayanan and Daniel Wolfenzon
NBER Working Paper No. 31107
April 2023
JEL No. G31,G32

ABSTRACT

We compare the investment of standalone firms across regions after a positive shock to the investment opportunities generated by a large-scale highway development project. We show that the standalones' investment sensitivity is lower in regions with a higher density of business groups in the local area. We investigate mechanisms driving our results and find support for a financing mechanism whereby banks allocate capital preferentially to group-affiliated firms in responding to the increase in credit demand. Overall, our study documents that business groups have spillover effects on standalone firms.

S. Lakshmi Naaraayanan
London Business School
Sussex Place
Regent's Park
United Kingdom
lnaaraayanan@london.edu

Daniel Wolfenzon
Graduate School of Business
Columbia University
Uris Hall, Room 808
3022 Broadway
New York, NY 10027
and NBER
dw2382@columbia.edu

I. Introduction

Business groups are present in many countries, and in some of them, they dominate the corporate landscape.[1] Concerns about the effect of business groups on the economy date back to at least the early twentieth century. Kandel, Kosenko, Morck, and Yafeh (2019) describe the presence and subsequent disappearance of business groups in the United States (henceforth, U.S.) between the 1920s and 1950s and document the government's concern that business groups wielded outsized market power and impeded competition. The U.S. is not the only case. Dau, Morck, and Yeung (2020) survey historical accounts for a broad cross-section of countries and argue that business groups' presence hampers economic growth.

More recently, in the 1990s, the South Korean conglomerates, or *chaebols*, were perceived to inhibit the growth of small- and medium-sized firms because, among other things, most of the finance available was directed to these business groups.[2] An even more recent example comes from Israel, where legislation was passed to limit the size and influence of business groups (Bebchuk, 2012; Haaretz, 2012; Times of Israel, 2013).[3]

The arguments advanced in the literature and policy debates suggest that business groups may impact the economy through their effect on non-affiliated firms. However, systematic empirical evidence on this issue is lacking. In this paper, we study this issue and examine whether business groups affect standalone firms' behavior, specifically, their investment.

Business groups can affect investments by standalone firms through a variety of channels. For example, if business groups have preferential access to bank capital, they can

[1] See Faccio and Lang (2002) and Claessens, Djankov, and Lang (2000) for evidence on the prevalence of business groups in Western European and East Asian countries, respectively.

[2] See, Financial Times (1998) for a discussion on these issues.

[3] Hamdani, Kosenko, and Yafeh (2020) describe similar attempts to dismantle business groups in other countries.

crowd out financing to standalone firms. Alternatively, if business groups are better at quickly seizing investment opportunities, they can increase industry capacity and reduce the attractiveness of further investments by standalone firms.

Because our focus is on studying the effect of business groups on standalone firms, the empirical methodology compares *standalone* firms in regions with high business group prevalence to *standalone* firms in regions with low business group prevalence. A key challenge in this comparison is the difficulty in controlling for investment opportunities. If these opportunities varied with business group prevalence, we would not be able to isolate the impact of business group prevalence from differences in investment opportunities (omitted variable bias). For this reason, instead of comparing the investment of standalone firms in regions with different levels of business group prevalence, our approach is to identify an investment opportunity shock that plausibly affects these regions equally and examine changes in investment by standalone firms.[4]

We focus on India, where there is significant variation in business group prevalence across regions and use a large-scale highway development project (called Golden Quadrilateral, henceforth GQ) as a shock to the investment opportunities for firms that lay along the road network. The project involved upgrades to the 5,800-kilometer highway system that connects the four major cities of Delhi, Mumbai, Chennai, and Kolkata, making it the fifth-longest highway in the world.[5] This program required significant investments by the government and represented an important infrastructure improvement for India.

The upgrade of the GQ road network led to improved inventory efficiency and input

[4]An alternative methodology would be to use exogenous variation in the level of business group prevalence. Indeed, such variation is possible to find. Larrain, Sertsios, and Urzúa I (2019) use industry shocks that lead to the breakup of business groups. Also, sometimes groups are partially dismantled due to a family feud. However, this type of variation cannot be used to address questions about the economy-wide role of business groups since it usually only affects a small set of firms and hence will have a negligible impact on the aggregate level of business group prevalence.

[5]In particular, it sought to upgrade highways to international standards of four or six-laned, dual-carriageway highways with grade separators and access roads. The road network connected as part of the GQ program represented 4% of India's highways in 2002, and the upgrade work raised this share to 12% by the end of 2006.

sourcing by manufacturing firms located along the GQ road network (Datta (2012)). In addition, the road upgrade also increased access to other regional markets for firms located along the network (Asturias, García-Santana, and Ramos (2016)). Indeed, we show that firms' total factor productivity increased after the road upgrade. We also show that there was a significant increase in the total investment even beyond infrastructure-related industries following the commencement of road upgrades. The magnitude of this increase is substantial, with the average firm increasing its investments by 6% of total assets. Taken together, the evidence suggests that road upgrade constituted a positive shock to investment opportunities for firms that lay along the road network.

Equipped with this shock, we turn to our main tests. We compare the investment by standalone firms around the shock as a function of the degree of business group prevalence in their local economies. We proxy for group prevalence by using the ratio of assets owned by all business group affiliated firms to assets owned by all firms in each region and split regions by business group shares. We find that the increase in investment is lower for standalone firms in high business group share regions compared to investments by standalone firms in low business group share regions.

This empirical strategy of comparing standalone firms across regions sidesteps issues related to the comparability of firms with different organizational structures. A vast literature in finance is interested in comparing standalone firms with business group firms (Faccio, Morck, and Yavuz, 2020; Faccio and O'Brien, 2020; Santioni, Schiantarelli, and Strahan, 2019). The concern with this approach is that characteristics that influence the choice of organizational form might also be correlated with the outcomes of interest. These papers mitigate this concern using propensity score matching and exogenously-failed control block transactions. Such an issue does not arise in our setting as we compare standalone firms in one region to another.

Admittedly, there are potential threats to our identification. First, standalone firms across these regions might be different. Second, beyond firm-level differences there could be regional differences across high and low business group share areas that might instead explain our findings. Lastly, the intensity of the investment opportunity shock itself might vary by region, which could drive the differential investment response by standalone firms across regions.

We address these concerns by comparing firm and regional characteristics, finding that they are similar across regions with varying levels of business group prevalence. To further alleviate concerns that regional differences other than business group prevalence might be operative, we show that our results are quantitatively similar in horse race regressions that control for important regional characteristics. To explore potential differences in the intensity of the shock, we gather survey evidence on the physical condition of the road, finding that ex-ante road quality was similar across regions with varying levels of business group prevalence. Further, we show that the average firm invests similarly in high and low business group share regions. Lastly, we show that stock price reactions to new plant announcements by standalone firms around the GQ upgrades are positive but similar across regions with varying levels of business group prevalence. Overall, these tests provide support for the identifying assumptions of our empirical strategy.

We next proceed to shed light on potential mechanisms driving the investment behavior of standalone firms. Our results on the average investment being similar, together with the baseline result on standalone investment, suggest a composition effect, with the investment of group-affiliated firms making up for the lower investment of standalone firms in high business group regions. We consider four mechanisms that can explain this pattern. First, business groups and standalone firms compete for factors of production, capital, and labor, with group-affiliated firms having an advantage over standalone firms in securing them. Indeed, we find evidence that banks preferentially allocate capital to business group firms. Using a novel hand-collected loan-level dataset, we show that banks with significant pre-existing lending relationships to business group firms reduce their supply of capital to standalone firms after controlling for any firm-level determinants of credit demand. This preference could be because it might be safer for banks to lend to group affiliates (Almeida, Kim, and Kim, 2015; Gopalan, Nanda, and Seru, 2007) or banks may find group affiliates more attractive than standalone firms as lending to one member in the group can generate demand for the banks' services from other members in the group. At the same time, we note that our data do not allow us to study changes in labor market conditions. Hence, it remains possible that the lower investment of standalone in high business group share regions is driven by their difficulty in finding labor required for the increased capital.

The second mechanism we consider is that group-affiliated firms in regions with greater business group prevalence crowd out demand for standalone firms' output in product markets. Investment by group-affiliated firms may increase industry capacity and reduce the attractiveness of further investment by standalone firms. To provide evidence for this mechanism, we focus on industries that do not rely on local demand, namely manufacturing and high-exporting industries, where crowding-out of demand is unlikely to be operative. We find that our main results hold in these subsamples and are quantitatively stronger. Together, these results suggest that crowding out of demand by group-affiliated firms is quantitatively unimportant, at least in these settings.

Third, another mechanism that could be operative is that group-affiliated firms can seize investment opportunities faster than standalone firms and that these opportunities have a "winner-takes-all" aspect to them. In such cases, investments by group-affiliated firms can crowd out investments of standalone firms in high business group share regions. To provide evidence on this mechanism, we utilize a regulation whereby manufacturing of certain products was reserved for production by smaller firms, likely standalone firms, and hence crowding out by business group affiliates was unlikely in these products. In our estimations, we find that the baseline effect is present even in these industries where business group firms are less likely to be allowed to invest. While we cannot completely rule out that in every industry group-affiliates crowd out investment opportunities of standalone firms, the fact that in settings with reduced scope for such crowding out, the results are equal or stronger than the baseline, suggests that crowding out of demand by group-affiliated firms is quantitatively unimportant.

Lastly, we consider a fourth mechanism whereby business groups have better political connections than standalone firms and hence can obtain preferential access to government contracts. Focusing on the subsample of infrastructure-related industries, we find suggestive evidence that the presence of politically connected business groups depresses standalone investments. However, we cannot establish whether political connections operate in other industries.

We perform several robustness tests to show that our baseline result of lower investment by standalone firms in high business group share regions holds up under different specifications. First, we show robustness to alternative definitions of our business group share measure. Second, we address the concern that firm exits might drive lower investment. While we do not observe exits in our sample, we use proxies for exit and show that dropping firms that stop reporting financial data and those with extremely large negative sales growth do not alter our baseline results. Finally, we also show that changes in the regional composition of firms, either through firm entry or through mergers and acquisitions, do not explain our findings.

Our paper is relevant to the large literature on business groups and conglomerates on several fronts. First, prior literature has often focused on examining conglomerates in isolation (Hoshi, Kashyap, and Scharfstein (1991); Gopalan, Nanda, and Seru (2007); Almeida, Kim, and Kim (2015)). These studies provide convincing evidence of the functioning and efficiency of internal capital markets. However, they are silent about the effect of business groups on standalone firms. Our study sheds light on a spillover in the economy wherein business groups inhibit the growth of standalone firms by reducing their investment.

Second, we provide empirical support for the prediction from Almeida and Wolfenzon (2006) that the presence of business groups reduces the supply of capital to standalone firms. However, the policy implications of our study are unclear. It could very well be that standalone firms are inefficient, and thereby any reallocation away from them may improve the capital allocation in the overall economy. Given this, more research is needed to ascertain whether business group prevalence can improve or hurt capital allocation in the economy.

In general, better knowledge of the mechanisms will help determine whether the aggregate effect of these spillovers is positive or negative. For example, if business groups reduce investments by standalone firms because productive group affiliates can seize investment opportunities faster, then the policy implication would be to promote business group presence. While our paper takes a modest step towards documenting specific mechanisms at play, further research can improve our understanding of the aggregate effects of the spillovers we document.

Lastly, our paper is also related to the literature examining the impact of highway infrastructure on local economic activity. Chandra and Thompson (2000) study the impact of U.S interstate highways and show that they have a differential impact on non-metropolitan areas across industries and affect the spatial allocation of economic activity. In the context of GQ, Datta (2012) find that firms in cities that lay along the routes of the upgrade benefited significantly from the improved highways. They find that firms had increased inventory efficiency due to lower transportation obstacles to production and access to efficient suppliers. Ghani, Goswami, and Kerr (2014) show that the upgraded GQ network substantially impacted the growth of manufacturing activity. While our findings are highly complementary, our goal is not to study the effect of GQ per se but use it as a shock to investment opportunity and examine the investment behavior of standalone firms as a function of the prevalence of business group firms.

II. Background about the Golden Quadrilateral

India has the second largest road network in the world.[6] National highways are critical to this road network and play a significant role in regional trade while carrying nearly half of the total road traffic volume. At the end of the 1990s, India's highway network was in a state of disarray marked by poor connectivity, sub-par road conditions, and congestion with limited lane capacity. Poor road surface conditions, frequent stops at state borders for tax collection, and increased demand from growing traffic all contributed to congestion with 25% of roads categorized as congested (World-Bank (2002)).

To tackle these issues, the Government of India (GoI) launched the National Highways Development Project (NHDP) in 1998 intending to improve the performance of the highway network. We study the upgrade of the 5,800-kilometer highway system called the Golden Quadrilateral (GQ) which connects the four major cities of Delhi, Mumbai, Chennai, and Kolkata, making it the fifth-longest highway in the world.[7] The project was initially approved in 1998, but many segments of the project started only as late as 2001. These delays in the start of construction led to differences in completion.[8] The

[6] It consists of expressways, national highways, state highways, major district, and rural roads. Taken together, these roads carry close to 65 percent of freight in terms of weight.

[7] The GQ work involved upgrading highways to international standards by incorporating features of high-quality highway systems such as expanded lane capacities, dual-carriageway highways with grade separators, over-bridges, by-passes, and access roads. This upgrade raised the share of highways to 12% of the road network by the end of 2006. In comparison, highways constitute about 5% of the road network in developed economies such as US and Japan and 13% in the United Kingdom (World-Road-Statistics (2009))

[8] The junior Highways Minister, Tushar Chaudhary told the Parliament that "Projects have been delayed mainly due to problems associated with land acquisition, shifting of utilities, obtaining environment and forest clearance, approval for a road over bridges, poor performance of some contractors due to cash flow constraints and law and order problems in some states."

construction was complete for a significant portion of the segments by the end of 2006, but minor work on additional phases of the project continued even as late as 2009.[9]

To complete the GQ upgrades, 128 separate contracts were awarded. Most of the construction involved public-private partnerships and cost was to be recovered by levying a cess of INR 1 on petrol and diesel. A significant portion of the funding came from the federal government while the remainder from multilateral financing agencies such as the Asian Development Bank (ADB), and the World Bank (WB).[10] Therefore, road construction by itself did not impose constraints on the banking system.

Figure I illustrates the *time* variation in the start year of the construction along the four major segments of the GQ road network, with the height of the bar corresponding to the number of sub-segments. As the figure shows, the bulk of construction is concentrated between 2000 and 2006, with no major differences in the timing of construction among the four segments. Figure II shows the *geographical* variation of various segments over the start year of construction. Panels (a) to (d) of the figure display the evolution of the road network over time. We see a significant increase in construction over this period. The median completion time across the 128 contracts was 2.3 years, and the median completed road length was 50 km. As we explain below, this variation in the commencement of road construction for cities located on the GQ road network is central to our empirical strategy.

The most direct benefit from upgraded connectivity is a significant reduction in transportation costs and improved market access for firms to other regional markets (Asturias, García-Santana, and Ramos (2016)).[11] Datta (2012) find that immediately after the upgrades *commenced*, there are improved inventory efficiency and input sourcing by manu-

[9] A significant portion of construction began in 2001, with a target completion date of 2004.

[10] The The federal government contributed about 60% of the financing, while the multilateral agencies contributed 20% and rest was raised through a variety of new public-private initiatives such as Build-Operate-Transfer (BOT), and equity sharing concessionaire agreements. For financing, the federal government created the Central Road Fund through the Central Road Fund Ordinance, 2000 in November 2000. The revenue accrued through levies would form part of the fund, which was used to finance the upgrade of highways.

[11] For other work related to market access, see Alder (2014)

facturing firms located along the GQ road network.[12] Ghani, Goswami, and Kerr (2014) show significant output growth, and entry in industries initially positioned along the GQ network. We also confirm improvements in inventory efficiency and total factor productivity for firms along the road network after the upgrade.

In addition to these studies, the benefit in terms of lower transport cost and ability to access new markets is also highlighted in a World Bank report (World-Bank (2000)):

> *The primary benefit of the project is a reduction in transport costs resulting from increased capacity, reduced bottlenecks, separation of local and through traffic in towns and improved pavements. This is directly linked to costs of goods and services, fares, ability to market local products and regional economic development.*

The popular press also commented on these benefits for firms located along the GQ network (Business-Today (2013)):

> *"'We have been able to serve customers faster than before,' he says. 'This has resulted in a higher number of repeat orders and our entry into newer markets such as Chennai and Bangalore.' "*

Thus, the upgrade of GQ road network is plausibly a positive investment opportunity shock that improved market access and reduced transportation costs for firms that lay along it.

III. Data and summary statistics

III.A. Data

GQ construction. We compile information on each of the 128 contracts from the annual reports of National Highway Authority of India (NHAI) from 1998-99 to 2013-2014 as well as from the Ministry of Roads, Transport and Highways (MORTH). These annual reports identified the project name for the highway stretch, the length of the highway

[12] For evidence on significant long-term economic benefits, see Khanna (2014)

stretch, the national highway number, the start date for the project, cost of the stretch, and financiers of the stretch.

In most cases, the name of the project indicated the start and end cities on a highway stretch along with highway number. In some cases, the project name was not clear or the city name could not be located. In such cases, we use information on the NHAI website for the highway project chainage and mapped to the preceding or succeeding highway stretch.

Firm financials. Our main data source is Prowess, a database maintained by CMIE. This dataset has been used by a number of prior studies on Indian firms, including Bertrand, Mehta, and Mullainathan (2002), Gopalan, Nanda, and Seru (2007), Lilienfeld-Toal, Mookherjee, and Visaria (2012), Naaraayanan and Nielsen (2021), and Gopalan, Mukherjee, and Singh (2016). Prowess contains annual financial data sourced from balance sheets and income statements for about 34,000 publicly listed and private Indian firms.

The data is of panel nature covering about 2000 to 6000 firms every year with assets plus sales of over INR 40 million. It contains additional descriptive information on the headquarter location, industry classification, the year of incorporation, and group affiliation. We adopt Prowess' group classification to identify whether a firm is affiliated to a business group or not.[13] This group affiliation has been used most notably in Khanna and Palepu (2000) and Bertrand, Mehta, and Mullainathan (2002). We extract data from the latest vintage of Prowess which is free from survivorship bias as highlighted by Siegel and Choudhury (2012).

We also take advantage of the granularity of the sales variable. Specifically, Prowess reports revenues and quantity for manufacturing firms at the product level, allowing us to

[13]According to Gopalan, Nanda, and Seru (2007), Prowess' broad-based classification is more representative of group affiliation than a narrow equity-based classification. We note that the very few firms change their group affiliation in the data.

perform sharper tests of our mechanisms and specifically to rule out concerns regarding rival investment opportunities. While the data is rich and provides physical quantities, we do not use it in the paper as they vary substantially in terms of the unit of measurement (e.g., weight, numbers, volume, etc.) within and across firms, thus making comparisons harder. These data are available due to the disclosure requirements imposed by the Companies Act 1956 and, thereafter, the Companies Act 2013.

New plant announcements. Data on new plant announcements is from the CapEx database maintained by CMIE. This dataset contains information on new plants announced in India since 1990. Specifically, it provides information on date of announcement, plant location, ownership, project cost, and industry classification. The information is obtained from multiple sources including annual reports, news articles and government press releases. The database is updated on a daily frequency and contains information on the entire project lifecycle whenever information is available. Typically, projects costing more than INR 100 million (approximately USD 2 Mmillion) are included in the database (Alok and Ayyagari, 2019).

Regional banking data. To study the response of aggregate regional lending to GQ upgrades, we use district-level data from the Reserve bank of India (RBI). Part of the data is downloaded from the RBI's data warehouse webpage, and the other part is hand-collected from the publication "Banking Statistics 1990 - 2016" available for download as a pdf file on the RBI webpage. Data on bank credit is drawn from the "Database on the Indian Economy" published by the RBI. A data series called "Quarterly Statistics on Deposits & Credit of Scheduled Commercial Banks" provides information on outstanding credit and the number of branches of all banks at various regional-levels (district-level and urban vs. rural levels). The data are reported on a quarterly basis.

Loan-level dataset. To study banks' credit supply decisions, we introduce a novel loan-level dataset that allows us to control for firm-level determinants of credit demand, a là

Khwaja and Mian (2008). To do so, we rely on the credit registry at the Ministry of Corporate Affairs (MCA), GoI. The MCA mandates registration of all secured lending as a condition for lenders to invoke their creditor rights. We scrape the data, which contains the firm's name, the name of the lender, the origination amount, the date of initiation of the loan, and when the lending relationship ends. To link this dataset to our baseline sample, we perform a time-intensive name matching exercise described in Appendix A. Appendix Table I provides the descriptive statistics for this merged sample.

World Bank Enterprise Survey. We gather regional information on road conditions for firms that lay on the GQ road network. We rely on the survey conducted by the WB just before the GQ upgrade, also used in Datta (2012). It focuses on a random sample of firms in the formal sector, stratified by industry, firm size, and location, that is representative of the non-agricultural economy (World-Bank, 2009).

Other datasets. We supplement these datasets with information on city and district population from the Population Census of 2001. We collect information on business registration and financial disclosure reporting from the MCA.

III.B. Final sample and summary statistics

From the overall Prowess sample of 1989 to 2016, we exclude all financial firms (NIC code: 641-663), firms owned by central and state governments, firms with negative values of total assets and sales, firms with leverage outside the [0,1] range, and observations with ratio of investment to lagged total assets greater than 1.[14] In addition, we exclude all firms operating in "other manufacturing industries" (NIC code: 321-329), "coke and refined petroleum products" (NIC code: 191-199), and "construction firms" (NIC code: 420-439). We do so to isolate the effect of GQ upgrades on firms that benefit from

[14] Firms with leverage greater than 1 were considered to be bankrupt in India until 2016.

market access as opposed to the actual road construction.[15] We exclude firms with sales growth exceeding 100% to avoid potential business discontinuities caused by mergers and acquisitions. Given that accounting data of very small firms are likely to be noisy, we exclude firms with capital, book assets, and sales with less than INR 2.5 million (around US$ 0.03 million) in the previous year.[16]

Table I reports descriptive statistics for firms in our sample. Panel A presents descriptive statistics for all firms, while panel B presents descriptive statistics for standalone firms in our sample. On average, standalone firms are smaller in size and younger than the average firm. However, standalone firms are similar in terms of other firm characteristics such as cash flow, profitability, investment, and debt.

IV. Empirical strategy

We study the investment behavior of standalone firms as a function of the prevalence of business groups in their local area. A key challenge in isolating the effect of regional variation in business group prevalence on standalone investment is to adequately control for all other determinants of investment. Specifically, these determinants can vary systematically across regions with different levels of business group prevalence. For this reason, instead of comparing the investment of standalone firms in regions with different levels of business group prevalence, our approach is to identify an investment opportunity shock that plausibly affects these regions equally and examine changes in investment by standalone firms around such a shock.

An alternative approach to studying this question would have been to exploit exogenous variation in business group prevalence. Indeed, such variation is possible to find.

[15] While we omit these industries in our main tests, we focus on infrastructure-related industries to tease out the mechanism in Section VI.D.

[16] Note that we include all available years with financial data for each firm. In our sample, on average, there are 7 observations per firm.

Larrain, Sertsios, and Urzúa I (2019) use industry shocks that lead to the breakup of business groups. Also, sometimes groups are partially dismantled due to family feuds. However, this type of variation cannot be used to address questions about the economy-wide role of business groups since it usually only affects a small set of firms and hence will have a negligible impact on the aggregate business group prevalence.

IV.A. Investment opportunity shock: GQ upgrades

We begin by presenting evidence that GQ road network upgrade is a plausibly exogenous shock to the investment opportunity of the firms located along the network. We estimate parametric models using *all firms* along the GQ road network and estimate the following equation:

$$\mathrm{y}_{ijcst} = \alpha_i + \beta \mathrm{PostGQ}_{ct} + \omega_{jt} + \theta_{st} + \eta_{ijcst} \tag{1}$$

where subscripts i and t refer to firm and year, respectively, and $PostGQ_{ct}$ is an indicator variable taking value 1 for all years including and after the commencement of GQ upgrade in the city (Datta, 2012). The subscripts j, c, and s refer to industry, city, and state, respectively. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level (Bertrand, Duflo, and Mullainathan, 2004). The parameter of interest is β, which measures the change in the outcome variables of firms in the cities that receive GQ compared to the yet-to-receive cities, conditional on the set of fixed effects.

We identify the firms that benefit from the commencement of the GQ upgrade based on the city of their headquarters and match these cities to highway project stretches. We code firms at both ends of a stretch to be treated at the start of the construction. A typical highway stretch in our sample connects two cities that are around 50 kilometers apart.[17]

[17] Such short distances imply that there are no major cities in the middle of the upgraded segments

After applying the sample selection criteria, which drops smaller firms, we are left with 110 stretches (out of the possible 128 stretches) connecting 44 cities with some economic activity. A few cities connect two or more stretches of the highway and in such instances we assign them the earliest start date among all stretches for such cities. Additionally, we treat the adjacent suburbs (Gurgaon, Faridabad, Ghaziabad, and NOIDA for Delhi; Thane for Mumbai) as part of the nodal city (Datta, 2012). Overall, the sample of firms in treated regions comprised 72% of total sales in India before the GQ upgrade.

To alleviate concerns that firms along the road network differ from firms located away, we restrict the sample to firms that eventually receive an upgraded road. Effectively, we employ a specification similar in spirit to a difference-in-differences strategy but exploit only the variation in the timing of construction of highway segments. Thus, at any point in time, the treated firms are those in cities that receive GQ, and control firms are those in cities that are yet to receive GQ.[18]

The empirical specification allows us to rule out concerns about location-specific and industry-specific effects that may differentially affect firms' investment policies. First, we include firm fixed effects to control for unobserved time-invariant firm characteristics. Second, we include industry-by-year fixed effects that control for time-varying industry shocks (e.g., technical innovation). Lastly, since the treatment varies within states, we also include state-by-year fixed effects to control for local economic confounds and general policies that affect firms (e.g., regional macroeconomic shocks).

Table II reports the results examining changes in total factor productivity (TFP), inventory efficiency, and investment around the commencement of GQ road upgrades. In column 1, we begin by examining the effect of GQ upgrades on productivity, which

and hence meaningful economic activity.

[18] We show robustness to recent concerns raised in the literature on staggered difference-in-differences designs by using the imputation estimator suggested in Borusyak, Jaravel, and Spiess (2022). In order to implement these estimations, we expanded our sample to include firms that did not lay along the GQ road network and hence were never treated.

we measure using the methodology outlined by Levinsohn and Petrin (2003). We find a significant increase in TFP among firms after the GQ upgrade, consistent with the idea that GQ road upgrades significantly raised the marginal product of capital and labor. Further, in column 2, we examine changes to inventory efficiency, which we measure using days sales of inventory, and find that firms that lay along the GQ road network experience a significant reduction around upgrades. This reduction is consistent with anecdotal accounts presented in Section II and Datta (2012) documenting improvements in input sourcing and inventory efficiency.

Lastly, in column 3, we show that there is an overall increase in investment among firms after the upgrade of the road network. The magnitude of this increase is substantial. Relative to the sample average, this represents an increase of 6%.

The interpretation of β in Equation 1 as the impact of GQ upgrades requires the assumption that the timing of the road construction is orthogonal to the investment opportunities of the firms that lay along the road network. In Appendix B, we confirm the validity of the identifying assumption by (i) ruling out pre-trends in investments, and (ii) showing that observable differences in firm characteristics cannot explain the timing of GQ road upgrade. As a recent survey by Roth and Sant'Anna (2020) points, such quasi-randomness supports the validity of parallel trends assumption. Further, we rule out concerns about staggered treatment adoption and heterogeneous causal effects using the imputation estimator suggested in Borusyak, Jaravel, and Spiess (2022).

Overall, we establish that GQ upgrades led to significant improvements in firm productivity and increased capital expenditures. Thus, the evidence suggests that we have a plausible shock to the investment opportunity for firms on the GQ road network.

IV.B. Standalone investment as a function of business group share

Equipped with the investment opportunity shock, we turn to our main tests, where we examine the extent to which the sensitivity of standalone firm investment to the GQ shock varies with regional business group prevalence. For each city, we compute the share of group-affiliated firms' assets in the year before the GQ road network upgrade. We then define *High BGS* as an indicator variable set to 1 if the share of assets is in the top quartile of the distribution.[19] Thus, this measure captures the prevalence of group-affiliated firms at each location.

We estimate the following equation:

$$\begin{aligned} \mathrm{y}_{ijcst} = \alpha_i + \beta_1 \mathrm{PostGQ}_{ct} + \beta_2 \mathrm{PostGQ}_{ct} \times \text{High BGS } + \omega_{jt} + \mu_{jt} \times \text{High BGS}, \\ + \theta_{st} + \gamma_{st} \times \text{High BGS} + \epsilon_{ijcst} \quad (2) \end{aligned}$$

For these tests, we restrict our attention to standalone firms because the paper aims to examine the association between regional prevalence of business groups and standalone investment.[20] For this reason, the coefficient of interest in Equation 2 is β_2, which measures the change in standalone investment around GQ road upgrades in high-business group share regions relative to low-business group share regions, conditional on the set of fixed effects. Put differently, β_2 captures the extent to which the sensitivity of standalone firm investment to the GQ shock varies across high business group share regions relative to low business group share regions.

[19] Appendix Figure II plots the distribution, specifically the mean and the standard deviation, of group-affiliated firms' asset share in each quartile, supporting the choice of the top quartile as the cutoff for the baseline measure. We also show that the results are robust to alternative cutoffs and using alternative definitions.

[20] In our sample, we have 4,177(1,856) unique firms (standalone firms) in *High BGS* regions and 805(394) unique firms (standalone firms) in *Low BGS* regions.

As our empirical specification focuses on standalone firms, the endogenous choice of organizational form does not bias the estimates, which is a significant hurdle for papers that compare standalone firms to group-affiliated firms. For example, one could worry that the investment opportunities vary by organizational form, e.g., perhaps group-affiliated firms have better networks and access to key market players allowing them to better capitalize on these opportunities. Given that we compare standalone firms across regions, our setting is devoid of such issues.

Furthermore, our empirical specification flexibly controls for the differential regional impact of industry-year and state-year specific effects across high and low business group share areas by interacting *High BGS* with the set of interactive fixed effects.[21] For example, the interaction of industry-year fixed effects with *High BGS* allows for the effect of technological shocks to vary for firms in high and low business group share regions.

V. Main results and identification challenges

V.A. Main results

Table III reports the main results of this study. The increase in investments is lower for standalone firms in high business group share regions relative to low business group share regions, with this difference being statistically significant. In terms of magnitude, we find that standalone investment around GQ upgrade for firms in the first three quartiles of business group share increase by 0.039 (coefficient on *PostGQ* in column 2), albeit statistically insignificant, while the standalone investment change for firms in the top quartile (coefficient on $PostGQ + PostGQ \times High\ BGS$ in column 2) is essentially zero. This effect on investment is economically sizeable and represents a 10% decrease in

[21] We do not include time-varying control variables such as cash flow, profitability, etc. in our empirical specification as they themselves may be affected by the treatment, rendering them "bad controls" (Angrist and Pischke, 2008).

investment relative to the average (panel B of Table I).

In Figure III, we plot the evolution of investment sensitivity in event-time for standalone firms in high business group share regions relative to low business group share regions. We note that all coefficients in the pre-period are close to zero, thereby supporting the assumption that there are no *differential* pre-trends in standalone investments across high and low business group share regions around the GQ upgrade. At the same time, the coefficient estimates including and after year 2 are large, negative, statistically significant, and persistent until ten years after the highway upgrade. In Appendix C, we document that these estimates are robust to using the imputation estimator suggested in Borusyak, Jaravel, and Spiess (2022), allowing us to rule out concerns about two-way fixed effects estimator providing biased estimates in cases when the treatment is staggered and in the presence of treatment effect heterogeneity (Goodman-Bacon, 2021; Roth, Sant'Anna, Bilinski, and Poe, 2022).

Collectively, the results suggest a lower sensitivity of standalone investment to the investment opportunity shock as a function of the regional prevalence of business groups.

V.B. Identification challenges

Admittedly, there are potential threats to our identification. First, differences in standalone firm characteristics across these regions might explain the differential investment response. Second, as business group shares are not randomly assigned, one may worry that regional differences might instead explain our findings. Lastly, the intensity of the shock itself might vary by region, thus leading to differences in investments. We address these concerns below in detail.

V.B.1. Do firm and regional differences drive the findings?

We begin by comparing *firm-level* characteristics of standalone firms across high and low business group share regions. Panel A of Table IV presents the results. We find no differences in the mean and median for profitability, investment, debt, cash flow, and assets. Importantly, since one of the main documented benefits of the road upgrade is to improve inventory efficiency, we also test whether firms are similar in this regard before the shock.[22] We find that this is the case. At the same time, we find a significant difference in the average (but not in median) TFP among standalone firms across regions, with standalone firms in *High BGS* regions having higher productivity compared to standalone firms in *Low BGS* regions. If anything, this difference would bias against finding a lower investment sensitivity in high business group share regions. Nonetheless, we address the general concern of differences in productivity more directly through horse race regressions, as explained below.

Another concern might be that regional differences instead explain our findings. To mitigate such a concern, we compare ex-ante *regional* characteristics across high and low business group shares in panel B of Table IV. We begin by comparing financial development across regions and find no statistical differences in the number of bank branches per capita or the fraction of listed firms. We also compare the physical infrastructure and find no difference in the rating that firm managers assign to roads or in the frequency with which they consider transportation an obstacle to growth. Finally, we compare labor market conditions. Before the GQ upgrade, a similar fraction of managers report facing constraints in hiring labor on a contractual basis, and there are no differences in the frequency of managers reporting labor as an obstacle to firm growth. Further, we do not find differences in the time to fill vacancies for managers at these firms.

[22]Inventory efficiency is measured by days sales of inventory. Fewer days sales of inventory suggest firms quickly convert inputs to sales. Such conversion can come about either due to a reduction in costs, switching to efficient suppliers, or access to new markets.

An alternative approach to mitigate concerns, that regional differences other than business group prevalence might operate, is to run horse race regressions that control for important regional characteristics that may instead explain our findings. In addition to *PostGQ* × *High BGS*, the empirical specification includes interactions between *PostGQ* and *High Listed Share*, *High Firm Age*, and *High TFP*. We use the same procedure as *High BGS* to construct these variables and define these characteristics using all firms and only standalone firms in panels A and B of Table V, respectively. Defining these characteristics using all firms (business group affiliates and standalones) accounts for the possibility that business group share proxies for other regional characteristics, and the definition based on standalone firms account for the possibility that standalone firms in high and low business group share regions are different. As in our main results (Table III), we focus on standalone firms and their investment in these regressions.

Controlling for these additional interactions does not affect the statistical and economic significance of *PostGQ* × *High BGS*. Specifically, interactions of *PostGQ* with *High Firm Age* and *High TFP* are statistically insignificant across both definitions and do not qualitatively affect the coefficient on *PostGQ* × *High BGS*.[23] Importantly, even though there was a significant difference in total factor productivity of standalone firms one year before GQ (Table IV), we note that controlling for interactions with TFP does not alter our baseline findings.

The coefficient on *PostGQ* × *High Listed Share* is negative and statistically significant in some specifications. Even in these cases, the coefficient on *PostGQ* × *High BGS* is essentially unchanged and remains statistically significant. The result that the inclusion of *PostGQ* × *High Listed Share* does not affect the significance and magnitude of our coefficient of interest is not surprising given that there is a very small correlation between business group share and the fraction of listed firms (ρ = -0.09), suggesting that they are

[23]In Appendix Table IV, we show that the results are qualitatively similar if we use firm size instead of firm age in these regressions.

orthogonal to each other.

Overall, these tests mitigate concerns that differences in either firm or regional characteristics drive the lower sensitivity of standalone investment to the investment opportunity shock.

V.B.2. Does the shock intensity vary by business group share?

A second potential threat to identification relates to the differential intensity of the shock across regions. For example, if the ex-ante road quality is poorer in some of these regions, then the GQ upgrade would constitute a larger shock to investment opportunities for the firms in those regions.

To rule out such a possibility, we gather evidence from the World Bank Enterprise Survey (WBES) on the quality of roads, which we show is similar ex-ante across high and low business group share in panel B of Table IV. Further, we read project documents from the NHAI and World Bank and found that the ex-post road quality, such as the number of lanes, road strength, and materials used, are similar for all stretches of the GQ upgrade.

Another implication of the shock intensity being equal across regions is that the average firm invests the same around GQ upgrades. To examine this, we estimate Equation 2, our primary test, on the entire sample of firms, reported in column 1 of Table VI, finding this is indeed the case. In column 2, for comparison, we show our baseline result that standalone firms invest *less* in high business group share regions. Moreover, in column 3, we focus on the sample of group-affiliated firms, finding that the average group-affiliated firm invests *more* in regions with a high business group share. This effect is economically small and statistically significant only at the 10% level.

Additionally, we use stock price reactions to new plant announcements by standalone firms to assess investors' views on the value of new investment around GQ upgrade. We obtain data on new plant announcement dates, their location, and the capital invested from the CapEx database maintained by CMIE. The database typically captures large plants, approximately USD 2 million. In panel A of Appendix Table V, we show that the stock price reactions to new plant announcements initiated after the GQ upgrade are positive and larger than before the GQ upgrade. This increase in value is consistent with improved investment opportunities for firms after the GQ upgrade. Importantly, this positive reaction is similar for standalone firms across high and low business group share regions. One interpretation of this evidence is that, while investors perceive the investment opportunities around GQ upgrades to be positive, they do not perceive them to vary by business group prevalence, suggestive of similar investment opportunities. Admittedly, the stock price reaction is also a function of the investment size. For this reason, we control for investment size in panel B and obtain essentially the same results.[24]

Further, as discussed in Section II, an important component of the investment opportunity shock is improved inventory efficiency and input sourcing (Datta (2012); Hesse and Rodrigue (2004); Li and Li (2013); Redding and Turner (2015); Shirley and Winston (2004)). In Appendix Table VI, we show that this is indeed the case finding that there are significant improvements to inventory efficiency. Importantly, we show that the improvement is similar for standalone firms across regions with varying levels of business group share. Admittedly, there could be other determinants of inventory efficiency beyond road quality. For this reason, we take these results as suggestive evidence for a similar investment opportunity shock.

Together, our evidence suggests that the intensity of the shock was similar across regions with varying levels of business group share, thus, mitigating concerns about such differences driving the lower sensitivity of standalone investment to GQ upgrades.

[24]The number of observations differs between the panels due to missing information on project costs. Note that we refrain from including project size directly in the regressions as a control variable as they are a function of the investment opportunity shock, rendering them "bad controls".

VI. Exploring potential mechanisms

So far, we have shown that standalone firms have lower investments in regions with a high business group share. In addition, we showed in Table VI, the average investment by firms in high business group share regions is similar to the average investment by firms in low business group share regions. These results suggest a composition effect, with group-affiliated firms making up for the lower investment of standalone firms in high business group regions.

One possible explanation for this pattern is that business groups and standalone firms compete for factors of production, with group-affiliated firms having an advantage over standalone firms. Another explanation for the lower investment by standalone firms in high business group share regions is that business group firms in these regions crowd-out demand for standalone firms' output in product markets. Moreover, it could be that group-affiliated firms are more adept at seizing investment opportunities sooner and that these opportunities have a "winner-takes-all" aspect to them. In such cases, group-affiliated firms can crowd out investments of standalone firms in high business group share regions. Lastly, the lower investment can instead be driven by business groups with better political connections than standalone firms, allowing them to obtain preferential access to government contracts. In the subsequent sections, we test these mechanisms and present the plausible assumptions and caveats for each of these mechanisms to be operative. Therefore, the evidence presented in these tests is suggestive and not conclusive.

VI.A. Factors of production

In this section, we focus on the allocation of bank capital as a possible driver of the lower investment sensitivity. First, we document an equal increase in bank lending around the GQ upgrade across high and low business group share regions. Second, we show that banks with significant lending exposure to group affiliates reduce lending to standalone firms, a result that holds after controlling for firm-level determinants of credit demand. Together, these results suggest that banks have a preference for lending to group affiliates at the expense of standalone firms.

We begin by examining whether banks directed lending differentially across high and low business group share regions around the GQ upgrade. To do so, we compile data from the RBI and compare aggregate bank lending across regions with varying business group prevalence.[25] We present results in Table VII. From columns 1 and 2, it is evident that around the GQ upgrades, districts along the GQ road network experienced an increase in overall bank lending. That is, banks respond to an increased demand for funds by allocating capital to regions that experience increased investment opportunities.[26] Importantly, columns 3 and 4 show no differential bank lending patterns across high and low business group share regions. These results show that while there is an increase in credit supply, such an increase is similar for regions with varying levels of business group prevalence.

Next, we test whether banks direct their scarce funds toward business group affiliates in response to the increase in demand for financing, thereby crowding out lending to standalone firms. A simple comparison of bank lending to group affiliates and standalone firms is insufficient to establish banks' preferences because it might be that these organizational forms differ in their demand for credit. For this reason, we control for firm-level determinants of credit demand using an empirical specification similar in spirit to Khwaja and Mian (2008). The idea of the test is to assess whether the *same* standalone firm borrowing from two different banks – one with significant lending to exposure

[25] Note that we focus on districts instead of cities for these tests, as the RBI only provides aggregated data at the district level.

[26] Our results are consistent with recent work showing that financing responds to large infrastructure investments and helps spur real economic outcomes (Agarwal, Mukherjee, and Naaraayanan, 2022; Das, Ghani, Grover, Kerr, and Nanda, 2019).

to group affiliates and the other without – borrows less from the more exposed bank after the shock. Specifically, we estimate the following loan-level equation using time-collapsed loans to standalone firms:

$$\Delta L_{ib} = \alpha_i + \beta_1 \text{Group exposure}_b + \eta_{ib} \tag{3}$$

where i stands for firm and b stands for bank. ΔL_{ib} is the *change* in the average loan amount to standalone firms five years after the GQ upgrade relative to five years before the GQ upgrade. The variable, *Group exposure*, is defined for each bank as the total lending to group affiliates before the start of the GQ upgrade. Importantly, the empirical specification includes firm fixed effects that control for firm-level determinants of credit demand. Further, group exposure and changes in lending might be correlated for each bank, and hence we conservatively cluster standard errors at the bank level (Khwaja and Mian, 2008).

As Prowess does not contain data on loan amounts lent to firms by banks and financial institutions, we resort to hand-collecting a novel loan-level dataset from the credit registry at the MCA, to implement this specification. The MCA mandates registration of all secured lending as a condition for lenders to invoke their creditor rights. The name match between Prowess and the loan-level dataset yields 2,430 loans to 302 unique firms from 140 lenders. Further, the empirical specification imposes a stringent requirement that the standalone firms borrow from multiple banks both before and after the GQ upgrade. Such a restriction leads to the final sample consisting of 163 loans to 17 unique firms from 15 unique lenders.

As we rely on firms borrowing from multiple lenders, we find that relative to the sample of firms in our baseline regressions, the firms in the loan-level dataset are larger, older, have higher cash flow and profitability, more likely to be listed, higher investment rates, and borrow more (Appendix Table I). We also note that there is substantial variation in the *Group exposure* amount as suggested by the coefficient of variation of 1.15.

Table VIII presents estimates from the loan-level regressions as laid out in Equation 3. Column 1 finds that banks with significant lending exposure to business group firms before the GQ upgrade reduced their lending supply to standalone firms after the investment opportunity shock. In column 2, we weight the regressions by the average firm size, measured before the GQ upgrade, to ensure that firm-level determinants correlated with firm size do not drive our estimates. In terms of economic magnitude, a one standard deviation above the mean increase in *Group exposure* amount leads to a 18.5 (19.1) percentage point reduction in the loan growth rate around GQ upgrade. Together, these results support the view that banks with significant pre-existing lending relationships with group-affiliated firms reduce their capital supply to standalone firms.[27]

Note that the empirical specification only uses information on location of the standalone firms to determine the timing of loan issuance relative to timing of GQ upgrades and not for classifying the location based on business group prevalence. As such, this test suggests that the average standalone firm, regardless of its location in a high or low business group share region, experienced a reduction in the supply of credit from exposed banks. Therefore, this specification alleviates concerns about unobserved characteristics of *High BGS* regions driving standalone firms' lower investment sensitivity. At the same time, by controlling for firm-level determinants of credit demand, these results are also not subject to the concern that the intensity of the investment opportunity shock differs across regions.

Our tests show that banks reduce the supply of credit to standalone firms. There are several reasons why banks preferentially allocate capital to group affiliates. One reason is that it might be safer to lend to group affiliates. As Gopalan, Nanda, and

[27]Note that the loan-level regressions can only speak to the mechanism for the matched sample. However, as this sample consists of large firm-bank pairs, we think that the reduction in the supply of capital is likely to affect smaller standalone firms even more which are lower in the pecking order for lending and not part of this matched sample.

Seru (2007) show, business group affiliates financially support member firms in financial distress with intra-group loans. Similarly, Almeida, Kim, and Kim (2015) show that group affiliates support member firms with positive investment opportunities through cross-equity investments. Moreover, banks may find group affiliates more attractive than standalone firms as lending to one member in the group can generate demand for the banks' services from other members in the group.

Lastly, we note that our data do not allow us to study changes in labor market conditions. Hence, it remains possible that the difficulty in finding the labor required for the increased capital drives the lower investment sensitivity of standalone firms.

VI.B. Product markets

Another potential explanation for the lower investment sensitivity is that group affiliates crowd out demand for standalone firms' output in product markets. As such, investments by group affiliates increase industry capacity and may reduce the attractiveness of further investment by standalone firms. Such an effect is more likely operative in industries that rely on local demand.

To explore whether this mechanism explains our main result, we focus on (i) manufacturing and (ii) high-exporting industries. In these sectors, firms rely on national and international demand, which are large relative to the size of local production. Given this, group affiliates' investments are less likely to crowd out demand for standalone firms' output.

We repeat our main tests (Table III) for these subsamples and present the estimates in Table IX. Column 1 (column 2) presents results for firms in the manufacturing sector (high-exporting industries). We define industries as *High Exporting* as those with a ratio of export earnings to sales above the median before the GQ upgrade. Across the two subsamples, we find an economically and statistically significant lower investment among standalone firms in high-business group share regions relative to low-business group share regions. Indeed, our main results hold in these subsamples and are quantitatively stronger, suggesting that, on average, crowding out of demand is not operative, at least in these settings.

VI.C. Rival investment opportunities

Another plausible mechanism could be that group-affiliated firms are more adept at seizing investment opportunities sooner and that these opportunities have a "winner-takes-all" aspect to them. In such cases, group-affiliated firms can crowd out investments of standalone firms in high business group share regions.

To test this mechanism, we utilize product-level information from Prowess and a regulation in India that effectively restricted certain products (henceforth, reserved products) from being manufactured by large firms.[28] Most group-affiliated firms, given their size, are prohibited by this regulation from producing and investing in this subset of products. Therefore, the crowding out of standalone firms' investment opportunities is less likely to be operative in this subset.[29]

In Table X, we focus on the subset of reserved products identified using the five-digit industrial classification. In columns 1 and 2, we focus on the sample of standalone firms whose main product is reserved by the regulation. Additionally, in column 3, we restrict the sample to reserved products in which standalone firms had a market share of above 90%. Again, our main result holds in these subsamples and is quantitatively stronger.

While we cannot completely rule out that in every industry group-affiliates crowd out investment opportunities of standalone firms, the fact that in settings with reduced scope

[28]The policy was specifically geared toward promoting small establishments and has been extensively studied in Martin, Nataraj, and Harrison (2017).

[29]We find that these reserved products are mostly produced by standalone firms, with their average market share being close to 90%. Comparing this to the market share of standalone firms in products that were never reserved, we find that standalone firms had an average market share of 38.5%.

for such crowding out, the results are equal or stronger than the baseline suggests that crowding out of demand by group-affiliated firms is quantitatively unimportant.

VI.D. Political connections

Lastly, our main result could be driven by business groups having political connections and using them to obtain contracts from the government (Khanna and Yafeh, 2007). Therefore, if group affiliates in high business group share regions wield outsized political power, standalone firms in these regions could be at a more considerable disadvantage leading them to obtain fewer government contracts, which, in turn, could explain the lower investment sensitivity of standalone firms in these areas.

To test this prediction, we focus on infrastructure-related industries since, in the period that we study, as the road is being built, many infrastructure projects are being allocated by local governments. This focus is motivated by prior work in economics suggesting rampant favortitism and corruption in infrastructure-related industries in emerging economies and especially at the time of award of contracts involving public procurements (Kenny, 2006; Lehne, Shapiro, and Eynde, 2018; Olken, 2009).

Appendix Table VII presents the results for the subsample of firms operating in infrastructure-related industries. Column 1 of panel A estimates our baseline regression for group-affiliated firms, showing no differential impact across regions with varying levels of business group prevalence. In column 2, we estimate the same specification but for the subsample of standalone firms and again find that increase in investment is not different across regions.

To sharpen these tests, we redefine business group share using only the largest 25 business groups as a proxy for their political influence (Fisman and Khanna, 2004). Estimates from panel B suggest that standalone investment is lower in areas with a significant presence of the 25 largest business groups, while investments by business group affiliates are larger in regions with a significant presence of the 25 largest business groups. These results highlight that the political connections mechanism is likely operative in infrastructure-related industries. However, we acknowledge that it is difficult to extrapolate this evidence to suggest that political connections also operates in other industries.

VII. Robustness checks

In this section, we explore the robustness of our findings and show that our baseline result of lower investment by standalone firms in high business group share regions holds up under various specifications. First, our empirical specification does not include time-varying control variables as they themselves may be affected by the treatment, rendering them "bad controls" (Angrist and Pischke, 2008). To assuage concerns, in Appendix Table VIII, we show that our results on lower investment by standalone firms are qualitatively similar when we include interactions of pre-treatment time-invariant firm characteristics with *PostGQ* in our empirical specification.

Next, we consider alternative definitions of our baseline measure, *High BGS*, defined using the top quartile of group-affiliated asset share in a city. Appendix Table IX presents these results. For ease of comparison, we report the coefficients from the baseline tests (column 2 of Table III). In column 2, we repeat our baseline tests without interacting the fixed effects with *High BGS*. Next, we alter the definition in two ways, (i) use a continuous business group share of assets in a city in column 3 and (ii) alternatively define the top quartile using the Hirschman Herfindahl Index (HHI) based on group-affiliated firms' sales at each location in column 4. To further examine whether the baseline results vary across the distribution of group-affiliated asset share, in columns 5 and 6, we interact *PostGQ* with quartiles and terciles of group-affiliated asset share in a city, respectively. Our results are qualitatively similar in all specifications and definitions, finding that greater business group prevalence is associated with a lower standalone investment.

Further, we address the concern that differential exits by standalone firms drive lower investment in high business group share areas, as *High BGS* may proxy for low productivity of standalone firms. While we do not directly observe firm exits, we conduct several tests to assuage the concern. First, in Appendix Table X, we examine whether a firm stopped reporting financial statements, finding that only 20 firms exited by this measure. Although this seems low, our sample consists of relatively large firms in the Indian context. Still, we show that our results are robust to dropping these exiting firms from our estimations. Second, we alternatively define exit as firms having extremely large negative sales growth. Appendix Table XI reports results from this exercise, finding that dropping such firms strengthens the coefficient estimate on *PostGQ* × *High BGS*. Third, in Appendix Table XII, we show that standalone firms are unlikely to be a target in mergers and acquisitions around GQ upgrades and the probability of being a target is similar across high and low business group share regions.

Lastly, we examine whether the lower investment by standalone firms results from changes in firm entry across regions. To do so, we examine firm entry at the regional level around GQ upgrades as a function of business group share. Appendix Table XIII finds an increase in firm entry after GQ upgrades. However, this increase is similar across high and low business group share areas. Note that the increase in the firm entry is consistent with the view that the GQ upgrade is a shock to investment opportunities.

Overall, we find robust evidence that standalone firms invest less in high business group share regions relative to low business group share regions.

VIII. Conclusion

We study whether standalone firms invest less in regions with greater business group presence. We use a recent large-scale highway development project in India as a shock to investment opportunities for firms that lay along the road network. We find that a higher density of business groups is associated with lower investment by standalone firms. Our results support a financing channel whereby demand for funds from group affiliates crowd out financing to standalone firms.

Our paper contributes to current debates on the economy-wide effects of business groups and, more broadly, ownership concentration. While we establish the existence of a spillover effect of business group affiliates on standalone firms, more research is required to pin down all the different mechanisms. Better knowledge of the mechanisms will help determine whether the aggregate effect of these spillovers is positive or negative. For example, if business groups crowd out financing for standalone firms and reduce the economy's allocation efficiency, the policy implication would be to dismantle business groups. Given the dominance of business groups worldwide, more research is needed to understand the mechanisms through which they affect the overall economy.

REFERENCES

Agarwal, Sumit, Abhiroop Mukherjee, and S Lakshmi Naaraayanan, 2022, Roads and Loans, *Review of Financial Studies* 36, 1508–1547.

Alder, Simon, 2014, Chinese roads in India: The effect of transport infrastructure on economic development, *Work. Pap., Univ. North Carolina, Chapel Hill* .

Almeida, Heitor, Chang-Soo Kim, and Hwanki Brian Kim, 2015, Internal capital markets in business groups: Evidence from the Asian financial crisis, *Journal of Finance* 70, 2539–2586.

Almeida, Heitor, and Daniel Wolfenzon, 2006, Should business groups be dismantled? the equilibrium costs of efficient internal capital markets, *Journal of Financial Economics* 79, 99–144.

Alok, Shashwat, and Meghana Ayyagari, 2019, Politics, State Ownership, and Corporate Investments, *Review of Financial Studies* hhz102.

Angrist, Joshua D, and Jörn-Steffen Pischke, 2008, Mostly harmless econometrics, in *Mostly Harmless Econometrics* (Princeton university press).

Asturias, Jose, Manuel García-Santana, and Roberto Ramos, 2016, Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India, *Journal of the European Economic Association* .

Bebchuk, Lucian, 2012, Corporate pyramids in the israeli economy: Problems and policies, *A Report Prepared for the Committee on Increasing Competitiveness in the Economy, http://mof. gov. il/Committees/CompetitivenessCommittee/FinalReport_ExpertOpinion2. pdf* .

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, How Much Should We Trust Differences-In-Differences Estimates?, *Quarterly Journal of Economics* 119, 249–275.

Bertrand, Marianne, Paras Mehta, and Sendhil Mullainathan, 2002, Ferreting out Tunneling: An Application to Indian Business Groups, *Quarterly Journal of Economics* 117, 121–148.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, 2022, Revisiting event study designs: Robust and efficient estimation, *Working Paper* .

Business-Today, 2013, Economic benefits of the golden quadrilateral project are already visible, *Business Today* .

Chandra, Amitabh, and Eric Thompson, 2000, Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system, *Regional Science and Urban Economics* 30, 457–490.

Claessens, Stijn, Simeon Djankov, and Larry H. P Lang, 2000, The separation of ownership and control in East Asian corporations, *Journal of Financial Economics* 58, 81–112.

Das, Abhiman, Ejaz Ghani, Arti Goswami Grover, William Robert Kerr, and Ramana Nanda, 2019, Infrastructure and Finance: Evidence from India's GQ Highway Network, *World Bank* .

Datta, Saugato, 2012, The impact of improved highways on Indian firms, *Journal of Development Economics* 99, 46–57.

Dau, Luis, Randall Morck, and Bernard Yeung, 2020, Corporate governance, business group governance and economic development traps, *NBER Working Paper* .

Faccio, Mara, and Larry H. P Lang, 2002, The ultimate ownership of Western European corporations, *Journal of Financial Economics* 65, 365–395.

Faccio, Mara, Randall Morck, and M Deniz Yavuz, 2020, Business groups and the incorporation of firm-specific shocks into stock prices, *Journal of Financial Economics* Forthcoming.

Faccio, Mara, and William O'Brien, 2020, Business groups and employment, *Management Science* Forthcoming.

Financial Times, , 1998, Small business living with the crumbs.

Fisman, Raymond, and Tarun Khanna, 2004, Facilitating development: The role of business groups, *World Development* 32, 609–628.

Ghani, Ejaz, Arti Grover Goswami, and William R Kerr, 2014, Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing, *Economic Journal* 126, 317–357.

Goodman-Bacon, Andrew, 2021, Difference-in-differences with variation in treatment timing, *Journal of Econometrics* 225, 254–277.

Gopalan, Radhakrishnan, Abhiroop Mukherjee, and Manpreet Singh, 2016, Do debt contract enforcement costs affect financing and asset structure?, *Review of Financial Studies* 29, 2774–2813.

Gopalan, Radhakrishnan, Vikram Nanda, and Amit Seru, 2007, Affiliated firms and financial support: Evidence from Indian business groups, *Journal of Financial Economics* 86, 759–795.

Haaretz, 2012, Committee recommends dismantling some of israel's business conglomerates.

Hamdani, Assaf, Konstantin Kosenko, and Yishay Yafeh, 2020, Regulatory measures to dismantle pyramidal business groups: Evidence the united states, japan, korea and israel, *European Corporate Governance Institute-Law Working Paper* .

Hesse, Markus, and Jean-Paul Rodrigue, 2004, The transport geography of logistics and freight distribution, *Journal of transport geography* 12, 171–184.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein, 1991, Corporate structure, liquidity, and investment: Evidence from Japanese industrial groups, *Quarterly Journal of Economics* 33–60.

Kandel, Eugene, Konstantin Kosenko, Randall Morck, and Yishay Yafeh, 2019, The great pyramids of america: A revised history of us business groups, corporate ownership, and regulation, 1926–1950, *Strategic Management Journal* 40, 781–808.

Kenny, Charles, 2006, *Measuring and reducing the impact of corruption in infrastructure*, volume 4099 (World Bank Publications).

Khanna, Gaurav, 2014, The road oft taken: The route to spatial development, *Working Paper* .

Khanna, Tarun, and Krishna Palepu, 2000, Is group affiliation profitable in emerging markets? an analysis of diversified Indian business groups, *Journal of Finance* 55, 867–891.

Khanna, Tarun, and Yishay Yafeh, 2007, Business groups in emerging markets: Paragons or parasites?, *Journal of Economic literature* 45, 331–372.

Khwaja, Asim Ijaz, and Atif Mian, 2008, Tracing the impact of bank liquidity shocks: Evidence from an emerging market, *American Economic Review* 98, 1413–42.

Larrain, Borja, Giorgo Sertsios, and Francisco Urzúa I, 2019, The effects of losing a business group affiliation, *Review of Financial Studies* 32, 3036–3074.

Lehne, Jonathan, Jacob N Shapiro, and Oliver Vanden Eynde, 2018, Building connections: Political corruption and road construction in india, *Journal of Development Economics* 131, 62–78.

Levinsohn, James, and Amil Petrin, 2003, Estimating production functions using inputs to control for unobservables, *Review of Economic Studies* 70, 317–341.

Li, Han, and Zhigang Li, 2013, Road investments and inventory reduction: Firm level evidence from china, *Journal of Urban Economics* 76, 43–52.

Lilienfeld-Toal, Ulf von, Dilip Mookherjee, and Sujata Visaria, 2012, The distributive impact of reforms in credit enforcement: Evidence from Indian debt recovery tribunals, *Econometrica* 80, 497–558.

Martin, Leslie A, Shanthi Nataraj, and Ann E Harrison, 2017, In with the big, out with the small: Removing small-scale reservations in india, *American Economic Review* 107, 354–86.

Naaraayanan, S. Lakshmi, and Kasper Meisner Nielsen, 2021, Does personal liability deter individuals from serving as independent directors?, *Journal of Financial Economics* 140, 621–643.

Olken, Benjamin A, 2009, Corruption perceptions vs. corruption reality, *Journal of Public economics* 93, 950–964.

Redding, Stephen J, and Matthew A Turner, 2015, Transportation costs and the spatial organization of economic activity, *Handbook of regional and urban economics* 5, 1339–1398.

Roth, Jonathan, and Pedro HC Sant'Anna, 2020, When is parallel trends sensitive to functional form?, *Working Paper* .

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, 2022, What's trending in difference-in-differences? a synthesis of the recent econometrics literature, *arXiv preprint arXiv:2201.01194* .

Santioni, Raffaele, Fabio Schiantarelli, and Philip E Strahan, 2019, Internal Capital Markets in Times of Crisis: The Benefit of Group Affiliation, *Review of Finance* .

Shirley, Chad, and Clifford Winston, 2004, Firm inventory behavior and the returns from highway infrastructure investments, *Journal of Urban Economics* 55, 398–415.

Siegel, Jordan, and Prithwiraj Choudhury, 2012, A reexamination of tunneling and business groups: New data and new methods, *Review of Financial Studies* 25, 1763–1798.

Times of Israel, 2013, New law makes business pyramids history.

World-Bank, 2000, Project appraisal document on a proposed loan in the amount of us $516 million equivalent to the government of india for the third national highways project .

World-Bank, 2002, India's transport sector: The challenges ahead, *World Bank Reports* 1–65.

World-Bank, 2009, World bank enterprise surveys and indicator surveys: Sampling methodology .

World-Road-Statistics, 2009, World road statistics 2009: Data 2002-2007 .

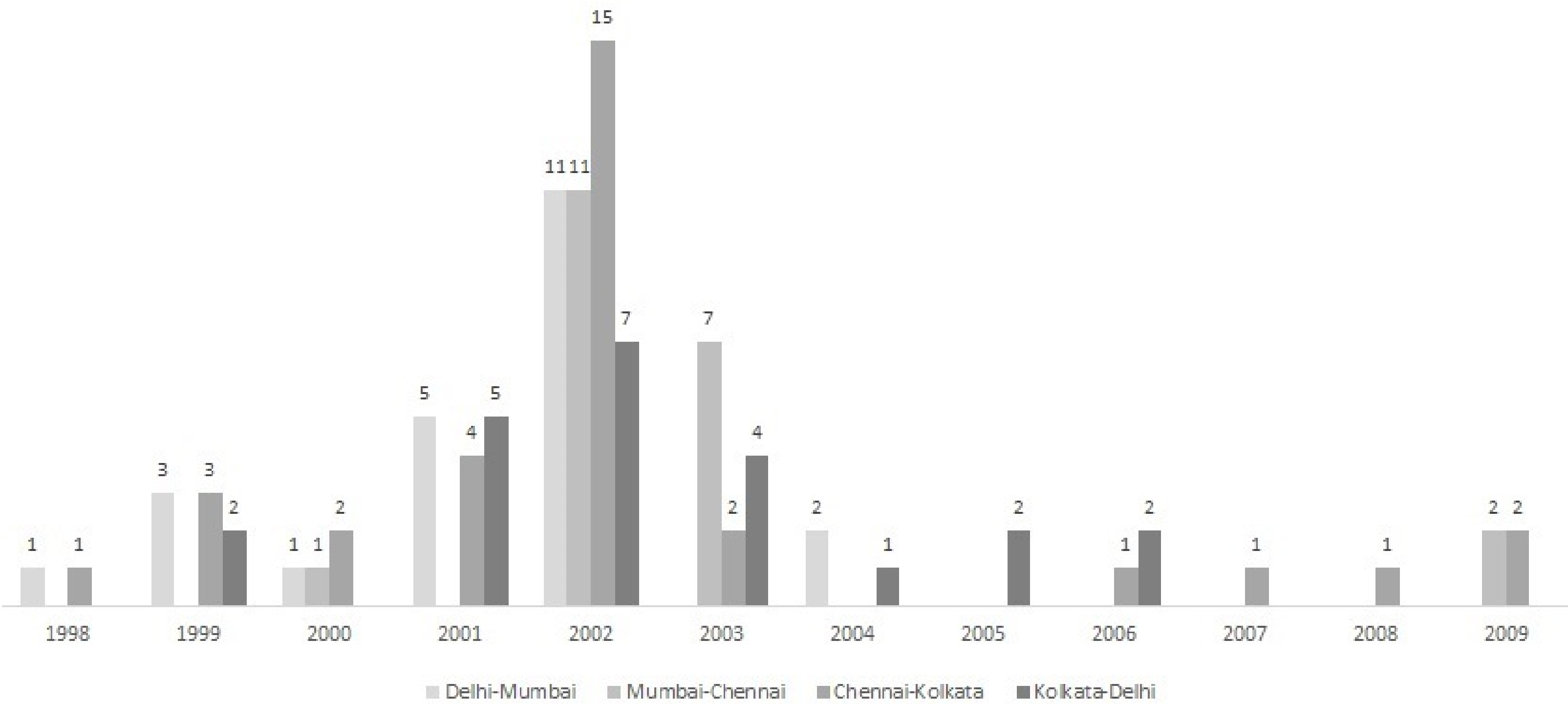


Figure I. Temporal variation in GQ construction

This figure illustrates the temporal variation in the commencement of construction of the four segments forming part of the Golden Quadrilateral (GQ), which connects the four nodal cities of Delhi, Mumbai, Chennai, and Kolkata. The height of each bar corresponds to the number of sub-segments that began construction each year. Data source: National Highway Authority of India.

(a) 2000

(b) 2001

(c) 2002

(d) 2004

Figure II. GQ construction evolution over time

This figure illustrates the spatial variation of segments at different points in time along the GQ network. The network is part of the 5,846 km stretch of the GQ connecting four nodal cities of Delhi, Mumbai, Chennai, and Kolkata. Map source: National Highway Authority of India.

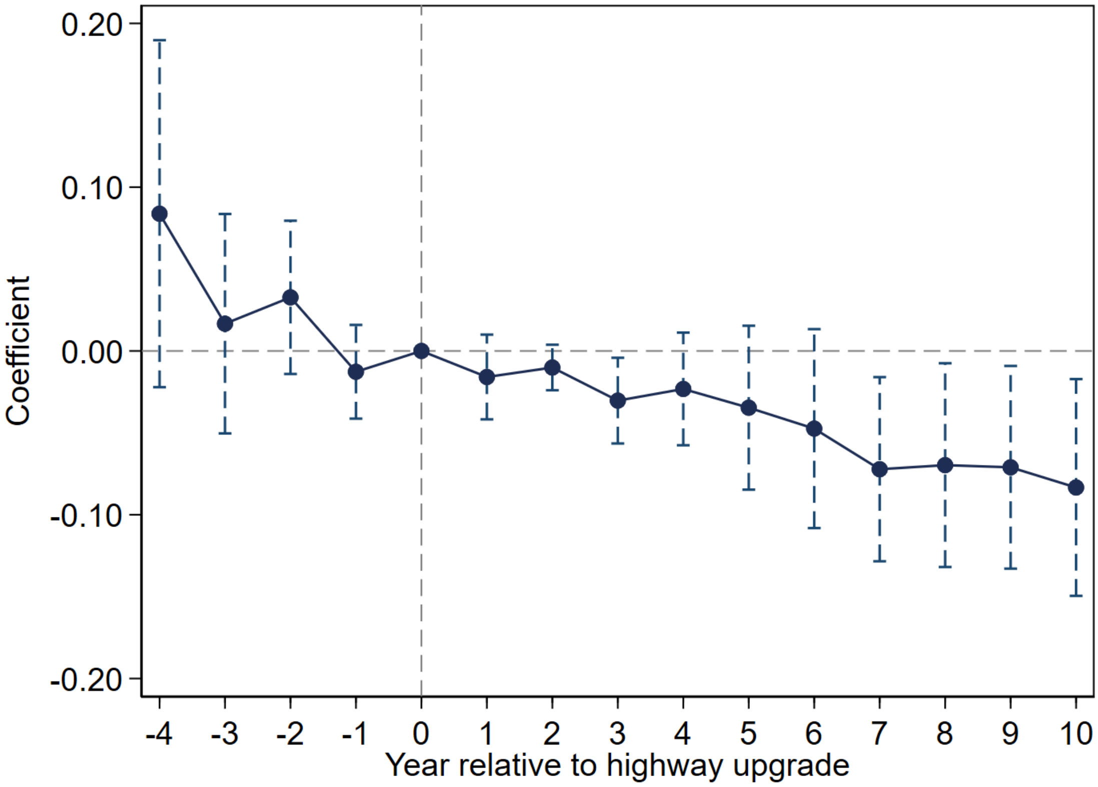


Figure III. Standalone investment by business group prevalence around GQ upgrades

This figure displays the dynamic coefficients (λ_k) and their corresponding 90% confidence intervals of the differential investment by standalone firms in high business group share regions relative to low business group share regions around the upgrade of the GQ road network. We estimate a fully dynamic specification that allows us to capture the dynamics of standalone firm investment relative to the year of commencement of GQ upgrade. Specifically, we estimate the following equation:

$$\text{Investment}_{ijcst} = \alpha_i + \sum_{k=-1}^{-4} \mu_k + \sum_{k=1}^{10} \mu_k + \sum_{k=-1}^{-4} \lambda_k \times \text{High BGS} + \sum_{k=1}^{10} \lambda_k \times \text{High BGS} + \text{High BGS} \times \theta_{jt} + \epsilon_{ijcst}$$

All coefficients are plotted relative to investment at k=0, which is normalized to zero. *High BGS* is an indicator variable set to 1 if the share of group assets is in the top quartile of the distribution. The sample is restricted to firms along the GQ network and covers the window of [-4,+10] around the commencement of the upgrade.

Table I. Summary statistics

This table reports the descriptive statistics of firm characteristics for our sample. Panel A reports the descriptive statistics for all firms while panel B reports the descriptive statistics for standalone firms in our sample. From the overall Prowess sample of 1989 to 2016, we exclude all financial firms (NIC code: 641-663), firms owned by central and state governments, firms with less than three years of data with positive values of total assets and sales, and drop observations with ratio of investment to lagged total assets greater than 1. All the financial variables are adjusted for inflation using the Wholesale Price Index (WPI) at 2010 constant prices. We also correct for changes in the financial reporting year by adjusting values for the number of months. To mitigate the effect of outliers, we winsorize all the ratios at 1% tails. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

	Panel A: All firms					
	N	**Mean**	**SD**	**P25**	**P50**	**P75**
Assets (INR millions)	24,709	8,176	73,751	290	919	3,075
Firm age (years)	24,709	30	21	16	23	35
Cash flow	24,709	0.06	0.11	0.00	0.06	0.12
Profitability	24,709	0.12	0.10	0.07	0.11	0.16
Listed	24,709	0.63	0.48	0.00	1.00	1.00
Investment	24,709	0.38	0.22	0.21	0.35	0.51
Debt	24,709	0.27	0.27	0.09	0.22	0.36
Total factor productivity	21,505	3.42	4.23	1.80	2.40	3.35

	Panel B: Standalone firms					
	N	**Mean**	**SD**	**P25**	**P50**	**P75**
Assets (INR millions)	15,842	2,593	11,873	236	657	1,893
Firm age (years)	15,842	27	19	16	22	31
Cash flow	15,842	0.05	0.11	0.00	0.05	0.11
Profitability	15,842	0.12	0.10	0.07	0.11	0.16
Listed	15,842	0.57	0.50	0.00	1.00	1.00
Investment	15,842	0.39	0.22	0.22	0.36	0.53
Debt	15,842	0.25	0.19	0.10	0.23	0.36
Total factor productivity	10,475	3.40	3.75	1.91	2.51	3.46

Table II. Productive efficiency and investment around GQ upgrades

This table examines changes in productive efficiency and investment for firms located along GQ around the upgrade of the road network. The dependent variable in column (1) is *TFP* is the total factor productivity which is estimated using the methodology outlined in Levinsohn and Petrin (2003). The dependent variable in column (2) is *Days sales of inventory*, defined as the ratio of ending inventory to cost of good solds multiplied by 365 and the dependent variable in column (3) is *Investment*, defined as the net capital expenditure divided by lagged total assets. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. All regressions include firm fixed effects and state-year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	TFP	Days sales of inventory	Investment
	(1)	(2)	(3)
PostGQ	0.146**	-3.961*	0.026**
	(0.067)	(2.350)	(0.012)
Fixed effects:			
Firm	Yes	Yes	Yes
Industry × year	Yes	Yes	Yes
State × year	Yes	Yes	Yes
Adjusted-R^2	0.79	0.63	0.70
Observations	21,494	21,053	24,709
Sample : All firms	Yes	Yes	Yes

Table III. Business group prevalence and standalone investment

This table presents estimates from regressions relating business group prevalence to standalone firms' investment around GQ upgrades. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects, High BGS × state × year, and High BGS × industry × year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment	
	(1)	(2)
PostGQ	0.038	0.039
	(0.029)	(0.028)
PostGQ × High BGS	-0.012***	-0.038***
	(0.000)	(0.012)
Fixed effects:		
Firm	Yes	Yes
High BGS × industry × year	No	Yes
High BGS × state × year	Yes	Yes
Adjusted-R^2	0.70	0.70
Observations	15,842	15,842
Sample : Standalone firms	Yes	Yes

Table IV. Pre-GQ firm and regional characteristics by business group share

This table compares the means (and medians in parentheses) of firm and regional characteristics as a function of the prevalence of business groups in the local area. Panel A presents firm characteristics while Panel B presents regional characteristics. Specifically, Panel A displays the means (medians) for standalone firms one year before the commencement of GQ upgrades. Column 1 displays mean (and median) for High business group share regions while Column 2 displays mean (and median) for Low business group share regions. Column 3 tests the difference in means (medians). *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. The firm characteristics we focus on are: *Firm size*, *Firm age*, *Cash flow*, *Profitability*, *Investment*, *Debt*, *Total factor productivity*, and *Days sales of inventory*. The regional characteristics we focus on are: *Bank branches*, *Fraction of listed firms*, *Bad roadwaWys*, *Obstacle to growth, transport*, *Labor constraint in contracting*, *Obstacle to growth, labor*, *Time to fill manager vacancy*, and *Average time to fill skilled worker vacancy*. *,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data sources: CMIE Prowess, Reserve Bank of India, and World Bank Enterprise Survey.

Panel A: Standalone firm characteristics

	High business group share	Low business group share	High - Low (1)-(2)
	(1)	(2)	(3)
Firm size	5.20	5.49	-0.29
	(5.05)	(5.10)	(-0.05)
Firm age	19.42	16.87	2.55
	(14.00)	(14.00)	0.00
Cash flow	0.06	0.06	0.00
	(0.06)	(0.06)	(0.00)
Profitability	0.08	0.11	-0.03
	(0.09)	(0.11)	(-0.02)
Investment	0.41	0.47	-0.06
	(0.39)	(0.49)	(-0.10)
Debt	0.25	0.28	-0.03
	(0.25)	(0.24)	(0.01)
Total factor productivity	3.48	2.71	0.77**
	(3.18)	(2.67)	0.51
Days sales of inventory	80.5	86.7	-6.1
	(78.8)	(90)	-11.2

Continued...

Panel B: Regional characteristics			
	High business group share	Low business group share	High - Low (1) - (2)
	(1)	(2)	(3)
1. Financial development			
Bank branches (per 100,000)	4.95	4.28	0.67
	(4.04)	(4.06)	0.26
Fraction of listed firms	0.70	0.73	-0.03
	(0.72)	(0.74)	-0.02
2. Physical infrastructure			
Bad Roadways (rating)	6.92	6.77	0.15
	(7.00)	(7.00)	0.00
Obstacle to growth, transport (1=yes)	0.26	0.34	-0.08
	(0.00)	(0.00)	0.00
3. Labor market conditions			
Labor constraint in contracting (1=yes)	0.07	0.07	0.00
	(0.00)	(0.00)	0.00
Obstacle to growth, labor (1=yes)	0.36	0.26	0.09
	(0.00)	(0.00)	0.00
Time to fill manager vacancy (weeks)	4	3	1
	(2)	(2)	0

Table V. Horse Race Regressions

This table presents estimates from horse race regressions that relate business group prevalence to standalone firms' investment. We consider the following covariates: *Listed share*, *Firm age*, and *TFP*. For each covariate, we define an indicator variable which is set to one if the specific characteristic of firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. Panel A defines the indicator based on all firms while Panel B defines the indicator using only standalone firms. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects, High BGS × state × year, and High BGS × industry × year fixed effects. As TFP is estimated for manufacturing firms, columns 4 and 5 restrict the sample to these industries. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

	Panel A: Definition using all firms				
Dependent variable	Investment				
	Baseline (1)	Listed share (2)	Firm age (3)	TFP (4)	All (5)
PostGQ	0.039	0.046*	0.020	0.002	-0.028
	(0.028)	(0.027)	(0.042)	(0.035)	(0.043)
PostGQ × High BGS	-0.038***	-0.084***	-0.106***	-0.105**	-0.170***
	(0.012)	(0.029)	(0.029)	(0.041)	(0.058)
PostGQ × High Listed Share (all firms)		-0.216**			-0.262***
		(0.095)			(0.083)
PostGQ × High Firm Age (all firms)			0.048		0.092
			(0.054)		(0.056)
PostGQ × High Firm TFP (all firms)				0.006	0.012
				(0.025)	(0.025)
Fixed effects:					
Firm	Yes	Yes	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.70	0.70	0.66	0.66
Observations	15,842	15,842	15,842	11,520	11,520
Sample : Standalone firms	Yes	Yes	Yes	Yes	Yes

Continued...

Dependent variable	Panel B: Definition using standalone firms				
	Investment				
	Baseline (1)	Listed share (2)	Firm age (3)	TFP (4)	All (5)
PostGQ	0.039	0.061**	0.005	0.007	0.105*
	(0.028)	(0.030)	(0.059)	(0.036)	(0.052)
PostGQ × High BGS	-0.038***	-0.099***	-0.095***	-0.110**	-0.116**
	(0.012)	(0.032)	(0.029)	(0.041)	(0.044)
PostGQ × High Listed Share (standalones)		-0.065**			-0.009
		(0.025)			(0.055)
PostGQ × High Firm Age (standalones)			0.051		0.075
			(0.055)		(0.088)
PostGQ × High Firm TFP (standalones)				-0.026	-0.060
				(0.031)	(0.036)
Fixed effects:					
Firm	Yes	Yes	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.70	0.70	0.66	0.66
Observations	15,842	15,842	15,842	11,520	11,520
Sample : Standalone firms	Yes	Yes	Yes	Yes	Yes

Table VI. Investment by firm type around GQ upgrades

This table presents estimates from regressions relating the prevalence of business groups on investment around GQ construction for different sample of firms. Column 1 focuses on all firms while Column 2 (Column 3) focuses on the sample standalone (Group-affiliated) firms. Across all columns, the dependent variable is *Investment*. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects, High BGS x state x year, and High BGS x industry x year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment		
Sample	All firms (1)	Standalone (2)	Business Group (3)
PostGQ	0.037*	0.039	0.024
	(0.021)	(0.028)	(0.057)
PostGQ × High BGS	-0.007	-0.038***	0.009*
	(0.005)	(0.012)	(0.005)
Fixed effects:			
Firm	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes
Adjusted-R^2	0.69	0.70	0.65
Observations	24,319	15,842	8,102

Table VII. Aggregate bank lending around GQ upgrades: District-level evidence

This table presents estimates from regressions relating the effect of business group prevalence on overall district-level lending. The dependent variable, *Log (1+credit)*, is defined as the natural logarithm of one plus the total credit disbursed in a district-year. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if a district consists of more than three cities with share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. All regressions include district fixed effects. Additionally, specifications in columns (2) and (4) include State × year fixed effects to control for local macroeconomic confounds. Standard errors are corrected for heteroscedasticity and auto correlation, and clustered at the district level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: Reserve Bank of India.

Dependent variable	Log (1+credit)			
	(1)	(2)	(3)	(4)
PostGQ	0.120**	0.125***	0.115**	0.104**
	(0.047)	(0.046)	(0.052)	(0.050)
PostGQ × High BGS			0.032	0.128
			(0.095)	(0.114)
Fixed effects:				
District	Yes	Yes	Yes	Yes
Year	Yes	No	Yes	No
State × year	No	Yes	No	Yes
Adjusted-R^2	0.97	0.98	0.97	0.98
Observations	6,862	6,862	6,862	6,862

Table VIII. Loan-level regressions: Bank lending to standalone firms

These regressions examine bank lending for the set of standalone firms borrowing at the time of GQ upgrade. All loans are time-collapsed into a single pre- and post-period of five years around the start of the GQ upgrade. The sample includes standalone firms that borrow from multiple banks. The dependent variable, ΔL_{ib}, is the change in the average loan amount to standalone firms five years after the GQ upgrade relative to five years before the GQ upgrade. The independent variable, *Group exposure*, is defined for each bank as the total lending to group affiliates before the start of the GQ upgrade. The empirical specification includes firm fixed effects that control for the firm-specific credit demand. Standard errors are corrected for heteroscedasticity and auto correlation, and clustered at the bank level. Standard errors are reported in parentheses.*, **, and *** indicate significance at 10%, 5% and 1% respectively. All variables are defined in Appendix Table XIV. Data source: Ministry of Corporate Affairs.

Dependent variable	Δ Log loan size	
	(1)	(2)
Group exposure	-0.185**	-0.191**
	(0.071)	(0.072)
Firm fixed effects	Yes	Yes
Weighted by firm size	No	Yes
Adjusted-R^2	0.432	0.431
Number of loans	163	163
Sample: Standalone firms	Yes	Yes

Table IX. Mechanism: Crowding out demand for standalone firms' output

This table presents estimates from regressions which rule out the alternative mechanism whereby group-affiliated firms crowd out demand for standalone firms' output in product markets. Column 1 focuses on the subsample of firms operating in manufacturing industries while Column 2 focuses on the subsample of firms operating in "High exporting" industries. We define industries as "High-exporting" as those with a ratio of export earnings to sales above the median before the GQ upgrade. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects and High BGS × state × year fixed effects, while column 2 additionally includes High BGS × industry × year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment	
Sample industries	Manufacturing	High-exporting
	(1)	(2)
PostGQ	-0.008	0.013
	(0.030)	(0.044)
PostGQ × High BGS	-0.081***	-0.049*
	(0.000)	(0.028)
Fixed effects:		
Firm	Yes	Yes
High BGS × industry × year	No	Yes
High BGS × state × year	Yes	Yes
Adjusted-R^2	0.66	0.69
Observations	11,521	9,081
Sample: Standalone firms	Yes	Yes

Table X. Mechanism: Rival investment opportunities

This table presents estimates from regressions which rule out the alternative mechanism whereby group-affiliated firms are adept at seizing investment opportunities sooner, therby crowding out investments by standalone firms in high business group share regions. For this test, we focus on the subsample of reserved products, defined as those that were restricted from being manufactured by large firms (Martin, Nataraj, and Harrison, 2017). The products are identified using the five-digit industrial classification. Columns 1 and 2 focus on all products while Column 3 focuses on products where standalone firms have a dominant market share (market share $\geq$ 90%). *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects, State × year, and Industry × year fixed effects. Additionally, column 1 interacts High BGS with State × year, and Industry × year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment		
	All products (1)	All products (2)	SA dominant products (3)
PostGQ	-0.013	-0.015	-0.260***
	(0.038)	(0.053)	(0.078)
PostGQ × High BGS	-0.152***	-0.171***	-0.236***
	(0.043)	(0.044)	(0.061)
Fixed effects:			
Firm	Yes	Yes	Yes
High BGS ×	Yes	No	No
Industry × year	Yes	Yes	Yes
State × year	Yes	Yes	Yes
Adjusted-R^2	0.66	0.66	0.63
Observations	2,544	2,657	298
Sample: Standalone firms	Yes	Yes	Yes

ONLINE APPENDIX

Business Group Spillovers

Appendix A Matching datasets

In this section, we discuss the matching between firm-bank pairs in Prowess to firm-bank pairs in the loan-level dataset from the credit registry at the MCA. We start with the sample of firms in Prowess and match them to firms in the loan-level dataset scrape between May 2021 to December 2021. We match them using the company identifier (CIN) provided by the MCA, yielding a match rate of 95%.

In this matched dataset, we keep loans starting from 1980 until 2016. We drop short maturity loans, those that are less than 3 years, and keep loan amounts larger than 10 million INR. Subsequently, we carry out cleaning and standardize names to merge with information on banks and financial institutions from Prowess. Note that this standardization on names is performed on both datasets i.e., Prowess and the loans. In the next step, we perform a fuzzy match of the names across the two datasets, yielding a match rate of 57% (out of 1,175 names in Prowess). This was a non-trivial task due to the fact that the loan-level sample included various financial institutions, including non-banks and private entities for which we had no information. As a result, we decided to only focus on banks and financial institutions reported in Prowess, resulting in a sample that covers larger firm-bank pairs.

In the last step, we merge the dataset to the sample of standalone firms that lay along the GQ road network. This merge yields 2,430 loans to 302 unique firms and 140 banks. Further, in our empirical specification, we require that standalone firms: (i) borrow from multiple banks, and (ii) borrow from multiple banks both before and after the GQ upgrades. Such a restriction leads the final sample to consist of 163 loans to 17 unique firms to 15 unique banks. Appendix Table I reports the descriptive statistics for firms and loan level characteristics for the matched sample.

Appendix B Discussion of identifying assumptions

In this section, we discuss in detail the identifying assumptions for interpreting β in Equation 2 as the causal impact of GQ upgrades. The identifying assumption is that the timing of the road construction is orthogonal to the investment opportunities for firms located along the GQ road network. We confirm the validity of this assumption by (i) ruling out pre-trends in investments and (ii) showing that observable differences in firm characteristics cannot explain the timing of GQ road upgrade.

We begin by assessing pre-trends in investments. Results are presented in event-time in panel A of Appendix Figure I. This figure confirms that treated firms indeed did not show any pre-trends, thus ruling out concerns that the increase in investment would have occurred regardless of the road upgrade.

Next, we investigate whether observable firm characteristics can predict the timing of road construction. To do so, we run Cox hazard rate regressions of the time to road upgrade. The explanatory variables include city-level averages (calculated over different periods) of investments, market concentration (based on sales), market share (based on sales), fixed assets, investment growth, sales growth, cash holdings, and profitability.

Appendix Table II presents the results. Panel A reports results using city-level averages computed using all firms while panel B reports results using city-level averages computed using only group-affiliated firms. All specifications include state fixed effects. The results suggest that observable differences in average firm characteristics cannot explain the timing of GQ road upgrade. While it is impossible to test whether the timing of the road upgrade is orthogonal to the (unobserved) investment opportunities, the results in this table provide comfort that at least the timing is orthogonal to a broad set of observable firm characteristics that are likely correlated with investments.

Lastly, in panel B of Appendix Figure I, we show robustness of event study results to recent concerns about the two-way fixed effect estimator providing biased estimates when treatment is staggered and in the presence of treatment effect heterogeneity (Borusyak, Jaravel, and Spiess, 2022; Goodman-Bacon, 2021; Roth, Sant'Anna, Bilinski, and Poe, 2022).

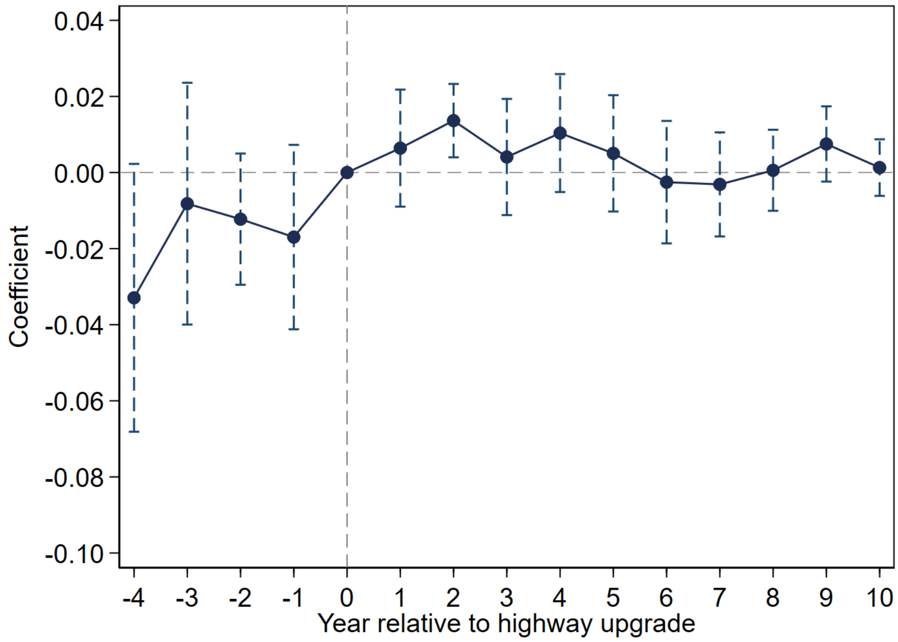
The imputation estimator and other recent econometric advances in the staggered treatment adoption literature propose using the "never treated" group as the control to obtain unbiased estimates of the treatment effect. As our baseline sample consists of all firms that are eventually treated, we identify the control group by expanding our sample to include firms that did not lay along the GQ road network. Further, as business group shares and organizational form are not randomly allocated, we choose to match each treatment firm in our sample to a control firm based on the following pre-GQ characteristics, one year before : (i) the level of business group share in the city, and (ii) the same organizational form (group-affiliated vs. standalone). Lastly, we also require that the control firms have the same (i) the level of investment, (ii) the level of leverage, and (iii) operate in the same state and two-digit industry, as the treatment firm. The figure confirms that (i) there are no pre-trends prior to the GQ upgrades, and (ii) the increase in investments start around two years into the GQ upgrades and are persistent up to five years after the GQ road upgrades.

Appendix C Robustness to fixed effects estimator

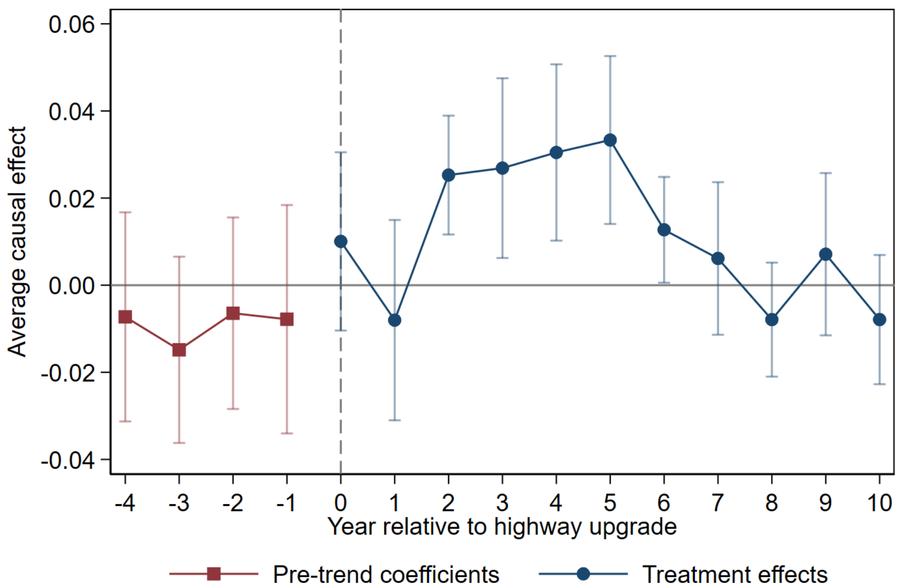
In this section, we examine robustness of our baseline results to recent concerns about the two-way fixed effects estimator providing biased estimates when treatment is staggered and in the presence of treatment effect heterogeneity (Borusyak, Jaravel, and Spiess, 2022; Goodman-Bacon, 2021; Roth, Sant'Anna, Bilinski, and Poe, 2022).

The imputation estimator and other recent econometric advances in the staggered treatment adoption literature propose using the "never treated" group as the control to obtain unbiased estimates of the treatment effect. As our baseline sample consists of all firms that are eventually treated, we identify the control group by expanding our sample to include firms that did not lay along the GQ road network. Further, as business group shares and organizational form are not randomly allocated, we choose to match each treatment firm in our sample to a control firm based on the following pre-GQ characteristics, one year before : (i) the level of business group share in the city, and (ii) the same organizational form (i.e., standalone). Lastly, we also require that the control firms have the same (i) the level of investment, (ii) the level of leverage, (iii) operate in the same state and two-digit industry, and (iv) operate in *High or Low BGS* regions, as the treatment firm.

In Appendix Table III, we present the results from two separate regressions for firms in *High BGS* and *Low BGS* regions. Each regression is akin to a difference-in-differences (DiD) estimation. The table confirms that there are no pre-trends for firms in either regions prior to the GQ upgrades. More importantly, there is a lower investment sensitivity to GQ upgrades for firms in *High BGS* regions relative to control firms (column 1) but for firms in *Low BGS* regions relative to control firms, we find a higher investment sensitivity to the GQ upgrades (column 2). These results are consistent with our baseline estimates from Figure III. Moreover, we note that the lower investment sensitivity in *High BGS* regions start in the immediate year around the GQ upgrades and are persistent up to ten years after the GQ road upgrades while the higher investment sensitivity in *Low BGS* regions are temporary.



(a) Fixed effects estimator



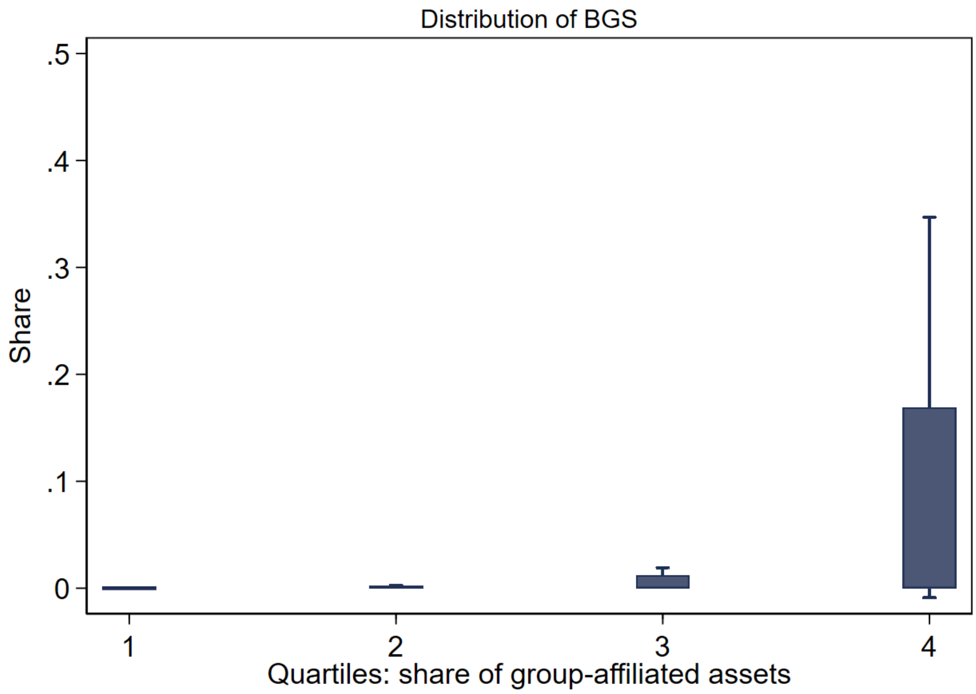
(b) Imputation estimator

Appendix Figure I. Firm investment around GQ upgrades

Panel A of the figure displays the dynamic coefficients (μ_k) and their corresponding 90% confidence intervals of investment by firms around the upgrade of the GQ road network. We use the fixed effects estimator to estimate a fully dynamic specification that allows us to capture the dynamics of firm investment relative to the year of commencement of GQ upgrade. Specifically, we estimate the following equation:

$$\text{Investment}_{ijcst} = \alpha_i + \sum_{k=-1}^{-4} \mu_k + \sum_{k=1}^{10} \mu_k + \theta_{jt} + \epsilon_{ijcst}$$

All coefficients are plotted relative to investment at k=0, which is normalized to zero. Panel B of the figure plots coefficients from the imputation estimator suggested in Borusyak, Jaravel, and Spiess (2022) and includes matched pair × year fixed effects. Details are in Appendix B.



Appendix Figure II. Distribution of BGS

This figure illustrates the variation the share of group-affiliated assets split by quartiles. The bars display the mean within each group while the lines present the standard deviation within each group.

Appendix Table I. Summary statistics, loan-level dataset

This table reports the descriptive statistics of firms matched to the Prowess sample. Panel A reports the loan-level characteristics while panel B reports the descriptive statistics for firms that borrow from multiple banks five years around the GQ upgrades. Following Khwaja and Mian (2008), we aggregate multiple loans of a firm from the same bank and collapse the data at the firm-bank pair level with two observations (pre and post) for each pair. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess and Ministry of Corporate Affairs.

	Panel A: Loan-level characteristics					
	N	**Mean**	**SD**	**P25**	**P50**	**P75**
Loan amount (INR millions)	163	147.2	383.4	26.0	70.0	136.6
Change in log lending	163	0.2	1.0	-0.4	0.2	0.8
Group exposure amount (INR billions)	163	55.68	64.51	8.47	39.25	126.19

	Panel B: Firm-level characteristics					
	N	**Mean**	**SD**	**P25**	**P50**	**P75**
Total assets	163	3,999	7,150	716	1,458	4,441
Firm age (years)	163	31.33	16.60	19.00	26.00	38.00
Cash flow	163	0.08	0.10	0.03	0.07	0.13
Profitability	163	0.13	0.08	0.10	0.13	0.16
Listed	163	0.60	0.49	0.00	1.00	1.00
Investment	163	0.46	0.20	0.31	0.45	0.60
Debt	163	0.29	0.18	0.18	0.28	0.37
Total factor productivity	163	2.98	2.17	1.87	2.29	3.30

Appendix Table II. Timing of GQ and pre-existing firm characteristics: Survival analysis

A Cox proportional hazards model is fitted to investigate the predictability of placement of GQ segments based on pre-existing firm characteristics. The explanatory variables include city-level averages (calculated over different periods) of investments, market share (based on sales), firm assets, investment growth, sales growth, cash holdings, and profitability. Panel A reports results using city-level averages computed using all firms while panel B reports results using city-level averages computed using only group firms. All specifications include state fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses. *,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

	Panel A. All firms			
	1994-97	1992-97	1990-97	1988-97
	(1)	(2)	(3)	(4)
Investment (City avg.)	-39.324	3.266	3.599	15.796
	(27.251)	(15.239)	(18.431)	(21.039)
Market share (City avg.)	-1.305	-0.689	-0.835	-0.697
	(3.277)	(2.172)	(2.163)	(2.130)
Sales growth (City avg.)	2.994	-0.768	-0.885	-0.592
	(2.857)	(1.517)	(1.482)	(1.506)
Investment growth (City avg.)	4.197	0.697	0.877	0.519
	(3.851)	(2.881)	(2.788)	(2.762)
Profitability (City avg.)	-11.911	-5.321	-5.411	-5.750
	(8.407)	(3.502)	(3.542)	(3.502)
Firm Size (City avg.)	-0.670	-0.052	-0.040	-0.050
	(0.742)	(0.336)	(0.315)	(0.306)
State fixed effects	Yes	Yes	Yes	Yes
χ^2 statistic	53.856	46.665	45.915	47.201
Log Pseudo likelihood	-75.639	-101.905	-101.856	-101.695
Observations	430	430	430	430

	Panel B. Group firms			
	1994-97	1992-97	1990-97	1988-97
	(1)	(2)	(3)	(4)
Investment (City avg.)	-2.262	-5.438	-5.236	17.734
	(20.788)	(25.234)	(30.136)	(41.819)
Market share (City avg.)	-1.775	-1.194	-1.379	-1.816
	(2.339)	(1.250)	(1.363)	(1.606)
Sales growth (City avg.)	1.891	0.213	-0.268	0.253
	(1.613)	(1.631)	(1.743)	(1.876)
Investment growth (City avg.)	-0.690	-0.999	-1.446	-1.757
	(1.427)	(1.060)	(1.232)	(1.516)
Profitability (City avg.)	7.711	6.770	9.246	9.189
	(12.036)	(6.473)	(7.509)	(7.431)
Firm Size (City avg.)	0.307	0.345	0.325	0.301
	(0.374)	(0.418)	(0.442)	(0.437)
State fixed effects	Yes	Yes	Yes	Yes
χ^2 statistic	72.366	51.023	65.776	78.016
Log Pseudo likelihood	-69.759	-70.971	-70.827	-70.784
Observations	364	364	364	364

Appendix Table III. Robustness to fixed effects estimator

This table presents estimates from the imputation estimator suggested in Borusyak, Jaravel, and Spiess (2022). Column 1 (column 2) presents coefficients for firms located in *High BGS (Low BGS)* regions compared to matched control firms from the "never treated" cities that did not lay along the GQ road network. $Before^k$ ($After^k$) is an indicator variable taking value 1 for all years leading upto (after) the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include matched pair × year fixed effects. Details are in Appendix C. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment	
	High BGS (1)	Low BGS (2)
$Before^{-4}$	-0.058	0.049
	(0.054)	(0.0369)
$Before^{-3}$	-0.045	0.059
	(0.041)	(0.044)
$Before^{-2}$	-0.009	0.071
	(0.020)	(0.0296)
$Before^{-1}$	0.004	0.154
	(0.019)	(0.0502)
$After^{0}$	-0.020***	0.096
	(0.009)	(0.0422)
$After^{1}$	-0.063***	-0.024
	(0.015)	(0.0244)
$After^{2}$	0.006	0.183**
	(0.015)	(0.0262)
$After^{3}$	-0.027*	0.143*
	(0.015)	(0.0361)
$After^{4}$	0.013	-0.049
	(0.011)	(0.0287)
$After^{5}$	0.009	-0.011
	(0.0121)	(0.047)
$After^{6}$	-0.085***	-0.016
	(0.0128)	(0.0348)
$After^{7}$	-0.017	-0.059**
	(0.0119)	(0.0393)
$After^{8}$	-0.061*	-0.07
	(0.0129)	(0.0473)
$After^{9}$	0.013	-0.011
	(0.0137)	(0.0476)
$After^{10}$	-0.033**	-0.017
	(0.012)	(0.0452)
Fixed effects:		
Matched pair × year	Yes	Yes
Observations	8,465	5,601
Sample: Standalone firms	Yes	Yes

Appendix Table IV. Horse race regressions, robustness to firm size

This table presents estimates from horse race regressions relating the effect of business group prevalence on standalone firms' investment. We consider the following covariates: *Listed share*, *Firm age*, and *TFP*. For each covariate, we define an indicator variable which is set to one if the specific characteristic of firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. Panel A defines the indicator based on all firms while Panel B defines the indicator using only standalone firms. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects, High BGS × state × year, and High BGS × industry × year fixed effects. As TFP is estimated for manufacturing firms, columns 4 and 5 restrict the sample to these industries. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

	Panel A: Definition using all firms				
Dependent variable	Investment				
	Baseline (1)	Listed share (2)	Firm size (3)	TFP (4)	All (5)
PostGQ	0.039 (0.028)	0.046* (0.027)	0.039 (0.029)	0.002 (0.035)	-0.028 (0.043)
PostGQ × High BGS	-0.038*** (0.012)	-0.084*** (0.029)	-0.094 (0.092)	-0.105** (0.041)	-0.170*** (0.058)
PostGQ × High Listed Share		-0.216** (0.095)			-0.262*** (0.083)
PostGQ × High Firm Size			0.016 (0.090)		0.003 (0.044)
PostGQ × High Firm TFP				0.006 (0.025)	0.012 (0.025)
Fixed effects:					
Firm	Yes	Yes	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.70	0.70	0.66	0.66
Observations	15,842	15,842	15,842	11,520	11,520
Sample: Standalone firms	Yes	Yes	Yes	Yes	Yes

Continued...

Dependent variable	Panel B: Definition using standalone firms				
	Investment				
	Baseline (1)	Listed share (2)	Firm size (3)	TFP (4)	All (5)
PostGQ	0.039 (0.028)	0.061** (0.030)	0.101* (0.053)	0.007 (0.036)	0.105* (0.052)
PostGQ × High BGS	-0.038*** (0.012)	-0.099*** (0.032)	-0.064** (0.029)	-0.110** (0.041)	-0.116** (0.044)
PostGQ × High Listed Share (standalones)		-0.065** (0.025)			-0.009 (0.055)
PostGQ × High Firm Size (standalones)			-0.075* (0.042)		-0.168*** (0.061)
PostGQ × High Firm TFP (standalones)				-0.026 (0.031)	-0.060 (0.036)
Fixed effects:					
Firm	Yes	Yes	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.70	0.70	0.66	0.66
Observations	15,842	15,842	15,842	11,520	11,520
Sample: Standalone firms	Yes	Yes	Yes	Yes	Yes

Appendix Table V. Stock price reactions to new plant announcements by standalone firms

This table shows stock price reactions to announcement of new plants by standalone firms around GQ upgrade as a function of business group share. Panel A presents results without controlling for project size while Panel B includes size decile fixed effects. Across both panels, the dependent variables are cumulative abnormal returns (CARs) and cumulative excess returns (CERs), and we use several event windows starting from one day before to one day after the announcement of a new plant. To calculate the abnormal returns, we assume a single-factor model, where beta is estimated using the data from the pre-event window. *Abnormal returns* are estimated as the difference between the return on a firm's stock and the return predicted by the capital asset pricing model (CAPM) with the S&P Nifty as the benchmark market portfolio. *Excess returns* are measured as the difference between the return on a firm's stock and the return on the benchmark S&P Nifty index. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. All regressions include firm and industry x year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. Data sources: CMIE Prowess and CapEx database.

	Panel A: Without controlling for project size							
	Cumulative abnormal returns				Cumulative excess returns			
Event window	(-1,1)	(-2,2)	(-3,3)	(-5,5)	(-1,1)	(-2,2)	(-3,3)	(-5,5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PostGQ	0.090***	0.120***	0.184***	0.218***	0.081***	0.102***	0.143***	0.133***
	(0.005)	(0.012)	(0.014)	(0.019)	(0.004)	(0.011)	(0.014)	(0.014)
PostGQ × High BGS	-0.005	0.032	0.021	0.030	0.015	0.052	0.035	0.073
	(0.027)	(0.035)	(0.023)	(0.060)	(0.033)	(0.038)	(0.026)	(0.080)
Fixed effects:								
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry × year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.26	0.24	0.19	0.21	0.23	0.21	0.16	0.16
Observations	1,759	1,759	1,759	1,759	1,759	1,759	1,759	1,759
Sample: Standalone firms	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B: Controlling for project size							
	Cumulative abnormal returns				Cumulative excess returns			
Event window	(-1,1)	(-2,2)	(-3,3)	(-5,5)	(-1,1)	(-2,2)	(-3,3)	(-5,5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PostGQ	0.117***	0.161***	0.226***	0.254***	0.092***	0.125***	0.161***	0.141***
	(0.015)	(0.020)	(0.024)	(0.032)	(0.016)	(0.021)	(0.026)	(0.033)
PostGQ × High BGS	-0.026	0.026	0.013	0.024	-0.015	0.051	0.030	0.064
	(0.019)	(0.026)	(0.028)	(0.065)	(0.025)	(0.030)	(0.026)	(0.074)
Fixed effects:								
Size decile	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry × year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.33	0.31	0.26	0.27	0.31	0.30	0.27	0.29
Observations	1,078	1,078	1,078	1,078	1,078	1,078	1,078	1,078
Sample: Standalone firms	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table VI. Standalone firms' inventory efficiency around GQ upgrades

This table presents estimates comparing days sales of inventory for standalone firms around the investment opportunity shock and as a function of business group prevalence in the local area. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the announcement of the GQ road network upgrades. All regressions include firm fixed effects, High BGS × state × year, and High BGS × industry × year fixed effects. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Days sales of inventory	
	(1)	(2)
PostGQ	-10.479**	-23.713**
	(4.621)	(9.332)
PostGQ × High BGS		-3.397
		(3.096)
Fixed effects:		
Firm	Yes	Yes
High BGS × industry × year	Yes	Yes
High BGS × state × year	Yes	Yes
Adjusted-R^2	0.63	0.63
Observations	10,845	10,845
Sample: Standalone firms	Yes	Yes

Appendix Table VII. Mechanism: Political connections in infrastructure-related industries

This table presents estimates from regressions examining political connections as a plausible mechanism. Panel A focuses on the baseline measure of business group prevalence while panel B focuses on a measure of business group prevalence based on the largest 25 business groups to proxy for political influence. In both panels, column 1 restricts the sample to group-affiliated firms while column 2 restricts the sample to standalone firms. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. *High BGS (Largest 25)* is an indicator variable set to one, if the share of assets of group-affiliated firms from that city that belongs to the 25 largest (by size) business groups, is in the top quartile in the year before the GQ road network upgrades. The sample includes firms operating in "other manufacturing industries" (NIC code: 321-329), "coke and refined petroleum products" (NIC code: 191-199), and "construction firms" (NIC code: 420-439). All regressions include firm and year fixed effects. Due to small number of clusters in these tests, the standard errors are corrected for heteroscedasticity and are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

	Panel A: Baseline	
Dependent variable	Investment	
	(1)	(2)
PostGQ	0.111	-0.095***
	(0.118)	(0.032)
PostGQ × High BGS	0.100	0.007
	(0.050)	(0.048)
Fixed effects:		
Firm	Yes	Yes
Industry × year	Yes	Yes
Adjusted-R^2	0.65	0.74
Observations	524	1,477
Sample	Group-affiliated firms	Standalone firms

	Panel B: Largest 25 groups	
Dependent variable	Investment	
	(1)	(2)
PostGQ	0.111	-0.023
	(0.118)	(0.071)
PostGQ × High BGS (Largest 25)	0.100	-0.080*
	(0.050)	(0.039)
Fixed effects:		
Firm	Yes	Yes
Industry × year	Yes	Yes
Adjusted-R^2	0.65	0.74
Observations	524	1,477
Sample	Group-affiliated firms	Standalone firms

Appendix Table VIII. Robustness: Controlling for pre-GQ firm characteristics

This table reports robustness to controlling for preGQ firm characteristics. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. In column 2, we interact *PostGQ* with whether the firm is *Listed* in the year before the GQ road network upgrades. In columns 3 and 4, we interact *PostGQ* with the median total factor productivity and natural logarithm of firm age before the GQ road network upgrades, respectively. All regressions include firm fixed effects, High BGS x industry x year fixed effects, and High BGS x state x year. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment				
	Baseline (1)	Listed (2)	TFP (3)	Firm age (4)	All (5)
PostGQ	0.039 (0.028)	-0.051 (0.064)	0.003 (0.050)	-0.048 (0.079)	-0.213 (0.143)
PostGQ × High BGS	-0.038*** (0.012)	-0.130** (0.063)	-0.102** (0.042)	-0.122* (0.067)	-0.327** (0.145)
PostGQ × Listed$_i$		0.097 (0.065)			0.153** (0.062)
PostGQ × TFP$_i$			0.005 (0.012)		0.014 (0.013)
PostGQ × Firm age$_i$				0.033 (0.024)	0.018 (0.034)
Fixed effects:					
Firm	Yes	Yes	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.65	0.62	0.65	0.62
Observations	15,842	7,491	5,274	7,491	5,274
Sample: Standalone firms	Yes	Yes	Yes	Yes	Yes

Appendix Table IX. Robustness: High Business Group Share definition

This table reports robustness for the business group share measure we use in our estimations. Column 1 reproduces the coefficients from baseline estimations in Table III. Note that in the baseline estimations *High BGS* is defined as an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. Column 2 repeats the same estimation, however, without the High BGS interaction with Industry × year and State × year fixed effects. Column 3 presents the interaction with the continuous measure while column 4 defines *High BGS* using Hirschman Herfindahl Index (HHI) based on group-affiliated firms' sales at each location. Column 5 presents outlines the quartile specification and lastly, column 6 presents the interaction with terciles of business group share. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment					
	Baseline (1)	No interaction with High BGS (2)	Continuous (3)	HHI (4)	Quartile (5)	Tercile (6)
PostGQ	0.039	0.022	2.719	0.017		
	(0.028)	(0.030)	(2.803)	(0.025)		
PostGQ × High BGS	-0.038***	-0.036**				
	(0.012)	(0.015)				
PostGQ × BGS (continuous)			-0.931**			
			(0.378)			
PostGQ × High BGS (HHI)				-0.072**		
				(0.030)		
PostGQ × Quartile_1					-0.036	
					(0.041)	
PostGQ × Quartile_2					0.177***	
					(0.041)	
PostGQ × Quartile_3					-0.008	
					(0.029)	
PostGQ × Quartile_4					-0.036**	
					(0.029)	
PostGQ × Tercile_1						-0.025
						(0.033)
PostGQ × Tercile_2						0.079***
						(0.026)
PostGQ × Tercile_3						-0.026*
						(0.014)
Fixed effects:						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Industry × year	Yes	Yes	Yes	Yes	Yes	Yes
State × year	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.70	0.70	0.71	0.70	0.70
Observations	15,842	15,842	15,842	13,097	15,660	15,842
Sample: Standalone firms	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table X. Robustness: Drop exiting firms

This table examines the robustness to dropping firms that exit the sample. Column 1 repeats the baseline estimation (Table III) while column 2 repeats the estimation on the subsample of firms exiting the sample if they stop filing annual reports or have been legally struck-off from the business register, identified using data from the MCA. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. All regressions include firm fixed effects, High BGS × industry × year fixed effects, and High BGS × state × year. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess and Ministry of Corporate Affairs.

Dependent variable	Investment	
	Baseline	Drop exiting firms
	(1)	(2)
PostGQ	0.039	0.039
	(0.028)	(0.028)
PostGQ × High BGS	-0.033**	-0.027**
	(0.012)	(0.012)
Fixed effects:		
Firm	Yes	Yes
High BGS × industry × year	Yes	Yes
High BGS × state × year	Yes	Yes
Adjusted-R^2	0.70	0.70
Observations	15,827	15,718
Sample : Standalone firms	Yes	Yes

Appendix Table XI. Robustness: Drop firms with extreme negative sales growth

This table examines the robustness to dropping firms that experience extreme negative sales growth. Column 1 repeats the baseline estimation (Table III) while other columns drop firms in the right tail of the sales growth distribution (10%ile in column 2, 5%ile in column 3, and 1% ile in column 4, respectively). *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. All regressions include firm fixed effects, High BGS × industry × year fixed effects, and High BGS × state × year. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Investment			
	Baseline (1)	Bottom 10 %ile (2)	Bottom 5 %ile (3)	Bottom 1 %ile (4)
PostGQ	0.039 (0.028)	0.029 (0.031)	0.029 (0.032)	0.041 (0.032)
PostGQ × High BGS	-0.038*** (0.012)	-0.105*** (0.030)	-0.099*** (0.019)	-0.039** (0.016)
Fixed effects:				
Firm	Yes	Yes	Yes	Yes
High BGS × industry × year	Yes	Yes	Yes	Yes
High BGS × state × year	Yes	Yes	Yes	Yes
Adjusted-R^2	0.70	0.75	0.73	0.71
Observations	15,842	9,774	12,892	15,210
Sample : Standalone firms	Yes	Yes	Yes	Yes

Appendix Table XII. Mergers and acquisitions of standalone firms

This table changes in merger and acquisition activity whereby standalone firms are target, as function of business group prevalence around GQ upgrades. Column 1 focuses on all years in the sample while column 2 focuses on acquisitions on or after the financial year 2000 due to limited data on transactions in the prior period. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. All regressions include firm fixed effects, High BGS × industry × year fixed effects, and High BGS × state × year. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	$Target_{SA}$	
	Full sample	≥2000
	(1)	(2)
PostGQ	0.041	0.051
	(0.038)	(0.045)
PostGQ × High BGS	-0.009	-0.020
	(0.039)	(0.046)
Fixed effects:		
Firm	Yes	Yes
High BGS × industry × year	Yes	Yes
High BGS × state × year	Yes	Yes
Adjusted-R^2	0.09	0.09
Observations	14,970	14,795
Sample : Standalone firms	Yes	Yes

Appendix Table XIII. Standalone entry at regional-level around GQ upgrades

This table reports examines entry by standalone firms around GQ road network upgrades as a function of business group share. The dependent variable in column 1 is the natural logarithm of new business incorporations at the city-level each year while the dependent variable in column 2 is the inverse hyperbolic sine of new business incorporations at the city-level each year. *PostGQ* is an indicator variable taking value 1 for all years including and after the GQ upgrade in the city. *High BGS* is an indicator variable set to one if the share of assets of group-affiliated firms from that city is in the top quartile in the year before the GQ road network upgrades. All regressions include city fixed effects, High BGS × industry × year fixed effects, and High BGS × state × year. Standard errors are corrected for heteroscedasticity and autocorrelation and clustered at the city level. Standard errors are reported in parentheses.*,**, and *** indicate significance at 10%, 5%, and 1% respectively. All variables are defined in Appendix Table XIV. Data source: CMIE Prowess.

Dependent variable	Log(incorporations)	IHS(incorporations)
	(1)	(2)
PostGQ	0.475**	0.397**
	(0.171)	(0.164)
PostGQ × High BGS	0.273	0.131
	(0.290)	(0.270)
Fixed effects:		
City	Yes	Yes
High BGS × industry × year	Yes	Yes
High BGS × state × year	Yes	Yes
Adjusted-R^2	0.10	0.10
Observations	967	967

Appendix Table XIV. Variable definitions

Variable	**Definition**	**Data source**
A. Firm characteristics		
Firm age	Firm i's age since incorporation.	CMIE Prowess
Cash flow	Ratio of cash flow from operations relative to book value of assets.	CMIE Prowess
Days sales of inventory	Ratio of ending inventory to cost of good solds multiplied by 365.	
Debt	Total outstanding debt from bank and financial institutions relative to book value of assets.	CMIE Prowess
High BGS	Indicator variable set to one if the share of group-affiliated firms' assets from that city is in the top quartile in the year before the GQ road network upgrades.	CMIE Prowess
Investment	Ratio of net fixed assets relative to book value of assets.	CMIE Prowess
Listed	Indicator variable set to 1 if the firm is listed on either the National Stock Exchange (NSE) or the Bombay Stock Exchange (BSE) by the financial year.	CMIE Prowess
Profitability	Earnings before interest, depreciation, taxes, and amortization relative to book value of assets.	CMIE Prowess
Return on assets	Profit after tax relative to book value of assets.	CMIE Prowess
Sales growth	Measured as the annual growth rate of sales.	CMIE Prowess
Size	Measured as the log of book value of assets.	CMIE Prowess
Total factor productivity	Estimation methodology as in Levinsohn and Petrin (2003). Details in Appendix B.	CMIE Prowess
B. Regional characteristics		
Average time to fill skilled worker (manager) vacancy	Average time in weeks to fill vacancy of a manager or a technician.	World Bank Enterprise Survey
Bank branches	Total number of bank branches scaled by the city population as recorded in the Population Census of 2001.	Reserve Bank of India
Bad roadways	Indicator set to 1 if the firm gives a rating of 1 or 2 (1 being worse and 10 being excellent) on the availability of road transport.	World Bank Enterprise Survey
Fraction of listed firms	Share of firms from that city that are listed on either the National Stock Exchange (NSE) or the Bombay Stock Exchange (BSE) by the year before the GQ road network upgrades.	CMIE Prowess
Labor constraint in contracting	Indicator set to 1 if the firm reports constraints in contracting labor.	World Bank Enterprise Survey
Loan amount	The natural logarithm of loan amount in Rs. million.	Ministry of Corporate Affairs
Obstacle to growth (transport)	Indicator set to 1 if the firm gives a rating of 2 ("Moderate obstacle"), 3 ("Major obstacle"), or 4 ("Very severe obstacle") on whether transportation is a problem for the operation and growth of their business.	World Bank Enterprise Survey
Obstacle to growth (labor)	Indicator set to 1 if the firm gives a rating of 2 ("Moderate obstacle"), 3 ("Major obstacle"), or 4 ("Very severe obstacle") on whether availability of skilled and educated Workers is a problem for the operation and growth of their business..	World Bank Enterprise Survey

Appendix D Total factor productivity: Estimation procedure

This section outlines the variables and their definitions that we use for estimating total factor productivity. All numbers are deflated using industry deflators to reflect real values. In estimating firm-level total factor productivity, we include firm size as a control variable.

Output: Value of total sales that includes income earned by the company from the sale of industrial goods as well as their raw materials, byproducts, stores and waste.

Capital: Gross fixed assets of a firm that includes both tangible assets, such as land, building, plant, and machinery, and intangible assets, such as goodwill assets, software, etc.

Labor: Compensation to employees that includes all cash and payments in kind made by a company to its employees.

Intermediate inputs: Combined value of raw materials, power and fuel consumptions. Raw materials are the sum of expenses on raw materials, stores, spares and tools used up by firms in the production process. Power and fuel include expenses made by the firms on power, fuel and water. The sum of these three variables is used as the proxy in the estimation of the production function.