Journal of International Economics 159 (2026) 104201

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Full length articles

Using satellite imagery to measure the impacts of new highways: An application to India[☆,☆☆]

Kathryn Baragwanath[a], Gordon Hanson[b,c], Amit K. Khandelwal[d,c,*], Chen Liu[e], Hogeun Park[f]

[a] University of Melbourne, Australia
[b] Harvard, United States of America
[c] NBER, United States of America
[d] Yale, United States of America
[e] NUS, Singapore
[f] World Bank, United States of America

A R T I C L E   I N F O

Dataset link: Using Satellite Imagery to Measure the Impacts of New Highways: An Application to India (Original data)

Keywords:
Satellite imagery
Infrastructure
Market access

A B S T R A C T

This paper integrates daytime and nighttime satellite imagery into a spatial general-equilibrium model to evaluate the returns to investments in new motorways. Our approach has particular value in developing-country settings in which spatially granular economic data are scarce. To demonstrate our method, we use publicly available imagery to evaluate India's road construction projects in the early 2000s. Estimating the model and evaluating welfare impacts only requires remotely-sensed data. We find that India's road investments improved aggregate welfare, particularly for the largest and smallest urban markets. Welfare gains were unevenly distributed across space, with the Golden Quadrilateral disproportionately benefiting large, already connected markets, while national highway upgrades delivered greater gains to smaller, more remote locations. More generally, within-district variation explains a large share of the overall spatial variance in welfare changes during this period, underscoring the value of high-resolution satellite imagery for capturing localized economic effects.

1. Introduction

This paper integrates daytime and nighttime satellite imagery to analyze the economic impacts of infrastructure projects. Our approach combines high-resolution imagery with analytical tools refined in the quantitative spatial economics literature (Redding and Rossi-Hansberg, 2017). Our goal is to provide a tractable framework for studying both the aggregate and highly local effects of infrastructure investments that would be applicable in contexts in which conventional data sources are limited or unavailable, for instance in many low-income countries.

Motivating our work is a multi-decade investment boom in transportation infrastructure in emerging and developing economies. Since the late 1990s, India has constructed 6000 km of new expressways to connect its largest cities as part of the Golden

☆ This article is part of a Special issue entitled: 'Trade and Uneven Development' published in Journal of International Economics.

☆☆ We thank the editor, Ben Faber, and two anonymous referees, Treb Allen, Devaki Ghose and participants at the World Bank/JIE Trade and Uneven Development Conference and the 2019 Banco de Mexico Conference on International Economics for valuable comments. We acknowledge funding from the Center on Global Transformation at UC San Diego, the World Bank, and the International Growth Centre, United Kingdom (Project 89448).

* Corresponding author at: Yale, United States of America.
E-mail address: amit.khandelwal@yale.edu (A.K. Khandelwal).

https://doi.org/10.1016/j.jinteco.2025.104201
Received 8 October 2024; Received in revised form 3 December 2025; Accepted 4 December 2025
Available online 13 December 2025

Quadrilateral (GQ) network, while upgrading 19,000 km of minor roads to national highways and building an extensive network of rural roads (ADB, 2017). India's investment boom has recently accelerated, with the country currently adding 10,000 km of highways *each year* (The Economist, 2022). China's efforts are even more ambitious, involving over 40,000 km of new motorways within China and a Belt-Road Initiative that is creating new transport links throughout Asia and Africa. Since 2000, the World Bank has devoted one-fifth of its total lending to transportation infrastructure, with much of this funding going to low-income countries (Woetzel et al., 2017).

A major challenge in studying these investments is the coarseness of conventional data. To evaluate impacts within a nation, one is typically left to work with census samples, which tend to appear with long lags and to limit geographic identifiers to municipalities or districts. Yet, in theory, the impacts and distributional consequences of investments are highly localized (Allen and Arkolakis, 2022). The high-resolution imagery we exploit lets us evaluate infrastructure investments from the national to the granular level. As a demonstration case, we study the expansion of India's highways during the late 1990s and 2000s. Because this episode has been studied using conventional data at lower spatial resolution (Datta, 2012; Ghani et al., 2014; Asturias et al., 2019; Alder, 2016), we are able to benchmark our results using satellite imagery with results using standard approaches.

Our analysis proceeds in four steps. First, we identify the footprints of local markets in India using the approach developed in Baragwanath et al. (2019), which applies an algorithm to define the geographic unit of analysis—a local market—as a cluster of adjacent pixels whose properties in daytime multi-spectral satellite imagery indicate the presence of builtup landcover. We measure economic activity in these granular markets using the intensity of nighttime lights captured by DMSP-OLS satellite sensors. The result is data on economic activity in 2001 and 2011 for 13,386 markets, which are either small towns unto themselves or the smaller cities that comprise a larger metropolitan area. Within the metropolitan borders of New Delhi, for example, we identify 579 local markets. In 2001, India's urban markets collectively contained 34.8% of the national population. As a comparison, World Development Indicators reports India's 2001 urbanization rate at 27.9%.

Next, we measure travel times between each pair of markets by digitizing official road maps for India in 1996, prior to the highway construction, and 2011, by which point most construction from this phase was complete.[1] The maps, which identify five categories of intercity roadways from national highways to simple roads, reveal that India's recent highway construction was much more extensive than the Golden Quadrilateral (GQ), the national artery connecting Delhi, Mumbai, Chennai, and Kolkata, and the primary focus of earlier studies. The GQ accounted for just one-fifth of new highway miles created by road upgrades or new construction between 1996 and 2011. A comprehensive analysis of India's road expansion should thus include the full extent of the intercity road construction during this period. Using Dijkstra algorithm to calculate the minimum travel time between each market pair, the new roads reduced travel times for 200–500 km distances on average by about 1 h.

Third, we demonstrate how to apply satellite imagery to an otherwise standard quantitative spatial model (Redding and Rossi-Hansberg, 2017; Allen and Arkolakis, 2023). The procedure requires calculating the market access function for each market in each year, using our data on market boundaries, nightlights, and bilateral travel times, and estimates of the trade cost elasticity and the elasticity of income with respect to nightlights (Henderson et al., 2012). We then estimate the impact of changes in market access on market income, accounting for endogenous placement of the roads using a least-cost-spanning-tree instrument (Faber, 2014). Finally, we use our parameter estimates to construct counterfactual changes in welfare based on alternative highway construction plans, which requires that we allocate population across our local markets in one year of our data. To do so, we use publicly available population rasters from the Gridded Population of the World for the year 2010.

We find that improved market access from road construction over 1996 to 2011 increased the net present value of Indian GDP by 1.66%. Nearly all of this increase comes from the construction of the GQ. Although the expansion of non-GQ national highways increased real GDP by 0.62%, this gain was offset by construction costs, leaving a small negative net impact. Impacts are far from uniform across markets: we uncover a monotonic pattern where the (initially) larger markets experienced larger gains than small and medium size markets. Welfare rose by 1.8% for markets in the first quartile in terms of size (as measured by the land area), followed by gains of 2.0% for markets in the second quartile, 2.0% for markets in the third quartile, 2.1% for markets in the fourth quartile, and, notably, of 2.2% and 3.3% for the top 5 percentile markets and top 0.1 percentile markets, respectively.

We then decompose the variance in welfare gains across different levels of spatial aggregation. 41.6% of the overall spatial variance in welfare effects is explained by within-district variation–this is variation that would be missed if the analysis were conducted solely at the district level. Satellite imagery thus reveals highly localized impacts of national infrastructure investment that would not be detectable using India's district-level administrative data.

When we disentangle the welfare impacts from the construction of the GQ versus the upgrading of national highways, we find opposing results. While the GQ disproportionately benefited the largest markets in terms of land area, the gains from national highway upgrades are the smallest among the largest markets. The GQ produced a welfare gain of 0.9% for the first quartile of markets, compared to a welfare gain of 2.6% for the top 0.1 percentile markets. By contrast, the expansion of national highways produced a welfare gain of 0.8% in the first quartile of markets, and a smaller 0.5% gain in the top 0.1 percentile markets.

These patterns are corroborated when we classify markets by population density, baseline remoteness, and agricultural intensity. Welfare gains from the NH network are concentrated in more remote locations, consistent with its role in extending access to previously disconnected markets. In contrast, the GQ primarily benefited larger, more urban markets which were initially closer to roads. Effects are largely invariant to agricultural intensity, suggesting that the primary mechanism operates through urban market integration rather than through improvements in agricultural surplus extraction.

[1] Surprisingly, we were unable to obtain digitized road maps for India in the 1990s or 2000s. We digitized official road maps using PDFs of the decadal Survey of India, with help from a digital mapping company.

Our work contributes to an expanding literature that applies satellite imagery to the study of roads and economic growth. Storeygard (2016) uses nightlights to estimate the impacts of road networks across 289 metropolitan regions in 15 Sub-Saharan Africa countries; and Jedwab and Storeygard (2020) examine spatial impacts of road construction in 3000 larger cities across 39 Sub-Saharan African countries. Relative to this work, our framework detects impacts for much finer geographic units and embeds the analysis in a quantitative spatial model; relative to broader literature in economics that uses satellite imagery (Henderson et al., 2012; Donaldson and Storeygard, 2016), we demonstrate how the *combination* of nightlight data, which captures the intensive margin of economic activity, with multi-spectral output from daytime images, which defines the extensive margin of regional markets, can be leveraged for analysis at very high spatial resolutions.[2]

Road construction in India has attracted much scholarly interest. In related work, Alder (2016) supplements administrative data with nightlights to study the impact of the GQ on Indian districts and finds that aggregate real GDP (net of construction and maintenance costs) would have been 0.8 percent lower in 2012 if the GQ had not been built.[3] Our framework, which arrives at similar aggregate impacts, reveals substantial within-district variation in gains. Other work finds positive impacts of Indian highways on manufacturing income and productivity (Datta, 2012; Asturias et al., 2019), and farm income Allen and Atkin (2022), while adjacent studies examine how the expansion of rural roads has affected Indian villages (Aggarwal, 2018; Asher and Novosad, 2020). Because we detect built-up towns and cities and not unilluminated hamlets, our approach is not well suited for the analysis of rural villages.[4]

Beyond the economic consequences of road construction in India, our work demonstrates how to apply satellite imagery to "off-the-shelf" spatial quantitative models, which have been developed and refined in recent years (Allen and Arkolakis, 2014; Ahlfeldt et al., 2015; Monte et al., 2018). In so doing, our goal is to provide a low-cost, tractable, and rigorous framework for spatial impact analysis in general equilibrium. It is difficult to imagine studying transnational infrastructure investments, such as the Belt-Road Initiative, using administrative data. Moreover, unlike conventional statistical sources, satellite imagery is available on a nearly real-time basis, which allows one, for instance, to evaluate India's current road investments. Naturally, our approach cannot assess outcomes for individual units within a market, such as firms or consumers. Overall, our results suggest that remotely-sensed data are useful for assessing national policies in data-poor environments.

2. Data

2.1. Detecting markets with satellite imagery

We use satellite imagery to construct markets by aggregating across adjoining pixels that meet given thresholds for economic activity. Baragwanath et al. (2019) develop a clustering algorithm that defines a boundary of a market as the set of contiguous or near-contiguous pixels that contain built-up economic activity as indicated by their day-time spectral properties. We apply this method to detect markets in India, and measure the intensive margin of economic activity of these markets through nightlight intensity. We summarize the procedure in Appendix A, and refer the reader to Baragwanath et al. (2019) for more details.

Market boundaries

To define the boundary of markets, we rely on builtup landcover from the Moderate Resolution Imaging Spectroradiometer (MODIS) 500 m resolution layer constructed by Friedl and Sulla-Menashe (2015), a publicly available dataset widely applied in remote sensing. We use their "Urban and Builtup" pixels to indicate builtup landcover for 2001 (Sulla-Menashe and Friedl, 2018). We assign adjoining pixels with builtup landcover to a cluster, and form the market boundary by combining proximate clusters whose boundaries are within 1 km of each other. Finally, in order to capture possible changes in market footprints over time, we draw 0.5 km buffers around the markets. This process creates 13,386 markets which cover 2.7% of India's surface area. Markets have mean and median areas of 6.5 km^2 and 3.4 km^2. The standard deviation across market sizes is 23.9 km^2, and the largest market – New Delhi – has a boundary of 1607.6 km.

Baragwanath et al. (2019) also generate larger "super-markets" based on an coarser buffers that nest disaggregate markets defined at smaller buffers: 1 km markets $\subset$ 4 km markets $\subset$ 8 km markets. These super-markets represent data-driven concentrations of economic activity, as opposed to those based on administrative district borders, which can be highly idiosyncratic in size. Relative to the 13,386 1 km markets, we identify 7,5394 km markets, and 3483 8 km markets. As a reference, the 2011 India Census recognizes 6171 "towns" that were home to India's 377 million urban residents (31% of total population). Table A.1 provides summary statistics for the 1 km, 4 km and 8 km markets. Our unit of analysis will be the 1 km markets, but we will also be interested in assessing welfare gains within and across the coarser market buffers in order to evaluate the analytical value of having observations at higher versus lower spatial resolutions.

To illustrate our approach, Fig. 1 shows local markets around the Ahmedabad district. In 2011, these markets contained 7.1 million inhabitants in the overall district (shown in gray) and 5.6 million inhabitants in the Ahmedabad metropolitan area, as

[2] The use of daytime satellite imagery in spatial economic analysis, which remains less common than the use of nightlights, was pioneered by Burchfield et al. (2006) in their study of urban sprawl.

[3] Alder (2016) studies 612 of India's 639 districts in 2001 compared to the 532 districts that contain detectable urban markets in our analysis.

[4] See Redding and Rossi-Hansberg (2017), Allen and Arkolakis (2023) for a discussion of the of quantitative spatial models and Coşar et al. (2021) for a recent application to Turkey.

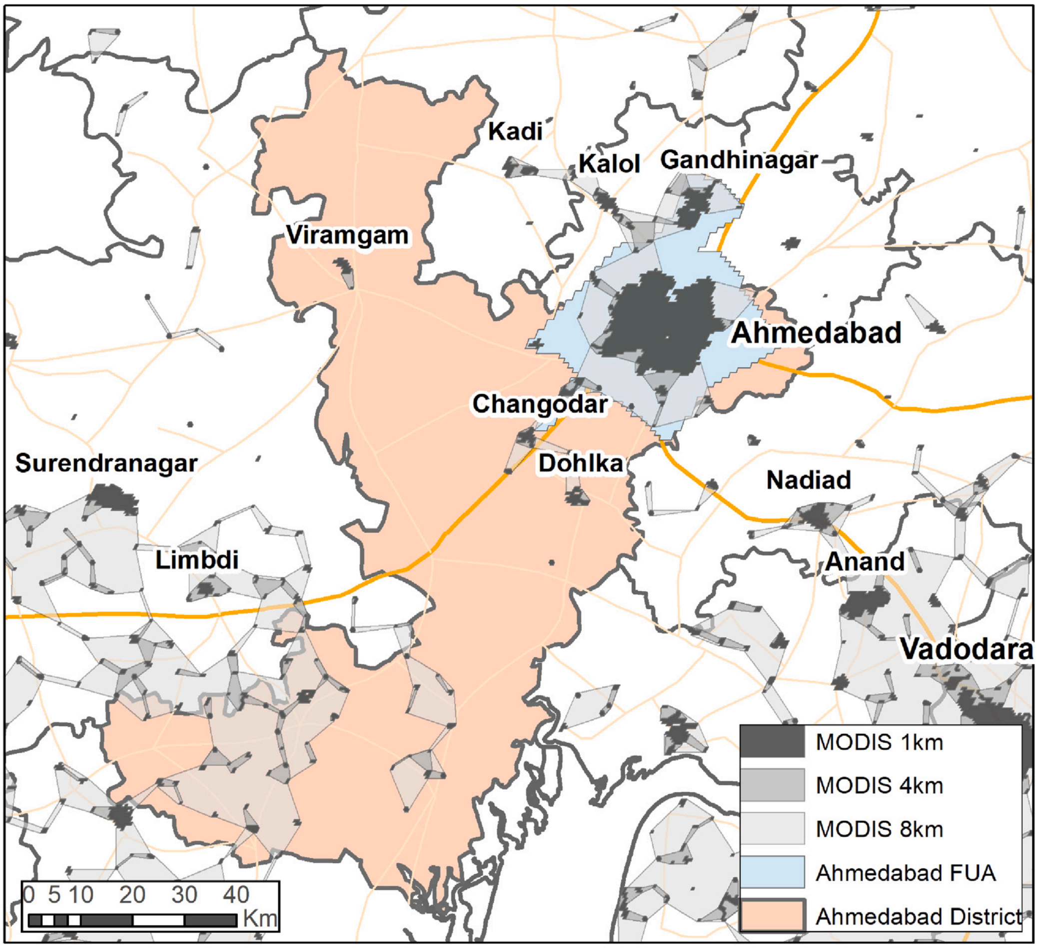


Fig. 1. Satellite-detected markets in Ahmedabad FUA, 2001.
Notes: The figure shows markets in the FUA and district of Ahmedabad. Major highways as of 2011 are shown in bold lines; minor roads as of 2011 are shown in faint lines. Within the figure, we observe 383 1 km markets, 183 4 km markets and 42 8 km markets.

measured by the Functional Urban Area (FUA) framework.[5] Local markets are shown in black. The main city of Ahmedabad is evident, as are its principal satellite cities of Gandhinagar, Kalol, Kadi, Changodar, and Dohlka, and many smaller satellite towns and cities, some of which lie inside and some of which lie outside district boundaries. Our method detects 69 markets within the district of Ahmedabad and a total of 383 1 km markets in the image.

Market economic activity

We measure economic activity within each market using the intensity of light emitted at night from the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS), which detects visible and near-infrared light (Dingel et al., 2019; Elvidge et al., 1999). DMSP-OLS pixels have a spatial resolution of 1 km. For each market, we identify the pixels that overlap with its borders and aggregate light intensity across these pixels to create a measure of economic activity for the market.[6] Appendix A discusses various approaches to address well-known issues with DMSP-OLS data: bottom-coding, top-coding, and blooming in nightlight data.

[5] We adopt the Functional Urban Area framework to define the area of a metropolis across India-defined as a city and its surrounding commuting zone-based on population density and commuting patterns, following the methodology developed by the EU Joint Research Centre (JRC) (Schiavina et al., 2019) (shown in light-blue).

[6] While we focus on urban economic activity, one limitation of this approach is that it does not pick up changes in GDP from agriculture, except through its byproduct of lights; for example, from income derived from agricultural activity that appears through nightlight activity and built-up areas.

Although we do not directly observe GDP at the level of our 1 km markets, we have GDP data at the district level from the Indian Planning Commission in 2001,[7] which allows us to determine the relationship between nightlights and GDP for these geographic units. Figure A.1 plots nightlights against district GDP (left panel), revealing a strong positive correlation. We take this as supportive evidence of our use of nightlight intensity to proxy for local economic activity. To estimate a precise connection between nightlights and economic activity, we assume the following relationship between log GDP, $\log y_{rt}$, log nightlight intensity, $\log \mathrm{N}_{rt}$,

$$\log y_{rt} = C + \alpha \log \mathrm{N}_{rt} + \upsilon_{rt} \tag{1}$$

where υ_{rt} is an i.i.d. disturbance associated with measurement error in the use of nightlights to capture GDP.

We estimate Eq. (1) using different sources of variation. First, we obtain Indian district-level data from 2001 and 2005 from Nielsen India, covering over 500 districts in each year. Due to the lack of a long panel for district-level GDP, we treat the data as cross-sectional and regress log GDP on log nightlights, controlling for state and year fixed effects. This yields a coefficient of 0.51 for log nightlights, a constant term of 3.26, and an R^2 of 0.84.

For state-level data, we have a longer panel covering the years 2001 to 2011 (Ministry of Statistics and Programme Implementation of India). We estimate Eq. (1) using log changes in GDP and nightlights across 31 states. This specification yields a coefficient of 0.10 for log nightlights, a constant of 1.19, and an R^2 of 0.26.

While panel data is often preferred for growth estimation, we are cautious in relying solely on short-difference specifications due to well-documented concerns about the low signal-to-noise ratio in DMSP-OLS nightlights in short panels (Michalopoulos and Papaioannou, 2013; Gibson et al., 2021; Bleakley and Lin, 2012). Given these limitations, we simply average the elasticity estimates from the cross-sectional district and panel-based state specifications.

We then use the estimated coefficients to project GDP for each 1 km market-year:

$$y_{rt} = \exp\left(2.22 + 0.303 \log \mathrm{N}_{rt}\right). \tag{2}$$

Market population

A key contribution of this paper is to demonstrate that satellite imagery can be used to estimate the elasticity of GDP to changes in market access. In order to convert this elasticity to welfare impacts, we further need an estimate of market population (in at least one year, either at the baseline or endline). This is because nightlights provide an estimate of GDP, whereas welfare in standard spatial models rests on changes in real GDP per person. Population data are therefore necessary to separately pin down market changes in population versus income.

Population data are publicly available across the globe in datasets such as the Gridded Population of the World - GPW (Center for International Earth Science Information Network - CIESIN - Columbia University, 2016).[8] GPW population data are available at a spatial resolution of 30 arc-seconds (approximately 1 km at the equator), for the years 2000, 2010, 2015, and 2020. We use data from year 2010 in our exercise.[9] We process the GPW data in Google Earth Engine to obtain the total number of people in each market by adding the population inside each pixel that overlaps with a given market. Table A.1 reports that the average number of people living inside a 1 km market increased from 8.9 thousand in 2000 to 10.4 thousand in 2010. As a validation, the largest 1 km market in our sample, (corresponding to New Delhi and covering 1,607.6 km^2) was assigned a population of 13.1 million in 2000 and 16.0 million in 2010, while the census reported population of New Delhi was 13.8 million in 2001 and 16.8 million in 2011.

2.2. *India's highway system*

2.2.1. *National highway development project*

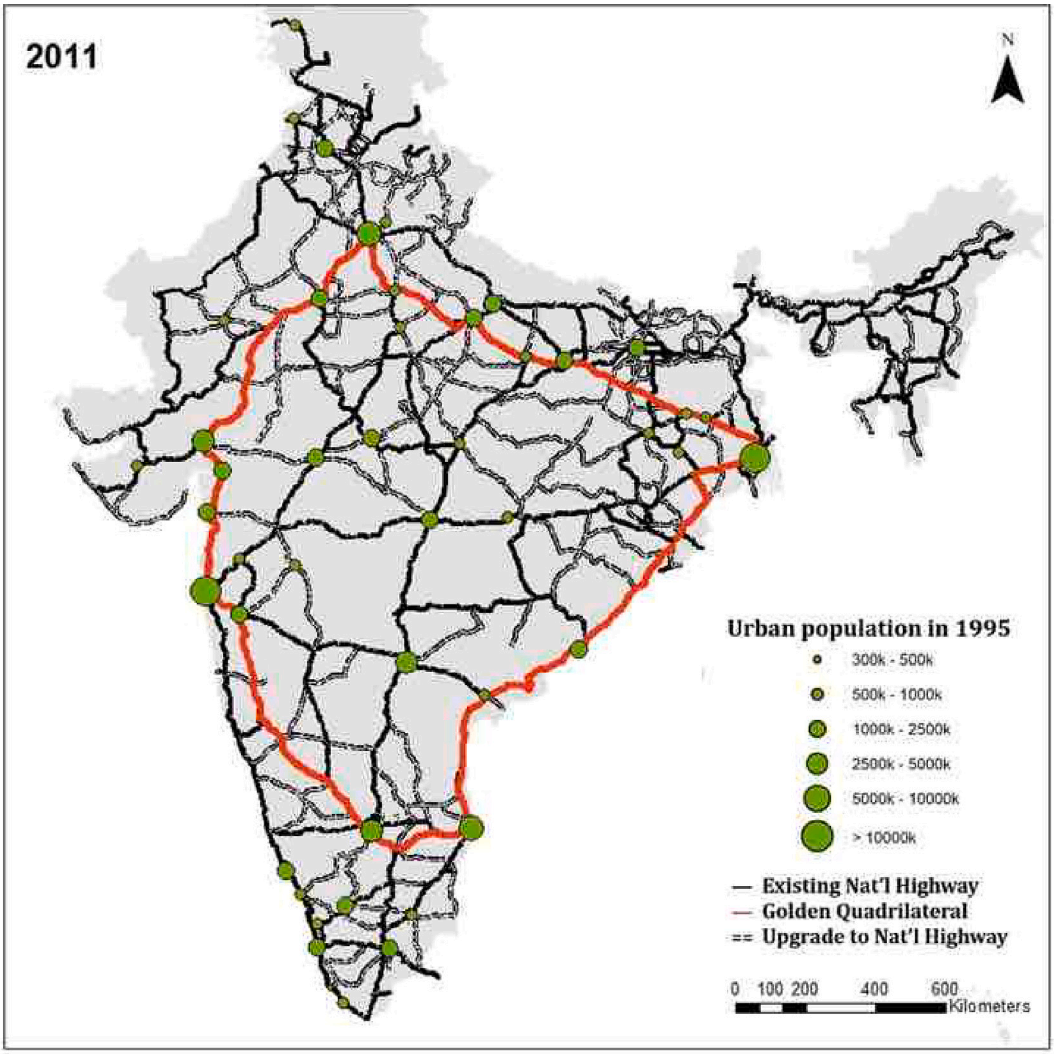
The Golden Quadrilateral was the centerpiece of India's National Highways Development Project (NHDP). Commencing in 2001 and reaching completion by 2012, the GQ upgraded existing roads to four or six-lane expressways in a circular route connecting New Delhi, Mumbai, Chennai, and Kolkata. The NHDP also upgraded many more roads to national highways and constructed other roadways de novo. To detect highway upgrades and new road construction, we digitized road maps from the Survey of India, which, prior to 2015, mapped intercity roads for all of India on a roughly decadal basis. We also digitized maps for 1996 and 2011, which span a period from five years before the NHDP to just before its end. Appendix B provides more details explaining the road digitization process and validation.

Fig. 2 shows the results of our digitization exercise. The GQ appears in red, new highways (whether upgrades or new construction) appear as dashed lines, and existing national highways appear in black. Other roads appear as faint lines. The newly constructed portions of the grid are more dense in India's north, east, and south, and less dense in the center. The figure also shows urban agglomerations with populations above 300,000. To a first approximation, the existing grid connected India's largest cities

[7] India does not produce GDP figures at the district level or sub-district level on a regular basis. We could only obtain estimates for 2001. Of India's 639 districts in 2001, 532 districts overlapped with one of our markets in 2001. Of these, we have GDP estimates for 326 districts.

[8] We choose GPW because it is the least modeled population dataset compared to alternatives like GHS-POP, GRUMP, and LandScan. While other datasets are developed using dasymetric mapping methods using other geographical data (e.g., built-up proportion, land cover), GPW relies solely on areal weighting. The primary advantage of disaggregating demographic variables using this method is avoiding potential conflicts with ancillary datasets, such as nightlights, that may be endogenous to more highly modeled surfaces.

[9] Previously we had used population data from the 2011 Indian Census rather the Gridded Population of the World (GPW). The results were similar to those presented in the current version. Specifically, we observed a total increase in the Net Present Value (NPV) of GDP by approximately 1.08% (compared to 1.60% in the current version), primarily driven by gains from the construction of the GQ, with a minor negative NPV attributable to NH expansion. The decision to transition to the GPW dataset was motivated by the desire to employ a model that leverages only public (and globally available) data.



Total Road Length by Road Type, India 1996 to 2011

Category of Road	1996	2011
Golden Quadrilateral		5,714
National Highway	29,422	49,172
All Weather Motorable	165,617	164,022
Motorable in Fair Weather	15,353	13,067
Where delay may occur	2,838	3,625
Total	213,230	235,600

Note: Table reports road construction estimates from digitized road maps from the 1996 and 2011 Survey of India

Fig. 2. Highway construction in India, 1996 to 2011.
Notes: The figure presents: (i) a map of India in 2011, in which lines represent highways (solid black for existing national highways, solid red for the Golden Quadrilateral, dashed black for upgrades to national highways), and green dots represent major cities (dot size is proportional to city population in 1995), and (ii) a table showing total road length by type of road in India in 1996 and 2011, in kilometers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(populations above 1 million), while new portions of the grid connected smaller cities. As shown in Fig. 1, within an individual region (in this case, Ahmedabad), there is considerable heterogeneity in access to roads according to market size. Fig. 2 further summarizes our measures of India's highway and intercity road construction. The GQ added 5714 km of new expressways, and national highways expanded by 19,750 km. Although the total increase in road mileage between 1996 and 2011 was relatively modest—the 22,371 km of additional roads represented only a 10% increase in total road length—the quality of roads improved significantly.[10] India's length of highway-quality roads increased by a substantial 86.5% over the period.[11]

2.2.2. Travel time between markets

The increase in road miles understates the full benefits of the construction through improved speed, as new highways have two to six lanes, whereas the roads being upgraded had one to two lanes. To translate road construction into changes in transport costs, we estimate changes in the time required to travel between each pair of markets in India (see Appendix B). Travel may occur along a portion of the GQ (four to six lanes, paved), a national highway (two to four lanes, paved), and lesser road types (one to two lanes, paved, semi-paved, or unpaved). As a starting point, we use estimates in Allen and Atkin (2016), which indicate average speeds of 20mph for the lowest-quality roads (roads where delays may occur) and 60mph for top-quality roads (GQ). For intermediate roadways, we assume speeds of 40mph for national highways, 30mph for all-weather motorable roads, and 20mph for fair-weather motorable roads.

Because the majority of markets are small, most do not have an identifiable road going directly through them. The median market in our data is 2.4 km away from an identifiable road in 1996 (interquartile range 3.9 km) and 2.2 km away from an identifiable road in 2011 (interquartile range 3.4 km). Because we cannot detect connecting roads to these markets, we assume that they follow the minimum distance from the market centroid to the nearest road, assuming a speed of 20mph.

Given assumed travel speeds, we estimate the minimum travel time between each pair of markets by implementing the Dijkstra algorithm (Dijkstra, 1959), which finds the shortest path from a starting node to a target node through a weighted graph, which is a data structure composed on a set of nodes and edges (Donaldson and Hornbeck, 2016; Donaldson, 2018; Kruskal, 1956) (see Appendix B). Between 1996–2011, Table B.1 indicates that travel times between markets 50–200 km apart fell by about 0.5 h, travel times between markets 200–500 km apart fell by about 1 h, and travel times between markets more than 500 km apart fell by more than 3 h.

3. Quantitative framework

We implement a standard spatial general equilibrium model, which embodies the market-access approach (see, e.g., Donaldson and Hornbeck, 2016; Bartelme, 2018). As recently noted by Allen and Arkolakis (2023), these spatial models are characterized by a system of regional consumer demand and labor supply equations. Consumer demand has a CES structure, and labor supply derives from workers' idiosyncratic preferences across locations.[12] Our model considers the spatial linkages in goods and worker mobility across these markets. However, it does not explicitly differentiate between workers' employment locations and residential choices, so commuting linkages are excluded from our analysis.

We present estimating equations and welfare formulas based on this framework. We denote the trade elasticity parameter as θ and the inverse labor supply elasticity as μ. For complete details, see Appendix C.

3.1. Equilibrium conditions in market access

The framework delivers two expressions that link log changes in wages (w) and population (L) to log changes in a region r's market access:

$$\hat{w}_{rt} = \varepsilon_w \hat{\Phi}_{rt} + \hat{\chi}^w_{rt} \qquad \hat{L}_{rt} = \varepsilon_l \hat{\Phi}_{rt} + \hat{\chi}^l_{rt} \tag{3}$$

where $\hat{x}_t \equiv \ln x_t - \ln x_{t-1}$ denotes log changes, ε_w and ε_l represent the reduced-form elasticities, each of which depends on structural parameters θ and μ. $\hat{\chi}^w_{rt}$ and $\hat{\chi}^l_{rt}$ are functions of unobserved regional amenities and productivity shocks.

The change in market access $\hat{\Phi}_{rt}$ is determined by:

$$\hat{\Phi}_{rt} = \ln\left(\sum_n \frac{w_{nt} L_{nt}}{\Phi_{nt}} \tau^{-\theta}_{rn,t}\right) - \ln\left(\sum_n \frac{w_{n,t-1} L_{n,t-1}}{\Phi_{n,t-1}} \tau^{-\theta}_{rn,t-1}\right) \tag{4}$$

where $\tau_{rn,t}$ is the symmetric iceberg trade cost between regions r and n at time t.[13] The level of market access at year t is solved from the following system of R equations,

$$\Phi_{rt} = \sum_n \frac{w_{nt} L_{nt} \tau^{-\theta}_{rnt}}{\Phi_{nt}}. \tag{5}$$

[10] Most of the improvements came from upgrading the quality of existing roads to highway-quality roads, not the construction of new roads that would add to total road length. This is why the 22,371 km of new roads does not coincide with the 5714 km added by the GQ plus the 19,750 km added by the expansion of national highways.

[11] Total highway-quality roads (which include the GQ and other national highways) in 2011 was equal to 5714 km + 49,172 km = 54,886 km while total highway quality roads in 1996 was equal to 29,442 km (see Fig. 2 for more details).

[12] Due to location-specific idiosyncratic tastes, in equilibrium, the real wage (or welfare) $\frac{w_n}{P_n}$ does not necessarily need to equalize across markets. It equalizes across markets as the labor supply elasticity ($1/\mu$) approaches infinity; in this case, idiosyncratic tastes are effectively the same across all workers and the model collapses to the setting where workers are perfectly mobile across markets.

[13] Given symmetric trade costs, consumer market access and firm market accesses are equal up to a factor of proportion, which simplifies the analysis (Donaldson and Hornbeck, 2016).

In order to compute market access for each region, Eq. (5) reveals that we need data for $w_{nt}L_{nt}$, which is local GDP predicted as a function of nighttime lights, and the iceberg trade cost composite, $\tau_{rnt}^{-\theta}$, which following the literature we assume is a function of travel times (see Redding and Rossi-Hansberg, 2017; Monte et al., 2018; Allen and Arkolakis, 2022):

$$\tau_{rn}^{-\theta} = \left(\text{travel time}_{rn} + \frac{\kappa \times (d_r + d_n)}{\text{speed}}\right)^{-\phi}. \tag{6}$$

Here, ϕ measures the elasticity of iceberg trade costs with respect to travel time. We set $\phi = 1.5$, following Redding and Rossi-Hansberg (2017). The variable d_r is the last miles traveled that connect market r to observed roadways, and d_n is the last miles traveled that connect market n to the observed roadways. We set $\kappa = 0.000621$ which converts meters into miles. We set last-mile travel speeds equal to 20mph.

3.2. The estimating equation

Assuming that GDP equals labor income, $\hat{y}_{rt} \equiv \hat{w}_{rt} + \hat{L}_{rt}$, Eq. (3) implies

$$\hat{y}_{rt} = \delta\hat{\Phi}_{rt} + \hat{\Gamma}_{rt}. \tag{7}$$

where $\delta \equiv \left(\varepsilon_w + \varepsilon_l\right)$ and $\hat{\Gamma}_{rt} = \hat{\chi}_{rt}^{w} + \hat{\chi}_{rt}^{l} + \hat{v}_{rt}$. For the dependent variable, Eq. (1) implies that the changes in GDP can be measured as

$$\hat{y}_{rt} = \alpha\hat{\mathrm{N}}_{rt} + \hat{v}_{rt}, \tag{8}$$

where we set $\alpha = 0.303$ based on estimation results for Eq. (2). The independent variable, the change in market access, $\hat{\Phi}_{rt}$, is derived by substituting y_{rt} for $w_{nt}L_{nt}$ in Eq. (5), solving the system of R equations in R unknowns in (5) for Φ_r at both time t and time $t-1$, and then taking the time difference between these two values.

3.3. Identification

The change in market access contains GDP and trade costs in period t, which embody endogenous responses to shocks between $t-1$ and t. This requires an instrument to identify δ in Eq. (7). We follow Faber (2014) by constructing an instrument based on a least-cost approach to connect markets:

$$\hat{\Phi}_{rt}^{\mathrm{mst}} = \ln\left(\sum_{r'} \frac{y_{n,t-1}}{\Phi_{nt}^{mst}} \tau_{rn,t}^{\mathrm{mst}}\left(\Omega_{\prime t-1}\right)^{-\theta}\right) - \ln\left(\sum_{n} \frac{y_{n,t-1}}{\Phi_{n,t-1}} \tau_{rn,t-1}^{-\theta}\right). \tag{9}$$

This instrument uses GDP at time $t-1$ to calculate market access in both time periods. Additionally, bilateral trade costs, $\tau_{rn,t}^{mst}\left(\Omega_{\prime t-1}\right)$, are constructed by a least-cost spanning tree algorithm to connect India's largest cities using information as of time $t-1$, denoted as $\Omega_{\prime t-1}$. See Appendix B for details. We use India's 180 largest cities in the baseline (cities with population above 100 thousand in 2000). We find similar estimates when using India's 100 largest cities (cities with a population above 350 thousand in 2000), as reported in Appendix D.

The value Φ_{nt}^{mst} is given by the solution to the $R \times R$ system of equations,

$$\Phi_{rt}^{\mathrm{mst}} = \sum_{n} \frac{y_{n,t-1}}{\Phi_{n,t}^{\mathrm{mst}}} \tau_{rn,t}^{\mathrm{mst}}\left(\Omega_{\prime t-1}\right)^{-\theta}. \tag{10}$$

The instrument maps roads a central planner would have constructed if the only objective was to connect all targeted points (i.e., India's 100 or 180 largest cities) in a network subject to global cost minimization (Faber, 2014). The identifying assumption for this instrument is that the location of non-targeted peripheral markets along the least cost spanning tree network affects changes of market-level economic outcomes only through new highway connections, conditional on district fixed effects.

3.4. Equilibrium changes in welfare

To perform counterfactual exercises, we solve the model in changes with 2011 as the baseline (Dekle et al., 2008).[14] We use $\hat{X}$ to denote the ratio between the counterfactual and the actual economy in 2011 of variable X (and henceforth suppress the subscript t). These exercises call for additional data on the population of each market and a parameter value for θ, which we obtain as described above.

Given data on local population L_r, the market access value Φ_r in 2011, our measure of market-level GDP, and counterfactual values of trade costs τ'_{rn}, we solve the counterfactual market access in changes $\hat{\Phi}_r$ from the following system of equations:

$$\hat{\Phi}_r\Phi_r = \left[\sum_n \hat{\Phi}_n^{\frac{1}{\mu}\left(\varepsilon_w + \frac{1}{\theta}\right)} \frac{L_n}{L}\right]^{-\frac{\theta}{\theta+1}} \left[\sum_n \frac{\hat{\Phi}_n^{\varepsilon_w+\varepsilon_l} y_n}{\hat{\Phi}_n\Phi_n} \tau'^{-\theta}_{rn}\right], \tag{11}$$

[14] Regarding whether to use the first or last period for the baseline, existing literature is agnostic. Caliendo and Parro (2015) calibrate the model to the beginning period (1996 in our case), whereas Adao et al. (2017) calibrate the model to the end period (2011 in our case).

Appendix C.6 explains the derivation. Briefly, in doing so, we first derive $\hat{w}_r$ and $\hat{L}_r$ as a function of changes in market access, $\hat{\Phi}_r$ as follows

$$\hat{w}_r = \hat{\Phi}_r^{\varepsilon_w} \left[\sum_n \hat{\Phi}_n^{\frac{1}{\mu}(\varepsilon_w + \frac{1}{\theta})} \frac{L_n}{L} \right]^{\frac{1}{\theta+1}}, \qquad \hat{L}_r = \hat{\Phi}_r^{\varepsilon_l} \left[\sum_n \hat{\Phi}_n^{\frac{1}{\mu}(\varepsilon_w + \frac{1}{\theta})} \frac{L_n}{L} \right]^{-1}. \tag{12}$$

These expressions also take into account the equilibrium changes in $\hat{U}$, a common shift that is absorbed by the regression constant but needs to be computed when evaluating the welfare impact.[15] Eq. (11) is then obtained by substituting $\hat{w}_r$, $\hat{L}_r$, and $\hat{U}$ as a function of $\hat{\Phi}_r$.

When solving for a given counterfactual equilibrium, we first solve the simultaneous equation system (11) to obtain changes in market access. Once we obtain $\hat{\Phi}_r$, we can then compute the changes in local wage and population using (12). Finally, we obtain the changes in the local price index and welfare as:

$$\hat{P}_r = \hat{\Phi}_r^{-\frac{1}{\theta}}, \qquad \widehat{\text{Welfare}}_r = \frac{\hat{w}_r}{\hat{P}_r}. \tag{13}$$

ε_w and ε_l are functions of the structural parameters μ and θ as

$$\varepsilon_w = \frac{\mu\theta - 1}{\theta\left[\mu\theta + \mu + 1\right]}, \qquad \varepsilon_l = \frac{2\theta + 1}{\theta\left[\mu\theta + \mu + 1\right]}. \tag{14}$$

Solving the model requires values for the trade elasticity on iceberg trade costs (θ) and the inverse labor elasticity (μ). We assume $\theta = 8$ (see Donaldson and Hornbeck, 2016; Allen and Arkolakis, 2022). Given our estimate of δ, we can recover the inverse labor supply elasticity μ as (see Appendix C.7 for details):

$$\mu = \frac{\varepsilon_w}{\delta - \varepsilon_w} \times \frac{2 - \delta}{1 - \delta}, \quad \text{where} \quad \varepsilon_w = \frac{1 - \delta}{\theta}. \tag{15}$$

3.5. Intuition

In our model, the inverse labor elasticity μ governs the extent of population relocation across markets, and determines the outcome only by affecting the magnitude of the common shifts, $\hat{U}$, see Eqs. (C.26) and (C.27) in Appendix C.6.

Two extreme cases are worth mentioning. The first is when the labor supply is elastic—i.e., when μ approaches one. The reduced-form parameters become $\varepsilon_l = (2\theta + 1)/(\theta^2 + 2\theta)$ and $\varepsilon_w = (\theta - 1)/(\theta^2 + 2\theta)$. As seen in Eq. (3), an improvement in market access translates into higher wages and higher city population, where impacts on population (ε_l) are larger than impacts on wages (ε_w), owing to labor supply being elastic. The second case is when the labor supply is inelastic—i.e., μ approaches infinity. In this case, Eq. (14) implies that the reduced-form parameters are $\varepsilon_l = 0$ and $\varepsilon_w = 1/(\theta+1)$. Naturally, all changes in market access are absorbed by changes in wages, with no impact on population.

4. Empirical analysis

In this section, we first present baseline estimates of δ from (7) using $\hat{\Phi}_{rt}^{\text{mst}}$ as an instrument and then discuss estimates under alternative model specifications.

4.1. Baseline elasticity estimates

Panel A of Table 1 presents estimation results for Eq. (7). Column 1 reports the OLS estimate, $\delta = 0.11$ (s.e. 0.02). This estimate rises to 0.29 (s.e. 0.11) when using IV, as seen in column 2. The IV results thus indicate an elasticity of GDP with respect to market access of 0.29. All models control for district-fixed effects and initial log GDP to account for mean reversion. We cluster the standard errors at the 4 km-buffer level. Given the inclusion of district fixed effects, markets that are the only unit within a district do not contribute to the estimation. When estimating the models, we exclude 894 markets which are contained within the Functional Urban Area (FUA) of the 180 target cities used in the least-cost spanning tree algorithm to construct the instrumental variable. This leaves us with 12,492 markets. Since we control for district fixed effects, markets that are the sole market within their district are excluded from the analysis. After dropping these singleton markets, the two-stage least squares (2SLS) model is estimated using the remaining 12,447 markets. Our estimate based on within-district variation across markets is smaller than Alder (2016), who estimates $\delta = 0.6$ (s.e. 0.23) using district-level variation.

To verify that our estimation results are not the by-product of longer-term trends in India's regional urban growth, we conduct a falsification exercise by regressing past nightlight changes for 1994 to 1999 on changes in market access for 2001 to 2011. In Column 3, the elasticity is small and statistically insignificant ($\delta = 0.08$, s.e. 0.14).

We conduct a chi-square test to assess whether the coefficients in columns (2) and (3) are statistically different from each other. The test yields a statistic of 4.48 and a p-value of 0.034, suggesting a systematic difference between the baseline and pre-trend estimates.

[15] The term $\hat{U}$ also corresponds to the "missing intercept" often referred in the literature.

Table 1
Baseline estimates.

	(1) OLS	(2) 2SLS	(3) Pre-trend
δ	0.11	0.29	0.08
	(0.02)	(0.11)	(0.14)
Observations	12,447	12,447	10,436
R^2	0.34	0.16	0.19
F		577.6	584.4
Fixed-effects	District	District	District

Notes: The table presents baseline estimates for the effects of changes in market access (due to improvements in India's road network) on market-level GDP as measured by nightlights. Columns 1 and 2 show the elasticity of GDP with respect to structural market access (in changes between 1996 and 2011), using OLS and 2SLS, respectively. Column 3 presents the pre-trend elasticity of GDP on structural market access, based on changes in GDP between 1994 and 1999. All models include district-level fixed effects. The instrumental variable for the 2SLS estimation is constructed based on connecting India's largest 180 cities using a least-cost spanning tree algorithm. Standard errors (clustered by 4 km markets) are in parentheses.

Table 2
2SLS estimates under different fixed-effects.

	(1) 2SLS	(2) 2SLS	(3) 2SLS
δ	0.35	0.21	0.27
	(0.07)	(0.08)	(0.12)
Observations	12,492	12,490	10,325
R^2	0.18	0.17	0.16
F	1796.6	1062.1	685.2
Fixed-effects	No	State	Super8

Notes: The table presents estimates for the effects of changes in market access (due to improvements in India's road network) on market-level GDP as measured by nightlights using alternative sets of regional fixed effects. We set $\theta = 8$. Column 1 shows estimates without fixed effects, column 2 includes state fixed effects, and column 3 includes 8km-market fixed effects. The instrumental variable is constructed based on connecting India's 180 largest cities using a least-cost spanning tree algorithm. Standard errors (clustered by 4 km markets) are in parentheses.

4.2. Variation within different geographic clusters

Our baseline estimates are identified from within-district variation. We compare the estimates for δ using variation within different levels of geographic clusters. Table 2 reports the IV estimates without any fixed effects in column 1, with state-level fixed effects in column 2, and 8 km-market fixed effects in column 3.[16]

When comparing across all markets, we find an elasticity of $\delta = 0.35$ (s.e. 0.07), which drops to $\delta = 0.21$ (s.e. 0.08) when comparing markets within the same states, and becomes $\delta = 0.27$ (s.e. 0.12) when comparing markets within the same 8 km markets. These estimates are comparable to our baseline estimates, although it appears that including only state fixed effects, the largest geographic unit, leads to a relatively notable underestimation of the effect. The similar and statistically significant coefficients identified across different levels of geographic variation underscore the stability of the positive relationship between GDP and market access that we estimate.

4.3. Alternative specifications

We perform robustness checks using alternative model specifications (for complete results, see Appendix D). First, we use a reduced-form market access measure (Donaldson and Hornbeck, 2016), defined as, $\Phi_{rt}^{\text{reduced-form}} = \sum_n y_{nt} \tau_{rn,t}^{-\theta}$. Results appear in Appendix Table D.1. Estimates are similar to our baseline results. Second, Table D.2 reports estimates using a least-cost spanning tree instrument based on connecting India's 100 largest cities, rather than its largest 180. Results are again similar to our baseline estimates. Finally, we construct market access using alternative values of the elasticity of iceberg trade costs with respect to travel time, ϕ. As a lower bound, we set $\phi = 1.3$, following Monte et al. (2018); as an upper bound, we set $\phi = 1.7$ following Allen and Arkolakis (2022). Appendix Table D.3 reports the estimates of δ. While the estimates of δ vary with the assumed value of ϕ, the results remain statistically significant across specifications.

[16] The number of observations varies across models because some markets drop out of the regression when they are the only unit inside a higher order geographic cluster. Thus, we end up with 12492 markets in the regression without fixed effects, 12490 markets in the regression with state fixed effects, and 10325 markets in the regression with 8 km-market fixed effects.

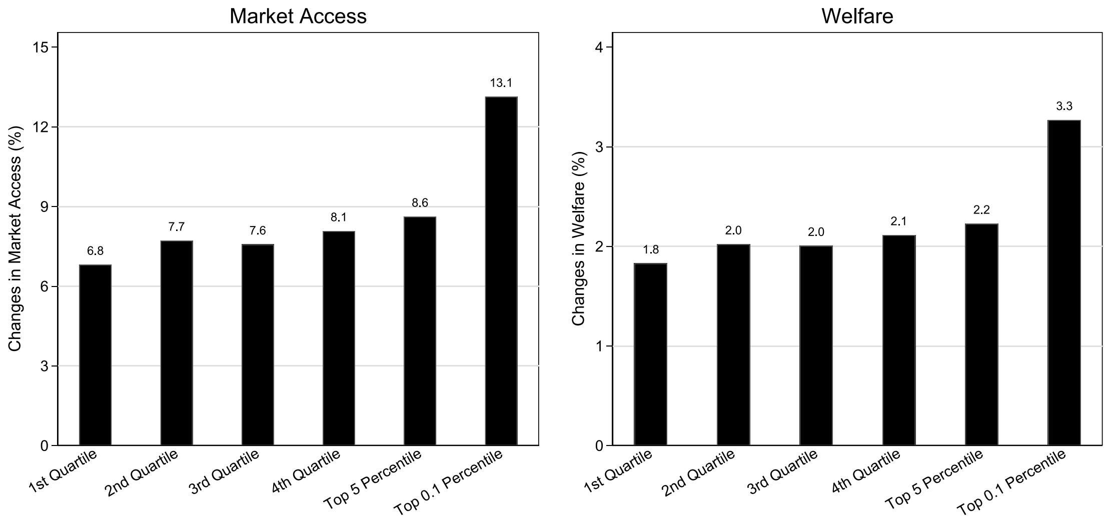


Fig. 3. (%) Change in market access and welfare by market size.
Notes: This figure shows the average change in market access and welfare by markets grouped according to size as measured by nightlight intensity for India's complete transportation upgrade (GQ plus national highway expansion). The 6 groups are the top 0.1 percentile markets, the top 5 percentile markets, and the four quartiles of markets, with the 4th quartile excluding the top 5 percentile markets.

We also present estimation results to examine whether the treatment effect varies across markets and, subsequently, whether these heterogeneous effects influence our welfare results. Specifically, Appendix Table E.1 reports estimates of δ using four different samples without district fixed effects, while Appendix Table E.4 includes district fixed effects and excludes markets that are the sole ones in their district. Although the estimated values of δ differ across samples, indicating heterogeneous elasticities, these differences yield welfare results that are very similar to our baseline (see discussion in Section 5 and Appendix E).

5. Welfare analysis

Given the assumed parameter value of $\theta = 8$ and the estimated parameter value of $\delta = 0.29$, Eq. (15) implies that the value of inverse labor supply elasticity as $\mu = 1.00$.[17] We use these parameter values and the model structure to evaluate the impact of India's highway build-up over 1996 to 2011. First, we evaluate the welfare effects of the full transportation network, and study the spatial distribution of these effects. These results focus on gross welfare gains from improved market access and do not account for the construction or maintenance costs of the roads, which we address separately in the cost–benefit analysis later in this section. Second, we decompose spatial variation in welfare changes into between and within-region components, which illustrate the benefits of working with granular satellite imagery. Third, we separately evaluate the effects of the construction of the GQ and of India's national highway expansion, which allows us to disentangle the impact of different types of road construction on welfare. Fourth, we evaluate the impacts of further upgrading of national highways and local roads. Finally, we incorporate the costs of road construction to perform cost–benefit analysis, and thus evaluate the net return of the infrastructure project as a whole. To implement the analysis, we take our model to the data in 2011, evaluate highway travel times at the level of those for 1996, and then consider alternative changes in India's highway infrastructure.

5.1. Effects of the full transportation upgrade on welfare by market size

We begin by considering changes in market access and welfare across markets resulting from the full transportation network build-up (i.e., the GQ construction and the expansion of national highways). We group markets into six categories based on the market area: the first three groups represent the first, second and third quartiles of markets (which together represent 25.2% of the population in 2000), the fourth group represents the fourth quartile of markets excluding the top 5 percentile markets (23.9% population in 2000), the fifth group represents the top 5 percentile markets (30.9% population in 2000), and the sixth group represents the top 0.1 percentile markets (basically, the top 12 largest markets) in terms of market area (13.9% population in 2000).

[17] The welfare impacts are insensitive to values of ϕ but are sensitive to values of θ. In our baseline calibration, we set $\theta = 8$. We also report the results when setting $\theta = 4$ (Simonovska and Waugh, 2014). Compared to our baseline estimates reported in Table 1, the welfare improvements become much larger as θ becomes smaller.

Table 3
Decomposing welfare changes.

	A. States and districts		B. Super-markets
Between state	16.8%	Between 8 km markets	60.8%
Within state	83.2%	Within 8 km markets	39.2%
Between district	41.6%	Between 4 km markets	23.1%
Within district	41.6%	Within 4 km markets	16.2%
		Between 2 km markets	10.8%
		Within 2 km markets	5.3%

Notes: The table presents a decomposition of changes in welfare between and within regional units in India for the complete transportation upgrade (GQ plus national highway expansion). The left panel decomposes aggregate welfare across and within Indian states and districts. The right panel performs an alternative decomposition using the 8 km, 4 km, and 2 km market boundaries detected from satellite imagery.

In the left panel of Fig. 3, which displays percentage changes in market access, it is evident that the gain in market access is greatest for India's largest urban areas. Market access rises by 13.1% for the top 0.1 percentile markets, 8.6% for the top 5 percentile markets, and 6.8%–8.1% for the remaining markets. The right panel of Fig. 3 displays percentage changes in welfare. Mirroring the market access results, welfare gains are most pronounced for the largest markets: welfare increases by 3.3% for top 0.1 percentile markets, 2.2% for the top 5 percentile markets, and 1.8%–2.1% for the remaining markets.

Fig. 4 displays the geographic distribution of welfare changes across markets. Panel (a) shows dispersion across all markets, while Panel (b) zooms in on the area surrounding New Delhi, and Panel (c) zooms in on the area surrounding Ahmedabad. Smaller welfare changes are represented by lighter shades of yellow; larger welfare changes are represented by darker shades of red. The thick blue line represents the GQ, while the thinner and lighter blue lines represent national highways.

Upon inspection of Fig. 4, it is clear that markets along the GQ display the largest welfare changes (shown in dark red), some markets along national highways also show large welfare changes, and less connected markets in the center of India display the smallest welfare changes (shown in pale yellow). Additionally, there is substantial variation in welfare changes across 1 km markets within the geographic regions that are typically the unit of analysis in conventional data. For instance, if we focus on Fig. 4b, which shows the New Delhi region, we see that the average welfare change for the 1 km market corresponding to New Delhi was 3.04%. The market with the smallest welfare change in panel (b) had a modest 0.54% increase, while the market with the largest change had a sizable 6.36% increase. Similarly, when we zoom in on Ahmedabad in Fig. 4c, we see that welfare changes for the market representing Ahmedabad city was 3.56%. The market with the smallest welfare change in the region had a 0.71% increase, while the market with the largest change showed a 7.34% gain. This variation in welfare impacts across markets within a region (5.82 percentage point difference for the New Delhi area, 6.63 percentage point difference for the Ahmedabad area) represents almost half of the difference between markets across all India (13.16 percentage points), a fact which underscores the analytic value of capturing welfare impacts in granular spatial data.

5.2. Decomposing welfare changes between vs. Within regions

Next, we decompose the variance of log changes in welfare into the sum of between and within-region components as follows:

$$\text{Total} = \text{Between} + \text{Within}. \tag{16}$$

where

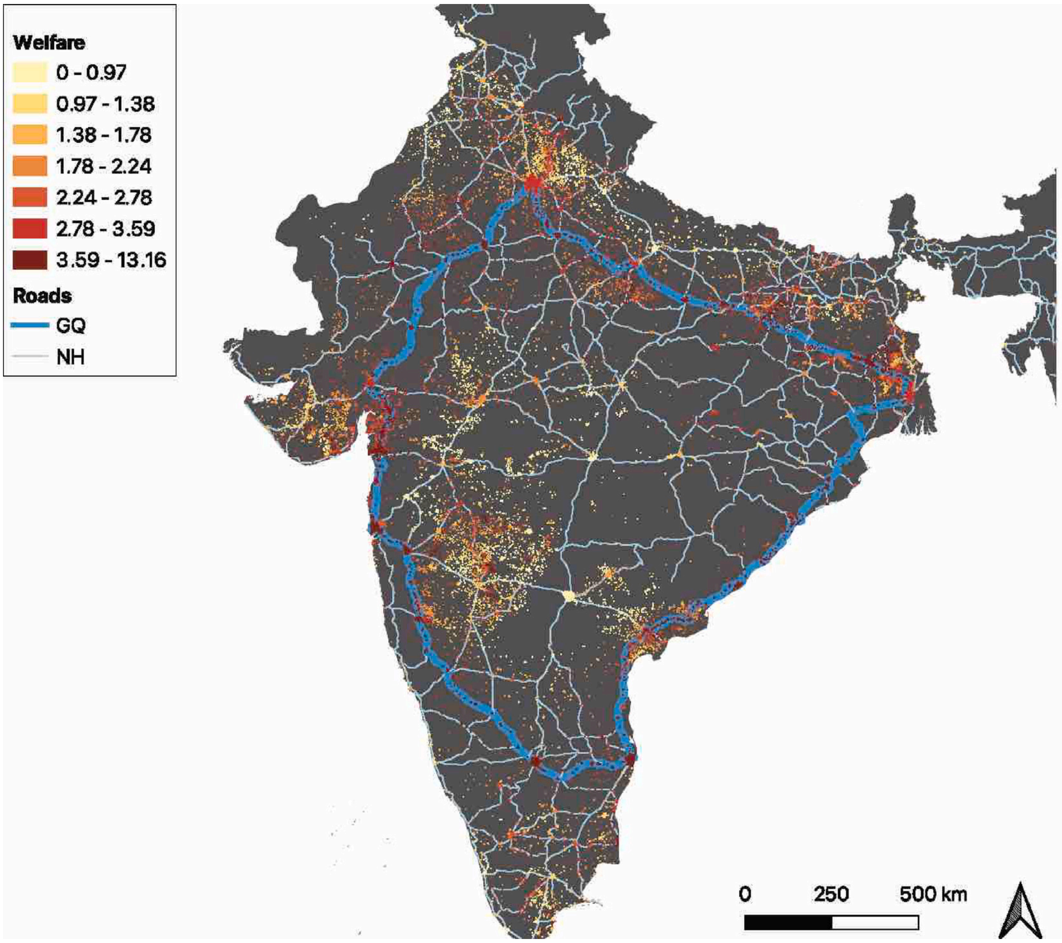
$$\text{Total} = \frac{1}{N}\sum_i (\text{W}_i - \bar{\text{W}})^2, \quad \text{Between} = \frac{1}{N}\sum_d N_d(\bar{\text{W}}_d - \bar{\text{W}})^2 \tag{17}$$

where W_i is change in log welfare in market i, $\bar{\text{W}}$ is the average of W_i across all markets, $\bar{\text{W}}_d$ is the average across markets within administrative unit d (e.g., state or district), N_d is the number of markets within d, and N is the total number of markets. The Within component is calculated as the difference between the Total and Between components. Across all 13,386 1 km markets, welfare changes have a mean of 2.04% and a standard deviation of 1.38%.

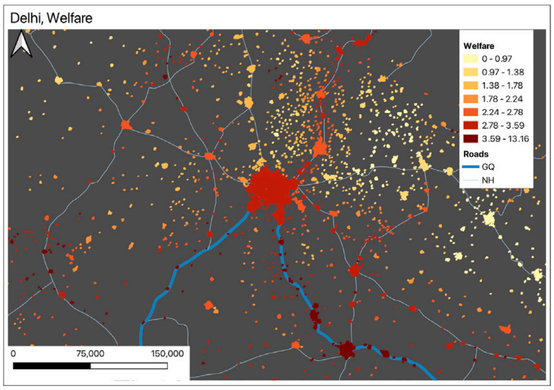
In Table 3, we display the variance decomposition. Panel A performs the decomposition by allocating 1 km markets into standard Indian administrative units-districts and states. The first row indicates that 16.8% of the overall spatial variance in welfare gains from India's highway infrastructure expansion is explained by between-state variation. The second row shows that 83.2% of the spatial variance is explained by variation across markets *within* states.

We further decompose the within-state variance using district administrative boundaries. Here again, much of the within-state variance is explained by the within-district component. While 41.6% of the within-state variance is driven by differences across districts (within states), 41.6% of the within-state variance is driven by variation in welfare across 1 km markets within districts. Thus, analysis performed at the district level may smooth over nearly half of the spatial variation in welfare impacts.

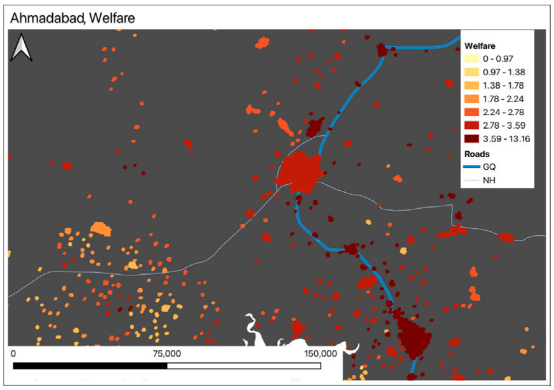
Panel B of Table 3 reports an alternative decomposition that uses the delineation of 8 km, 4 km, and 2 km markets. Recall that the satellite imagery detects 3483 8 km markets in India, and the decomposition shows that 60.8% of overall spatial variance in welfare changes is driven by between-8 km market variation. This leaves the remaining 39.2% of overall variance to be explained by within-8 km market variation, again indicating large levels of heterogeneity in the impacts of roads at the sub-district level. Of this



(a) All of India



(b) New Delhi



(c) Ahmedabad

Fig. 4. Geographic distribution of welfare changes.
Notes: This figure shows the geographic distribution of welfare changes across all urban markets in India for the complete transportation upgrade (GQ plus national highway expansion). Smaller welfare effects are represented by lighter shades of yellow; larger welfare effects are represented by darker shades of red. The thicker blue line represents the GQ; the lighter blue lines represent national highways. Panel (a) shows the geographic distribution for all markets in India, Panel (b) shows effects for the area around New Delhi, and Panel (c) shows effects for the area around Ahmedabad. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

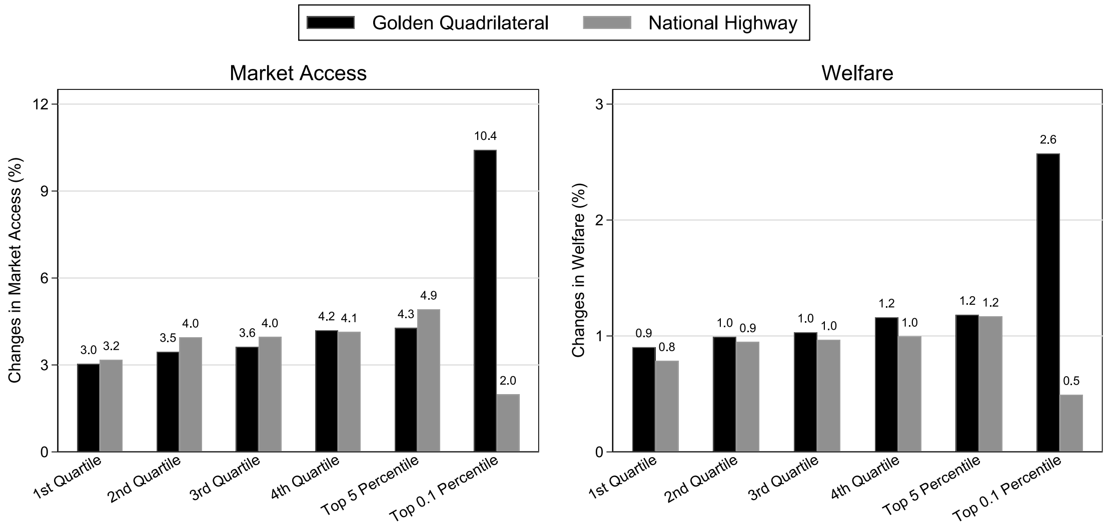


Fig. 5. Market access and welfare impacts by highway type and market size.
Notes: This figure shows the average change in market access and welfare by markets grouped by size based on one scenario in which only the GQ is built (darker bars) and another scenario in which only national highways are upgraded (lighter bars). The 6 groups are organized according to their market area (top 0.1 percentile markets, top 5 percentile markets, and four quartiles, with the 4th quartile excluding the top 5 percentile markets).

39.2% variance, 23.1% is driven by within-8 km but across-4 km market variation. These patterns reinforce the idea that analysis conducted at the level of administrative boundaries – which do not overlap directly with transactions in urban environments – may smooth over variation in the impacts of changes in market access.

Finally, to analyze the underlying economic forces driving this variance decomposition, Appendix Table F.1 presents a variance decomposition of changes in market access. The decomposition closely mirrors the pattern uncovered above: for both Panels A and B, we find that a large share of the overall spatial variance is explained by within-district variation, further indicating that our welfare results are almost entirely driven by changes in market access and underscoring the value of high-resolution satellite imagery for capturing localized economic effects.

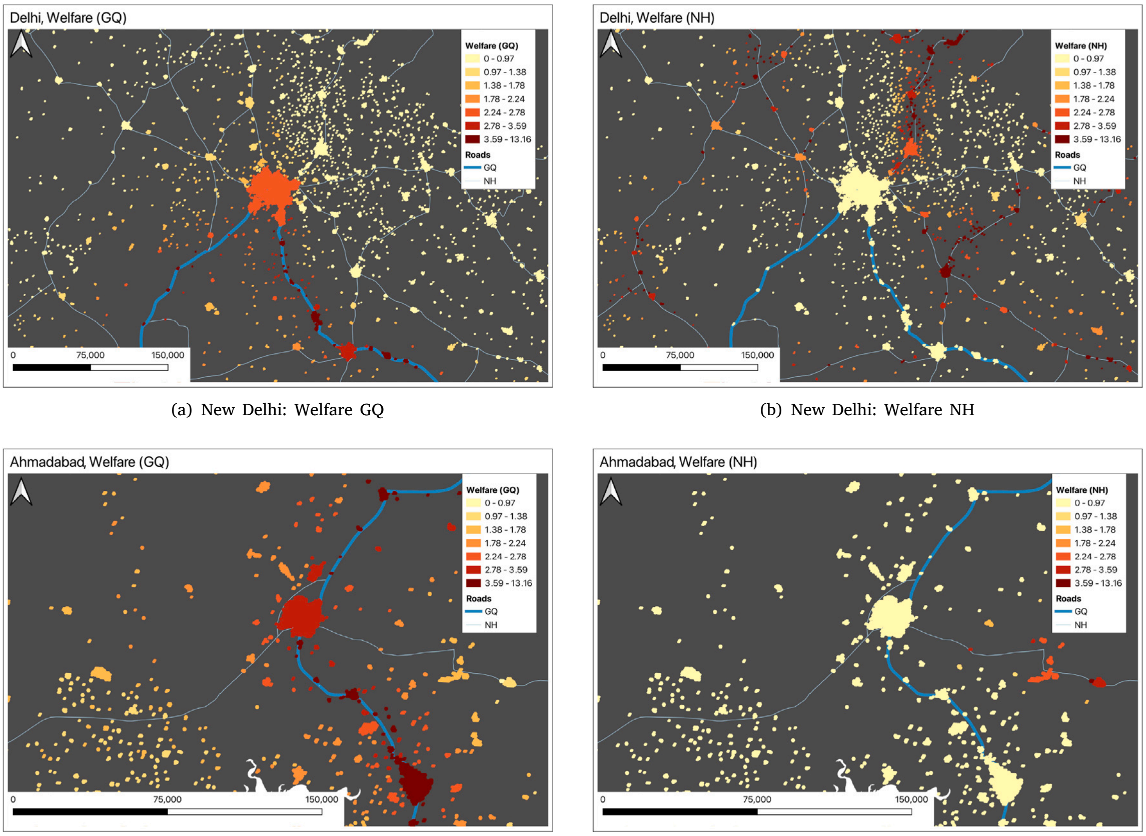
5.2.1. Heterogeneous elasticities

While the results presented above use a constant value of $\delta = 0.29$ across all markets, we also examine whether the treatment effect varies across markets and, if so, whether these heterogeneous treatment effects influence our welfare results. Specifically, in Appendix E, we estimate δ using four different samples, one for each quartile of market area. We then use the estimated reduced-form effects to compute the implied heterogeneous structural parameters and recompute the counterfactual welfare effects. We find the results under heterogeneous elasticities are very similar to the baseline. The indication is that our welfare and inequality results are driven by differential changes in market access, whereas the equilibrium adjustments of labor relocation play a small role.

5.3. Disentangling the effects of the GQ and national highways

So far, we have analyzed the impacts of India's entire transportation network (the GQ plus national highways). While the effects of the GQ have been previously analyzed using district-level data (Alder, 2016), we are not aware of a welfare comparison between the GQ and national highways. Evaluating the relative economic benefits of the two—the first of which connected India's very largest cities and the second of which connected India's next tier of cities—provides insight into how alternative strategies to expand road networks affect well-being. Worthy of note in this comparison is that the expansion of national highways (in kilometers of lanes built) was larger than that of the GQ (see Section 2.2), where the cost of the national highway buildup was almost double that of the GQ (see Appendix Table B.2).

We compare a counterfactual in which only the GQ is built against one in which only national highways are expanded. Fig. 5 displays percentage changes in market access (left panel) and welfare (right panel) across markets grouped according to size. The GQ had similar welfare impacts as the national highway expansion for most markets, except the top 0.5 percentile in terms of land area. The GQ-national highway difference is most pronounced for the very largest markets. Whereas the difference in welfare gains between the GQ and national highways ranges from 0.1% (1st quartile of markets by size) to 0.06% (top 5 percentile), the difference for the largest markets, the top 0.1 percentile, is significantly larger at 2.1%. This difference is driven by the substantial increase in market access generated by the GQ for these largest markets in terms of size. Overall, the GQ alone would have led to a 1.2% increase in welfare, while the national highway expansion alone would have led to a 1.0% increase in welfare.



(a) New Delhi: Welfare GQ

(b) New Delhi: Welfare NH

(c) Ahmedabad: Welfare GQ

(d) Ahmedabad: Welfare NH

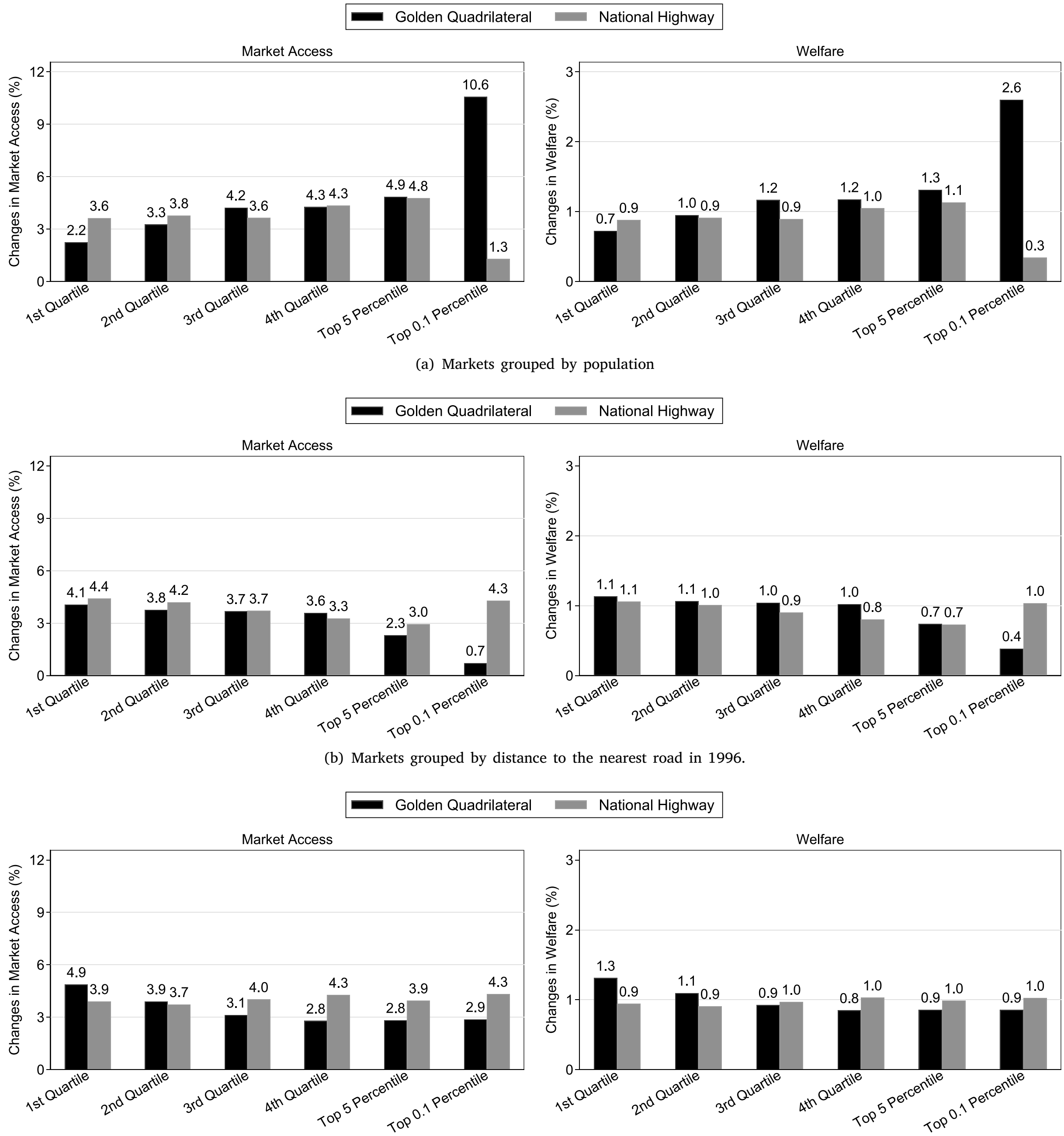
Fig. 6. Geographic distribution of welfare from the GQ and the NH.
Notes: This figure shows the geographic distribution of changes in welfare from the GQ and the national highway expansion (NH) in markets around New Delhi (top two panels) and Ahmedabad (bottom two panels). Smaller welfare changes are represented by lighter shades of yellow while larger welfare changes are represented by darker shades of red. The thicker blue line represents the GQ, while the lighter blue lines represent national highways. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

To further explore these differences, Fig. 6 displays geographic variation in welfare improvements due to the GQ (left panels) and national highways (right panels) for New Delhi (top panels) and Ahmedabad (bottom panels). As expected, welfare gains for the GQ are concentrated in markets that lie close to the construction of the GQ, whereas welfare improvements from national highways accrue mostly in more remote markets that lie along the construction of the network. Overall, the absolute increases in welfare from the GQ tend to be larger than those of national highways, as evidenced by a higher presence of lighter shades of yellow in the right panels of Fig. 6. Figures F.2 and F.3 in Appendix F show the geographic distribution of welfare from GQ and NH for all of India, which mirrors the patterns uncovered in Fig. 6.

We further explore heterogeneity in the effects of the GQ and NH expansions by classifying markets based on three alternative characteristics: population size, distance to the nearest road (as of 1996), and agricultural intensity. We replicate Fig. 5, but now group markets by these variables to examine how gains in market access and welfare vary across different types of locations.

Markets are grouped by population using gridded population data, by remoteness using distance to the nearest road in the 1996 network, and by agricultural intensity using the proportion of cropland pixels within an 8 km buffer. Cropland is identified using the MODIS MCD12Q1 Land Cover Type product, based on the UMD classification scheme, which defines croplands as areas with at least 60% of the land actively cultivated.

Fig. 7 shows three key patterns. First, the results for population closely mirror the results by market land area in Panel (a): larger markets as defined by population size were disproportionately benefited by the GQ, and this effect is especially concentrated on the largest 0.1 percentile, while these same markets saw negligible effects from the NH expansion. Second, markets that were more remote in 1996 experienced larger gains from the NH network than from the GQ, and derived only limited benefits from the GQ



(a) Markets grouped by population

(b) Markets grouped by distance to the nearest road in 1996.

(c) Markets grouped by agricultural intensity

Fig. 7. Market access and welfare impacts by highway type and market size.
Notes: Each panel shows the average change in market access and welfare for six market groups under two highway scenarios: one with only the Golden Quadrilateral (black bars) and another with only National Highway upgrades (gray bars). The six groups are defined as four quartiles, the top 5 percentile, and the top 0.1 percentile based on each characteristic.

alone (see Panel B, where top 0.1 percentile indicates the most remote markets). This is consistent with the idea that the NH provided access to less connected regions, while the GQ mostly benefited already large and more urban markets. Third, Panel (c) shows little variation in effects across markets grouped by agricultural intensity, suggesting that the gains in market access and welfare are more closely linked to urban characteristics than to agricultural production—consistent with the mechanisms emphasized by our model.

Table 4
Cost-benefit calculations by different roads.

	All roads	GQ	NH
Overall Gain in GDP (NPV)	2.90	2.25	0.62
NPV of Construction Costs	0.39	0.13	0.26
NPV of Maintenance Costs	0.86	0.28	0.57
NPV of Total Costs	1.24	0.40	0.83
Net Gain in GDP (NPV)	1.66	1.84	−0.21

Notes: The table presents a cost–benefit analysis of the improvement in the road network, which disentangles the overall gain in GDP and the costs of road construction and maintenance. We differentiate the effects of all roads (column 1), GQ construction only (column 2), and national-highway construction only (column 3). GQ counterfactual: keep the road structure as in 2011, but lower the speed of GQ roads to 40mph (speed of national highways); NH counterfactual: keep the road structure as in 2011, but reduce the speed of newly built national highways to 20mph (speed of other roads). All values are the NPV values computed over a 50-year time horizon using a 5% discount rate, and then normalized by the NPV of the 2011 GDP.

5.4. Cost-benefit of India's highway investment

Until this point, we have focused on welfare effects without consideration of the costs incurred in the construction of these roads. We next perform a cost–benefit analysis to evaluate the net return of building roads, explicitly considering the costs of construction and maintenance. Throughout this section, we focus on overall gains, and we do not explicitly model the geographic distribution of the costs or net benefits.[18] We calculate the net present value (NPV) of the increase in GDP from the entire road-buildup program, considering benefits from welfare improvements and costs of construction and maintenance (World Bank Group, 2017a,b,c,d,e,f,g).

To annualize construction costs, we conduct cost–benefit analysis in terms of NPV (Tsivanidis, 2023). NPV values are computed over a 50-year time horizon using a 5% discount rate. Panel B of Table 1 displays the NPV estimates for gross GDP, costs, and net GDP. Since all NPVs are normalized by the NPV of GDP in 2011, they can be interpreted as the annualized percentage of GDP.

We compare the actual economy with the observed road network in 2011, our benchmark scenario, to three counterfactual scenarios. Column 1 reports the impacts of the entire road network, which we obtain by comparing the benchmark scenario to an economy with no highway construction. Column 2 isolates the effects of the GQ by comparing our benchmark scenario to a counterfactual economy with the 2011 road network, except for the GQ upgraded sections (i.e., GQ speeds equal regular highway speeds). Column 3 reports the effects of only expanding national highways, which we obtain by comparing our benchmark scenario to a counterfactual economy with the 2011 road network that has the upgraded national highway sections removed (i.e., national highway speeds equal regular motorway speeds).

The first row of Table 4 reports that India's highway investment during this period increased annual gross GDP by 2.90%. Column 2 indicates that the contribution of the GQ to overall welfare is 2.25%, or 77.3% of the overall increase. The expansion of non-GQ national highways increased the NPV by 0.62%.[19] Row 2 shows the NPV of the total costs, which has two parts. The first is the construction cost.[20] Table B.2 details the overall construction costs in 1996 USD, which translates to an NPV of $129.93 billion in 2011. The annual maintenance costs are assumed to be 12% of the total construction costs (Allen and Arkolakis, 2014); row 4 reports the associated NPV. Row 5 reports the NPV of GDP net of the costs, equaling the difference between rows 1 and 2. We estimate that the GQ delivered a net GDP increase of 1.84%, but national highways, by contrast, do not appear to have delivered a net benefit, which is surprising given the large expansion in this road type reported in Fig. 2.

Our results are similar, although slightly larger in magnitude, to those estimated by Alder (2016), who finds that the construction of the GQ led to gross and net GDP gains of 1.30% and 0.80%, respectively. This suggests that granular high-resolution satellite data can similarly aggregate estimates uncovered from (district-level) nightlight data alone. Additionally, one necessarily requires our granular spatial to evaluate the distribution of welfare changes inside districts.

6. Conclusion

Many developing country policymakers appear to take it as an article of faith that building new highways will raise national income per capita. They act on these beliefs by dedicating substantial public funds to transportation infrastructure, often with the encouragement of international financial institutions and countries eager to finance such investments. With standard administrative data, analyzing the economic impacts of highway construction in the typical lower-income country would be infeasible until a decade or more after project completion. We provide a method that permits rapid and low-cost evaluation of infrastructure investments.

[18] It must be noted that if the financial burden of road construction were disproportionately borne by certain areas, the market-level welfare changes reported in previous sections – which reflect only gross benefits – could overstate the true welfare gains in those locations.

[19] Note that, because the model is non-linear, the joint effect of GQ and NH does not necessarily equal the sum of the two individual effects. The differences reflect the interactive effects between the two, which are small.

[20] The construction cost is from the World Bank's Project Appraisal Document and the Implementation Completion and Results Report. The average construction cost is calculated based on data from 11 transportation projects that were executed in India between 2008 and 2017.

Analysts need only freely available satellite imagery, to evaluate impacts on GDP, and a baseline or endline year of population data, to perform additional welfare analysis. With the ongoing digitization of road maps globally (e.g., OpenStreetMap) and the availability of high-resolution population data through such resources as Gridded Population of the World, our method should have wide applicability.

In the much studied Indian context, our analysis delivers similar estimates of aggregate welfare changes to those that use conventional data. This is to be expected, considering that the market access approach is well-tested. Even for India, however, our approach delivers insights that analyses using conventional, lower-resolution data cannot. Our approach reveals the distributional consequences of new infrastructure across spatial markets that vary enormously in size and in their proximity to new roads and highways. We expect that such results would be important to policymakers, considering that both efficiency and equity considerations weigh in how large projects are evaluated and sold to the public.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jinteco.2025.104201.

Data availability

Using Satellite Imagery to Measure the Impacts of New Highways: An Application to India (Original data) (Mendeley Data)

References

Adao, R., Costinot, A., Donaldson, D., 2017. Nonparametric counterfactual predictions in neoclassical models of international trade. Am. Econ. Rev. 107 (3), 633–689.

ADB, 2017. Meeting Asia's Infrastructure Needs. Asian Development Bank.

Aggarwal, S., 2018. Do rural roads create pathways out of poverty? Evidence from India. J. Dev. Econ. 133, 375–395.

Ahlfeldt, G.M., Redding, S.J., Sturm, D.M., Wolf, N., 2015. The economics of density: Evidence from the berlin wall. Econometrica 83, 2127–2189.

Alder, S., 2016. Chinese roads in India: The effect of transport infrastructure on economic development. Available At SSRN 2856050.

Allen, T., Arkolakis, C., 2014. Trade and the topography of the spatial economy *. Q. J. Econ. 129 (3), 1085–1140.

Allen, T., Arkolakis, C., 2022. The welfare effects of transportation infrastructure improvements. Rev. Econ. Stud. 89 (6), 2911–2957.

Allen, T., Arkolakis, C., 2023. Economic activity across space: A supply and demand approach. J. Econ. Perspect. 37 (2), 3–28.

Allen, T., Atkin, D., 2016. Volatility and the Gains from Trade. Tech. Rep., National Bureau of Economic Research.

Allen, T., Atkin, D., 2022. Volatility and the gains from trade. Econometrica 90 (5), 2053–2092.

Asher, S., Novosad, P., 2020. Rural roads and local economic development. Am. Econ. Rev. 110 (3), 797–823.

Asturias, J., García-Santana, M., Ramos, R., 2019. Competition and the welfare gains from transportation infrastructure: Evidence from the golden quadrilateral of India. J. Eur. Econ. Assoc. 17 (6), 1881–1940.

Baragwanath, K., Goldblatt, R., Hanson, G., Khandelwal, A.K., 2019. Detecting urban markets with satellite imagery: An application to India. J. Urban Econ.

Bartelme, D., 2018. Trade Costs and Economic Geography: Evidence from the US. mimeo Univ. Michigan.

Bleakley, H., Lin, J., 2012. Portage and path dependence. Q. J. Econ. 127 (2), 587–644.

Burchfield, M., Overman, H., Puga, D., Turner, M., 2006. Causes of sprawl: A portrait from space. Q. J. Econ. 121 (2), 587–633.

Caliendo, L., Parro, F., 2015. Estimates of the trade and welfare effects of NAFTA. Rev. Econ. Stud. 82 (1), 1–44.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2016. Gridded population of the world, version 4 (GPWv4): Population count.

Coşar, A.K., Demir, B., Ghose, D., Young, N., 2021. Road capacity, domestic trade and regional outcomes. J. Econ. Geogr. 22 (5), 901–929.

Datta, S., 2012. The impact of improved highways on Indian firms. J. Dev. Econ. 99 (1), 46–57.

Dekle, R., Eaton, J., Kortum, S., 2008. Global rebalancing with gravity: Measuring the burden of adjustment. IMF Staff. Pap. 55 (3), 511–540.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1 (1), 269–271.

Dingel, J.I., Miscio, A., Davis, D.R., 2019. Cities, lights, and skills in developing economies. Working Paper 25678, National Bureau of Economic Research.

Donaldson, D., 2018. Railroads of the raj: Estimating the impact of transportation infrastructure. Am. Econ. Rev. 108 (4–5), 899–934.

Donaldson, D., Hornbeck, R., 2016. Railroads and American economic growth: A market access approach. Q. J. Econ. 131 (2), 799–858.

Donaldson, D., Storeygard, A., 2016. The view from above: Applications of satellite data in economics. J. Econ. Perspect. 30 (4), 171–198.

Elvidge, C.D., Baugh, K.E., Dietz, J.B., Bland, T., Sutton, P.C., Kroehl, H.W., 1999. Radiance calibration of DMSP-OLS low-light imaging data of human settlements. Remote Sens. Environ. 68 (1), 77–88.

Faber, B., 2014. Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. Rev. Econ. Stud. 81 (3), 1046–1070.

Friedl, M., Sulla-Menashe, D., 2015. MCD12q1 MODIS/Terra+Aqua land cover type yearly L3 global 500m SIN grid V006. distributed by NASA EOSDIS Land Processes DAAC.

Ghani, E., Goswami, A.G., Kerr, W.R., 2014. Highway to success: The impact of the golden quadrilateral project for the location and performance of Indian manufacturing. Econ. J. 126 (591), 317–357.

Gibson, J., Olivia, S., Boe-Gibson, G., 2021. Which night lights data should we use in economics, and where? J. Dev. Econ. 149, 102602.

Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring economic growth from outer space. Am. Econ. Rev. 102 (2), 994–1028.

Jedwab, R., Storeygard, A., 2020. The average and heterogeneous effects of transportation investments: Evidence from Sub-Saharan Africa 1960–2010. Working Paper 27670, National Bureau of Economic Research.

Kruskal, J.B., 1956. On the shortest spanning subtree of a graph and the traveling salesman problem. Proc. Amer. Math. Soc. 7 (1), 48–50.

Michalopoulos, S., Papaioannou, E., 2013. National institutions and subnational development in Africa. Q. J. Econ. 129 (1), 151–213.

Monte, F., Redding, S.J., Rossi-Hansberg, E., 2018. Commuting, migration, and local employment elasticities. Am. Econ. Rev. 108 (12), 3855–3890.
Redding, S.J., Rossi-Hansberg, E., 2017. Quantitative spatial economics. Annu. Rev. Econ. 9, 21–58.
Schiavina, M., Moreno-Monroy, A., Maffenini, L., Veneri, P., 2019. GHS-FUA R2019A - GHS functional urban areas, derived from GHS-UCDB R2019A (2015). PID: http://data.europa.eu/89h/347f0337-f2da-4592-87b3-e25975ec2c95.
Simonovska, I., Waugh, M.E., 2014. The elasticity of trade: Estimates and evidence. J. Int. Econ. 92 (1), 34–50.
Storeygard, A., 2016. Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa. Rev. Econ. Stud. 83 (3), 1263–1295.
Sulla-Menashe, D., Friedl, M.A., 2018. User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product. 1, p. 18, USGS: Reston, VA, USA.
The Economist, 2022. India is getting an eye-wateringly big transport upgrade.
Tsivanidis, N., 2023. Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio.
Woetzel, J., Garemo, N., Mischke, J., Kamra, P., Palter, R., 2017. Bridging infrastructure gaps: Has the world made progress. McKinsey Co. 5.
World Bank Group, 2017a. P050649: India - Tamil Nadu road sector project. World Bank.
World Bank Group, 2017b. P067606: India - uttar pradesh state roads project. World Bank Group.
World Bank Group, 2017c. P069889: India - mizoram state roads project. World Bank Group.
World Bank Group, 2017d. P070421: India - karnataka state highways improvement project. World Bank Group.
World Bank Group, 2017e. P072539: India - kerala state transport project. World Bank Group.
World Bank Group, 2017f. P073776: India - allahabad bypass project. World Bank Group.
World Bank Group, 2017g. P077856: India - lucknow-muzaffarpur national highway project. World Bank Group.