From Battlefield to Marketplace: Connectivity, Industrialization, and Spatial Convergence in the Greater Mekong Sub-Region*

Manabu Nose†
Yasuyuki Sawada‡
Tung Nguyen§

May 8, 2022

Abstract

We evaluate heterogeneous impacts of a large-scale highway construction project in the Greater Mekong Sub-Region that connects areas between the mountainous China-Vietnam border and industrial hubs in the northeastern Vietnam. We apply the market access approach to granular geo-coded highway network data and estimate aggregate impacts of highways, accounting for their spillover effects through production input-output linkages. We exploit the variation in topological and meteorological conditions to estimate the arguably causal impact of market potential growth on the agglomeration of firms and workers. We find that expansion of market potential triggered clustering of manufacturing firms in both core and peripheral cities. The polycentric development reflects spillovers of the GMS impact through supply chain linkages. Real wage also grew significantly faster in rural peripheries with increasing share of manufacturing employment. The overall findings suggest that the interregional highway facilitated the economy-wide industrialization and income convergence across space.

Keywords: spatial structural transformation, market access, treatment spillover, agglomeration, core-periphery.

J.E.L. classification codes: O14, O18, O22, R12, R32, R58

---

*Authors thank Yuhei Miyauchi, Soyoung Kim, participants at the JEA Spring meeting, WEAI and IIPF annual conferences and the ADB LEAP seminar for their useful comments. The project is supported by the ADB and the Joint Research Program No. 970 at the Center for Spatial Information Science, The University of Tokyo. The opinions expressed in this paper are the authors' and do not reflect the views of the IMF, the Executive Boards, or managements.

†International Monetary Fund. manabu.nose@gmail.com
‡The University of Tokyo. sawada@e.u-tokyo.ac.jp
§Hitotsubashi University

1 Introduction

The Greater Mekong Sub-region (GMS) Economic Cooperation Program is a large-scale economic development program covering countries along the Mekong River in Southeast Asia which has played a critical role in transforming the region from battlefields in the 1970's into marketplaces through enhancing positive spillover effects of regional cooperation and integration (RCI).[1] The governments and aid agencies have supported the program by providing large-scale funding for transport infrastructure to facilitate cross-border trade, investments, and inclusive growth. While the GMS program has been regarded as one of the most successful RCI programs in the policy arena (Asian Development Bank (2020)), to the best of our knowledge, there has been no rigorous impact assessment of the program, uncovering possible mechanisms behind such success. Particularly, it is imperative to uncover the program's resulting pattern of spatial development quantitatively because, theoretically speaking, it is unclear whether it can lead to monocentric or polycentric development (Krugman (1991)).

This paper aims at bridging this important lacuna by estimating the program's heterogeneous impacts utilizing newly-created granular geospatial data sets. We closely examine how infrastructure investments can unleash the region's market potential, facilitate industrial transformation, and promote production chains between rural periphery and urban core regions through interregional trade cost reductions. The paper places a particular focus on the GMS's North-South corridor: the two regions around the China-Vietnam border and the northeastern Vietnam. Since the old national road that links China's Kunming city with Vietnam's two provincial capitals (i.e., Hanoi and Haiphong) became obsolete and overloaded, the governments decided to invest in rehabilitating the road. The Vietnamese government invested about 1.4 billion USD on upgrading the 244 km old road from Lao Cai (the border city with China) to Hanoi into a highway with multiple lanes. The highway construction started in 2009, completed and opened in September 2014. In parallel, the extension of the new highway to Haiphong was completed in 2015. In addition, the expansion of the transport network to Lan Song (a border city with China's Guangxi Zhuang Autonomous region) has been partially undertaken as of 2022.

Theoretically speaking, the construction of the interregional transport network significantly improves connectivity and facilitates the exchange of goods and people by reducing transport costs. The improvement in infrastructure connectivity, in turn, increases local

[1] The GMS program comprises Cambodia, the People's Republic of China (PRC, specifically Yunnan Province and Guangxi Zhuang Autonomous Region), the Lao People's Democratic Republic (Lao PDR), Myanmar, Thailand, and Viet Nam, covering an area of 2.6 million square kilometers and a combined population of more than 330 million.

market potential, changing landscape of economic activities over the space. Economic geography models predict non-linear relationship between connectivity infrastructure and the agglomeration of activities: improved connectivity can strengthen both *centripetal* forces to central (core) cities and *centrifugal* forces which may expand the city fringe, spreading out workers to the suburban (periphery) area (Krugman (1991); Baum-Snow (2007)). Indeed, the existing empirical findings are mixed: the transport investment contributed to the diffusion of economic activities and suburbanization of the urban landscape in several specific contexts (Jedwab and Storeygard (2020); Baum-Snow (2007); Baum-Snow et al. (2017)) while other studies show the centralization of the system of cities (Faber (2014)). When the transport investments connect large metropolitan areas and small peripheral cities like the case in China's national highway system (Faber (2014)), ex-ante spatial unevenness in local market size, amenity conditions, and the degree of transport network externality may lead to the asymmetric impact of the GMS program on urban development (Allen and Arkolakis (2014)).

In this paper, we undertake a hybrid approach of two empirical strategies using multi-layered data sets from households, firms, and communes in Vietnam as well as satellite imageries. We quantify, first, how highway network improvements affect each district's market access and, second, estimate how the enhanced market access, in turn, promoted industrial development differently between core and periphery areas. As for the first analysis, similar to Donaldson and Hornbeck (2016), we employ a reduced-form model of market potential (MP), i.e., the Harris index (Combes and Gobillon (2015)) as a function of infrastructure access. We match high-resolution satellite data with ground-level firm panel data to describe formally the (non-parametric) relationship between spatial development of the market potential.[2] As the second analysis, given that the China-Vietnam border is mountainous, we exploit the variation in topological conditions as an instrumental variable for the change in market potential so that we can estimate the aggregate causal impact running from better highway network to the agglomeration of firms and workers over space.[3]

To preview our empirical results, there are two findings emerged. First, our empirical results show that the MP index significantly expanded within 20km from the GMS highway, with faster growth by around 40% in peripheral cities where local market was initially underdeveloped. The effect declines as the distance from the highway increases and disappears in remote districts which are located more than 70km away from the

[2] Donaldson and Storeygard (2016) reviews recent applications of satellite data in Economics. Our satellite-based analysis closely relates to Burchfield et al. (2006) and Gibbons et al. (2019).

[3] As reported in the appendix, we also account for treatment spillovers by estimating the program's treatment effect separately from local externality based on the methodology of Miguel and Roland (2011).

highway. Second, however, using commune-level panel data, we find an agglomeration of manufacturing firms in both core and peripheral cities which experienced larger improvement in market access due to the GMS investment. While the system of cities was *monocentric* to the capital city (Hanoi) 20-30 years ago, the reduction in transport costs has contributed to *polycentric* urban development, inducing the decentralization of manufacturing activities to peripheries. At the same time, rural peripheries experienced significant labor reallocation away from agriculture to manufacturing sector. Our analysis using granular geocoded data provides a new evidence of the employment shift to manufacturing industries due to the reduction in transport costs in the area near the highways. We believe this is our novel contribution because the role of internal transportation in shaping structural transformation has received less attention in the literature.[4]

Theoretically, the polycentric core-periphery development could emerge when transport cost reduction promotes agglomeration of firms near the highway, while creating spillover benefits to suburban areas through production chains (Krugman and Venables (1995)). To explore this possible mechanism empirically, we estimate the heterogeneity of GMS highway's treatment effect by supply chain linkages. We measure the input-output linkage by the Leontieff inverse and apply the generalized difference-in-difference design with continuous treatment (Lindo et al. (2020)). At both core and peripheral regions, we find that manufacturing firms with strong production linkages agglomerate near the highway and create jobs. Number of firms and employments also significantly increased in second-tier peripheral cities which have supply chain linkage to the GMS highway through national highways. Based on firm performance, we also find rapid increase in real wages and firm productivity in the rural peripheries, showing larger marginal gains of the interregional highways in the remote area. While this is in line with the findings from the interregional transportation infrastructure in the sub-Saharan Africa (Jedwab and Storeygard (2020)), our findings also highlights a potentially important role of labor market adjustments in generating polycentric development. Based on the overall findings, we can derive two important implications: First, the interregional GMS highway facilitated the economy-wide, large-scale industrialization; and, second, the highway construction led to significant spatial income convergence between core and peripheral cities.

The rest of this paper is organized as follows. Sections 2 and 3 introduce our data and background. In section 4, we estimate the GMS impact on market potential, followed by the project impact on local market development. 5 estimates the treatment spillover of the GMS highway. Finally, Section 6 provides concluding remarks.

[4]Except a few recent works such as Fajgelbaum and Redding (2021), Nagy (2020), and Trew (2020). Fajgelbaum and Redding (2021) found that Argentina's integration into the world economy in the late-19th century triggered structural change due to a spatial Balassa-Samuelson effect.

2 Data

The paper employes multi-layered data sets from households, firms, and communes in Vietnam. The master data are based on the Vietnam Household Living Standard Survey (VHLSS) and Vietnam Enterprise Survey (VES), respectively, for household and firm information. The VHLSS is a biennial survey which covers around 45,000 households each round. The data constructed from the VHLSS contain information on population, housing value, and human capital stock captured by average years of schooling. The VHLSS covers both formal and informal workers, but the location information is based on the living place rather than working place. The VHLSS data also collect additional information on household expenditures including expenditures on housing and education for the smaller set of respondents, covering about 9,000 households.

The VES covers all registered firms with more than a certain number of employees and a sample of smaller firms where the size threshold depends on each survey year and the province where the firm is located (see Nose (2020)). We geocoded each firm's location at the commune level for the spatial analysis. Two censuses in 2011 and 2016 cover the universe of registered firms, which is the main dataset of our empirical analysis. We compute most of our main variables such as employment, sales per worker, and profit margins of registered firms using the VES data.

The master dataset covers about 11,160 communes and 700 districts in eight regions and we use the subset of these communes in three northern regions. Given data inconsistency in tracking state-owned enterprises (SOEs; including collectives) in the annual enterprise survey, we confine our attention to private registered firms and excludes SOEs. We believe that exclusion of SOEs in our analysis will also improve the identification of the impact of the GMS highway construction as the change in SOE performance has been affected by on-going SOE reforms since 2011.

3 Background

Our focus is to understand how recent infrastructure investments caused formation of the particular spatial redistribution of economic activities in the northern Vietnam. Raw data exhibit high concentration of population and firms around core cities such as Hanoi and HCMC, whereas the rural area in the northwestern (NW) and northeastern (NE) national borders remain dominated by agriculture. As Table 1 shows, population size and the number of establishments are particularly large in central cities in Red River Delta (RRD) and South East regions. We can confirm this spatial pattern at finer grid level based on the distribution of lights-at-night (NTL) data (Figure 1). We look at the monthly average Day-Night Band (DNB) radiance values of the NASA Visible and Infrared Imaging Radiometer Suite (VIIRS) satellite image. The line layer shows the geo-coded location of the GMS highways connecting Kunming, Lao Cai, and Hanoi which is further linked to Lang Son and Haiphong. The NTL image shows an expansion of lit area over years, which potentially reflects the impact of investments under the GMS program. The satellite image also reveals spatial diffusion of manufacturing activities to periphery cities along the GMS transport network.

Figure 2 reveals large socio-economic disparities in the northern Vietnam at an earlier year of 1999. In general, the northwestern (mountainous region) and northeastern Vietnam show initially high poverty rate and inequality with lower access to basic infrastructure. The living standard is higher in the metropolitan area near Hanoi and Haiphong.

3.1 GMS Highway Construction Program

Before the highway construction, international container cargos from Kunming were often sent to Guangzhou province in China rather than being sent via the Kunming - Lao Cai - Hanoi/Haiphong route (Figure 1). Although bilateral trade between Vietnam and China had increased rapidly over the last decade, the nearest Chinese port is 1,053km away from Kunming. As the transportation cost was extremely high, Vietnamese companies tended to use small trucks or railroad instead of the existing national road.

Under such a situation, the construction of the GMS expressway was expected to halve cargo travel time from more than 2 days to 1 day by shortening travel distance because Vietnam's main port at Haiphong is only 813km away from Kunming with better road pavement. The highway was estimated to reduce the travel time of passenger traffic from at least 10 hours to about 3.5 hours. The traffic volume reached 18,000-19,000 passenger car unit (PCU) per day in late 2017, which was far beyond the level of near 4,000 PCU per day travelled by the national road before 2015 (GMS Secretariat (2018)). The final segment of the Kunming - Lao Cai - Hanoi/Haiphong route is the most modern highway in Vietnam with total investment of more than 2 billion USD under Built-Operate-Transfer (BOT) scheme. The modern expressway runs in parallel to the old national road, which was built during French colonial period and upgraded in 1998. The highway is designed to have six travel lanes and two emergency lanes, allowing speed of 120km per hour, much faster than 80km per hour outside and 50km per hour inside residential area in the old national road.

The Hanoi - Lang Son (HN-LS) expressway is expected to better connect the northeastern region to the Guangxi region in China and form part of the GMS eastern corridor (Figure 1). The expressway consists of three sections: Hanoi - Bac Giang (50.9km), Bac Giang - Lang Son (63.1km), and Chi Lang - Huu Nghi (Lang Son, 43.3km). Of these, the first two sections were built under BOT contracts and were finished in 2016 and 2019, respectively.

The volume of interprovincial passenger travel along the Hanoi - Lao Cai (HN-LC) section in Figure 1 has significantly increased. In five provinces along the highway, annual growth rates of passenger travels by road increased from 5.4% in 2011-2014 to 6.6% in 2015-2018. Meanwhile, in other provinces within the northeastern and northwestern regions where the highway did not pass, the traffic volume decreased from 7.0% to 6.4%. Similarly, the volume of passenger traffic increased rapidly in provinces along the HN-LC section.

In contrast, provinces along the Hanoi - Haiphong (HN-HP) section experienced faster growth in freight transportation. This reflects the fact that transportation firms concentrate more in larger cities (such as Hanoi and Haiphong) than cities in mountainous areas.

3.2 Agglomeration Patterns of Firms: Core vs. Periphery

In this paper, we define core cities as all districts within Hanoi and Haiphong, the two central provinces in the north.[5] Districts within 100km from Hanoi centroid are defined as peripheral cities, and the remaining other cities are categorized as "2nd-peripheral cities".

Population density is significantly high at core cities. The last two columns of Table 1a show that population size grew fast in the peripheral cities at the northwestern Vietnam. Industrial activities were also highly concentrated around Hanoi and HCMC. Figure 3 shows a rapid increase in the number of registered firms in the districts near the core cities. The strong agglomeration effect at core cities may partially reflect its relatively high skill endowment.

The industrial composition differs between core and peripheral cities. In core cities, about two-third of firms registered in the service sector and the share of service employments accounted for one-third of total employment. In peripheral cities, manufacturing firms accounted for almost 25% of total registered firms while the share of manufacturing employment increased from 56% in 2011 to 62% in 2016. In 2nd-peripheral cities, most of firms registered in construction which can account for about 50% of total employment. In districts along the border with PRC, only small number of manufacturing firms

[5] In Vietnam, central provinces, or cities directly under central government, includes Hanoi, Haiphong, Danang, Can Tho, and HCMC. These cities are socio-economic center in the region. Haiphong is smaller core city close to Hanoi, so that we take the centroid of Hanoi as the origin for core-periphery definition.

existed, while construction firms dominated larger market share and engage in building infrastructure.

In terms of firm share, all districts in northern regions have experienced disproportionately-large market entry of service firms entering the market. In terms of employment, Figure 4 shows a considerable shift from agricultural and construction sectors to manufacturing and service sectors in peripheral cities. The growth of manufacturing employment was particularly high in peripheral cities at 66% in northeast and 51% in northwest regions. Meanwhile, core cities in RRD region have experienced relative shifts in workers from the manufacturing toward the service sector.

Along the GMS highways, districts within 50km access experienced expansions of manufacturing firms and employments. In absolute term, firm's market entry was not necessarily large along the GMS highway. However, the market share of manufacturing firm increased near the highway while decreased in other districts. The share of manufacturing employment similarly increased by 4.5 percentage points near the highway while decreased by 3.4 percentage points in other districts. In districts out of 50km distance buffers, the share of service employments increased. This implies the agglomeration of manufacturing firms and employments towards districts near the highway.

4 Empirical Analysis

Given the background, we estimate the aggregate impact of the GMS program's investments in transport infrastructure projects. A community not only gain benefits from the direct access to highway, but also from improved accessibility to neighboring markets through the road network indirectly. We take advantage of geo-coded GMS highway network data and follow the "market access" approach (Donaldson and Hornbeck (2016)) to account for the GMS program's treatment spillovers. As derived from Eaton and Kortum (2002) model (hereafter, EK model), goods and factor markets interdependency across communes are represented by market access. Thus, direct and indirect effects of the highway are reflected in the changes in market access.

The sub-section 4.1 defines our measure of the market potential and industrial specialization which followed by the sub-section 4.2 examining how the construction of the GMS highway network affected local market size expansion. Then, sub-section 4.3 estimates GMS-induced market potential effects on spatial distributions of firms and workers. We also quantify how the agglomeration of economic activities created the welfare gains on firms' production and workers' wages.

4.1 Measurement of Market Access

Potential gains from GMS investments materialize through improvements in market accessibility by reducing travel time. We characterize the market access based on "market potential" which captures not only the size of own market, but also accounts for the market size in the distant markets where firms trade with. Market potential expresses effective density as the distance weighted sum of the employment of all districts within an accessible distance. The index is higher with stronger localization effects. Assuming the iceberg cost of trade, the Harris market potential index (MP) for district c at year t is defined as density of workers within an accessible distance (Combes and Gobillon (2015)):

$$MP_{ct} = \sum_{l \neq c} \frac{den_{lt}}{d_{cl}}, \tag{1}$$

where den_{lt} is employment density in district l at year t and d_{cl} is the shortest distance between People Committee's offices of district c and l. We calculate the shortest distance for all pairwise combinations of districts and keep neighboring districts within 50 km for each district.

In Vietnam, the market potential has concentrated in core cities within RRD and SE regions which expanded over time (Figure 5). From 2011 to 2016, we found faster increase in MP index in districts near the GMS highways, especially those along the HN-LS highway. All districts within 50km from GMS highways experienced an expansion of market potential, while the MP index decreased in mountainous districts. Thai Nguyen and Bac Giang provinces benefited the most from the highways where an increase in the MP index is greater than 50% (red color). While Bac Giang is an important midpoint in the HN-LS expressway, Thai Nguyen is close to both HN-LS and HN-LC highways. The locational advantage led to an expansion in the market potential. The MP index increased faster in the NE region than in the NW region, demonstrating larger project impacts in the NE region.

4.2 GMS Impact on Local Market Development

4.2.1 Description using high-resolution satellite images

Before discussing our formal empirical analysis, we begin with two sets of stylized facts about the agglomeration and the decentralization of economic activities using decennial high-resolution satellite images in 1990, 2000, 2010, and 2019.

We obtain the land use and land cover maps by applying image processing technique to satellite images of Landsat. The satellite image archives provided on the Google Earth

Engine enable us to track the growth and changes around the target areas in the past decades (Miyazaki et al. (2019)). We acquire cloud-free Landsat 5 and 8 imageries to prepare training data for the supervised classification of land cover mapping using a neural network algorithm. For mapping geographical extent of urban growth with high spatial accuracy, we look at the monthly average DNB radiance values of the VIIRS nighttime light. The spatial resolution of the VIIRS is 15 arc-seconds (about 500m grid) which is available after April 2012. [6]

Figure 6 depicts the long-term urban development along the GMS highway in the northern Vietnam over around three decades at the outset of the *Doi Moi* reform in 1990 to 2019. We define the land covers with four classes (waterbody, vegetation, bare land, and built-up area) and compute the percentage of built-up area in each commune. The built-up area had been less than 10% in most of communes until early 2000 except for the center of Hanoi. Urban development had accelerated since 2000, resulting in an increase in built-ups and the diffusion of settlements to peripheral cities near the coast (Haiphong and Ha Long) and toward the border with China (Lao Cai, Yen Bai). After 2010, the significant improvement in transportation network connectivity contributed to further expansion of the built-up areas along the GMS highway corridor. We also observe that the urban fringe has sprawled further down to the remote peripheral areas like the border with Laos (Hoa Binh, Son La) through an existing Asian Highway 13 (AH-13) corridor.

In Figure 7, dynamics of the spatial distribution of the VIIRS nighttime light (NTL) shows a consistent picture: (a) urban core cities near Hanoi and industrial hubs at the northeastern Vietnam experienced faster growth in economic activities; and (b) city fringe has spread out to peripheries near the highway network.

The Landsat imageries reveal that the metropolitan areas can be characterized as *monocentric*, particularly in the past, suggesting stronger agglomeration economies due to Marshall's local externalities in information sharing and/or externalities arising from demand and production linkage spillovers. In line with the system of cities model (Henderson (1974)), the urban core near Hanoi specializes in sectors with stronger agglomeration economies. With higher wages and land prices, the land development tended to be compact. On the other hand, the monocentric city model also predicts that lower transport costs within a city will induce scattered development (sprawling) in sectors where trade and commuting are easy with public transport. As the figures show, like the U.S. in 1970s (Burchfield et al. (2006)), the suburban areas outside Hanoi have become more *polycentric*

[6] Gibson et al. (2021) discusses the superiority of using VIIRS as a proxy for GDP in both urban and rural cities. Spatial inequality tends to be understated when the Defense Meteorological Satellite Program (DMSP) is used.

with decentralization of economic activities to peripheries. The extent of decentralization varies across cities and industries depending on the strength of agglomeration economies (Glaeser and Kahn (2001);Glaeser and Kahn (2004)).

4.2.2 Quantifying the extent of local market expansion

Built on the findings from the satellite images, this subsection estimates the magnitude and the spatial scale of the GMS project impacts on local market development in terms of market accessibility. Given the non-linearity of the project impact, we apply the Robinson (1988)'s semi-parametric regression model:

$$\ln MP_{drt} = x_{drt}\beta + m(GMSdist_{dr}) + \epsilon_{drt}, \tag{2}$$

where $m(\bullet)$ is a non-parametric component as a function of $GMSdist$, which is the Euclidian distance between the People Committee (PC) office of district d in region r (the economic center of each district) and the nearest interchange point of the GMS highway. The parametric part is composed of variables x_{drt}, including provincial dummies and the distance to the capital city, Hanoi. Figure 8 shows the relationship between the change in MP index and the distance to the GMS highways with the 95% confidence interval. For the non-parametric estimates of the MP index, we report results separately for the RRD region (which mainly covers core districts; the middle panel) and for the combined NE and NW regions (mainly the first and second peripheral districts; bottom panel). In each panel, the results on the left show four-year difference in each index between 2007 and 2011 (before the completion of the GMS highways), while the right charts show the four-year difference between 2012 and 2016 (after the project completion).

The left chart in Figure 8(a) shows that market grew equally around 40% in the core areas, while Figure 8(b) shows faster growth in the MP index in districts far from the highways' locations, reflecting long-term catch-up trend in the peripheral areas. The situation reversed after the completion of GMS highways. Over the period 2012-2016, the market potential expanded by around 20% for places within 20km and the positive effect gets smaller as the location moves away from the highway. The effect decreases to 15% in 40-60km and becomes zero in 70-100km. The market potential expanded even faster in the NE and NW regions at nearly 50% within 20km from the highway. This shows that the GMS significantly affected the market potential of places within 70km at different magnitude with potentially large agglomeration effects.

4.3 Main Empirical Specification

In the presence of agglomeration economies, the GMS transport projects would have an asymmetric impact on production growth across districts depending on spillover effects. To uncover the patterns empirically, we postulate the following fixed effect (FE) model with the lagged effects of the market potential:

$$\ln y_{ct} = \beta_0 + \beta_1 \ln MP_{ct-1} + X_{ct}\gamma + \mu_c + \tau_t + \varepsilon_{ct}, \quad (3)$$

where y is the commune-level outcomes such as the number of enterprises, real sales per worker, and profit margins. X_{ct} are commune-level control variables, such as the share of female workers and workers covered by social insurance, the share of employment in industrial zones, total trade values (imports and exports), and the distance from the PC office to the nearest GMS highway interchange point.

Besides market potential, the EK model predicts that the equilibrium outcome will also be affected by each commune's productivity and land supply. FE model, at least partly, purges time-invarying unobserved city factors and controls for time-varying district-specific local productivity shocks, both may simultaneously affect market potential and firm outcomes. The heterogeneous impact on the core and peripheral areas (defined earlier) is also examined by splitting the sample into districts in RRD region (core districts; panel A in Table 2) and in NE and NW regions (peripheral districts; panels B and C).

4.3.1 Fixed effect estimation results

Agglomeration through firm entries First, we investigate the extent of agglomeration economies from an angle of the nexus between the market potential and firm entry. In Table 2, results for the total sample are presented in columns 1-2 where we can easily verify the close connection between the market potential and the number of firms; and columns 3-6 report results for the sub-samples in manufacturing and service sectors, respectively. Service sector is defined in three categories: transport and related services (column 4), wholesale and retail trade (column 5), and other services (column 6). Columns 7-9 report results by the firm size: small firms (with less than 10 employees), medium-size firms (with greater than 10 but less than 100 employees), and large firms (with greater than 100 employees).

In columns 1-2, we compare the results with district-year FE and with commune FE. In the RRD region (Panel A), 1% increase in lagged market potential led to 34.2% increase in the number of firms at the district level. The effect is smaller at 11.8% for the periphery area (Panel B), which suggests stronger agglomeration forces at play in the core area.

The negative coefficient of the distance to the highway indicates that firms significantly agglomerate near the highway. With commune-level FE, the agglomeration effect of market potential significantly decreases. Yet, the agglomeration effect is still significant and large at around 12-14%. Even at the small spatial scale within communes, local market expansion has driven concentration of firms at the core district. The difference between the district- and commune-level results mainly captures the reallocation of firms between communes within a district.

In columns 3-6, we observe faster growth in the number of manufacturing firms in the core area at 11.6%. In both core and periphery areas, a presence of industrial zones tends to attract more manufacturing firms. In the core area, the number of registered firms grew significantly faster in the wholesale and retail service sector at 10.3%. In addition, results by the firm size (in columns 7-9) shows that an expansion of market potential attracted many small firms in the core region while larger firms increased as the market potential expanded in the peripheral regions.

The results indicate that the market expansion, partly driven by the GMS infrastructure investment, contributed to firms' agglomeration in which the core districts have increasingly shifted toward basic service trade industries. At the same time, the core areas broadened the industry base in an inclusive manner by promoting entry of small-sized enterprises. In contrast, peripheral districts seems to attract relatively large firms particularly in industrial zones.

Firm performance The reduction in intranational trade costs, induced by the GMS highway projects, is expected to create productivity gains on firms through various channels. In larger markets, increased competition from trade leads to higher productivity through within-industry resource reallocations (Melitz (2003)) as well as improved production efficiency by narrowing product scope (Abeberese and Chen (2022)) or reoptimizing capital-labor mix within firms. In addition, the agglomeration of firms would enhance firms and worker productivity, directly or in part through local spillovers between firms and workers (Baum-Snow et al. (2021)).

Table 3 and 4 collects FE estimates of the core parameter in Equation 3, β_1, for firm real sales per worker and profit margins by sectors and firm size, respectively, aggregated at the commune level. Firm's financial performance greatly fluctuate year-on-year which will be better captured in the FE specification than the long-difference specification.

In columns 1-4 of Table 3, FE estimates indicate insignificant effect on real sales per worker of manufacturing and service firms in the core cities. In the peripheral regions, on the other hand, an increase in market potential triggered significant increase in real sales per worker by 27.7% for manufacturing and by 23.7% for wholesale and retail trade services. In the peripheral region, significant positive effect on firms' sales for manufacturing and trade sectors implies that these firms tended to be more productive as number of manufacturing firms increased and local market has become more industrialized with larger demand.

Columns 5-8 of Table 3 shows that profit margins in the core cities significantly increased with the market potential in all sectors. In the core districts, profit margins increased by 16.6 percentage point (p.p.) for manufacturing, 22 p.p. for transport and related services, 26.3 p.p. for wholesale and retail services, and 79.2 p.p. for other services. Note that these patterns might be driven by firms' cost reduction behavior in response to reinforced competition due to firm entries. In the peripheral NW region, the GMS program impact on profit margins is significant only for manufacturing firms . The effect is large at 10.6%, reflecting that better infrastructure connectivity induced industrialization in the remote area at the northwestern border with China.

Table 4 reports the heterogeneity of the market potential effect by firm size. The effect on sales is negative, albeit statistically insignificant, among small firms that locates in the core districts, which seems consistent with our observation that small firms faced stronger competition due to the market expansion. The positive market potential effect on profit margins for small firms in the core area (column 7) reflects that small service firms cut production costs, mainly labor costs, to survive (columns 4-6).

In the peripheral districts, real labor costs grew fast with the elasticity of 0.17, which reflects faster wage growth and job creation mainly by the small-scale enterprises. Besides productivity growth, the asymmetric coefficients of the market potential variable on wage may also suggest factor price equalization across local labor market either through direct population movements or indirect goods and services trad. These results imply that the labor market adjustments enable wage rates to converge between the urban core and left-behind rural areas.

4.4 Instrumental Variable Regression

While the FE model using annual data addresses the omitted variable bias to some extent, there could be remaining bias arising from endogeneity problems. To mitigate such potential issues, we extend estimation of Equation 3 to IV regression that estimates the arguably causal effect of market potential growth on firm and employment growth over 5 years in the First-Difference (FD) specification.

The challenge to estimate the causal effect of highways is that, first, randomized experiments are almost impossible to design and implement by nature; and, second, actual roads are built in growing places or areas with development prospects to meet the traffic demand, or lagging areas to accelerate inclusive development, which suggests possible existence of reversed causality. We minimize this potential bias in estimation due to non-random allocation of highways and the induced market expansion in the inconsequential unit approach (Gibbons et al. (2019); Ahlfeldt and Feddersen (2018); Redding and Turner (2015); Chandra and Thompson (2000)) that complements the instrumental variable (IV) identification.[7] We restrict our sample to districts within 100km from the GMS transport network. In this way, we estimate effects from variations in the intensity of treatment (the change in market potential induced by the GMS highway) within the buffer zone of 100km from the project site. The identification comes from comparison of units experiencing larger and smaller changes in market potential, amongst the sub-sample that are all close to the road.

Previous literature has proposed various IV strategies to tackle the endogenous placement of infrastructure based on: (a) historical transport network (Baum-Snow et al. (2017); Duranton and Turner (2012); Michaels (2008)), (b) counterfactual least-cost-path network (Faber (2014)), (c) straight-line connections (Banerjee et al. (2020); Ghani et al. (2016)), (d) non-local changes to the road network (Jedwab and Storeygard (2020)), and (e) geographic barriers like land gradient and elevation (Dinkelman (2011); Duflo and Pande (2007)). The topological instrument has not been applied to transportation infrastructure, but certainly predicts the region's market potential growth by affecting the cost of firms' market entry and workers' occupational mobility.

Our IV strategy follows the last strand of related literature discussed above by exploiting the exogenous variation in topological and meteorological conditions in the northern Vietnam. As Figure 9 shows, topological and meteorological conditions significantly differ between flat hinterland of the Red River (core districts) and mountain plateaus in NE and NW regions that are inhabited by tribal groups (peripheral districts). The mountain plateaus are irregular in elevation and very rugged. Temperature and weather within RRD region are also diversified according to terrains and seasons. The initial geographic conditions would affect the amenity and create natural barriers for expanding local markets. The distance between each district's center and Hanoi also captures fixed costs of trade, which affects market potential. By exploiting the exogenous initial geographic variations, we will present the following first-difference IV regression:

[7] Even under the ideal setting with random road placement, local market expansion will be affected by the region's geographic location in the transportation network (Borusyak and Hull (2020)). Omitted variable bias may arise for the difference in unobserved productivity and amenity condition between the core and periphery regions. We plan to apply the recentered IV using counterfactual highway network to further improve our identification once the exact timing of the GMS highway construction could be recorded.

$$\Delta \ln MP_{ct} = \delta_0 + Z_{ct_0}\delta_1 + X_{ct_0}\delta_2 + \Delta\nu_{ct} \quad (4)$$

$$\Delta \ln y_{ct} = \beta_0 + \beta_1 \Delta \ln MP_{ct} + X_{ct_0}\gamma + \Delta\varepsilon_{ct}$$

where we use two census rounds in 2011 and 2016 to compute the first-differences; y is the number of enterprise and the sector share of employments at the commune-level; and X_{ct_0} is control variables including the initial share of female workers, workers with social security, and the employment share in industrial zones before the highway construction in 2011. We also fix the 2011 employment size in each commune to better identify the market potential growth due to the highway construction.

We estimate this model of first-difference IV model separately for the core and periphery subsamples. By doing so, we believe we can address non-linearlity arising from differentiated centrifugal and centripetal forces.

A set of our IVs, Z_{ct_0}, comprises 1999 district-level variations in historical geographic conditions (as a proxy for city amenities), including the elevation level, average temperature, amount of sunshine and precipitation, and the number of U.S. bombs, all taken from Miguel and Roland (2011). All of them are interacted with the distance to the GMS highway. Only IVs that are significantly relevant to $\Delta \ln MP$ are used for each sub-sample.

4.4.1 IV regression results

We verify the robustness of the evidence of firm agglomeration in subsection 4.3.1 using an alternative method of the first-difference (FD) two-stage least squares (2SLS) specification with refined data by restricting sample to districts within 100km from the GMS transport network. We also examine whether the expansion of market potential contributed to the changes in the industrial composition, i.e., the share of sector employment.

Agglomeration of firms The OLS and 2SLS estimation results are reported in Table 5. Table 6 provides the first-stage regression result. There are five main findings. First, results in Table 5 confirms large entries both in the core and periphery which can be seen from rapid increase in the number of registered firms in both cities. In columns 1 and 2, the magnitude of the market potential effect gets larger when the MP index is instrumented with topological variables.

Second, The first-stage estimates (in Table 6) confirms that the market potential is significantly affected by topological and meteorological conditions in a reasonable manner. The Kleibergen-Paap F statistics support that our IVs are valid. Yet, we also find that the preferred sets of independent variables are asymmetric between the core and periphery districts, suggesting the existence of complicated non-linealiries in mechanisms behind agglomeration economies.

Third, in consistent with the literature (Redding and Turner (2015); Jedwab and Storeygard (2020)), larger estimated coefficients from IV than those from OLS imply that the highway network was more likely to be built toward lagging areas in the northern Vietnam where the changes in market access have larger impacts. This significant downward bias may also be the result of measurement error in the market potential index.

Fourth, columns 3 and 4 show the 2SLS estimates by sector in which we can verify that the agglomeration effect is significant for the manufacturing sector in both core and peripheral districts. 1% increase in MP results in 1.2% increase in manufacturing in the core districts, while the agglomeration effect on manufacturing firms is 0.7% in the peripheries.

Finally, columns 5-7 show the result by the firm size. Medium to large firm significantly increased in the core, while both small and large firms increased in the peripheral areas.

Sectoral employment share Table 7 reports the estimation results on commune-level employment share of each sector. Similar to the agglomeration pattern of firms, employment significantly shifted towards the manufacturing sector in the peripheral cities, suggesting major structural transformation from agriculture to manufacturing in the periphery. The pace of industrialization seems faster in the peripheries than the core region. 1% expansion of the market potential drove an increase in the share of manufacturing workers by 0.3p.p (significant at 95%) in the NE and NW regions.

In the core cities, while FE results in Table 2 indicates an increase in the number of service firms, we also observe sectoral labor reallocation away from the wholesale and retail service sector. This may reflect that many trade service firms operate in small-scale or informal, and workers were absorbed to labor intensive industries.

4.5 Effects on Night Lights and Built-ups

Better market access could improve city welfare via lower prices and higher wages, which may attract population with higher outputs. To examine this hypothesis in our context, we employ novel satellite-based data, i.e., changes in (a) Night Lights (NTL) as a proxy for overall output and (b) the proportion of built-up areas captured by Landsat imageriesthat reflects human settlements and businesses. As discussed in 4.2.1, the VIIRS from 2012-2019 and the Landsat images from 2010-2019 are used for the estimation.

Columns 1-5 in Table 8 report results from NTL separately for core (Panel A) and peripheral regions (Panel B). FE estimate of the market potential coefficient shows much larger elasticity at 0.4 for the core city than the peripheries in level. Columns 2-3 look at the market access effect on NTL growth over 5 years (since the GMS highway construction) in OLS and 2SLS models, showing the same pattern of fast output growth in core cities. Columns 4-5 look at the change in NTL for the communes located within 50km from the GMS highway after the highway construction over five years (column 4) and eight years (column 5). The output grew significantly faster in areas close to the highway in both core and peripheral areas, suggesting significant program impacts on economic output, especially in later yearss.

Columns 6-10 report the estimates on changes in human settlements and businesses. We compute the proportion of built-up areas for each commune using ArcGIS zonal calculations based on our land-use classification map over decades (see Figure 6). To confirm the descriptive findings in the Figure 6, the estimate shows that 1% increase in market potential raises human settlements by 0.13 p.p mainly in core areas (column 6). In growth term, over five years since the GMS highway construction, better market potential significantly increased settlements only in peripheral cities (columns 7-8). However, if we look at communities near the highway, there was a significant and broader expansion in residential areas, with slightly faster growth in core cities (at 6.8 p.p) than peripheral cities (2.7 p.p). We believe that these results are largely consistent with a hypothesis that the GMS highway construction program facilitate economic development in both the core and periphery.

5 Mechanism of Polycentric Industrialization: Intermediate Production Spillovers

In theory, the market integration through lower transport costs facilitates industrial concentration to central regions with larger market at the expense of peripheral regions. The polycenetric development after the upgrade in inter-regional highways shows the co-existence of agglomeration of firms to core areas as well as diffusion of production to peripheral areas. This would be a reflection of alternative mechanisms at play beyond the market size (home market effect), which the canonical Krugman (1991) and Fujita et al. (1999) model cannot explain.

Krugman and Venables (1995) considers the interaction between transport costs and trade in differentiated intermediate goods, which creates externalities by linking firms and consumers. The development of supply chains network makes peripheral region, with higher transport costs but cheaper initial wages, more attractive. As backward linkage with upstream industries and forward linkage with downstream customers tighten, positive intermediate production spillovers would reshape the core-periphery pattern. As transport costs drop further, the model predicts that the industrial location could diffuse to peripheral regions, which results in convergence in core-periphery income disparity.

5.1 Generalized Difference-in-Difference

This section estimates the direct effect of GMS highway on firms and employments, accounting for intermediate production spillovers. We use the generalized difference-in-difference (DID) design, which exploits within-commune variation over time using two-way fixed effects (Lindo et al. (2020); Callaway et al. (2021)).

The treatment cannot be defined locally at point as transportation infrastructure affects the neighborhood along road network. Thus, our event study deals with two continuous treatments: (a) different exposure to transport cost reduction by physical locational distance and (b) input-output linkages through transport network (Ellison et al. (2010)). Specifically, we postulate the following empirical model:

$$y_{ct} = \beta \mathbb{1}[Dist_c \leq d] \times POST + X_{ct}\gamma + \mu_c + \tau_t + \lambda_{dt} + \varepsilon_{ct} \tag{5}$$

where y_{ct} is the number of firms and employment in commune c for year t. $Dist$ is the distance from the GMS highway, measuring accessibility to the highway for firms locating in each commune. The distance threshold is discretized into d=20km and 50km. The variable, $POST$, takes one for the post-GMS construction years ($t \geq 2015$). μ_c, τ_t, and λ_{dt} are commune and year fixed effects, and district-level trend prior to the start year of the GMS highway upgrades. X_{ct} is commune-level covariates. Standard errors are clustered at the district level to account for serial correlation and the correlation across communes.

The identification comes from changes in firm outcomes for communes with small changes in the exposure to highway as a counterfactual for the changes in outcomes that would have been observed for communes with larger changes in the exposure.

The effect of highway construction would be heterogeneous depending on the degree of intermediate goods transactions and the location of each commune from the highway. Following Li et al. (2017), one way to measure the supply chain linkage is to use the input-output (IO) table that reports each industry's input value from other industries. Let A denote the matrix of IO coefficient a_{ij}'s:

$$A_{ij} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} \\ a_{21} & a_{22} & \cdots & a_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i,2} & \cdots & a_{ij} \end{pmatrix}$$

As a measure the degree of IO linkages, we define the Leontief inverse of the IO matrix A as $L \equiv (I - A)^{-1}$. We use the OECD's harmonized IO table, covering 45 sectors, which is available every year until 2018. This allows us to measure the change in intermediate production linkages. Then, we define the intermediate production multiplier (treatment variable T_{ct}) as the average of L_{st} for each commune c in year t using the initial share of each industry s's sales as the weight:

$$T_{ct} = \sum_{s} L_{st} \frac{Sales_{cst}}{\sum Sales_{ct}} \tag{6}$$

Figure 10 plots the spatial correlation between the growth in production multiplier (x-axis) and the number of manufacturing and trade firms (y-axis) from 20010 to 2016. Descriptively, the correlation is strongly positive, which underscores the importance of the depth of supply chains in explaining the agglomeration in Vietnam.

5.1.1 Baseline regression

As a baseline, we present estimates with the log of production multiplier (ln(multiplier)) as a linear term and discretize physical distance into a set of indicator variables, including an indicator for being 0-20 km from the highway and an indicator for being 20-50km from the highway.

In terms of the distance, semi-parametric estimates in subsection 4.2.2 reveal that the districts within 20km distance buffer from the GMS highway experienced rapid expansion in the market potential, while the market expansion decreases to zero above 50km from the highway. Based on this, we use 20km and 50km as the boundary of the highway's treatment zone.

Results in tables 9 and table 10 indicate that relative to communes located in remote areas (more than 50km away from the highway as the reference group), firms' agglomeration is particularly strong in communes close to the highway. In core regions, the effect is the largest in the immediate neighborhood of the highway (within 20km). For trade business (e.g., retail, food and accommodation) which serves consumers, the agglomeration effect remains strong within 50km distance buffer, reflecting demand spillover.

Similarly, the effect is positive in peripheral regions where the positive effect remains strong in communes within 20-50km from the highway. This suggest some spatial spillover of the GMS impact to the outskirt places which are connected to the GMS highway through local transport network. Several second-tier cities (such as Hoa Binh, Son La, and Thai Nguyen), being connected to the GMS highway through main national highways, have experienced the expansion of manufacturing production (e.g., food processing, textile) and trade services. The wider geographic scale of agglomeration in the peripheral region seems consistent with the satellite-based evidence in Figure 6 as well as the market expansion in peripheral areas further away from the GMS highway as our first-stage estimate shows in Table 6.

5.2 Heterogeneous treatment by supply chain linkages

The target causal parameter β (GMS highway impact) would be larger if firms in the commune are strongly interconnected to firms located in the same or different communes within accessible markets.

Next, we aim to estimate the average causal response (ACR) to a unit change in production multiplier T. For each distance group, we estimate ATT for communes with treatment (IO linkages) level $j \in T$ relative to adjacent communes with treatment level $j-1 \in T$. Regardless of physical distance, the impact of GMS would differ by each commune's particular production structure in local supply chains. Like Lindo et al. (2020), one way to recover the ACR for treatment group j is to estimate ATT for a given treatment level j (i.e., the interaction term β):

$$y_{ct} = \beta \mathbb{1}[Dist_c \leq d] \times POST \times T_{ct} + \beta_2 T_{ct} + \gamma X_{ct} + \mu_c + \tau_t + \lambda_{dt} + \varepsilon_{ct} \tag{7}$$

As in a DID design, we need a parallel trend assumption. In the continuous treatment DID design, we compare the heterogeneity in gains from the treatment, i.e., the ATT for treatment group j and a group just below it $j-1$ in the supply chain (Callaway et al. (2021)). The identification needs equal ATT, i.e., communes at the different location of the supply chain should be similarly treated by the GMS investment.

In reality, each commune would have heterogeneous expected returns from the GMS by moving up or down the industrial layer of Vietnam's supply chain due to several factors, such as an initial industrial composition and the level of industrial development in each locality. If firms choose locations based on the difference in an expected treatment effect, this could potentially bias our estimates for a particular ACR group. To minimize the selection bias, our approach is to control for pre-trend term (the growth of number of firms or employment before the GMS investment) in all our regressions 5 and 7.

5.2.1 Results of heterogeneous treatment effect

In Figure 11, we plot the coefficient β of the ACR on treatment group j. The effect is allowed to be different across each percentile of the distance from the GMS highway. The ACR on the number of firms and employments are reported for core and periphery regions, separately.

Agglomeration of firms For the firm agglomeration effect (Figure 11), the ACR results for the core region (left panel) show significant agglomeration of firms in the communes within 50km from the GMS highway for both manufacturing and trade sectors. The agglomeration effect is stronger as the proximity to the GMS highway gets closer. Within 10-30km from the highway, the ACR on manufacturing (left-upper panel) is slightly above 10% which declines to zero for places outside 60km. Similarly, the agglomeration effect on the number of trade firms (left-bottom panel) is the strongest within 10km, but the positive effect extends to the areas at the outskirt of the core cities. This demonstrates the formation of small trade start-ups in surrounding areas due to the positive demand spillover from the central business areas.

At the peripheral region (right panel), the relationship between the ACR and the distance from GMS appears to be highly non-linear. However, the point estimate clearly shows even stronger agglomeration effect on manufacturing business at 20% within 10km from the highway (right-upper panel). We also observe strong positive externality in communes located 50-60km away from the GMS highway. This is likely be driven by strong agglomeration and local input-output spillovers occurred in several large second-tier cities. Indeed, when second-tier cities are excluded, the effect gets flattened in remote areas as similarly found in the core region. Strong agglomeration for manufacturing firms located near the GMS highway and the spillover effect through local transport network seem the salient feature of industrialization in the peripheral region.

Generally speaking, large ACR near the highway supports positive externality due to input-output linkage. The vicinity of GMS highway benefited directly from lower transport costs, which gets much bigger due to production spillovers if local industries are tightly connected in local supply chains through the input-output linkages. The production externality is generally weaker in the peripheral region except for large secondary cities located along the highways, though the direct impact of the GMS highway is equally large on average.

Job creation Figure 12 shows the GMS effect on commune's total employment through job creation. Similar to the effect on firms, we find significant treatment externality on manufacturing employments in both core and periphery regions near the GMS highway. Within 20-30km from the highway, the ACR on manufacturing and trade job creation is sizable at 20-30% for core region and 50% for peripheral regions, which is even larger than the agglomeration effect on firms.

Similar to the result on firms, the relationship between the multiplier effect on employment and the distance from the GMS highway is non-linear in the peripheral regions. The job creation seems local: particularly large multiplier effect on job creation in communes within 10km from the highway, while positive externality also gets significant in places 50-60km away from the GMS highway as previously found in Figure 11. The relation gets smoother when second-tier cities are excluded, suggesting that the non-linearity is driven by several big cities in peripheral areas.

For trade business, the spillover effect on job creation remains large, extending to communes outside 50km from the highway, which reflects demand spillovers from downstream customers regardless of the distance from the GMS highway.

The overall findings from the generalized DID estimation supports that the GMS project facilitated significant economy-wide industrialization. In core region, production externality on manufacturing firms and employments concentrates in communes near the highway, which shapes typical core-periphery pattern. For service trade, positive externality diffuses more equality over space possibly through consumers' demand linkages. In the peripheral regions, the spillovers appear to cover wider geographic scale due to production linkages via local transport networks.

6 Conclusions

This paper evaluated the GMS programs' heterogeneous effects on the market potential, the agglomeration of firms and workers, and industrial development between the core and peripheral cities in the northern Vietnam over the last decade.

The improved accessibility due to the GMS transport investments had significantly expanded market potential of the treated districts within 20km from the highway, through which wages and firm profits were enhanced. The market potential grew faster by around 40% in peripheral cities where the local market was initially underdeveloped.

We found strong agglomeration economies among the manufacturing firms along the highway in both core and peripheral cities which experienced large market expansion due to the GMS transportation network. The reduction in transport cost has also induced the decentralization of manufacturing activities to peripheries, contributing to emergence of polycentric urban development. Rural peripheries also experienced significant labor reallocation through which the share of manufacturing employments significantly increased near the highway.

The agglomeration increased firm productivity disproportionately more in the rural peripheries, showing large marginal effects of the GMS highway construction in the remote area. The significant productivity growth in the peripheries were driven by a rapid real wage growth through industrialization. At the metropolitan area, better market access significantly increased firms' profit margins, although small businesses needed to reduce labor costs to survive the intense market competition. The overall findings suggest that the interregional GMS transport investment (a) facilitated the economy-wide industrialization through firm entries and (b) led to significant spatial income convergence between urban core and rural peripheral cities.

In the next step, we need to uncover mechanisms behind the observed spatial development patterns based on the reduced-form empirical models. Besides locational fundamentals or the "first-nature" geography *a la* Redding (2020), it requires us to formally model the locational choice of firms in response to transport improvement in a multi-sector setup. This would allow us to structurally interpret the interplay between transport investment, market access, and spatial structural transformation. With the microfounded spatial general equilibrium model, we plan to quantify the main channels of the spatial equilibrium and welfare effects of the GMS network (Allen and Arkolakis (2014); Redding and Rossi-Hansberg (2017)). We believe the spatial equilibrium analysis will help us decompose the welfare effect into the Ricardian channel, factor reallocation, policy complementarity, institutional barrier, or path dependence. This is one of the most important tasks for our research in the future.

References

Abeberese, A. B. and M. Chen (2022). Intranational trade costs, product scope and productivity: Evidence from India's Golden Quadrilateral project. *Journal of Development Economics 156*, 102791.

Ahlfeldt, G. M. and A. Feddersen (2018). From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography 18*(2), 355–390.

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *Quarterly Journal of Economics 129*(3), 1085–1140.

Asian Development Bank (2020). *Asia's Journey to Prosperity: Policy, Market, and Technology Over 50 Years*. Manila, Philippines: Asian Development Bank.

Banerjee, A., E. Duflo, and N. Qian (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics 145*, 102442.

Baum-Snow, N. (2007). Did highways cause suburbanization? *Quarterly Journal of Economics 122*(2), 775–805.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics 99*(3), 435–448.

Baum-Snow, N., N. Gendron-Carrier, and R. Pavan (2021). Local productivity spillovers.

Borusyak, K. and P. Hull (2020). Non-random exposure to exogenous shocks: Theory and applications.

Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner (2006). Causes of sprawl: A portrait from space. *Quarterly Journal of Economics 121*(2), 587–633.

Callaway, B., A. Goodman-Bacon, and P. H. Sant'Anna (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.

Chandra, A. and E. Thompson (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics 30*(4), 457–490.

Combes, P.-P. and L. Gobillon (2015). The empirics of agglomeration economies. *Handbook of Regional and Urban Economics 5*, 247–348.

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review 101*(7), 3078–3108.

Donaldson, D. and R. Hornbeck (2016). Railroads and American economic growth: A "market access" approach. *Quarterly Journal of Economics 131*(2), 799–858.

Donaldson, D. and A. Storeygard (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives 30*(4), 171–98.

Duflo, E. and R. Pande (2007). Dams. *Quarterly Journal of Economics 122*(2), 601–646.

Duranton, G. and M. A. Turner (2012). Urban growth and transportation. *Review of Economic Studies 79*(4), 1407–1440.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Ellison, G., E. L. Glaeser, and W. R. Kerr (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review 100*(3), 1195–1213.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from China's national trunk highway system. *Review of Economic Studies 81*(3), 1046–1070.

Fajgelbaum, P. and S. J. Redding (2021). Trade, structural transformation and development: Evidence from Argentina 1869-1914.

Fujita, M., P. R. Krugman, and A. Venables (1999). *The spatial economy: Cities, regions, and international trade.* MIT press.

Ghani, E., A. G. Goswami, and W. R. Kerr (2016). Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing. *Economic Journal 126*(591), 317–357.

Gibbons, S., T. Lyytikäinen, H. G. Overman, and R. Sanchis-Guarner (2019). New road infrastructure: The effects on firms. *Journal of Urban Economics 110*, 35–50.

Gibson, J., S. Olivia, G. Boe-Gibson, and C. Li (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics 149*.

Glaeser, E. L. and M. E. Kahn (2001). Decentralized employment and the transformation of the American city.

Glaeser, E. L. and M. E. Kahn (2004). Sprawl and urban growth. *Handbook of Regional and Urban Economics 4*, 2481–2527.

GMS Secretariat (2018). Assessment of gms economic corridors - Viet Nam. Report.

Henderson, J. V. (1974). The sizes and types of cities. *American Economic Review*, 640–656.

Jedwab, R. and A. Storeygard (2020). The average and heterogeneous effects of transportation investments: Evidence from sub-Saharan Africa 1960-2010. *Journal of the European Economic Association.*

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 483–499.

Krugman, P. and A. J. Venables (1995). Globalization and the inequality of nations. *Quarterly journal of economics 110*(4), 857–880.

Li, Z., M. Wu, and B. R. Chen (2017). Is road infrastructure investment in China excessive? evidence from productivity of firms. *Regional Science and Urban Economics 65*, 116–126.

Lindo, J. M., C. K. Myers, A. Schlosser, and S. Cunningham (2020). How far is too far? new evidence on abortion clinic closures, access, and abortions. *Journal of Human resources 55*(4), 1137–1160.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *Review of Economics and Statistics 90*(4), 683–701.

Miguel, E. and G. Roland (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics 96*(1), 1–15.

Miyazaki, H., H. Bhushan, and K. Wakiya (2019). Urban growth modeling using historical landsat satellite data archive on google earth engine.

Nagy, D. K. (2020). Hinterlands, city formation and growth: Evidence from the U.S. westward expansion. University of Pompeu Fabra.

Nose, M. (2020). Minimum wage, market imperfection, and production efficiency: Panel event study in Vietnam.

Redding, S. J. (2020). Trade and geography. NBER Working Paper No. 27821.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*(1), 21–58.

Redding, S. J. and M. A. Turner (2015). Transportation costs and the spatial organization of economic activity. *Handbook of Regional and Urban Economics 5*, 1339–1398.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica 56*, 931–954.

Trew, A. (2020). Endogenous infrastructure development and spatial takeoff in the first industrial revolution. *American Economic Journal: Macroeconomics 12*(2), 44–93.

Table 1: Descriptive Statistics

(a) Change in district population: 1999-2018

8 economic regions	Number of districts in 2018	Population in 1999 (Mean)	Population in 2009 (Mean)	Population in 2018 (Mean)	Pop. Growth: 1999-2009 (Mean)	Pop. Growth: 2009-18 (Mean)
1. Red River Delta	128	139,334	152,956	194,308	0.10	0.22
2. North East	80	85,662	86,557	128,127	0.05	0.35
3. North West	59	59,368	69,103	111,505	0.19	0.48
4. Central North	85	115,953	116,684	151,873	0.02	0.27
5. Coastal Central	83	98,106	105,411	144,718	0.10	0.36
6. Central Highlands	61	66,469	83,710	127,587	0.31	0.41
7. South East	71	142,124	198,060	257,512	0.32	0.25
8. Mekong River Delta	132	122,281	129,868	145,460	0.07	0.09
Total	**699**	**109,816**	**122,393**	**160,069**	**0.13**	**0.27**

Source: Population Census 1999 and VHLSS 2018

(b) Change in the number of enterprises: 2011-2016

8 economic regions	Total enterprises		Manufacturing		Service	
	2011	2016	2011	2016	2011	2016
1. Red River Delta	104,022	157,278	16,188	23,747	74,641	113,271
2. North East	9,011	12,472	1,551	2,191	5,022	7,446
3. North West	5,013	7,144	614	774	2,570	3,981
4. Central North	18,565	26,158	2,198	3,052	11,506	16,803
5. Coastal Central	24,120	37,647	3,517	4,844	15,935	25,826
6. Central Highlands	8,512	12,921	964	1,206	5,284	8,943
7. South East	129,168	212,442	22,854	33,275	91,096	155,419
8. Mekong River Delta	27,169	37,366	4,762	6,071	17,048	24,584
Total	**325,580**	**503,428**	**52,648**	**75,160**	**223,102**	**356,273**

Source: VES 2011 and 2016

Table 2: FE Estimates: The Market Potential Effect on the Agglomeration of Firms

	Log (number of firms)								
	All		Sector				Size		
	(1)	(2)	(3) Manufacturing	(4) Service: trans.	(5) Service: trade	(6) Service: other	(7) Small	(8) Medium	(9) Large
A. RRD region									
Log MP_{t-1}	0.342***	0.137***	0.116***	0.022	0.103***	0.120***	0.182***	0.008	0.040
	(0.033)	(0.027)	(0.030)	(0.025)	(0.031)	(0.030)	(0.033)	(0.031)	(0.025)
Emp. Share in industrial zone$_{t-1}$	0.108**	0.048	0.123***	0.054	0.028	0.064*	0.073**	0.021	0.142***
	(0.051)	(0.030)	(0.043)	(0.042)	(0.036)	(0.038)	(0.034)	(0.046)	(0.045)
Log distance to GMS highway	-0.067**								
	(0.030)								
N	13,266	13,227	13,227	13,227	13,227	13,227	13,227	13,227	13,227
N(districts)	125								
N(communes)		2292	2292	2292	2292	2292	2292	2292	2292
B. NE & NW regions									
Log MP_{t-1}	0.118***	0.124***	0.097**	0.052	-0.001	-0.056	0.075	0.100**	0.083***
	(0.030)	(0.042)	(0.041)	(0.036)	(0.041)	(0.040)	(0.048)	(0.047)	(0.028)
Emp. Share in industrial zone$_{t-1}$	-0.075	0.173***	0.172***	0.129***	0.176***	0.203***	0.183***	0.131**	0.296***
	(0.096)	(0.051)	(0.058)	(0.049)	(0.049)	(0.054)	(0.062)	(0.064)	(0.057)
Log distance to GMS highway	-0.151***								
	(0.039)								
N	6,259	6,168	6,168	6,168	6,168	6,168	6,168	6,168	6,168
N(districts)	109								
N(communes)		1157	1157	1157	1157	1157	1157	1157	1157
District - Year FE	YES	NO	NO	NO	NO	NO	NO	NO	NO
Commune FE	NO	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the level of fixed effect, are reported in the bracket. Other control variables include the shares of female and insured workers, and total regional trade values.

Table 3: FE Estimates: The Market Potential Effect on Firm Production (By Sector)

	Log (sales per worker)				**Profit margins**			
	(1) Manufacturing	(2) Service: trans.	(3) Service: trade	(4) Service: other	(5) Manufacturing	(6) Service: trans.	(7) Service: trade	(8) Service: other
A. RRD region								
Log MP_{t-1}	0.093	0.089	-0.060	0.124	0.166***	0.220***	0.263***	0.792***
	(0.085)	(0.089)	(0.079)	(0.111)	(0.060)	(0.060)	(0.060)	(0.157)
N	10,139	6,581	11,134	7,019	10,112	6,584	11,140	6,957
N(communes)	1879	1308	2018	1389	1876	1308	2019	1382
B. NE & NW regions								
Log MP_{t-1}	0.277*	0.185	0.237*	-0.095	0.106*	0.046	0.031	0.083
	(0.165)	(0.186)	(0.127)	(0.190)	(0.060)	(0.048)	(0.029)	(0.078)
N	3,460	1,989	3,825	1,788	3,457	1,987	3,830	1,756
N(communes)	693	420	760	377	692	422	760	370
Commune FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the level of fixed effect, are reported in the bracket. Other control variables include the shares of female and insured workers, and total regional trade values.

Table 4: FE Estimates: The Market Potential Effect on Firm Production (By Size)

	Log (sales per worker)			**Log (real labor costs)**			**Profit margins**		
	(1) Small	(2) Medium	(3) Large	(4) Small	(5) Medium	(6) Large	(7) Small	(8) Medium	(9) Large
A. RRD region									
Log MP_{t-1}	-0.074	0.038	0.034	-0.669***	-0.335***	0.025	0.314***	-0.001	0.021
	(0.068)	(0.076)	(0.080)	(0.080)	(0.057)	(0.043)	(0.054)	(0.055)	(0.027)
N	12,443	10,162	4,193	12,555	10,184	4,186	12,432	10,155	4,191
N(communes)	2207	1935	849	2217	1938	851	2205	1933	849
B. NE & NW regions									
Log MP_{t-1}	0.119	-0.083	0.006	0.168**	0.073	0.257	0.058	-0.077**	0.001
	(0.125)	(0.128)	(0.212)	(0.075)	(0.074)	(0.203)	(0.047)	(0.036)	(0.061)
N	5,158	3,632	1,031	5,287	3,648	1,043	5,129	3,613	1,027
N(communes)	1011	742	226	1038	745	229	1010	738	226
Commune FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the level of fixed effect, are reported in the bracket. Other control variables include the shares of female and insured workers, and total regional trade values.

Table 5: FD Estimates: The Market Potential Effect on the Agglomeration of Firms

	Δ Log(number of firms)						
	OLS	2SLS					
	(1) All	(2) All	(3) Manufacturing	(4) Service	(5) Small	(6) Medium	(7) Large
A. RRD region							
Δ Log MP	0.449*	1.085*	1.168***	0.849	0.651	1.281***	0.682**
	(0.257)	(0.558)	(0.442)	(0.569)	(0.620)	(0.456)	(0.309)
Log employment in 2011	-0.015*	-0.015*	-0.001	0.011	0.020**	-0.037***	0.005
	(0.008)	(0.008)	(0.007)	(0.009)	(0.009)	(0.008)	(0.005)
N	2,197	2,197	2,197	2,197	2,197	2,197	2,197
Kleibergen-Paap F statistics		11.88	11.88	11.88	11.88	11.88	11.88
Hansen's J-statistics (p-value)		0.0012	0.0031	0.0340	0.0262	0.0077	0.0150
B. NE & NW regions							
Δ Log MP	0.226**	0.596**	0.686**	0.214	0.528*	0.301	0.135*
	(0.106)	(0.238)	(0.267)	(0.214)	(0.273)	(0.209)	(0.078)
Log employment in 2011	0.043***	0.040***	0.010	0.056***	0.091***	-0.050***	0.003
	(0.009)	(0.010)	(0.009)	(0.010)	(0.012)	(0.009)	(0.009)
N	1,038	970	970	970	970	970	970
Kleibergen-Paap F statistics		15.76	15.76	15.76	15.76	15.76	15.76
Hansen's J-statistics (p-value)		0.5619	0.3747	0.5782	0.2134	0.1306	0.0014
Region FE	YES	YES	YES	YES	YES	YES	YES

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the district level, are reported in the bracket. First differences are computed from two firm census rounds (the VES 2011 and 2016). Other control variables include the initial shares of female and insured workers, and the employment share in industrial zone in 2011. Instrumental variables are topological condition in 1999, including elevation level, average temperature and amount of sunshine and precipitation, and the distance from the highway.

Table 6: FD Estimates: First-stage regression

	Δ Log MP	
	(1)	(2)
	RRD	NE & NW
Land elevation in 1999 x Log distance to GMS	-0.023	-0.022***
	(0.015)	(0.006)
Avg. temperature in 1999 x Log distance to GMS	0.082***	
	(0.028)	
Avg. sunshine in 1999 x Log distance to GMS	-0.075***	
	(0.017)	
Avg. precipitation in 1999 x Log distance to GMS	-0.028***	
	(0.006)	
Number of U.S bombs x Log distance to GMS		0.020***
		(0.004)
Log distance to GMS highways	-0.015**	0.048***
	(0.007)	(0.017)
N	2,380	1,185
R-squared	0.477	0.374
Region FE	YES	YES

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the district level, are reported in the bracket. All instruments except the distance to GMS highway are in z-score.

Table 7: FD Estimates: The Market Potential Effect on Sector Employment Share

	Δ Share of sector employments				
	2SLS	2SLS			
	(1) Manufacturing	(2) Service	(3) Service: trans.	(4) Service: trade	(5) Service: other
A. RRD region					
Δ Log MP	0.255	-0.128	0.229*	-0.496**	0.139
	(0.216)	(0.208)	(0.125)	(0.213)	(0.100)
Log employment in 2011	-0.013***	0.027***	0.001	0.025***	0.001
	(0.004)	(0.005)	(0.002)	(0.004)	(0.003)
N	2,162	2,162	2,162	2,162	2,162
Kleibergen-Paap F statistics	11.46	11.46	11.46	11.46	11.46
Hansen's J-statistics (p-value)	0.2718	0.0561	0.1598	0.0061	0.1635
B. NE & NW regions					
Δ Log MP	0.302**	-0.104	0.019	-0.057	-0.066
	(0.148)	(0.118)	(0.033)	(0.089)	(0.109)
Log employment in 2011	-0.015**	0.030***	-0.005*	0.028***	0.006**
	(0.007)	(0.008)	(0.003)	(0.008)	(0.003)
N	893	893	893	893	893
Kleibergen-Paap F statistics	14.11	14.11	14.11	14.11	14.11
Hansen's J-statistics (p-value)	0.3680	0.8310	0.9526	0.7891	0.4575
Region FE	YES	YES	YES	YES	YES

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the district level, are reported in the bracket. First differences are computed from two firm census rounds (the VES 2011 and 2016). Other control variables include the initial shares of female and insured workers, and the employment share in industrial zone in 2011. Instrumental variables are topological condition in 1999, including elevation level, average temperature and amount of sunshine and precipitation, and the distance from the highway.

Table 8: The Effects on Night-time Light and Built-ups

	Night-time Light (NTL)					The proportion of built-up area (LC)				
	(1) Log NTL	(2) ΔLog NTL 2011-16	(3) ΔLog NTL 2011-16	(4) Log NTL 2011-16	(5) Log NTL 2011-19	(6) Land Cover	(7) Δ LC 2010-15	(8) Δ LC 2010-15	(9) LC 2010-15	(10) LC 2010-19
A. RRD region										
Log MP_{t-1}	0.399***					0.127***				
	(0.033)					(0.031)				
Δ Log MP		0.362*	0.790***				-0.027	0.087		
		(0.189)	(0.235)				(0.044)	(0.089)		
Within 50km × Post				0.112**	0.172***				0.033*	0.068***
				(0.043)	(0.043)				(0.019)	(0.005)
N	13,266	2,197	2,197	4,712	4,712	2,248	2,193	2,193	4,702	4,679
Kleibergen-Paap F statistics			11.66					11.99		
B. NE & NW regions										
Log MP_{t-1}	0.058***					0.016				
	(0.013)					(0.010)				
ΔLog MP		0.085	0.092				0.047***	0.063*		
		(0.058)	(0.105)				(0.017)	(0.037)		
Within 50km × Post				0.055	0.123***				0.001	0.027***
				(0.039)	(0.038)				(0.005)	(0.005)
N	6,259	1,038	970	2,614	2,614	1,084	1,038	970	2,614	2,614
Kleibergen-Paap F statistics			14.86					15.76		
Specification	FE(OLS)	FD(OLS)	FD(2SLS)	DID	DID	FE(OLS)	FD(OLS)	FD(2SLS)	DID	DID
District - Year FE	YES			YES	YES	YES			YES	YES
Region FE		YES	YES				YES	YES		

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the district level, are reported in the bracket. Annual night-time light is computed as the medium-value of monthly VIIRs data. First-difference (FD) and the difference-in-difference (DID) regressions use two data points from pre- and post-GMS highway construction period. Same IVs as in Equation 5 are used for columns (3) and (8).

Table 9: Generalized DID Estimates: GMS's Agglomeration Effect

	Log # of firms					
	RRD region			NE & NW regions		
	(1) Manufacturing	(2) Service: transport	(3) Service: trade	(4) Manufacturing	(5) Service: transport	(6) Service: trade
$\mathbb{1}[Dist \leq 20km]$	0.148***	0.181***	0.175***	0.105***	0.109***	0.159***
	(0.017)	(0.018)	(0.019)	(0.022)	(0.023)	(0.029)
$\mathbb{1}[20km < Dist \leq 50km]$	0.096***	0.061**	0.167***	0.123***	0.068***	0.151***
	(0.031)	(0.027)	(0.036)	(0.030)	(0.023)	(0.031)
Log IO multiplier	0.050***	0.042***	0.062***	0.057***	0.071***	0.053***
	(0.008)	(0.008)	(0.009)	(0.017)	(0.012)	(0.015)
N	15,087	15,087	15,087	6,825	6,825	6,825
Commune FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Pre-trend	YES	YES	YES	YES	YES	YES

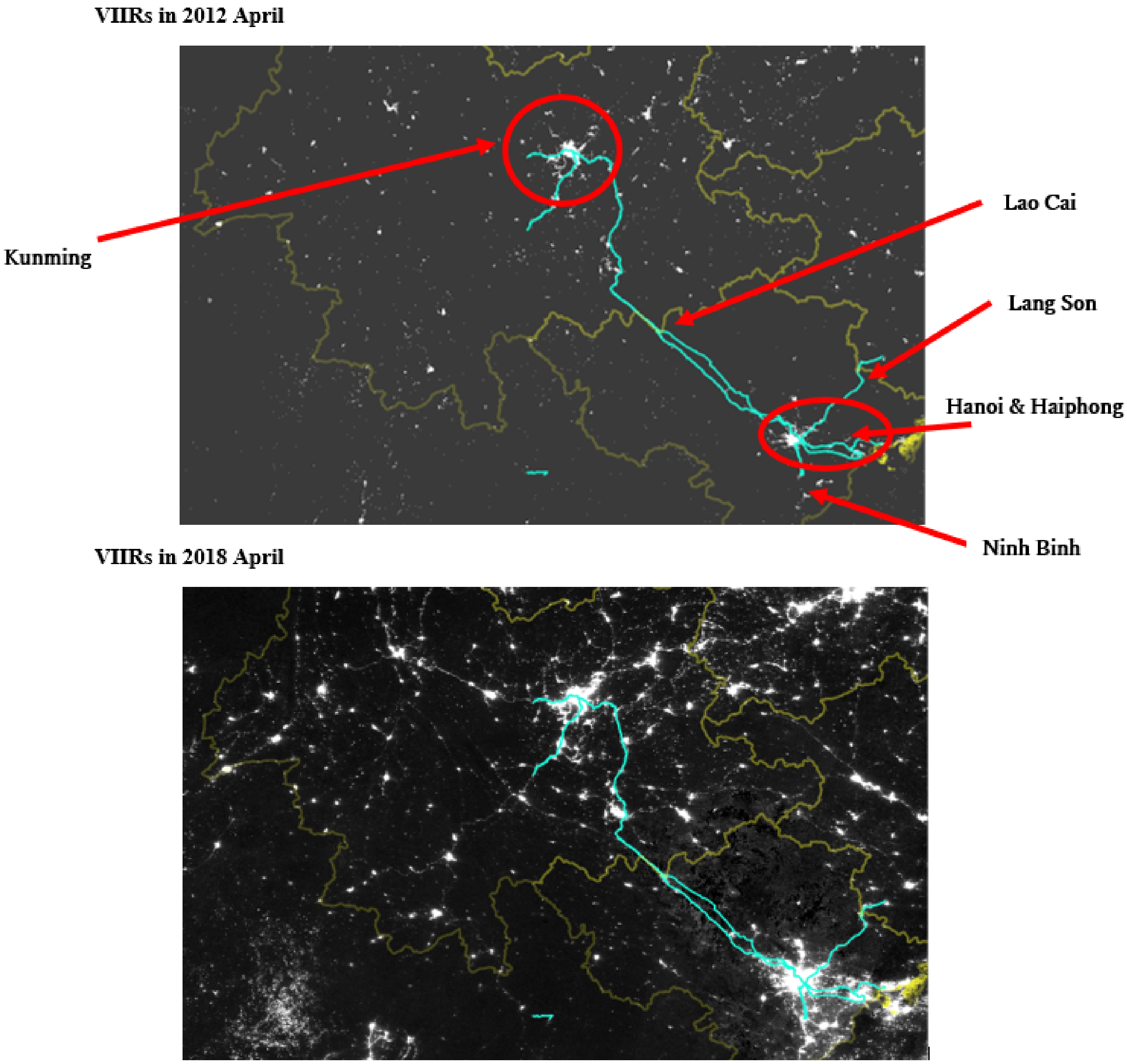
Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the district level, are reported in the bracket.

Table 10: Generalized DID Estimates: GMS's Effect on Job Creation

	Log employment					
	RRD region			NE & NW regions		
	(1) Manufacturing	(2) Service: transport	(3) Service: trade	(4) Manufacturing	(5) Service: transport	(6) Service: trade
$\mathbb{1}[Dist \leq 20km]$	0.185***	0.300***	0.208***	0.157***	0.196***	0.243***
	(0.033)	(0.030)	(0.024)	(0.059)	(0.054)	(0.063)
$\mathbb{1}[20km < Dist \leq 50km]$	0.181***	0.084**	0.192***	0.185***	0.170***	0.221***
	(0.043)	(0.040)	(0.047)	(0.061)	(0.054)	(0.041)
Log IO multiplier	0.118***	0.103***	0.133***	0.206***	0.219***	0.126***
	(0.030)	(0.021)	(0.021)	(0.057)	(0.033)	(0.038)
N	15,087	15,087	15,087	6,825	6,825	6,825
Commune FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Pre-trend	YES	YES	YES	YES	YES	YES

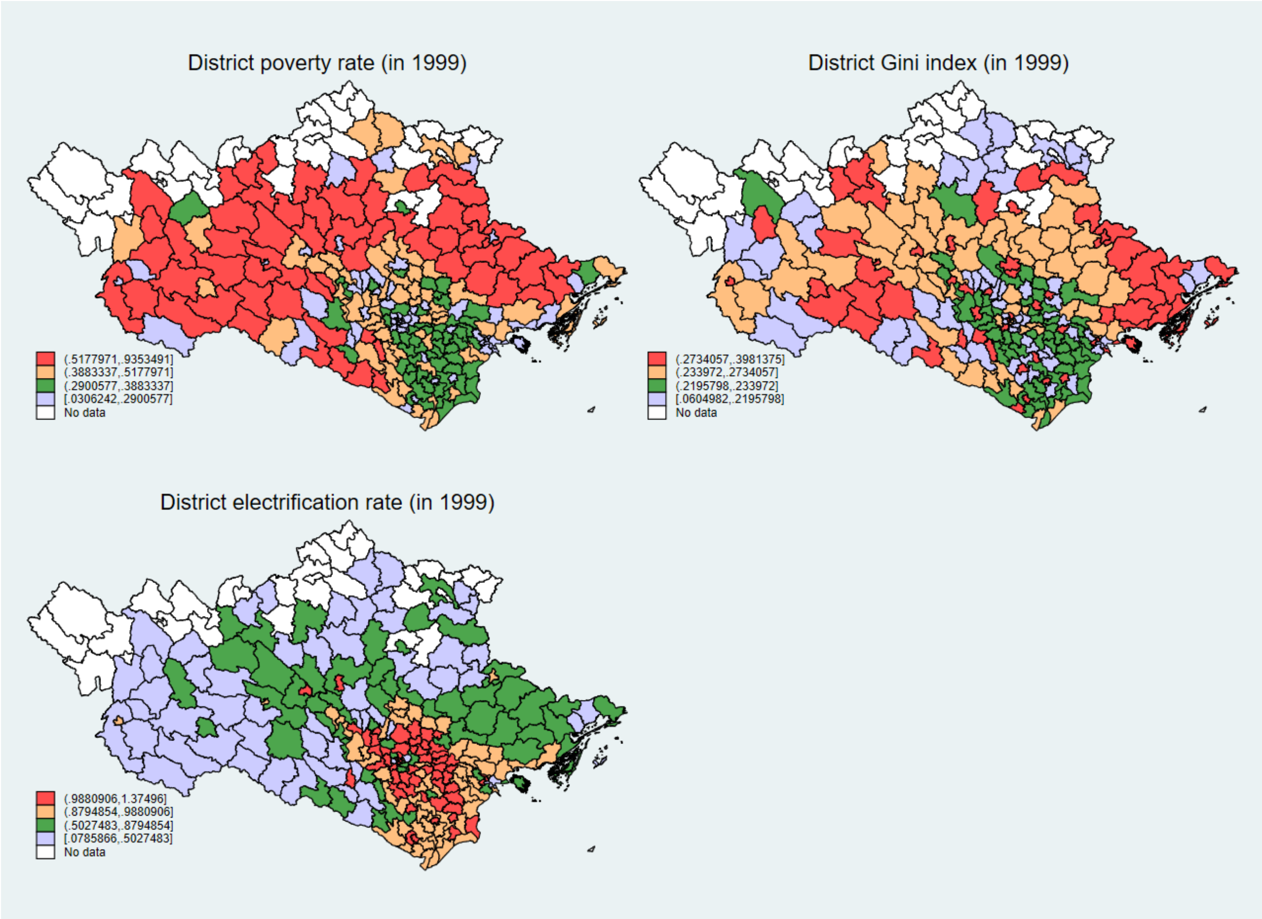
Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the district level, are reported in the bracket.

Figure 1: VIIRS Nightlight around the GMS Corridor



Note: light blue line locates the GMS highway network.

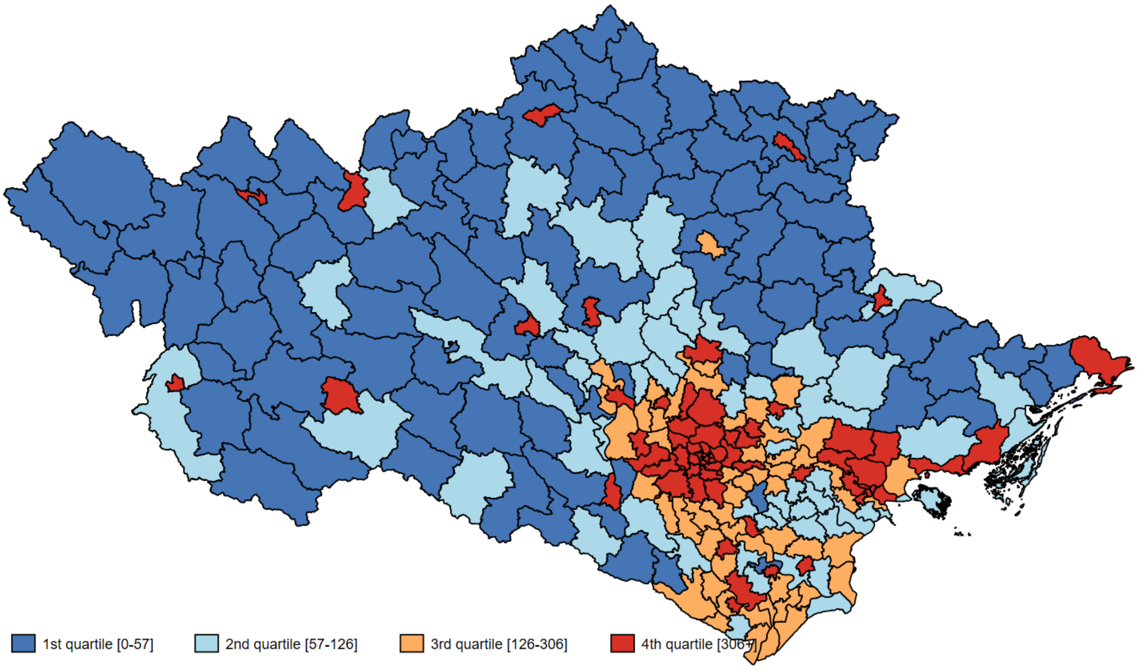
Figure 2: Poverty Rate, Inequality, and Access to Electricity in 1999



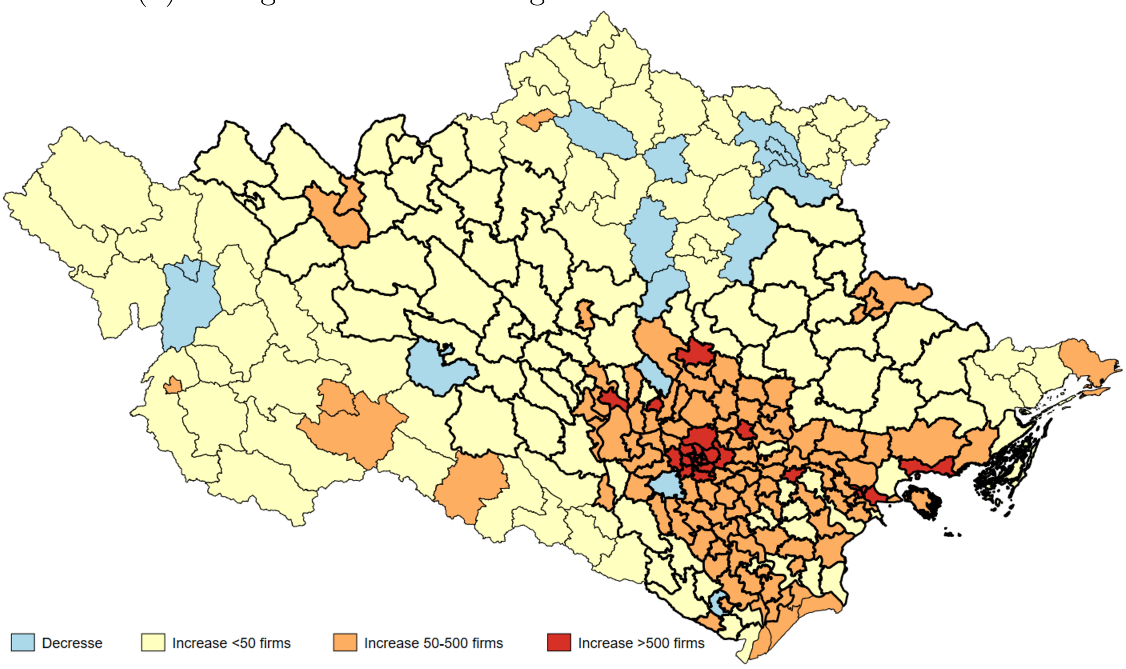
Source: Miguel and Roland (2011)

Figure 3: Spatial Distribution of the Number of Registered Firms

(a) The Number of Registered Firms in 2011

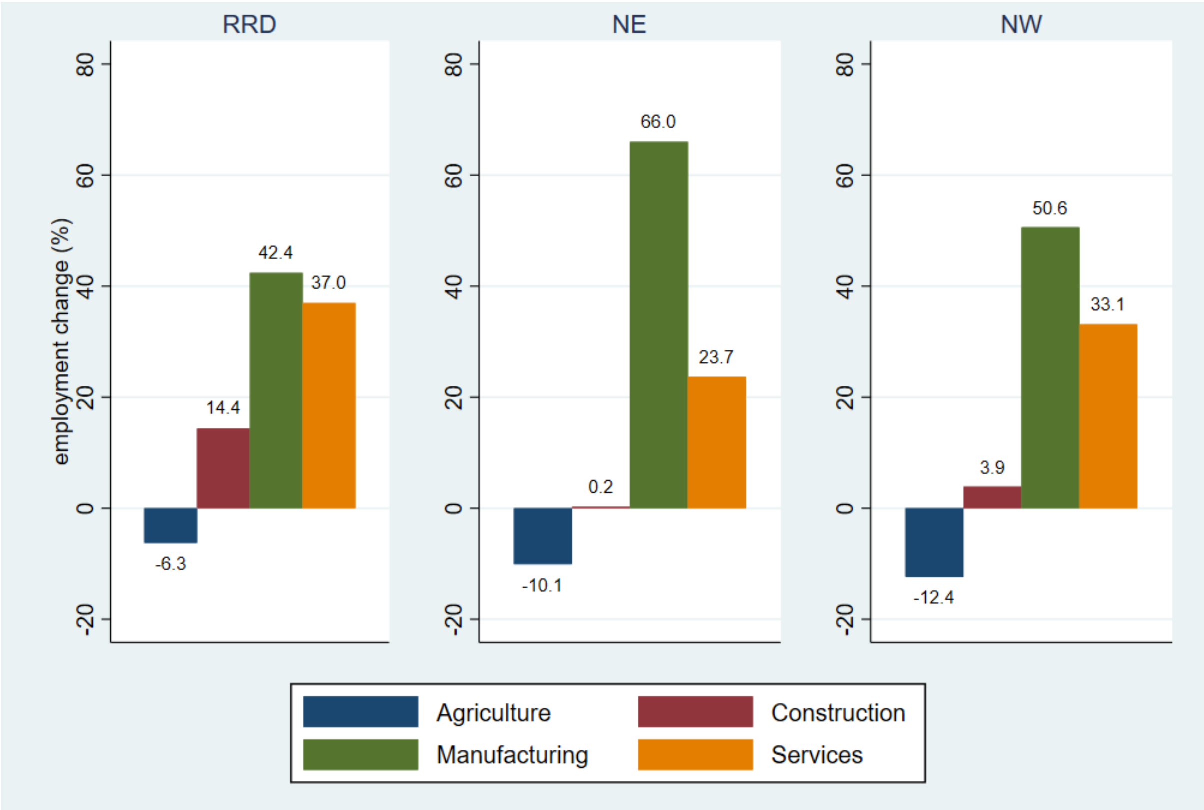


(b) Change in Number of Registered Firms from 2011 to 2016



Note: Data is the number of registered firms recorded in the VES data at the district level. Panel (b) shows five-year change in absolute terms with bold borders that indicate districts located within 50km from the GMS highways.

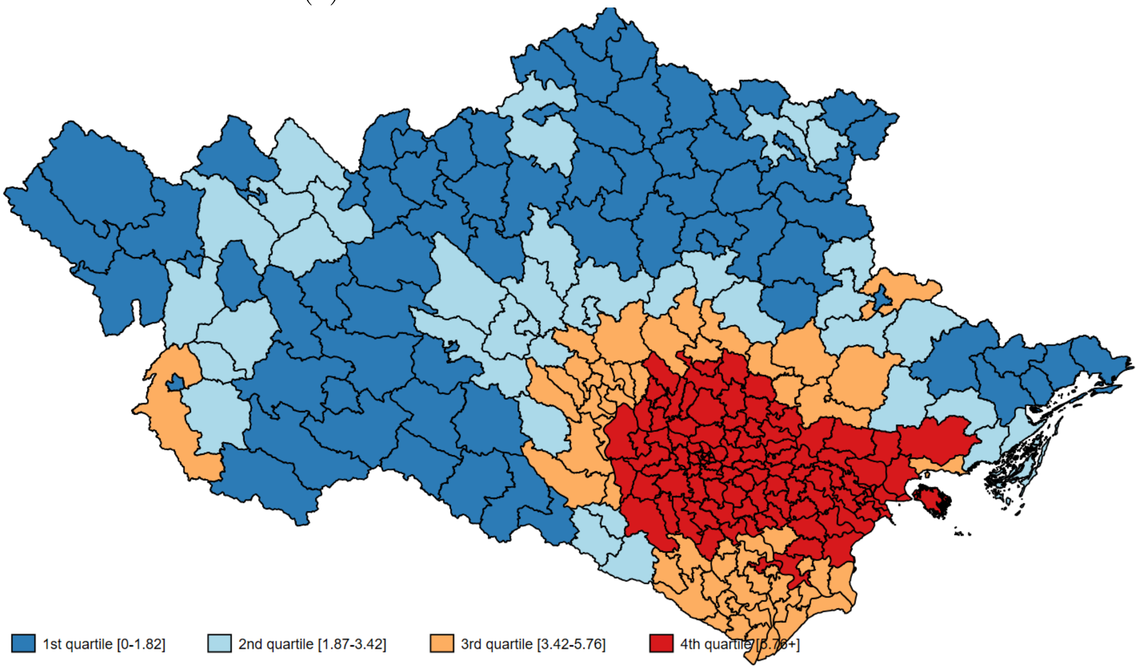
Figure 4: Employment Growth Rate by Sector



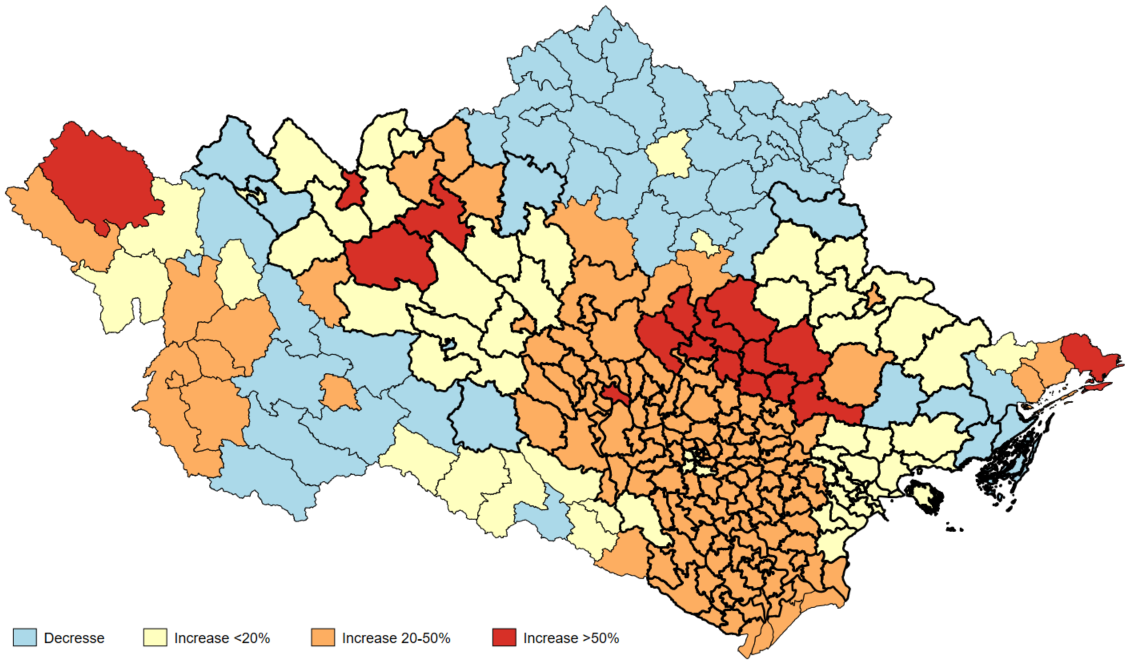
Note: Total district-level employment in each sector is calculated using VES data. Average employment growth rate from 2011 to 2016 is calculated using district's employment level in 2011 as the weight.

Figure 5: Spatial Distribution of the MP Index

(a) The Level of the MP index in 2011



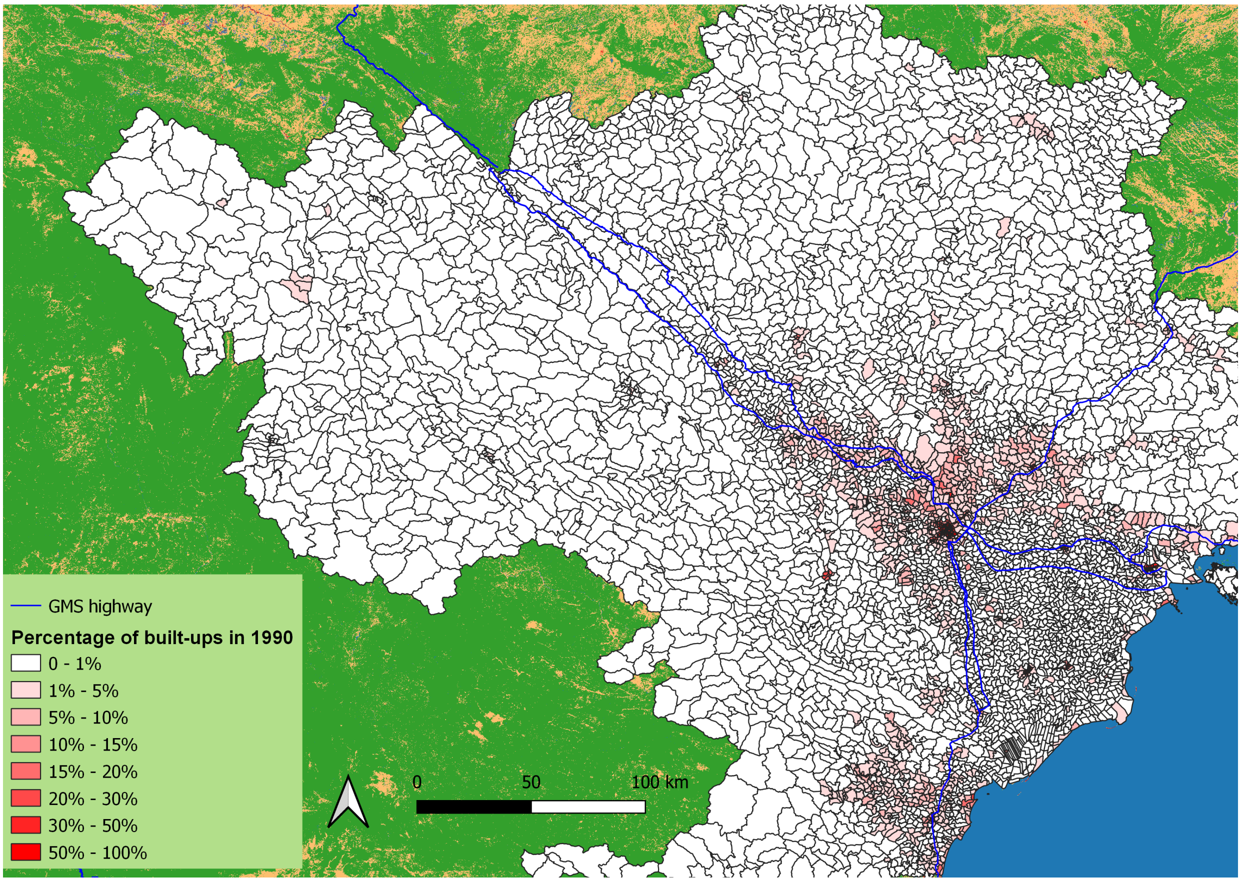
(b) Change in the MP Index from 2011 to 2016



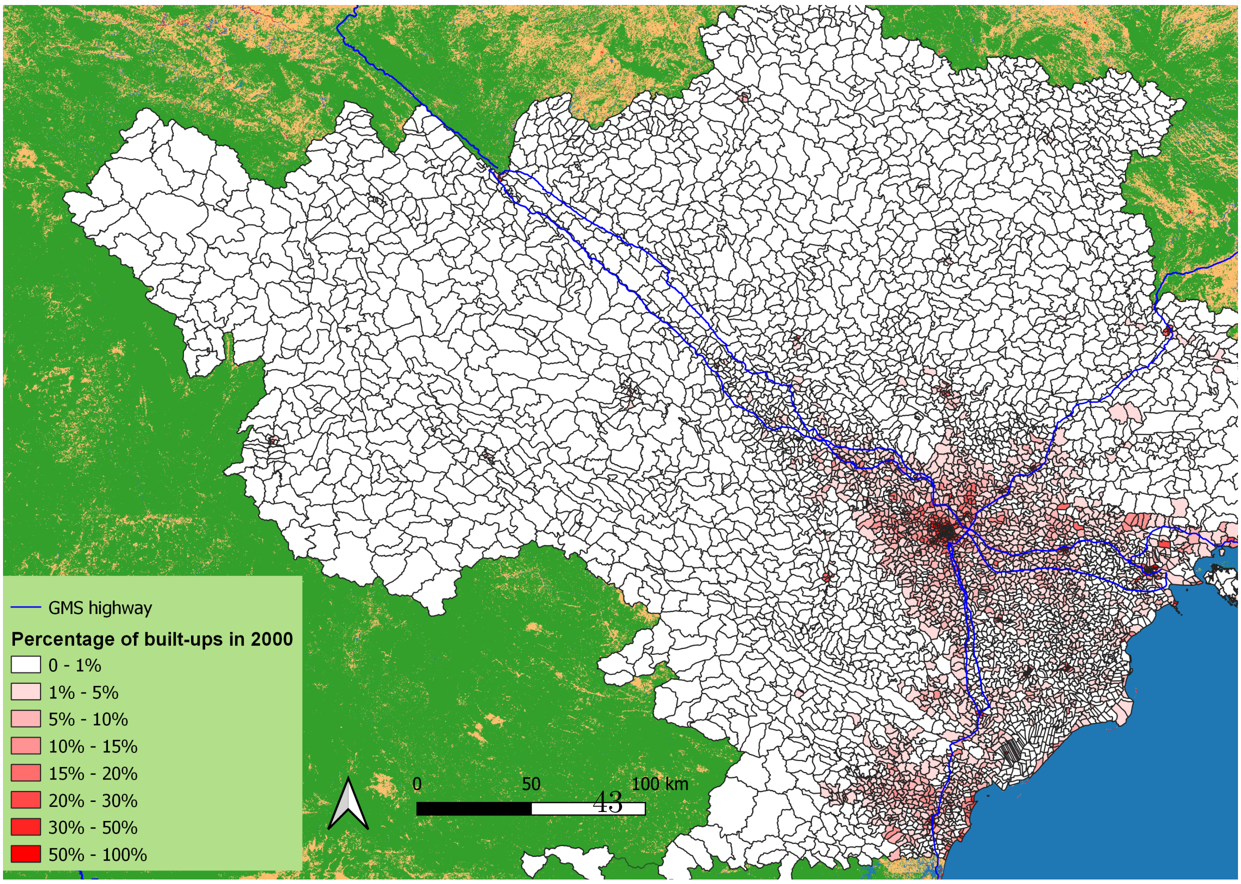
Note: MP index is computed using total district employment of registered firms. In panel (b), bold borders indicate districts that are located within 50km from the GMS highways.

Figure 6: Proportion of Built-ups within Communes: in 1990, 2000, 2010, and 2019

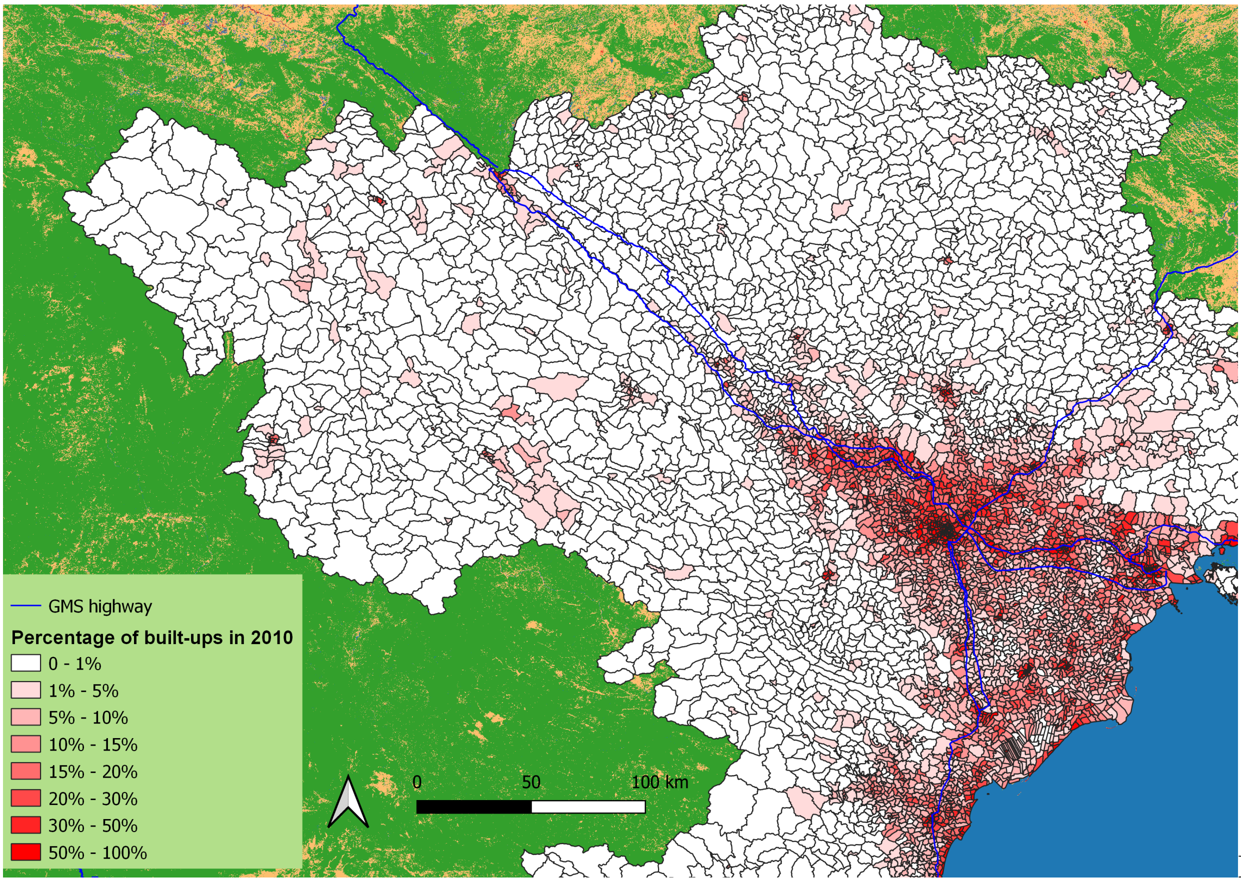
In 1990



In 2000



In 2010



In 2019

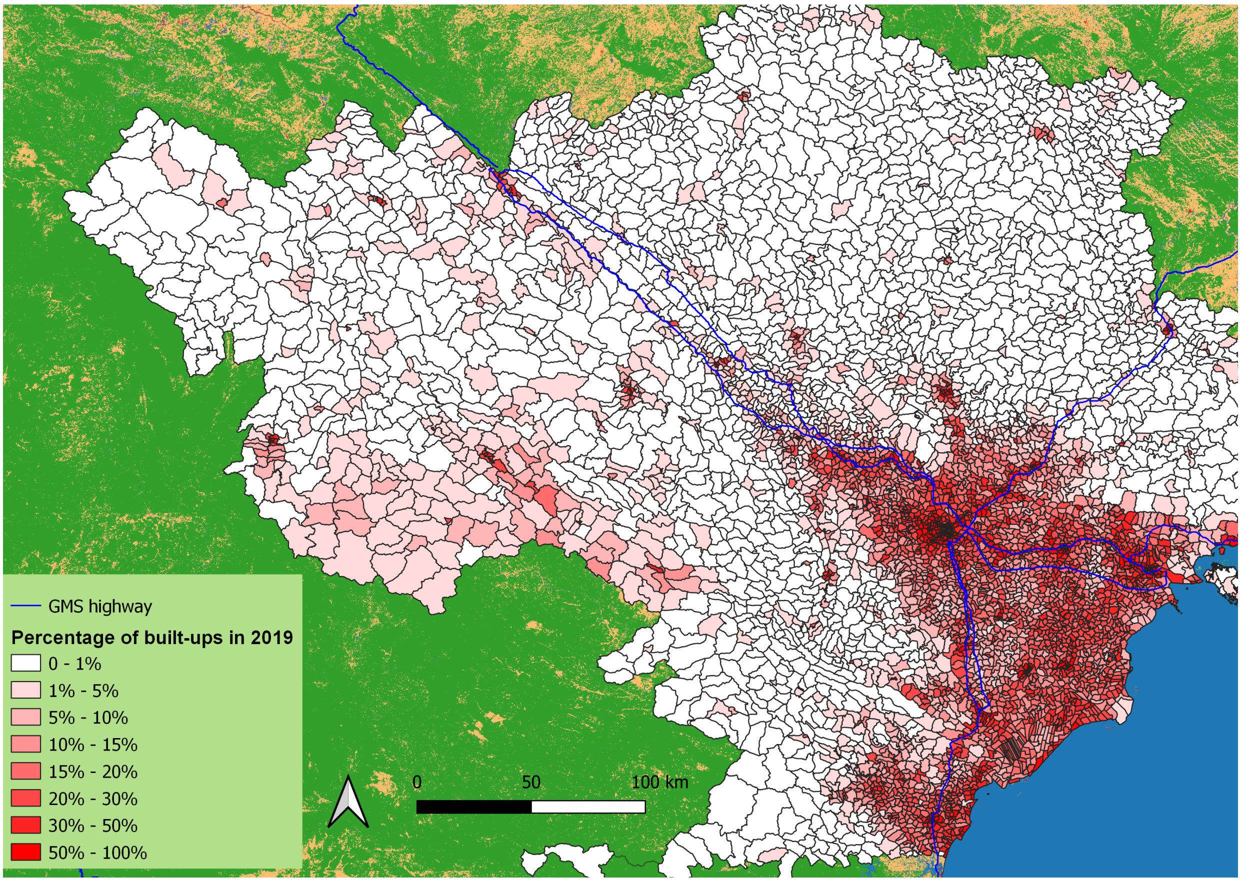
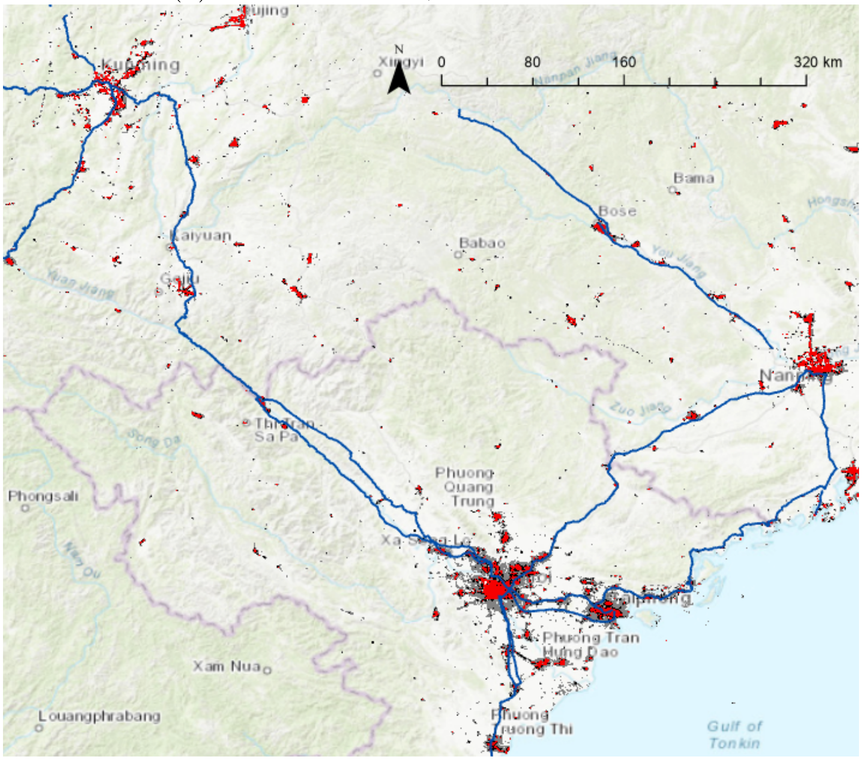
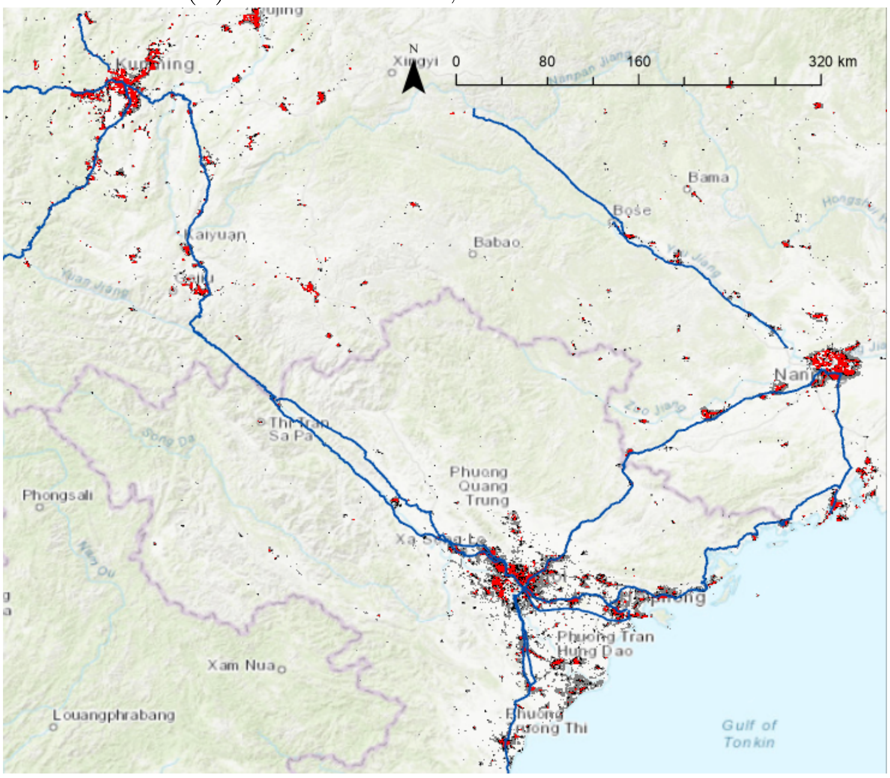


Figure 7: Distribution of the Change in VIIRS Nighttime Light (NTL), 2012-2019

(a) NTL difference, from 2012 to 2016



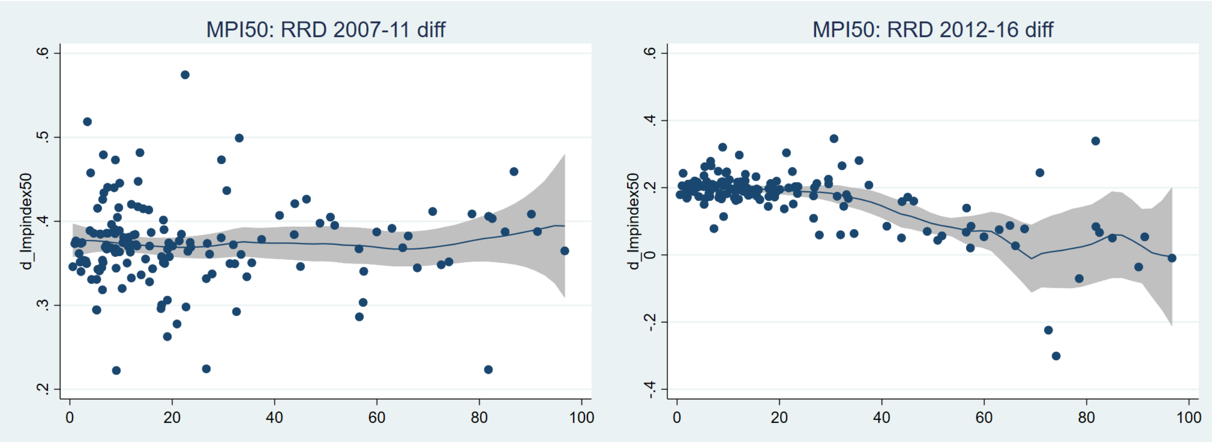
(b) NTL Difference, from 2016 to 2019



Note: NTL changes: significant increase (red), moderate increase (gray), no change or decrease (no color)

Figure 8: Change in the MP Index and the Distance from the GMS Corridor

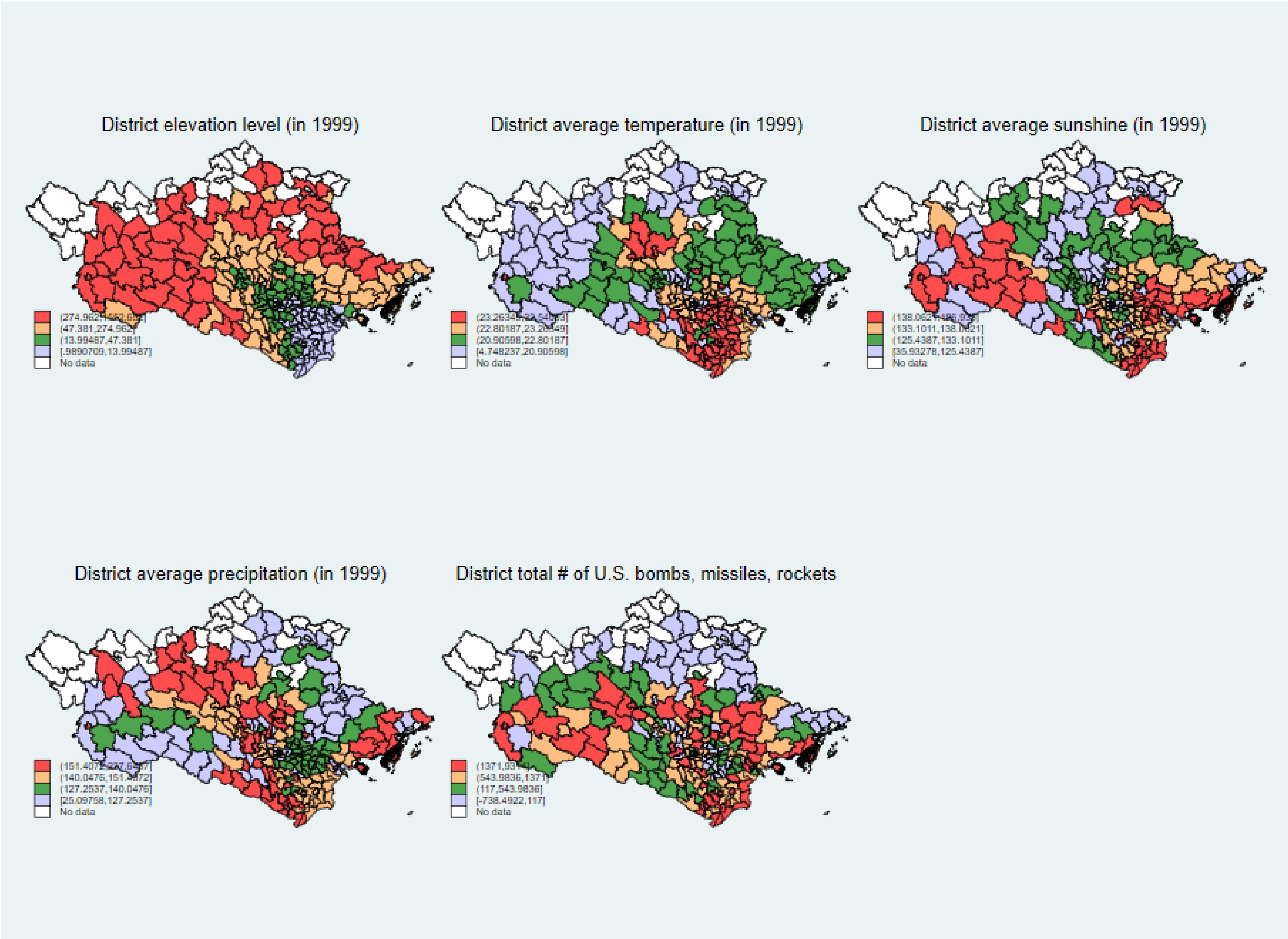
(a) The MP Index for RRD Region, 2007-2011 (left) and 2012-16 (right)



(b) The MP Index for NE NW Regions, 2007-2011 (left) and 2012-16 (right)

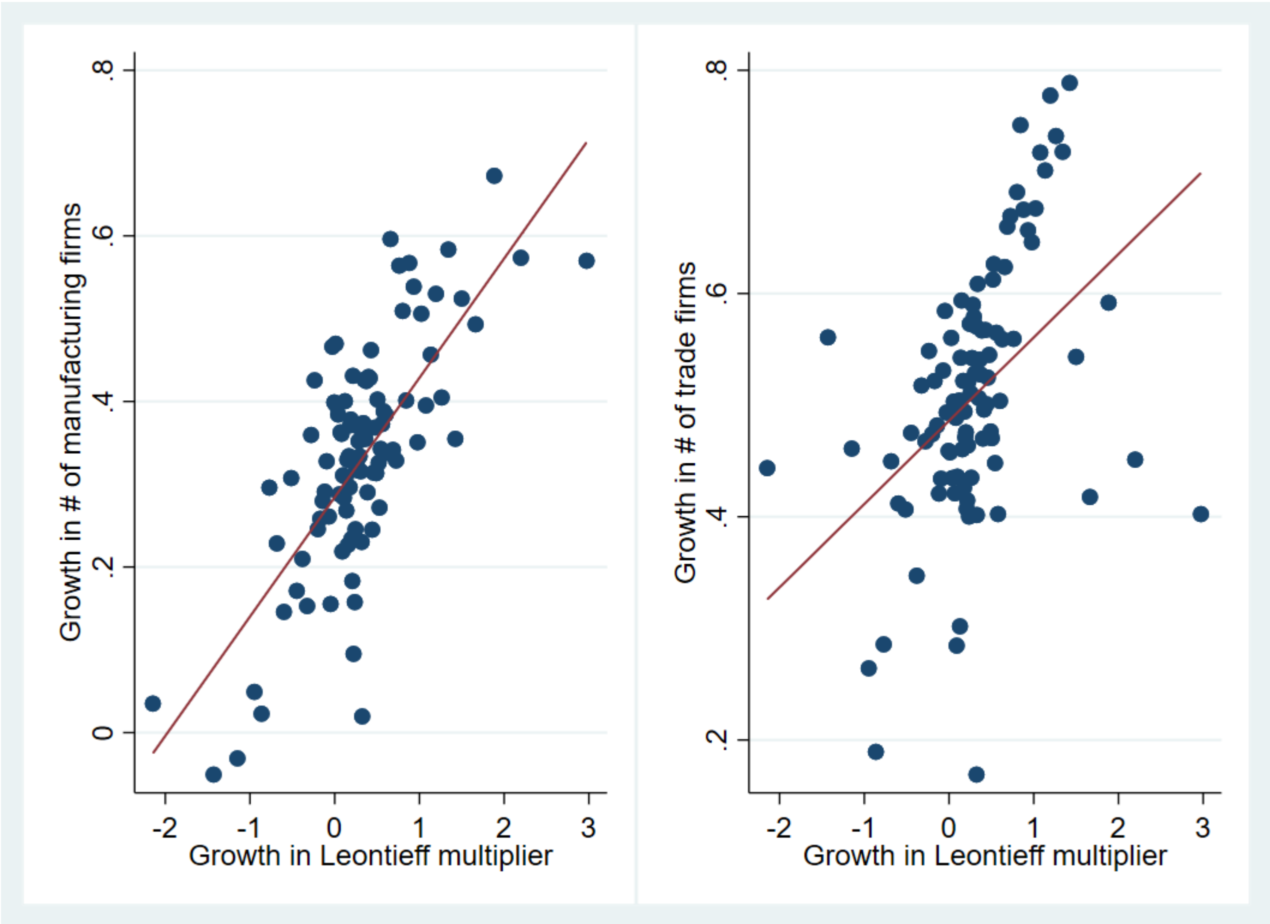
Note: Epanechnikov kernel is used. Provincial dummy and the distance to capital (Hanoi) are controlled in the semi-parametric regression.

Figure 9: Spatial Variation in Urban Amenities across Districts



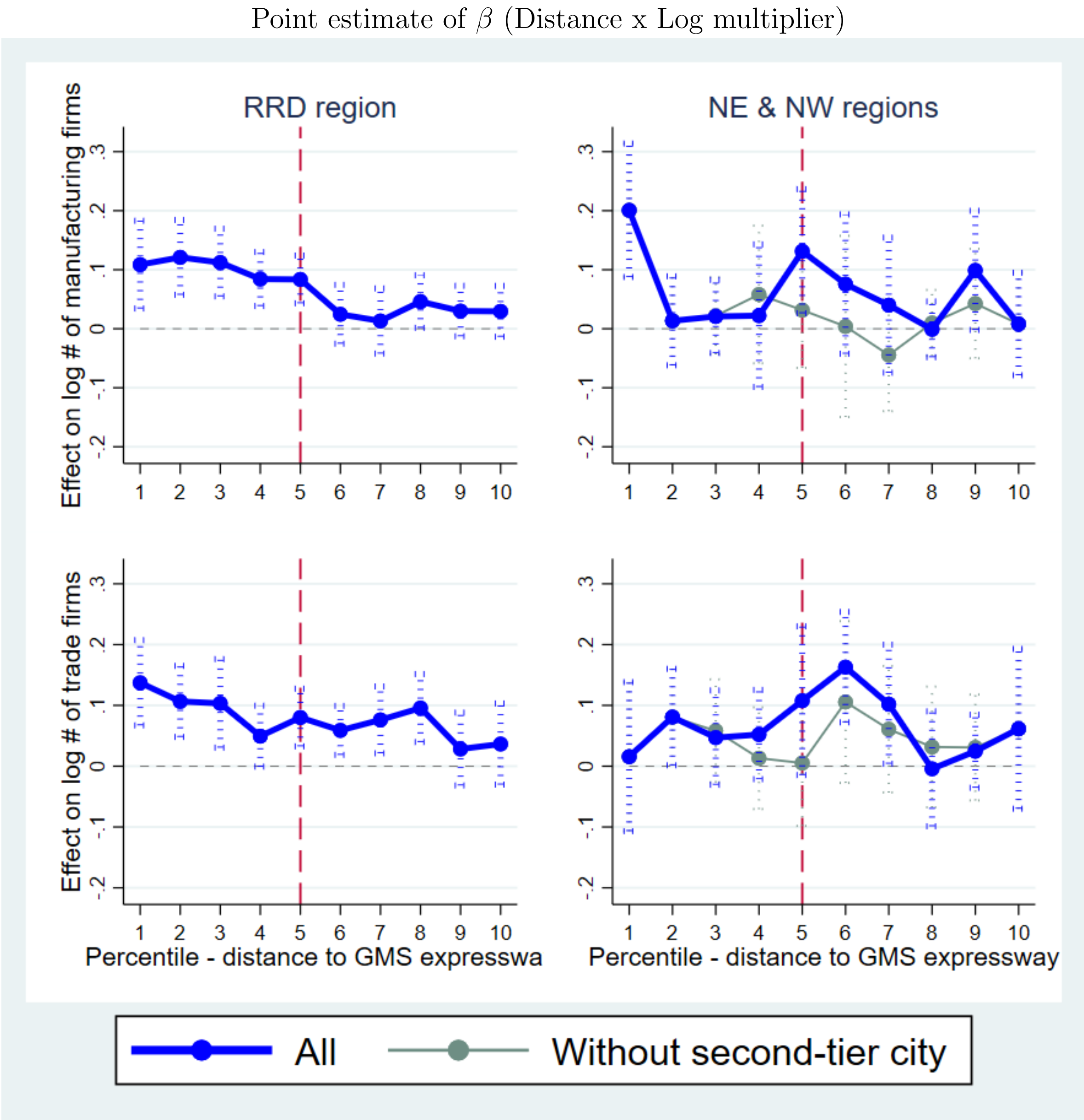
Source: Miguel and Roland (2011)

Figure 10: Spatial Correlation: IO linkage Growth vs. Agglomeration, fron 2010-2016



Source: VES2010-16, OECD Input-output Table

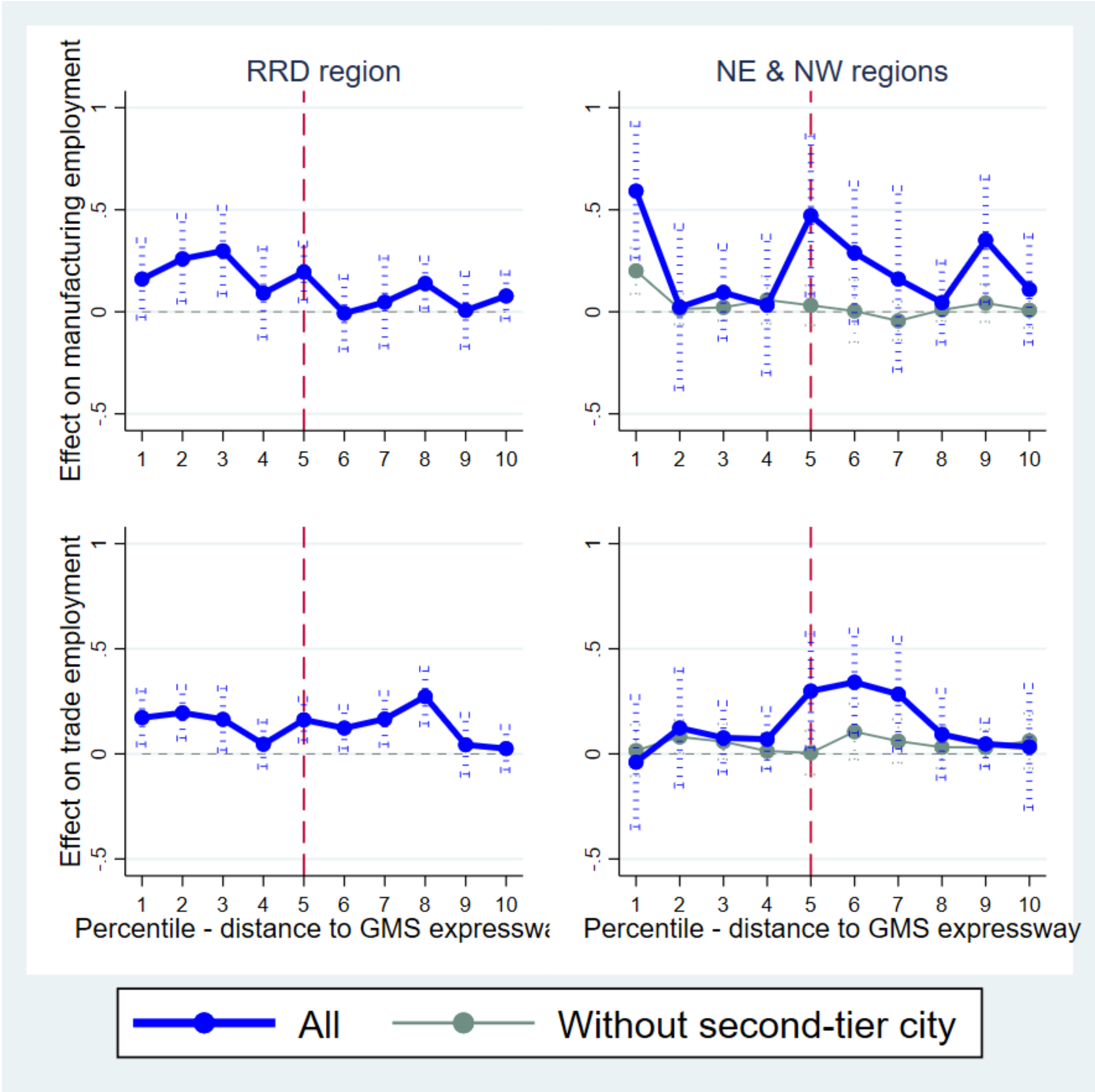
Figure 11: Leontieff multiplier effect on firm agglomeration



Source: VES2010-16, OECD Input-output Table

Figure 12: Leontieff multiplier effect on employment

Point estimate of β (Distance x Log multiplier)



Source: VES2010-16, OECD Input-output Table