Highways and Handsets: Infrastructure Expansion and Service-Led Structural Change*

Anmol Agarwal Gaurav Chiplunkar Aishwarya Kekre

March 2026

Abstract

This paper examines how transport and communication infrastructure expansion differentially shapes the *direction* of structural change in developing economies. Empirically, we study the causal impact of a rapid expansion in the roads and cellular infrastructure in India on its service-led growth. To interpret this evidence, we embed communication costs in a canonical multi-sector spatial model of trade. Quantitative estimates indicate that India's infrastructure expansion generated substantial welfare gains, accounting for 76% of the real income growth over this time period. Importantly, the gains were asymmetric: the impact from cellular network expansion was much larger than roads, highlighting the central role of communication in generating service-led structural transformation.

Keywords: Transport, Mobile Internet, Employment, Structural Transformation.

*Chiplunkar *(corresponding author)*: Darden School of Business, University of Virginia. Email: ChiplunkarG@darden.virginia.edu. Agarwal and Kekre: Department of Economics, University of Virginia.

1. INTRODUCTION

While structural transformation remains one of the clearest markers of economic development, its path is no longer as straightforward as the canonical move from farms to factories. Many developing countries are witnessing the reallocation of labor out of agriculture and into services, instead of the traditional manufacturing-led industrialization. India in the 2000s is a striking example of this pattern, where rapid economic growth for over a decade has been predominantly service-led (Fan et al., 2023b).

In this paper, we argue that infrastructure is central to understanding this pattern. Structural transformation is shaped not only by sectoral productivity growth, but also by the frictions that govern whether firms and households can access markets, whether goods can move across space, and whether workers and producers can coordinate at low cost. Roads relax one set of frictions and determine how quickly goods travel. Communication networks, on the other hand, determine how easily firms and workers coordinate, search, hire, and trade. If roads and communication relax different frictions, then they may generate distinct patterns of structural transformation, such as a service-led growth instead of a manufacturing-led one.

This distinction matters for both economics and policy. A large literature studies how falling trade costs reallocate activity across space and sectors,[1] while another shows how communication technologies reduce information frictions, improve market functioning, and expand opportunity.[2] However, little overlap between these literatures has meant that we know little about a basic question: when transport and communication infrastructure expand together, which frictions matter most for labor reallocation, and why could they reallocate workers into services rather than manufacturing? This is especially important in low-income settings, where policy often treats "infrastructure" as a single object even though different infrastructure systems may relax very different constraints.

We study this question in the context of India during the 2000s, a setting that is unusually well suited to it. In a span of a decade, India experienced the rapid rollout of two major infrastructure systems: a large expansion of highways and rural roads, and a rapid diffusion of GSM cellular coverage (see Figure 1). These changes also occurred at a time when India grew rapidly, reallocated labor out of agriculture, but did so primarily into services rather than manufacturing (Fan et al., 2023b).

[1] See Donaldson (2018); Asher and Novosad (2020); Asturias et al. (2019); Vogel et al. (2024); Morten and Oliveira (2024) for examples.

[2] See Jensen (2007); Aker (2010); Goyal (2010); Allen (2014) for examples.

To examine this question, we construct novel district-level measures of road connectivity and cellular coverage between 1999-2009, and combine them with representative household and plant-level data during this period. On the transport side, we digitize the road network and construct bilateral travel times across districts from 1996 to 2011, similar to Allen and Atkin (2022). On the communication side, we construct novel data on GSM coverage for each district between 2000 to 2011. We combine these infrastructure measures with nationally representative household surveys from the National Sample Survey, as well as establishment-level data from the Annual Survey of Industries. This allows us to trace how changes in transport and communication infrastructure map into district-level market access, workers' sectoral employment, real incomes, and hence aggregate welfare.

Our first result is empirical. Expansions in both road access and GSM access lead to economically large declines in agricultural employment and corresponding increases in services employment, with little effect on manufacturing employment. The worker-level results imply that the average increase in road access between 1999 and 2009 reduced the probability of agricultural employment by 10.5 percentage points. The corresponding increase in GSM access reduced that probability by 16.5 percentage points. These patterns are robust to district and year fixed effects, flexible controls for baseline district characteristics, and instrumental-variable strategies that exploit terrain-driven variation in optimal road access and topography-driven differences in GSM rollout. The reduced-form evidence therefore points to a central empirical fact of the paper: both physical connectivity and communication access result in structural transformation, but can potentially reallocate workers from agriculture into services, instead of manufacturing.

The last pattern is not obvious ex ante. Better roads might have been expected to favor manufacturing by lowering shipping costs for tradable goods. Better communication might also have been expected to ease coordination in manufacturing supply chains. Yet manufacturing barely responds, while services expand sharply. The answer, we argue, lies at the intersection of infrastructure, trade costs, communication frictions, and non-homothetic demand where services are locally supplied (non-tradable) but increasingly demanded as incomes rise. Infrastructure expansion lowers the cost of accessing markets and organizing production, but the equilibrium incidence of those gains depends on where demand expands, which sectors are tradable, and how local labor markets absorb workers leaving agriculture. We show robustly that these forces — in the Indian context — resulted in a service-led, not manufacturing-led, structural transformation.

To interpret these patterns and quantify their aggregate implications, we develop a multi-

sector spatial model of structural transformation with both transportation and communication frictions. In the model, roads reduce the cost of moving tradable goods across districts, while GSM reduces communication frictions that affect both trade and production. Communication infrastructure reduces frictions in two distinct ways: it lowers the cost of transacting with buyers across space, and it lowers within-district hiring frictions in production. These wedges affect delivered prices, sectoral expenditure shares, and labor demand across agriculture, manufacturing, and services. We discipline the model using worker-level employment data, district-level expenditure shares, and establishment-level estimates of how travel times and GSM coverage affect firm production costs and prices.

Our second result is quantitative. Using the exact-hat algebra, we find that the combined expansion of roads and GSM between 1999 and 2009 raises real income by 40.7% for the average district and can account for roughly 76% of the observed increase in India's real GDP per capita over the period. Most strikingly, the welfare effects are highly asymmetric across infrastructure types. Holding all else fixed, road expansion alone generates an average welfare gain of 3.45%. By contrast, reducing communication frictions through GSM expansion generates an average welfare gain of 36.5%. We further unpack the GSM effect by separating the output-side communication wedge from the input-side hiring wedge. Even under conservative variants that limit some of the service-sector advantage and make the comparison with roads more restrictive, communication infrastructure continues to generate substantially larger gains. This is not simply because GSM expanded more quickly, although it did. It is that communication infrastructure is important for services, relaxes multiple margins at once, and those margins have first-order equilibrium consequences for sectoral reallocation and economic growth.

The model also helps rationalize why manufacturing does not expand much even as agriculture contracts. The key mechanism is not just that rising incomes shift demand toward services, though that force is present. More fundamentally, improved infrastructure changes the way districts meet tradable demand. Agriculture and manufacturing are both tradable sectors, so when transport and communication frictions fall, districts can source more tradable goods from other, more efficient locations, rather than expanding local production one-for-one. That makes better connectivity qualitatively different from a pure local productivity shock. It raises real income, but it also exposes local tradable producers to stronger inter-district competition. Agriculture contracts sharply because its expenditure share falls with income and because many lower-productivity districts can now source agricultural goods more cheaply from elsewhere. Manufacturing demand is relatively more protected, but not enough to absorb agricultural workers.

Services are different. They are non-tradable and must be produced locally when income rises. Once infrastructure lowers the cost of tradables and raises real income, local demand for services expands; because those services cannot be imported from other districts, local labor must move into service production. Communication expansion amplifies this force further because it lowers within-district frictions that directly affect service-sector production as well. So the service-led transition in the model is not merely a reduced-form Engel-curve story. It comes from the interaction of non-homothetic demand, the tradability of agriculture and manufacturing, the non-tradability of services, and the fact that improved connectivity allows tradable demand to be satisfied through spatial reallocation of production rather than local labor absorption. That mechanism, we think, is central for understanding why many developing economies can experience substantial structural transformation without a classic industrialization experience.

Lastly, we also shed light on why infrastructure gains are uneven across space. Districts with lower initial productivity in tradable sectors gain more from improved connectivity because they are more reliant on purchases from other districts and therefore benefit more when trade and communication frictions fall. More educated and more urbanized districts, which appear to be more productive ex ante, gain less for the same reason. These heterogeneity patterns help connect the paper's reduced-form and quantitative results: infrastructure does not simply raise activity everywhere. It changes who can access outside markets, who can source from elsewhere, and where the gains from lower frictions are concentrated.

We explore two key extensions that reinforce the central message, rather than overturn it. We first ask whether the relatively modest welfare effects of roads are driven by the particular pattern of highway expansion observed in the data. However, even when we replace the realized road network with a more efficient, geography-implied highway network, the gains from lower travel times rise only modestly and remain far below the gains from lower communication costs. In a second extension, we relax the assumption of no labor mobility across districts. Allowing workers to move toward locations with higher real wages changes the spatial distribution of gains, and indeed amplifies welfare dispersion across districts, but it does not alter the paper's core conclusions: communication infrastructure remains quantitatively more important than roads, and structural transformation continues to take a service-led rather than manufacturing-led approach. Taken together, these extensions show that our main findings are not artifacts of a particular road-building pattern or of a no labor mobility benchmark. They reflect a deeper feature of the environment as alluded to earlier.

Related Literature. Our primary contribution is to the macro-development literature on structural transformation (see Gollin and Kaboski (2023) for a review). Closest to our paper is Fan et al. (2023b), which shows that productivity growth in services was a key driver of India's service-led structural transformation between 1987 and 2012. We document a similar pattern for 1999-2009, but show that infrastructure expansion played a central role. Our quantitative exercise suggests that it accounts for 76% of the observed increase in real GDP per capita over this period. To our knowledge, this is the first study to examine how the joint expansion of two major infrastructure systems—roads and cellular networks—shapes labor reallocation in a large developing economy. In India, Asher and Novosad (2020) show that rural roads shifted labor from farm to non-farm work, while Kaboski et al. (2024) find mixed evidence on structural change following highway expansion in China and India. For communication infrastructure, Blumenstock et al. (2020) and Azam et al. (2024) show that cellular expansion affects employment margins such as self- versus wage employment. We add to this literature by providing both empirical and quantitative evidence on how infrastructure expansion affects structural change and welfare in India.

We contribute to the literature on the effects of mobile phone expansion. Allen (2014) shows that mobile phone coverage affects trade by reducing information frictions. Jensen (2007), Aker (2010), and Goyal (2010) similarly show that 2G expansion reduces price dispersion by easing information frictions. In the spirit of this literature, we incorporate the channel through which communication frictions affect price discovery for producers in output markets, as well as distortions in the hiring of inputs, and examine the implications for a key development outcome, structural transformation. Hjort and Tian (2021) provides an excellent review of the literature on the economic effects of cellphones.

Lastly, our work is closely related to a large body of research evaluating the effects of changes in road infrastructure on economic outcomes (see Gonzalez-Navarro et al. (2023) for a recent review). A substantial literature examines the welfare effects of highway expansion in developing countries such as India, Brazil, China, and Turkey (Vogel et al., 2024; Morten and Oliveira, 2024; Fan et al., 2023a; Coşar et al., 2022). In the Indian context, Vogel et al. (2024) and Asturias et al. (2019) show that the Golden Quadrilateral highway network generated aggregate welfare gains. Ghani et al. (2016) also document gains in manufacturing output and employment arising from the Golden Quadrilateral. We build on this literature by embedding both road and cellular expansion in a multi-sector gravity model of trade to study how changes in physical and communication trade costs shape sectoral labor transitions during the decade of rapid infrastructure expansion in India.

The remainder of this paper is organized as follows: Section 2 provides details on the infrastructure expansion in India, describes the data used in this analysis and discusses drecriptive evidence. Section 3 outlines the model. Section 4 discusses the reduced-form evidence. Section 5 describes the estimation strategy for the model. Section 6 discusses the results and counterfactuals, while Section 7 concludes.

2. Transport and Cellphone Network Expansion

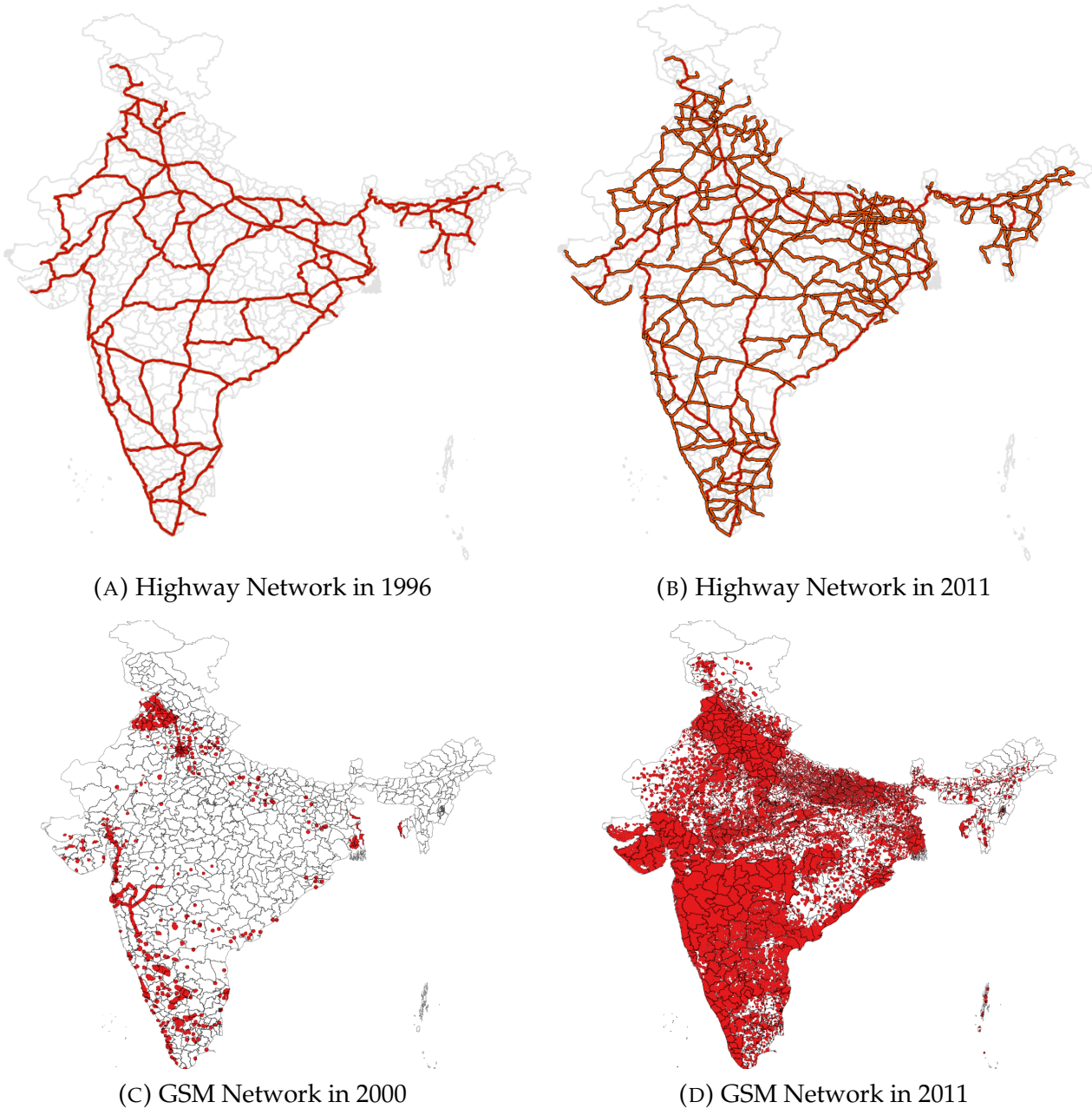
We begin by describing the expansion patterns in transportation and communication. In Section 2.1, we first summarize the expansion of the road network and its implications for bilateral travel times. Section 2.2 then describes the rollout of GSM coverage and our district-level measure of communication access. Finally, Section 2.3 introduces the household and firm data that we use for our analysis and presents descriptive evidence linking infrastructure expansion to structural transformation across districts.

2.1. Roads, Travel Time, and Market Access

Expansion of Road Network in India. Road transport in India in the 1990s was slow and costly. A 1,500-km truck journey took four to five days, compared with roughly one day in the United States (Asturias et al., 2019), reflecting the limited reach and quality of the road network (Figure 1(a)). Two major policy initiatives launched around 2000 reshaped this landscape. First, the National Highways Development Project, initiated in 2001, upgraded long-distance road connectivity. Its first phase, the Golden Quadrilateral, linked Delhi, Mumbai, Chennai, and Kolkata, while subsequent phases expanded the North-South and East-West corridors. Second, the Pradhan Mantri Gram Sadak Yojana, launched in 2000, sought to connect previously unconnected rural villages through all-weather roads. By the end of the decade, these investments had substantially expanded and upgraded India's road infrastructure (Figure 1(b)).

Travel Time Between Districts. To quantify changes in road connectivity, we follow Allen and Atkin (2022) and construct district-level bilateral travel times using digitized road maps published by the Government of India for 1996, 2004, and 2011 (Appendix Figure A.1). We first classify the road network into five categories: (i) national and state highways; (ii) all-weather motorable roads; (iii) fair-weather motorable roads; (iv) roads subject to delays; and (v) other roads. We then assign each category an average speed using bench-

FIGURE 1: Expansion of Road and Cellphone Networks in India



(A) Highway Network in 1996

(B) Highway Network in 2011

(C) GSM Network in 2000

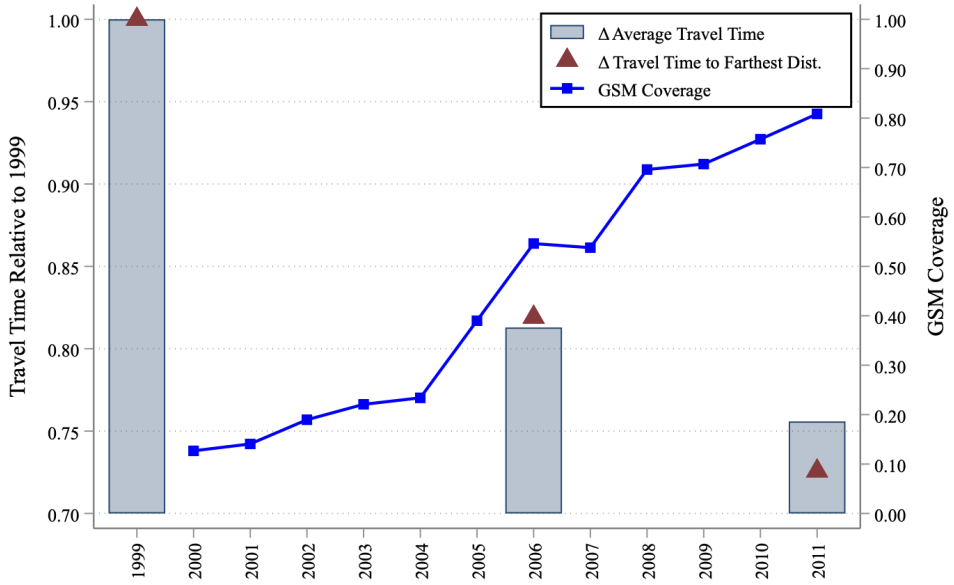
(D) GSM Network in 2011

Notes: Panels (a) and (b) show digitized maps of India's highway network in 1996 and 2011. Panels (c) and (d) show district-level GSM coverage in 2000 and 2011.

marks from the literature (Allen and Atkin, 2022; World Bank, 2022; Vogel et al., 2024).[3] We hold these speeds fixed over time, so changes in travel times are driven by changes in the

[3]Following the literature, we assign speeds of 30 mph to national and state highways, 20 mph to all-weather motorable roads, and 10 mph to all other road types. Moreover, since road maps are available only for 1996, 2004, and 2011, we linearly interpolate bilateral travel times for the intermediate years.

FIGURE 2: Travel Time and GSM Coverage Between 1999-2011



Notes: The figure reports changes relative to 1999. Blue bars show the average travel time across district pairs, red triangles show the travel time to the farthest district, and blue squares show average district-level GSM coverage.

extent and composition of the road network rather than by changes in assumed vehicle speeds. We then use Dijkstra's algorithm to compute the shortest travel time between the centroids of all district pairs.

Appendix Figure A.2 shows that travel times fell sharply between 1996 and 2011. For example, trips of less than 500 km took about 7 hours on average in 1996, whereas trips of more than 2,000 km took over 40 hours. By 2011, the corresponding travel times had fallen to roughly 5 hours and 30 hours. Figure 2 summarizes these changes relative to 1999. Average bilateral travel time fell by about 20% between 1999 and 2006, with a further decline of roughly 6 to 10 percentage points between 2006 and 2011. The time required to reach the farthest district follows a similar pattern. These reductions are economically large and suggest that road expansion substantially widened the set of locations that could be reached within commercially relevant time horizons.

2.2. WIRELESS CELLULAR NETWORK EXPANSION

Mobile-phone infrastructure also expanded rapidly during the 2000s (Figure 2). Although wireless technology was introduced in the mid-1990s, coverage initially remained concen-

trated in major urban centers (Figure 1(c)). A series of policy changes then accelerated nationwide rollout. The New Telecom Policy of 1999 introduced a revenue-sharing licensing regime that improved the financial viability of mobile operators and facilitated private entry. The Unified Access Service License, introduced in 2003, allowed providers to offer fixed-line and mobile services under a common license. Rural expansion was supported by the creation of Bharat Sanchar Nigam Limited in 2000 and by subsidies through the Universal Service Obligation Fund. Taken together, these supply-side reforms produced a dramatic expansion in GSM coverage, which was close to universal by the end of the decade (Figure 1(d)).

On the demand side, network expansion and falling prices generated a parallel surge in adoption. As shown in Appendix Figure A.4, India had only 36 million telephone subscriptions in 2000, of which just 10% were mobile. By 2005, the number of telephone connections had risen to 140 million, with mobile connections accounting for roughly two-thirds of the total. By 2011, total subscriptions exceeded 900 million, and 97% were mobile. The supply expansion we study therefore coincided with a broad-based diffusion of mobile technology among households and firms.

Measuring Expansion in GSM Coverage. We use geospatial data on 2G (GSM) network coverage from 2000 to 2011 assembled by the GSM Association and used in prior work (Guriev et al., 2021; Manacorda and Tesei, 2020; Mensah, 2021; Chiplunkar and Goldberg, 2022). The data consist of 1×1 km grid cells indicating whether a location had GSM coverage in a given year. Following this literature, we combine these coverage maps with sub-district population counts from the 2001 Census to construct a population-weighted measure of district-level GSM coverage:

$$GSM_{dt} = \sum_{s \in d} \frac{\text{Pop}_{s,2001}}{\text{Pop}_{d,2001}} \times 1(GSM)_{st}.$$

This measure captures the share of a district's population living in locations with network access. As Figure 2 shows, average district-level GSM coverage rose from about 10% in 2000 to more than 80% by 2011.

Taken together, the road and mobile expansions generated large and distinct changes in connectivity. Roads reduced the time required to move goods across space, while GSM expansion lowered communication frictions and broadened access to market information. This joint transformation motivates both our empirical strategy and the model that follows: we use the differential expansion of transport and communication infrastructure across

districts to study how reductions in physical and communication frictions reallocated labor across sectors during a decade of rapid structural change in India.

2.3. Data on Employment Outcomes

National Sample Surveys. Our primary source of employment data is the National Sample Survey (NSS), a repeated cross-sectional household survey conducted by the Government of India. The NSS provides nationally representative information on employment, demographics, and household characteristics at roughly five-year intervals. We use four rounds spanning our period of interest: 1999–2000 (55th round), 2004–05 (61st round), 2007–08 (64th round), and 2009–10 (66th round).

The NSS reports detailed individual-level information on employment and demographics. In particular, it records sector of employment—agriculture, manufacturing, or services—and employment status—employer, self-employed, wage employed, or unpaid household worker. It also provides a rich set of worker and household characteristics, including age, education, marital status, land ownership, and rural or urban residence. We restrict attention to working-age individuals between ages 15 and 65. We harmonize district boundaries over time and exclude smaller states and territories with limited data coverage, leaving a final sample of 394 districts defined at 1999 boundaries.[4]

Annual Survey of Industries. To estimate key model parameters, we also use the Annual Survey of Industries (ASI), an annual plant-level survey covering all large manufacturing plants and a representative sample of smaller establishments. The ASI reports detailed information on output quantities and values, input use, expenditure components, factor payments, revenues, location, and industry. Although the ASI includes plant identifiers, the panel files do not report district identifiers, which are central to our analysis. We therefore work with repeated cross-sections of establishments.

2.4. Descriptive Evidence

We conclude the section with descriptive evidence linking infrastructure expansion to sectoral reallocation. Figure 3 plots the relationship between district-level employment shares and two measures of connectivity: GSM coverage and road market access. Higher GSM coverage and better road market access are both associated with lower agricultural employ-

[4]These include Jammu & Kashmir, the North-Eastern states (Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura), and the island territories of Andaman & Nicobar and Lakshadweep. Together, they account for only 3.8% of India's 1999 population.

ment shares and higher services employment shares.[5] There appears to be little correlation with employment in the manufacturing sector. Furthermore, Appendix Figure A.5 shows that these patterns are not driven solely by cross-sectional differences: they also appear within districts as infrastructure expands over time.

3. MODEL

This section develops a parsimonious spatial model that links infrastructure (both transport and communication) to structural transformation. The model has three building blocks. We first specify household demand across sectors and across district varieties within tradable sectors. We then show how roads and GSM enter firm costs and delivered prices through distinct transportation and communication wedges. Finally, we define equilibrium and highlight the objects that will discipline the empirical analysis and the quantitative exercises in subsequent sections.

The economy consists of several locations (districts), indexed by $\{i, j\} \in \mathcal{D}$, that trade with each other. There are three sectors $k \in \{A, M, S\}$ representing agriculture, manufacturing, and services. We assume that agriculture and manufacturing are tradable across districts, while services are non-tradable, that is, produced and consumed locally. Each district produces a unique variety in agriculture and manufacturing, and consumers have preferences over varieties from all districts. A representative household composed of identical workers L_i in district i supplies labor inelastically at wage w_i. Workers are perfectly mobile across sectors within a district but cannot move across districts in the baseline model; Section 6.5.2 discusses the implications of relaxing this assumption.

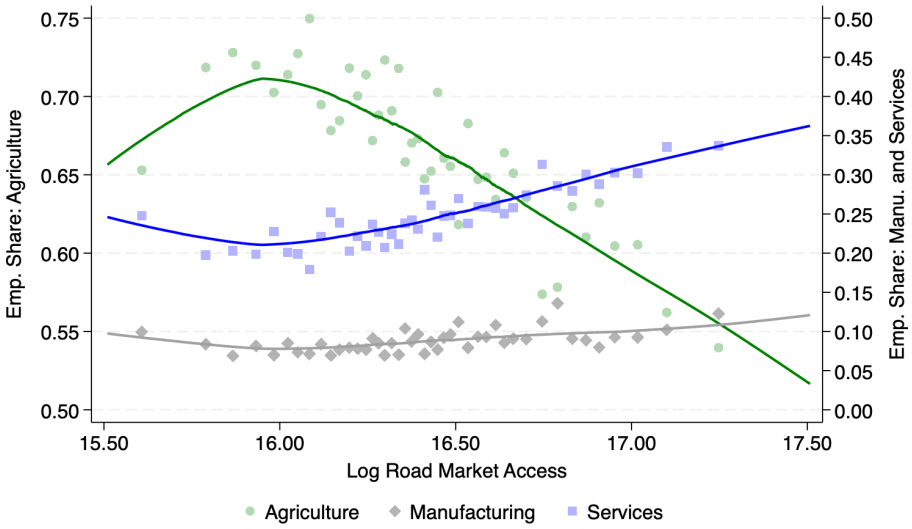
For notational clarity, X_{kij} denotes a variable of interest X (for example, a price or cost) in sector-k, when district i buys from district j. Variables without bilateral linkages, such as sectoral price indices, are denoted by X_{ki}.

3.1. CONSUMER DEMAND

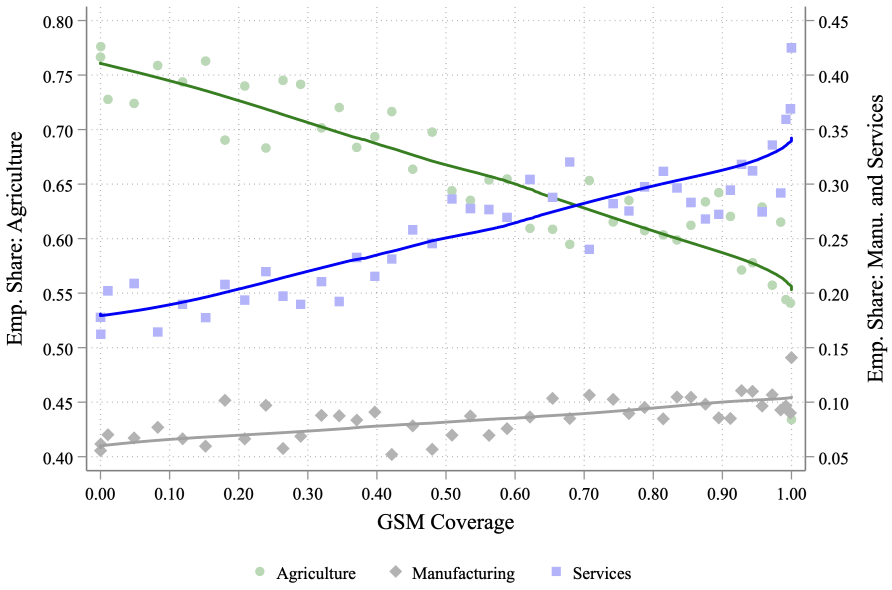
Consumer Preferences. Preferences are nested. Households first allocate expenditure across sectors and then across varieties within each tradable sector. A representative household in district i consumes a final good C_i that aggregates sectoral composites C_{ki} according

[5] In the spirit of Donaldson and Hornbeck (2016), we construct road market access (RMA) for origin district o as $RMA_o = \sum_{d \in \mathcal{D}} (1 + time_{od})^{-0.15} \times population_d$. Section 5.1 describes how this measure maps to the model and parameterization.

FIGURE 3: Correlation Between Infrastructure and Sectoral Employment



(A) Log Road Market Access



(B) Cellular GSM Coverage

Notes: The figure uses NSS rounds from 1999 to 2009 to plot nonparametric relationships between district-level sectoral employment shares and log market access (Figure a) and GSM coverage (Figure b).

to the following non-homothetic CES structure:

$$\sum_{k \in A,M,S} \eta_k^{\frac{1}{\sigma}} \left(\frac{C_{ki}}{C_i^{\varepsilon_k}} \right)^{\frac{\sigma-1}{\sigma}} = 1, \qquad \sigma \in (0,1) \tag{1}$$

where $\eta_k > 0$ are sectoral preference shifters, σ is the elasticity of substitution, and ε_k is the degree of non-homotheticity across sectors (e.g., $\varepsilon_k = 1$ yields homothetic CES preferences). Within each tradable sector k, consumers in district i have preferences over varieties from all districts $j \in \mathcal{D}$ given by:

$$C_{ki} = \left[\sum_{j \in \mathcal{D}} c_{kij}^{\frac{\sigma_k - 1}{\sigma_k}} \right]^{\frac{\sigma_k}{\sigma_k - 1}}, \qquad \sigma_k > 1$$

where σ_k is the standard elasticity of substitution across varieties within each sector.

Sectoral Prices and Expenditure Shares. The aggregate price index in district i, denoted by P_i, and the tradable-sector price indices, denoted by P_{ki}, are given by:

$$P_i = \left[\sum_k \eta_k P_{ki}^{1-\sigma} \left(\frac{I_i}{P_i} \right)^{(\varepsilon_k - 1)(1-\sigma)} \right]^{\frac{1}{1-\sigma}} \tag{2}$$

$$P_{ki} = \left[\sum_{j \in \mathcal{D}} p_{kij}^{1-\sigma_k} \right]^{\frac{1}{1-\sigma_k}}, \quad \forall k \in \{A, M\} \tag{3}$$

where I_i is total income in district i and p_{kij} is the price paid by consumers in district i for the sector-k variety produced in district j. Appendix B.1 provides the derivations. The corresponding expenditure share on sector k, denoted x_{ki}, is:

$$x_{ki} = \eta_k \left(\frac{P_{ki}}{P_i} \right)^{1-\sigma} \left(\frac{I_i}{P_i} \right)^{(\varepsilon_k - 1)(1-\sigma)} \tag{4}$$

Thus, ε_k governs how a sector's expenditure share responds to real income. If $\varepsilon_k > 1$, the expenditure share on sector k rises with income; if $\varepsilon_k < 1$, it falls. For two sectors k and k', the ratio of expenditure shares is:

$$\frac{x_{ki}}{x_{k'i}} = \left(\frac{\eta_k}{\eta_{k'}} \right) \left(\frac{P_{ki}}{P_{k'i}} \right)^{1-\sigma} \left(\frac{I_i}{P_i} \right)^{(\varepsilon_k - \varepsilon_{k'})(1-\sigma)} \tag{5}$$

As in a homothetic CES environment, relative expenditure shares depend on preference shifters ($\eta_k / \eta_{k'}$) and substitution induced by relative sectoral prices ($P_{ki}/P_{k'i}$). With non-homothetic preferences, they also depend on an income effect governed by the difference in income elasticities, $\varepsilon_k - \varepsilon_{k'}$. Within tradable sectors, the share of district i's sectoral ex-

penditure allocated to varieties from district j is:

$$s_{kij} = \left(\frac{p_{kij}}{P_{ki}}\right)^{1-\sigma_k}, \quad \forall k \in \{A, M\} \tag{6}$$

3.2. TRANSPORT, COMMUNICATION, AND FIRM PRODUCTION

Each sector k in district i is populated by a representative firm with linear technology $Y_{ki} = z_{ki}L_{ki}$, where z_{ki} is sectoral productivity and L_{ki} is labor input. Roads and communication infrastructure affect the economy through different margins: transportation costs govern the movement of goods across districts, while communication costs affect both production and trade.

Transportation Costs. Following the trade literature, shipping goods from district j to district i in sector k requires paying a sector-specific iceberg transportation cost $\tau^T_{kij} \geq 1$. Transportation costs are symmetric, so $\tau^T_{kij} = \tau^T_{kji}$, and there is no physical transport cost within a district, so $\tau^T_{kii} = 1$.

Communication Costs. A large literature shows that communication technology mitigates frictions in hiring, input sourcing, and buyer-seller matching. In our model, communication costs therefore play a dual role: they affect firms' production costs and the cost of transacting across districts. Firms incur a communication cost $\delta_i \geq 1$ when hiring labor, so the effective marginal labor cost is $\delta_i w_i$. Firm profits are $(z_{ki}p_{ki} - \delta_i w_i)L_{ki}$, which implies the optimal producer price:

$$p_{ki} = \frac{\delta_i w_i}{z_{ki}} \tag{7}$$

Net of the hiring distortion, firm revenue is $\widetilde{R}_{ki} = w_i L_{ki}$. Communication costs also affect trade. Analogous to transportation costs, we assume an iceberg communication cost in selling one unit produced in j to consumers in i, denoted by $\tau^C_{ij} \geq 1$, with $\tau^C_{ij} = \tau^C_{ji}$. In contrast to transportation costs, however, we do not impose $\tau^C_{ii} = 1$, so communication frictions can matter even within a district. This feature is important for services, which are locally consumed but may still be communication intensive.

Consumer Prices. The price paid by consumers in district i for a variety produced in district j, denoted by p_{kij}, is:

$$p_{kij} = \tau^T_{kij}\tau^C_{ij}p_{kj} = \tau^T_{kij}\tau^C_{ij}\left(\frac{\delta_j w_j}{z_{kj}}\right), \quad \forall k \in \{A, M\} \quad \text{and} \quad p_{Sii} = \tau^C_{ii}\left(\frac{\delta_i w_i}{z_{Si}}\right) \tag{8}$$

Transportation costs do not matter for services because services are non-tradable by assumption. Communication costs can nevertheless affect services through within-district frictions because τ^C_{ii} need not equal one. Combining Equations (6) and (8), the share of expenditure by consumers in district j on a sector-k variety produced in district i, denoted by s_{kji}, is:

$$s_{kji} = \left[\frac{\tau^T_{kji}\,\tau^C_{ji} \times (\delta_i \frac{w_i}{z_{ki}})}{P_{kj}}\right]^{1-\sigma_k}, \quad \forall\, k \in \{A, M\} \tag{9}$$

This implies that total net revenue earned by a firm in district i and sector k is:

$$\widetilde{R}_{ki} = \sum_{j\in\mathcal{D}} \widetilde{R}_{kji} = \sum_{j\in\mathcal{D}} s_{kji}\, x_{kj}\, I_j \quad , \forall k \in \{A, M\} \quad \text{and} \quad \widetilde{R}_{Si} = x_{Si} I_i \tag{10}$$

Because services are non-tradable, net service-sector revenue equals local expenditure on services.

3.3. EQUILIBRIUM

Given the parameter vector $(\sigma, \{\eta_k, \varepsilon_k, \sigma_k\}_{\forall k})$, sectoral productivities $(\{z_{ki}\}_{\forall k,i})$, hiring costs $(\{\delta_i\}_{\forall i})$, bilateral transportation and communication costs $(\{\tau^T_{kij}\}_{\forall k,i,j}, \{\tau^C_{ij}\}_{\forall i,j})$, and labor supplies $(\{L_i\}_{\forall i})$, an equilibrium is a vector of wages and prices $\{\{w_i\}_{\forall i}, \{p_{kij}\}_{\forall k,i,j}\}_{\forall i}$ such that the following market-clearing conditions hold:[6]

1. **Labor Market Clearing:** In every district i, the total labor supplied equals the sum of the labor employed across the three sectors i.e., $L_i = \sum_k L_{ki}$.

2. **Output Market Clearing:** The total (net) revenue of a firm operating in district i and sector k equals total expenditure on these goods by consumers. For the tradable sectors A and M, this implies $w_i L_{ki} = \sum_{j\in\mathcal{D}} s_{kji}\, x_{kj}\, I_j$, where x_{kj} and s_{kji} are given by Equations (4) and (9) respectively. Since services are non-tradable, $w_i L_{Si} = x_{Si} I_i$.

3. **Aggregate Output Market Clearing:** The total (net) value of total production in the

[6] Appendix B.2 lists the full set of equilibrium conditions.

economy of sector k equals its total consumption, or $\sum_i w_i L_{ki} = \sum_i \sum_j s_{kji} x_{kj} I_j = \sum_i x_{ki} I_i$.

4. Empirical Strategy

The above model provides a stylized spatial framework to interpret how infrastructure plays a role in sectoral employment (and hence welfare). We now map the model's mechanisms into an empirical design. First, we construct measures of road- and GSM-driven market access that map directly to the bilateral expenditure-share equation in the model. We then describe the baseline worker-level specification, the identification strategy for transport and communication infrastructure, and the main reduced-form results on sectoral labor reallocation.

4.1. Measures of Market Access

Our model delivers a natural measure of market access in the spirit of Donaldson and Hornbeck (2016). Relative to a standard gravity setting, our environment includes both transportation frictions and communication frictions, each of which affects delivered prices and therefore expenditure shares. Improvements in market access can therefore arise through two distinct channels: reductions in travel time due to road expansion and reductions in communication frictions due to GSM rollout. Equation (9) provides the mapping from these frictions to bilateral expenditure shares. Because both wedges matter simultaneously, the market-access gain associated with one type of infrastructure must be defined holding the other fixed. Accordingly, we construct road-driven market access by holding GSM coverage at its 1999 level, and GSM-driven market access by holding road connectivity at its 1999 level.

We exploit cross-district variation in road construction and GSM rollout between 1999 and 2009 to study how improvements in physical and communication infrastructure affected workers' sectoral allocation. Our baseline specification is:

$$Y_{i(kd)t} = \beta_1 RoadAccess_{dt} + \beta_2 GSMAccess_{dt} + \beta_3 X_{it} + \gamma_t Z_{dt_0} + \alpha_d + \alpha_t + \varepsilon_{it} \qquad (11)$$

The dependent variable $Y_{i(kd)t}$ is an indicator for whether worker i in district d at time t is employed in sector $k \in \{A, M, S\}$. X_{it} is a vector of worker characteristics, including age, age squared, an indicator for completion of primary education, gender, marital status,

urban residence, and land ownership.[7] Z_{dt_0} is a vector of baseline district characteristics in 1999—the primary-education share, the urban population share, log night lights, and baseline employment shares in manufacturing and services—interacted with year fixed effects. These interactions allow districts with different initial socioeconomic conditions to follow flexible time paths. The specification also includes district fixed effects α_d and year fixed effects α_t. Our key explanatory variables are the road- and GSM-driven measures of market access described above, $RoadAccess_{dt}$ and $GSMAccess_{dt}$. Standard errors are clustered at the district level.

4.2. Instrumenting for Infrastructure Expansion

Although the baseline specification absorbs time-invariant district characteristics, aggregate shocks, and heterogeneous trends based on initial conditions, infrastructure expansion may still be endogenous to local economic change. Districts with greater growth potential or more capable local administrations may have received infrastructure earlier. In that case, within-district changes in GSM coverage or market access would reflect underlying economic dynamics rather than plausibly exogenous variation in supply-side rollout. To address this concern, we construct two distinct instrumental variables, one for road access and one for GSM access. Both instruments exploit geographic and engineering constraints that shaped the supply-side rollout of infrastructure independently of short-run district labor-market conditions.

Instrumenting for Road Access: Our road-access instrument is motivated by the idea of an optimal transportation network implied by geography. We begin with the 2000 global accessibility map developed by the European Commission and the World Bank (Nelson, 2008), which assigns a travel cost to each 1-km-by-1-km pixel in India (Appendix Figure A.6(a)). These costs are constructed from spatial data on topography, terrain, environmental conditions, and political boundaries.[8] Combining this raster with the 1996 road network, we compute the minimum spanning tree connecting all district centroids. The resulting network is the lowest-cost way to connect districts given preexisting roads and physical geography alone. It therefore captures the road network that would arise absent other distortions to infrastructure placement (Appendix Figure A.6(b)).

For each district, we compute optimal road market access, RMA_d^*, based on this geography-driven network and define $\widetilde{RMA_d}$ as the ratio of actual road market access in 1999, RMA_d^{99},

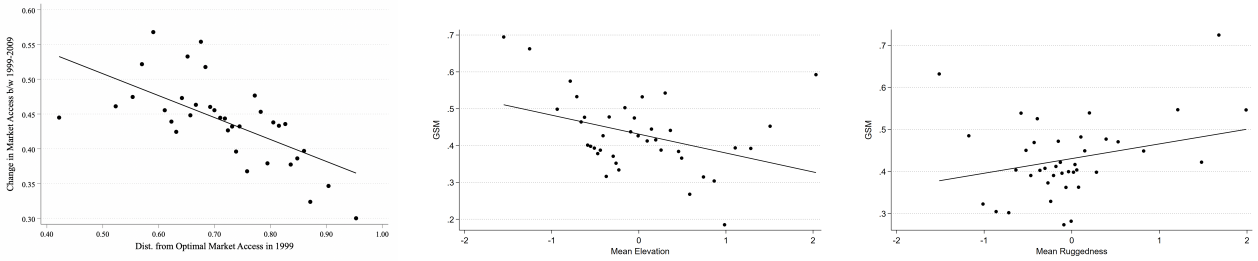
[7] For land ownership, we follow the Agriculture Census of India and classify marginal landowners as those owning less than one hectare of land and others as those owning more than one hectare.
[8] Construction details are available here.

to RMA_d^*. Our instrument is then $RMA_{dt}^{IV} = \widetilde{RMA}_d \times t$. The intuition is straightforward: districts that were initially underconnected relative to their geography-determined benchmark, i.e., farther away from their optimal access in 1999, should experience faster growth in road access as the observed network expands. Figure 4(a) confirms this.

Instrumenting for GSM Coverage: We instrument for changes in GSM coverage using time-invariant district topography. The key idea is that elevation and ruggedness affect both the cost and engineering feasibility of cellular rollout, but not necessarily in a monotonic way. Higher elevation may raise construction and maintenance costs, slowing rollout, yet it may also improve line-of-sight and signal reach once towers are installed. Rugged terrain similarly raises deployment costs but can also alter propagation conditions in ways that make the relationship with coverage nonlinear. These offsetting forces suggest that a flexible topography-based instrument is more appropriate than a single linear control.

To capture this flexibility, we place districts into five bins based on baseline elevation and five bins based on baseline ruggedness, and interact these bin indicators with time. The resulting instruments allow GSM coverage to follow different time paths across topographic categories while remaining anchored in predetermined geographic characteristics. Figures 4(b) and 4(c) illustrate the implied nonlinearity. Higher-elevation districts exhibit systematically lower GSM coverage over time, whereas more rugged districts tend to exhibit higher coverage, but neither relationship is close to linear.

FIGURE 4: First Stage Relationship in Road Access and GSM Coverage



(A) Market Access (B) GSM and Elevation (C) GSM and Ruggedness

Notes: Panel (a) plots changes in road access between 1999 and 2009 against the district-level road instrument constructed from the ratio of actual to optimal road access in 1999. Panels (b) and (c) plot GSM coverage against standardized elevation and ruggedness, respectively, controlling for the other topographic measure.

Because both the endogenous regressors and the instruments vary at the district-year level, identification comes from differences in how infrastructure expands across districts over

time. We estimate the following first-stage system:

$$\begin{aligned} \textit{Road Access}_{dt} = \beta_1^R \widetilde{RMA}_d * t + \sum_{k\in[2,5]} \zeta_k^R \mathbb{1}(\textit{Elev.Bin} = k) * t + \sum_{k\in[2,5]} \kappa_k^R \mathbb{1}(\textit{RuggedBin} = k) * t \\ + \delta_t^R Z_{dt_0} + \alpha_d^R + \alpha_t^R + \varepsilon_{dt}^R \\ \textit{GSM Access}_{dt} = \beta_1^G \widetilde{RMA}_d * t + \sum_{k\in[2,5]} \zeta_k^G \mathbb{1}(\textit{Elev.Bin} = k) * t + \sum_{k\in[2,5]} \kappa_k^G \mathbb{1}(\textit{RuggedBin} = k) * t \\ + \delta_t^G Z_{dt_0} + \alpha_d^G + \alpha_t^G + \varepsilon_{dt}^G \end{aligned} \tag{12}$$

Appendix Table A.1 reports the corresponding first-stage estimates. The optimal-road instrument strongly predicts road access, with an F-statistic of 15, while elevation and ruggedness have little predictive power for road access (F-statistic of 1.96). The reverse pattern holds for GSM access: elevation and ruggedness strongly predict GSM expansion, with an F-statistic of 12, while the optimal-road instrument has limited power for GSM access (F-statistic of 4.40). These results suggest that the two instruments capture distinct sources of variation in the rollout of roads and cellular coverage.

4.3. RESULTS

Table 1 reports our main empirical findings. Column (1) reports the baseline specification without controls for baseline district socioeconomic trends. Column (2) adds the full set of controls. In Column (3), we weight individuals using survey weights to make the sample more nationally representative. Column (4) simplifies the construction of the Road Access and GSM Access variables. In particular, we remove the role of initial GSM levels in the construction of Road Access and the role of initial travel times in the construction of GSM Access. Finally, Column (5) replaces GSM Access with GSM levels directly, providing a more general specification.[9]

Panel A shows that expansions in both road access and GSM access lead to statistically significant declines in the probability of agricultural employment across specifications. Panels B and C show that these declines are mirrored primarily by increases in service-sector employment rather than manufacturing employment. To interpret the magnitudes, consider our preferred specification in Column (2). For the average individual, road access rises by 0.14 log points between 1999 and 2009, so the estimated coefficient of -0.75 implies a 10.5-percentage-point decline in the probability of agricultural employment. Over the same

[9]For example, the model discusses the direct role of GSM in impacting district-level employment outcomes through production channels as opposed to market access.

TABLE 1: **Transport, Communication, and Sectoral Employment**

	(1)	(2)	(3)	(4)	(5)
Panel A. Agriculture					
Road Access	-0.539*	-0.748**	-1.153**	-0.721**	-1.132**
	(0.325)	(0.361)	(0.527)	(0.350)	(0.481)
GSM Access	-0.130***	-0.166***	-0.221***	-0.166***	
	(0.044)	(0.047)	(0.056)	(0.047)	
GSM					-0.255***
					(0.092)
Panel B. Manufacturing					
Road Access	-0.015	-0.031	0.401	-0.053	0.025
	(0.205)	(0.225)	(0.330)	(0.222)	(0.251)
GSM Access	0.020	0.041	0.042*	0.038	
	(0.026)	(0.028)	(0.024)	(0.027)	
GSM					0.046
					(0.042)
Panel C. Services					
Road Access	0.555*	0.779**	0.752	0.774**	1.107**
	(0.324)	(0.374)	(0.485)	(0.369)	(0.457)
GSM Access	0.109**	0.125***	0.179***	0.128***	
	(0.044)	(0.046)	(0.046)	(0.047)	
GSM					0.209**
					(0.086)
Observations	590,324	590,324	590,324	590,324	590,324
Individual Controls	Y	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y	Y
Socio-Economic Controls	N	Y	Y	Y	Y

Notes: The table uses all NSS survey rounds between 1999 and 2009. The dependent variable is an indicator for whether an individual works in agriculture (panel A), manufacturing (panel B), or services (panel C). All specifications include district and year fixed effects and worker controls for age, age squared, completion of primary education, gender, marital status, urban residence, and land ownership. Column (2) adds baseline district characteristics in 1999—the primary-education share, urban population share, log night lights, and baseline manufacturing and services employment shares—interacted with year fixed effects. Column (3) uses survey weights. Column (4) simplifies the construction of road and GSM access by removing the role of initial GSM and travel-time levels, respectively. Column (5) replaces GSM access with GSM levels directly. Standard errors are clustered at the district level. $*$ denotes $p < 0.1$, $**$ denotes $p < 0.05$, and $***$ denotes $p < 0.01$.

period, GSM access rises by 0.99 log points, implying a 16.5-percentage-point decline in the probability of agricultural employment, with the reallocation again occurring primarily toward services. The reduced-form evidence therefore points to economically meaningful

effects of both road and communication infrastructure on structural transformation, with somewhat larger implied effects for GSM expansion.

Appendix Table A.2 further decomposes employment within each broad sector into self-employment, wage employment, unpaid work, and employers. Two patterns stand out. GSM-driven market access is associated with movement out of self-employment and unpaid work in agriculture and into wage employment in services. Road-driven market access is associated with movement out of wage and unpaid agricultural work and into self-employment and wage employment in services, although these estimates are less precisely measured.

Taken together, the reduced-form evidence suggests that both transportation and communication infrastructure matter for structural transformation, and that both primarily reallocate labor out of agriculture and into services. The quantitative model in the next sections is designed to discipline these patterns, decompose the underlying mechanisms, and quantify the aggregate welfare consequences of infrastructure expansion.

5. Model Estimation

We now turn to calibrating the model. We begin (in Section 5.1) by parameterizing preferences and how transportation and communication costs relate to observed travel times and GSM coverage. We then estimate the key elasticities governing how these wedges affect delivered prices and production costs (Section 5.2). These parameter estimates discipline the counterfactual exercises in the next section.

5.1. Parameterization

Consumer Preferences. We set the demand-side parameters $\{\sigma, \{\eta_k, \varepsilon_k, \sigma_k\}_{\forall k}\}$ using values from the literature. Following the macro-development literature, we set the elasticity of substitution across sectors to $\sigma = 0.4$.[10] Following Tombe (2015), we set the within-sector elasticities of substitution across varieties to $\{\sigma_A, \sigma_M\} = \{4.06, 4.63\}$ in agriculture and manufacturing, respectively. We take the income elasticities across agriculture, manufacturing, and services from Comin et al. (2021), setting $\varepsilon_k = \{0.37, 1, 1.49\}$. Equation (5) shows that only relative values of ε_k matter for sectoral expenditure shares, so we normalize $\varepsilon_M = 1$ as in Comin et al. (2021). The implied ranking, $\varepsilon_A < 1 < \varepsilon_S$, captures the

[10] Several papers using CES demand estimate this elasticity between 0.4 and 0.5, including Heblich et al. (2024); Fajgelbaum and Redding (2022); Bah (2007); Rogerson (2008); Duarte and Restuccia (2010).

empirical pattern that agriculture's expenditure share falls with income while services' expenditure share rises. Finally, because the counterfactual analysis uses exact-hat algebra, we do not need to recover the levels of the preference shifters η_k; observed district-level expenditure shares x_{ki} absorb these terms.

Transportation and Communication Costs. Identification of the transportation and communication cost parameters across districts requires district-to-district trade flows, which we do not observe in the data. However, Donaldson (2018) and Asturias et al. (2019) develop a novel method to address this limitation. Intuitively, it is based on the observation that the production of certain goods is heavily concentrated in one district, but demand is spatially dispersed across the country. These can be used to generate proxy trade-flows for these products across districts, to then identify transportation costs.

We implement this strategy in four steps using detailed product-level data from the ASI. The ASI reports both quantities and expenditures for all 5-digit products produced and purchased by an establishment.[11] First, we construct establishment-level output and input prices for each product. Second, we identify a set of products $\mathcal{M}$ that are highly concentrated in a single district j, defined as products for which more than 95% of national sales come from establishments located in district j. Third, for two districts i and j, we assume that transportation and communication costs are related to travel times (T_{ij}) and GSM coverage (G_i, G_j) as follows:

$$\tau^T_{kij} = (1 + T_{ij})^{\theta_T} \quad \text{and} \quad \tau^C_{ij} = (1 + C_{ij})^{\theta_C} \tag{13}$$

where $C_{ij} = 1 - G_i G_j$, and θ_T and θ_C capture the elasticities of transportation and communication costs with respect to travel time and GSM coverage, respectively. A reduction in C_{ij} requires better GSM coverage in both districts i and j, rather than in only one district. We also impose $\tau^C_{ij} = 1$ when both districts have full GSM coverage, so communication frictions vanish when $C_{ij} = 0$.[12]

Finally, we assume that GSM mitigates hiring frictions through a parallel wedge, $\delta_i = (1 + C_{ii})^{\theta_L}$, where θ_L captures the elasticity of input frictions with respect to within-district

[11] To maintain consistency with the empirical analysis, we use ASI rounds for 1999, 2004, 2007, and 2009. Although the ASI contains both panel and repeated cross-sectional data, district identifiers are not available in the panel files for confidentiality reasons, so we use repeated cross-sections.

[12] We do not allow the transportation elasticity to differ across tradable sectors, although recent work suggests that $\theta_A > \theta_M$ because agricultural goods are more perishable. We show that the results are robust to higher values of θ_A. Note also that own-district travel time satisfies $T_{ii} = 0$, whereas $C_{ii} = 1 - G_i^2$. Thus, while $\tau^T_{ii} = 1$, τ^C_{ii} depends on own-district GSM coverage and need not equal one.

communication costs.

5.2. Estimates of Cost Elasticities

Estimates for θ_T and θ_C. Combining Equations (8) and (13), the price paid in district i for buying a product $m \in \mathcal{M}$ that is manufactured in district j is given by:

$$ln\ p_{im(j)} = \theta_M\ ln(1 + T_{ij}) + \theta_C\ ln(1 + C_{ij}) + \ln p_{m(j)} \tag{14}$$

We estimate the empirical counterpart of Equation (14):

$$ln\ p_{mft} = \alpha_{mt} + \theta_M\ ln(1 + T_{ijt}) + \theta_C\ ln(1 + C_{ijt}) + X_{ft} + \nu_{mft} \tag{15}$$

where p_{mft} is the price paid by establishment f in district i and year t for product m, which is produced predominantly in district j. Product-year fixed effects α_{mt} absorb changes in production costs and producer prices for product m over time, as well as other time-varying shocks to the origin district. We also control for establishment characteristics, including log total employment, to allow buyer size to affect pricing power. As in the worker-level analysis, we instrument both main explanatory variables. Specifically, we instrument $\ln(1 + T_{ijt})$ with the ratio of bilateral travel time in 1999 to the MST-implied travel time, $(T_{ij,1999}/T_{ij}^{MST})$, interacted with a time trend, and we instrument $\ln(1 + C_{ijt})$ using average elevation and ruggedness in district i, also interacted with a time trend.[13]

Table 2 reports the results. Column (1) shows the OLS estimates. Both coefficients are positive and statistically significant. Column (2) presents the IV estimates, which are larger in magnitude and remain precisely estimated, with $\hat{\theta}_T = 0.11$ and $\hat{\theta}_C = 0.60$. Columns (3)–(6) assess robustness to additional controls. Column (3) includes rural-year fixed effects to allow pricing patterns to differ between rural and urban establishments over time. Column (4) adds 3-digit industry-year fixed effects to absorb broad industry shocks. Column (5) allows for within-state pricing differences between origin and destination districts, capturing state-border frictions such as border delays, regulation, and tax differences. Column (6) excludes 1999, when GSM coverage was close to zero in most districts and exhibited little cross-district variation. Across specifications, the estimates remain stable and statistically significant. We therefore set $\{\theta_T, \theta_C\} = \{0.15, 0.55\}$ in the baseline calibration.

[13] Because $\alpha_{m(j)t}$ absorbs time-varying shocks to the origin district, including changes in G_j, the instrument primarily captures exogenous variation in C_{ij} driven by changes in G_i.

TABLE 2: **Estimates of θ_T and θ_C**

	Ln(Price$_{mft}$)					
	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
$\ln(1+T_{ijt})$	0.06***	0.11***	0.11***	0.13***	0.20***	0.12***
	(0.01)	(0.04)	(0.04)	(0.04)	(0.06)	(0.04)
$\ln(1+C_{ijt})$	0.26**	0.60**	0.58**	0.51*	0.57**	0.61**
	(0.11)	(0.28)	(0.28)	(0.29)	(0.28)	(0.29)
Prod.-Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Rural-Year F.E.	No	No	Yes	No	No	No
Ind.-Year F.E.	No	No	No	Yes	No	No
Same State F.E.	No	No	No	No	Yes	No
Sample	All	All	All	All	All	Excl. 1999
N	21600	20202	20202	20059	20202	19207
# Prod.	1505	1447	1447	1433	1447	1365

Estimates for θ_L. We now estimate θ_L, which governs how GSM coverage mitigates hiring and input frictions within a district. From the model, $p_{ki}y_{ki} = \delta_i(w_i L_{ki})$. We parameterize δ_i analogously to τ^C, setting $\delta_i = (1+C_i)^{\theta_L}$ with $C_i = 1 - G_i^2$. Before taking this implication to the data, note that the model uses labor as the only factor of production, whereas ASI establishments also use capital and materials. We therefore estimate the following reduced-form relationship in the ASI:

$$\ln \text{Sales}_{ft} = \alpha_i + \alpha_t + \theta_L \ln\big(1 + (1 + C_{it})\big) + \ln \mathbf{X}_{ft} + \nu_{ft} \tag{16}$$

where $\mathbf{X}_{ft}$ denotes expenditure on a vector of inputs, including labor, capital, and materials. District and year fixed effects absorb time-invariant local characteristics and aggregate shocks. Endogeneity remains a concern: districts with faster growth or better growth prospects may simultaneously experience more rapid GSM rollout and stronger sales growth.

Table 3 reports several specifications designed to address these concerns. First, as in the worker-level analysis, we instrument GSM availability using elevation and ruggedness interacted with a time trend. Column (2) yields an estimate of $\theta_L = 0.22$. Column (3) adds 3-digit industry-year fixed effects, so identification comes from comparing establishments within the same industry-year that differ in GSM growth because of location. Column (4) further adds state-year fixed effects to absorb state-specific growth patterns and policy

TABLE 3: **Estimate of θ_L**

	Ln(Sales$_{ft}$)					Res. Ln(Sales$_{ft}$)
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(1 + C_{it})$	0.04**	0.22**	0.17*	0.20**	0.20**	0.20**
	(0.02)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Dist. FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No
Ind.-Year FE	No	No	Yes	Yes	Yes	Yes
State-Year FE	No	No	No	Yes	Yes	Yes
Rural-Year FE	No	No	No	No	Yes	No
K.P. F-Stat.		357.40	322.79	599.65	579.48	599.65
N	70653	70653	70644	70644	70644	70644
Reg.	OLS	2SLS	2SLS	2SLS	2SLS	2SLS

changes correlated with GSM rollout. Column (5) adds rural-year fixed effects to allow rural and urban establishments to follow different time paths. Finally, to align the estimating equation more closely with the model, Column (6) residualizes log sales with respect to capital and material expenditures and then regresses the residual on $\ln(1 + C_{it})$ and the establishment wage bill, with the same fixed effects. Across Columns (2)–(6), the estimate of θ_L remains close to 0.2 and statistically significant. We therefore set $\theta_L = 0.2$ in the baseline model.

6. QUANTIFYING THE AGGREGATE IMPACTS OF TRANSPORT AND COMMUNICATION INFRASTRUCTURE

With estimates of the key trade and communication elasticities, we now turn our attention to using them to quantify the aggregate consequences of infrastructure expansion. We first describe how we invert the model using exact-hat algebra and recover the baseline moments required for the counterfactuals. We then assess the model's fit, quantify the effects of roads and GSM on structural transformation and welfare, and study why gains differ across districts. We conclude with robustness exercises that clarify why communication frictions matter so much quantitatively and how the results change when labor is allowed to move across space.

6.1. MODEL INVERSION

Let y denote the value of a variable of interest (e.g., employment share, sectoral price, etc.) observed in some 'baseline' equilibrium (say, 1999), and y' denote its value in a counterfactual equilibrium. $\hat{y} \equiv y'/y$ then denotes the relative change in y. In Appendix B.3, we show that our model structure allows us to express all key variables of interest in changes (or "hats"). However, there are two key moments that need to be recovered from the data at baseline, namely the sectoral (x_{ki}) and bilateral expenditure shares (s_{kij}). We now discuss them below.

Sectoral Expenditure Shares in a District (x_{ki}). We take 1999–00 as the baseline equilibrium to which the model is calibrated. The 1999–00 NSS round reports total weekly earnings, days worked in the previous week, and industry of employment for each worker. Using survey weights, we compute the average daily wage w_i, sectoral employment L_{ki}, and total employment $L_i = \sum_{k \in \{A,M,S\}} L_{ki}$ for each district.

For recovering sectoral expenditure shares, we combine empirical data moments with their model-implied counterparts. Since services are not traded, the total expenditure on services in district i equals the total cost of producing services there, net of the labor hiring wedge. This implies that:

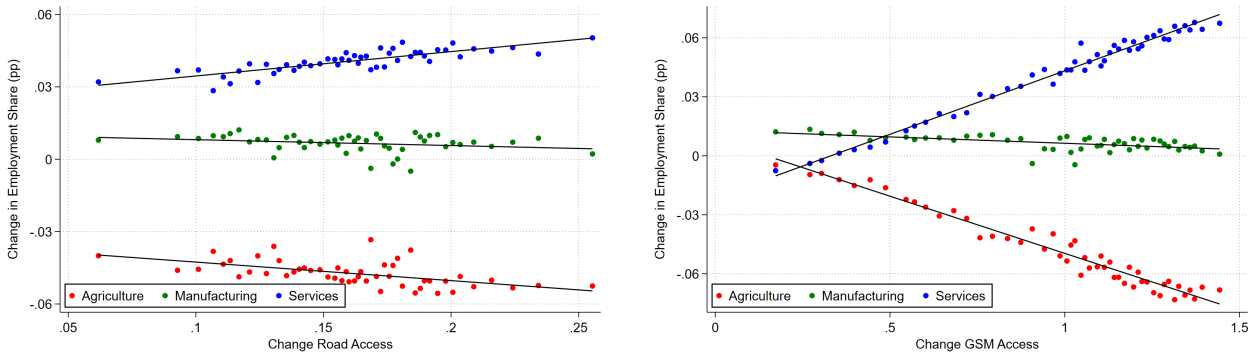
$$x_{Si} = \frac{L_{Si}}{L_i}. \tag{17}$$

Accordingly, we set the services expenditure share x_{Si} equal to the observed services employment share in each district. To recover expenditure shares in agriculture and manufacturing, we use the 1999–00 NSS consumer-expenditure survey, which records item-level household expenditures on agricultural, manufacturing, and services commodities over the previous month. Aggregating across households, we compute a district's manufacturing expenditure as a share of tradable expenditure, $x^T_{Mi} \equiv \frac{x_{Mi}}{x_{Mi}+x_{Ai}}$. Given x_{Si} from Equation (17), we recover the implied manufacturing expenditure share as $x_{Mi} = x^T_{Mi}(1 - x_{Si})$ and then define agricultural expenditure as $x_{Ai} = 1 - x_{Si} - x_{Mi}$.[14]

Bilateral Expenditure Shares (s_{kij}). Bilateral trade flows are not observed for India during this period, so they must be inferred from the model. We recover them using the following fixed-point algorithm:

[14] We rescale x_{Mi} by a common factor A so that aggregate manufacturing goods-market clearing holds nationally, that is, $\sum_{i \in \mathcal{D}} w_i L_{Mi} = \sum_{j \in \mathcal{D}} x_{Mj} w_j L_j$. Given the calibrated shares in manufacturing and services, aggregate goods-market clearing in agriculture, $\sum_{i \in \mathcal{D}} w_i L_{Ai} = \sum_{j \in \mathcal{D}} x_{Aj} w_j L_j$, then holds automatically.

FIGURE 5: Change in Employment Share: Road vs Communication Market Access



(A) Road Market Access (B) Communication Market Access

Notes: Panel (a) plots model-implied changes in district employment shares against changes in road market access, controlling for GSM market access. Panel (b) reports the analogous relationship for GSM market access, controlling for road market access.

(i) Initialize a guess for manufacturing productivity in district j, $\{z_{Mj}\}_{j\in\mathcal{D}}$.

(ii) Given z_{Mj}, compute p_{Mj}, and hence the bilateral manufacturing expenditure shares s_{Mji} using Equation (9).

(iii) From Equation (10), the total demand for manufacturing goods produced in district i can therefore be given by: $\widetilde{R}_{Mi} = \sum_{j\in\mathcal{D}} s_{Mji}\, x_{Mj}\, w_j L_j$, which from the zero-profit condition, should equal $w_i L_{Mi}$.

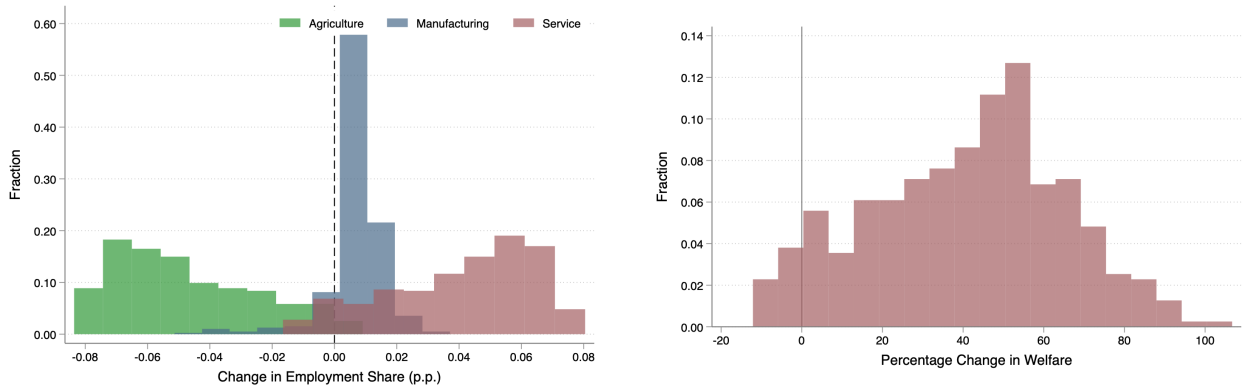
(iv) We update $\{z_{Mi}\}_{i\in\mathcal{D}}$ and iterate steps (ii)–(iv) until convergence.

This procedure yields district-level manufacturing productivity, z_{Mi}, and bilateral manufacturing expenditure shares, s_{Mij}, in the baseline equilibrium. We apply the same algorithm to recover $\{z_{Ai}, s_{Aij}\}$ for agriculture.

6.2. MODEL VALIDATION

We first assess whether the model can replicate the key reduced-form patterns in the data. Panel (a) of Figure 5 relates district-level changes in sectoral employment shares to changes in road market access, controlling for changes in GSM market access. Panel (b) presents the analogous relationship for GSM market access, controlling for road market access. The model reproduces the central empirical pattern: districts with larger expansions in either roads or GSM experience larger declines in agricultural employment and larger increases in services employment, with little response in manufacturing employment. The model also implies that GSM expansion accounts for the larger share of the predicted structural transformation, reflecting the much greater increase in communication market access over this period. We return to this decomposition when quantifying welfare gains by channel.

FIGURE 6: Impact on Sectoral Employment Shares and Real Income



(A) Change in Sectoral Employment Shares (B) Change in Real Income

Notes: Figure (a) shows the distribution (across districts) of the change in employment shares (in percentage points) between 1999 and 2009 for agriculture (green), manufacturing (blue) and services (red). Figure (b) shows the distribution (across districts) of the percentage change in real income, our measure of welfare. In both figures, the dotted vertical line is at zero.

Figure 6(a) plots the distribution of district-level changes in sectoral employment shares. Consistent with India's aggregate structural transformation, agriculture employment declines in nearly all districts, accompanied by rising services employment and only modest increases in manufacturing. For the median district, the agriculture employment share falls by 5.3 percentage points, the manufacturing share rises by 0.7 percentage points, and the services share rises by 4.6 percentage points. Overall, the model matches both the cross-district patterns associated with road and GSM expansion and the aggregate reallocation away from agriculture.

6.3. IMPACT OF INFRASTRUCTURE EXPANSION

We now use the model to quantify how the expansion of roads and GSM shaped the distribution of real income across districts. Figure 6(b) shows substantial heterogeneity in welfare gains. The unweighted average gain is 40.7%, while the employment-weighted average is 42.38%. The 25th and 75th percentiles are 21.6% and 57%, respectively. Although most districts gain from infrastructure expansion, the most adversely affected district experiences a welfare decline of 12.1%. Between 1999 and 2009, India's official real GDP per capita rose by 55.6%.[15] The model therefore implies that the combined expansion of roads and GSM can account for roughly 76% of observed growth over this decade.

Policymakers in developing countries often face binding budget constraints, so scaling up

[15]Data from the World Bank are available here.

one type of infrastructure may come at the expense of investment in another. It is therefore important to assess which frictions matter more, quantitatively, in shaping prices and real incomes. In our framework, the answer depends on two considerations.

First, the two frictions differ in how much they can be compressed at scale. By 2011, GSM coverage in India was close to universal in the sense that most of the population lived within reach of a tower, suggesting that communication frictions were relatively easy to reduce economy-wide. By contrast, even after substantial highway investment, long-distance travel times remained large, so transport frictions were harder to eliminate. Second, prices and trade flows differ in their sensitivity to each wedge, and that sensitivity is governed by the corresponding elasticities in the model. To illustrate, consider a district pair with bilateral travel time of 5 hours and GSM coverage of 25% in both districts in 1999. Using our estimates of θ_T and θ_C, halving travel time reduces the transportation component of the bilateral price wedge from roughly 30% to 20%, whereas doubling GSM coverage in both districts reduces the communication component from roughly 44% to 36%. Because both wedges enter nonlinearly, the quantitative effect depends on initial conditions. In particular, price wedges are more sensitive to reductions in travel time when baseline travel times are already low, while the marginal effect of GSM expansion depends on the product $G_i G_j$.

Table 4 reports average welfare changes across counterfactual exercises that vary one set of frictions at a time while holding the others fixed. When we allow only for road expansion—implemented as a reduction in transportation costs τ^T while holding GSM coverage at its 1999 level—the model implies an average welfare gain of 3.45%. This magnitude is broadly in line with recent estimates of highway expansion in other developing countries, including Brazil, China, and Turkey (Morten and Oliveira, 2024; Fan et al., 2023a; Coşar et al., 2022). In contrast, holding the road network fixed at its 1999 level and reducing only communication-related frictions (τ^C and δ) through GSM expansion yields an average welfare gain of 36.5%. Communication frictions therefore account for a large share of income growth between 1999 and 2009. Travel-time reductions matter, but their model-implied contribution is an order of magnitude smaller than that of declining communication frictions.

To further isolate the role of communication frictions, we decompose GSM-driven changes into an output-price channel and an input-cost channel and evaluate each component separately. A reduction in τ^T is most directly comparable to a reduction in the communication-linked output friction τ^C, since both operate through delivered prices of tradable goods. In the baseline model, however, τ^C also affects prices in the non-traded services sector. Hold-

TABLE 4: Decomposing Average Welfare Changes

Channel	Unweighted (%)	Employment Weighted (%)
	(1)	(2)
All Frictions (τ_T, τ_C, δ)	40.70	42.38
Only Roads (τ_T)	3.45	3.26
Only GSM (τ_C, δ)	36.51	38.37
Only GSM on Outputs (τ_C)	27.77	29.37
Only GSM on Tradable Outputs (τ_C)	17.81	19.15
Only GSM on Inputs (δ)	6.78	6.95
Only Output Frictions (τ_T, τ_C)	31.78	33.21

Notes: The above table shows the unweighted (Column 1) and employment-weighted (Column 2) percentage change in welfare due to various channels.

ing τ^T and δ fixed, reducing τ^C in line with observed GSM expansion still delivers a large average welfare gain of 27.8%. To make the comparison with roads more conservative, we also consider a version of the model in which communication-linked price wedges are removed from services, so that both τ^T and τ^C affect only tradable sectors. Even in this restricted setting, lowering τ^C implies an average welfare gain of 17.9%, well above the gains from lower travel-time costs. Finally, isolating the input-side channel by reducing the communication-linked labor friction δ while holding τ^T and τ^C fixed yields an average welfare gain of 6.78%.

Taken together, the decomposition shows that GSM-driven reductions in communication frictions generate substantial welfare gains through both output and input channels, and that their quantitative importance is robust to alternative modeling assumptions.

6.4. HETEROGENEITY IN WELFARE GAINS

Because the model features multi-sector trade and cross-district productivity differences, two districts experiencing similar infrastructure expansion need not realize similar welfare gains. We therefore study how baseline conditions shape the distribution of gains.

Sectoral Productivity. We first examine the role of initial productivity in tradable sectors. Panels A and B of Figure 7 show that, conditional on changes in road and GSM market access, districts with higher initial productivity in agriculture and manufacturing experience smaller welfare gains. In other words, holding the pace of infrastructure expansion fixed, lower-productivity districts in the tradable sectors benefit more from improved connectivity.

The intuition is straightforward. Baseline productivity determines how reliant a district is on purchases from other districts (Appendix Figure A.8).[16] More productive districts can supply tradable goods at lower cost (Equation 8), so they rely less on outside suppliers. When trade frictions decline, lower-productivity districts, which depend more heavily on imports, experience larger declines in local prices than high-productivity districts. At the same time, demand for goods produced by high-productivity net-exporting districts rises, which puts upward pressure on wages there.[17] The net result is that lower-productivity districts gain more from improved connectivity.

Initial Education and Urbanization Rates. We next examine how welfare gains vary with baseline education and urbanization. Panels C and D of Figure 7 plot district-level welfare gains against the 1999 primary-education share and urbanization rate, controlling for changes in road and GSM market access. More educated and more urbanized districts exhibit smaller welfare gains. Although the model does not explicitly include education or urbanization, both are plausibly associated with higher sectoral productivity. Appendix Figure A.9 supports this interpretation by showing a strong positive correlation between model-implied productivity in tradable sectors and baseline education and urbanization.[18] Consistent with the mechanism above, more educated and more urbanized districts appear to be more productive, less reliant on tradable imports, and therefore less exposed to the gains from declining trade frictions.

6.5. Robustness and Model Extensions

We conclude with two exercises that clarify the quantitative results. First, we ask whether the relatively modest role of roads reflects the realized pattern of highway construction rather than an intrinsic feature of transport costs. Second, we allow labor to move across districts and assess how spatial mobility changes both average welfare gains and their distribution.

6.5.1 Why Does Communication Matter More than Transport Costs for Welfare?

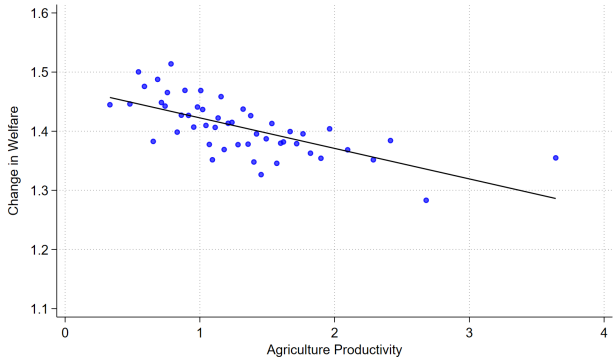
Our baseline results emphasize the central role of communication frictions in generating welfare gains, relative to the more modest contribution of lower travel times. One natural

[16] Because the model is closed, 'imports' refer to purchases from other districts within India.

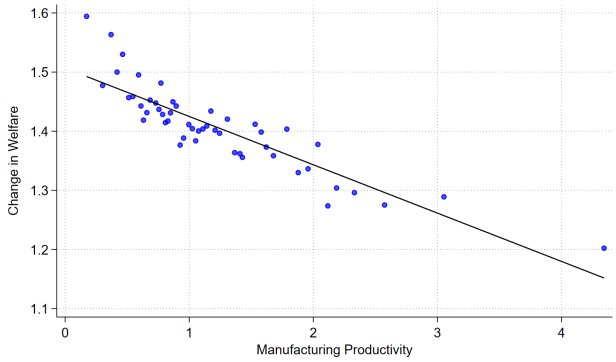
[17] Appendix Figure A.8 plots these opposing price and wage responses against initial productivity, controlling for changes in road and GSM access. In both tradable sectors, the direct price effect for low-productivity districts dominates the offsetting wage effect.

[18] These correlations provide reassurance that the inferred productivities are economically meaningful, up to scale, and offer an additional validation check for the model.

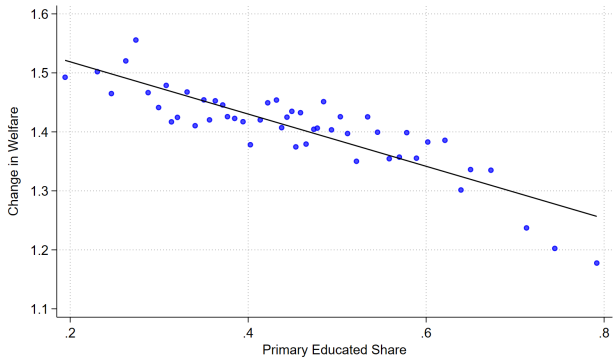
FIGURE 7: Heterogeneity in Welfare by Initial Productivity, Education, and Urbanization



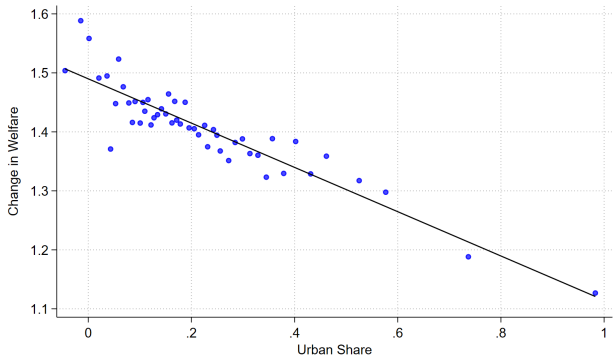
(A) Agricultural Productivity



(B) Manufacturing Productivity



(C) Education Share



(D) Urbanization Rate

Notes: The above figures show the change in welfare (relative to the benchmark) differentially by a district's initial productivity in agriculture (Figure a), manufacturing (Figure b), primary educational share (Figure c), and urbanization rate (Figure d).

interpretation is that GSM expanded much more dramatically than the highway network during this period.

To sharpen this comparison, we construct a counterfactual highway network using the same MST algorithm employed in the empirical analysis. This network minimizes travel time between district pairs given geographic constraints. Recomputing the counterfactual with the travel times implied by this optimal network raises the average welfare gain from road expansion to 4.68%, compared with 3.45% in Table 4. Thus, the realized highway expansion generated welfare gains about 25% smaller than those under an optimally constructed network. Even so, the implied gains remain far below the 17.81% welfare gain from lowering communication-linked output frictions alone. The dominance of communication frictions is therefore not simply an artifact of inefficient road placement.

6.5.2 Labor Mobility

Consistent with broad patterns of limited internal migration in India, the baseline model shuts down labor mobility across districts. We now relax that assumption. Following Redding (2016), we rewrite worker ω's utility from living in district i as:

$$\widetilde{U}_i(\omega) = b_i(\omega) U_i \tag{18}$$

where $U_i \equiv \frac{w_i}{P_i}$ is real income in district i, and $b_i(\omega)$ is a worker-district-specific idiosyncratic preference shock. We assume these shocks are independently drawn across workers and locations from a Fréchet distribution, $F(b) = e^{-B_i b^{-\chi}}$, where B_i captures district amenities and $\chi > 1$ governs the dispersion of idiosyncratic location preferences. Under this assumption, the probability of choosing district i is:

$$\zeta_i = \frac{(B_i w_i / P_i)^{\chi}}{\sum_{j \in \mathcal{D}} (B_j w_j / P_j)^{\chi}} \tag{19}$$

Intuitively, the probability of locating in district i rises with the amenity-adjusted real wage in that district and falls with the amenity-adjusted real wage elsewhere. Expected utility is equalized across locations in spatial equilibrium. We set the migration elasticity, χ, to 0.68, following Fan et al. (2023b), who estimate this parameter for India over a similar period.

The implications of labor mobility depend on the preference structure. With homothetic preferences, mobility tends to compress real-income differences across space because workers move toward districts with larger wage gains, expanding local labor supply and attenuating those gains. With non-homothetic preferences, however, migration also raises local income in receiving districts and therefore increases demand for locally provided services. This reinforces local labor demand and can offset, or even dominate, the equalizing effect of labor reallocation. As a result, labor mobility can amplify regional inequality in real income rather than reduce it.

In the counterfactual that allows for labor mobility, the average welfare gain is 39.7%, only slightly below the benchmark without mobility (Table 4). The decomposition by channel is also similar: transportation accounts for 3.61% and communication for 17.3% of average welfare gains. The main difference is distributional. Labor mobility substantially amplifies spatial inequality in real income. The worst-affected district now experiences a welfare loss of 42%, compared with 12.1% in the no-mobility benchmark, while the highest-gain district experiences a 116% gain, compared with 106% without mobility.

7. Conclusion

This paper studies how the joint expansion of roads and cellular networks shaped structural transformation in India during the 2000s. We show empirically that improvements in both road access and GSM access moved workers out of agriculture and into services, with little effect on manufacturing employment. We then develop and quantify a multi-sector spatial model in which roads lower transportation costs and cellular coverage lowers communication frictions in both trade and production. The model matches the reduced-form evidence and implies that the combined expansion of roads and GSM can account for a large share of India's real income growth over the period.

Our quantitative results also point to a sharp asymmetry across infrastructure types. Although road expansion generated positive welfare gains, the gains from reducing communication frictions were substantially larger. The reason is not simply that GSM expanded rapidly, but that communication infrastructure relaxes multiple margins at once and, in an economy with non-homothetic demand and non-tradable services, does so in a way that reinforces service-led structural transformation. More broadly, the paper shows that infrastructure matters not only for the *pace* of structural transformation, but also for its *direction*. When connectivity improvements allow tradable demand to be met through spatial reallocation while rising local income increases demand for non-tradable services, labor need not move from farms to factories. It may instead move from farms to services.

These findings have broader implications for how we think about development policy. In many low- and middle-income countries, transport and communication investments are often discussed under the common heading of infrastructure. Our results suggest that this aggregation can obscure economically important differences. Distinct infrastructure systems relax distinct frictions, and those differences can have first-order consequences for labor allocation, welfare, and spatial inequality. Understanding the interaction between infrastructure, market integration, and sectoral reallocation may therefore be central to explaining why structural transformation increasingly takes a service-led form in much of the developing world.

REFERENCES

Aker, Jenny C, "Information from markets near and far: Mobile phones and agricultural markets in Niger," *American Economic Journal: Applied Economics*, 2010, *2* (3), 46–59.

Allen, Treb, "Information frictions in trade," *Econometrica*, 2014, *82* (6), 2041–2083.

_ **and David Atkin**, "Volatility and the Gains from Trade," *Econometrica*, 2022, *90* (5), 2053–2092.

Asher, Sam and Paul Novosad, "Rural roads and local economic development," *American economic review*, 2020, *110* (3), 797–823.

Asturias, Jose, Manuel García-Santana, and Roberto Ramos, "Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India," *Journal of the European Economic Association*, 2019, *17* (6), 1881–1940.

Azam, Mehtabul, M Shahe Emran, and Forhad Shilpi, "ICT Skills and Labor Market Outcomes in India: Evidence from Cell Tower Expansion," Technical Report, IZA Discussion Papers 2024.

Blumenstock, Joshuaa, Niall Keleher, Arman Rezaee, and Erin Troland, "The impact of mobile phones: Experimental evidence from the random assignment of new cell towers," *Background paper, Innovations for Poverty Action, New Haven, CT*, 2020.

Chiplunkar, Gaurav and Pinelopi Koujianou Goldberg, "The employment effects of mobile internet in developing countries," Technical Report, National Bureau of Economic Research 2022.

Comin, Diego, Danial Lashkari, and Martí Mestieri, "Structural change with long-run income and price effects," *Econometrica*, 2021, *89* (1), 311–374.

Coşar, A Kerem, Banu Demir, Devaki Ghose, and Nathaniel Young, "Road capacity, domestic trade and regional outcomes," *Journal of Economic Geography*, 2022, *22* (5), 901–929.

Donaldson, Dave, "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American economic review*, 2018, *108* (4-5), 899–934.

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

Duarte, Margarida and Diego Restuccia, "The role of the structural transformation in aggregate productivity," *The quarterly journal of economics*, 2010, *125* (1), 129–173.

Fajgelbaum, Pablo and Stephen J Redding, "Trade, structural transformation, and development: Evidence from Argentina 1869–1914," *Journal of political economy*, 2022, *130* (5), 1249–1318.

Fan, Jingting, Yi Lu, and Wenlan Luo, "Valuing domestic transport infrastructure: A view from the route choice of exporters," *Review of Economics and Statistics*, 2023, *105* (6), 1562–1579.

Fan, Tianyu, Michael Peters, and Fabrizio Zilibotti, "Growing like India—the unequal effects of service-led growth," *Econometrica*, 2023, *91* (4), 1457–1494.

Ghani, Ejaz, Arti Grover Goswami, and William R Kerr, "Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing," *The Economic Journal*, 2016, *126* (591), 317–357.

Gollin, Douglas and Joseph P Kaboski, "New views of structural transformation: insights from recent literature," 2023.

Gonzalez-Navarro, Marcos, R David Zarate, R Jedwab, and N Tsivanidis, "Land transport infrastructure," *VoxDevLit*, 2023, *9* (1), 3.

Goyal, Aparajita, "Information, direct access to farmers, and rural market performance in central India," *American Economic Journal: Applied Economics*, 2010, *2* (3), 22–45.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya, "3g internet and confidence in government," *The Quarterly Journal of Economics*, 2021, *136* (4), 2533–2613.

hadj M Bah, El, "A three-sector model of structural transformation and economic development," 2007.

Heblich, Stephan, Stephen J Redding, and Yanos Zylberberg, "The distributional consequences of trade: Evidence from the grain invasion," Technical Report, National Bureau of Economic Research 2024.

Hjort, Jonas and Lin Tian, "The economic impact of internet connectivity in developing countries," *Annual Review of Economics*, 2021, *17*.

Jensen, Robert, "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector," *The quarterly journal of economics*, 2007, *122* (3), 879–924.

Kaboski, Joseph P, Jianyu Lu, Wei Qian, and Lixia Ren, "Melitz meets Lewis: The impacts of roads on structural transformation and businesses," Technical Report, National Bureau of Economic Research 2024.

Manacorda, Marco and Andrea Tesei, "Liberation technology: Mobile phones and political mobilization in Africa," *Econometrica*, 2020, *88* (2), 533–567.

Mensah, Justice Tei, "Mobile phones and local economic development: A global evidence," *Available at SSRN 3811765*, 2021.

Morten, Melanie and Jaqueline Oliveira, "The effects of roads on trade and migration: Evidence from a planned capital city," *American Economic Journal: Applied Economics*, 2024, *16* (2), 389–421.

Nelson, A, "Estimated travel time to the nearest city of 50,000 or more people in year 2000. Global Environment Monitoring Unit-Joint Research Centre of the European Commission, Ispra Italy," *URL: http://bioval. jrc. ec. europa. eu/products/gam/(accessed 30/07/2011)*, 2008.

Redding, Stephen J, "Goods trade, factor mobility and welfare," *Journal of International Economics*, 2016, *101*, 148–167.

Rogerson, Richard, "Structural transformation and the deterioration of European labor market outcomes," *Journal of political Economy*, 2008, *116* (2), 235–259.

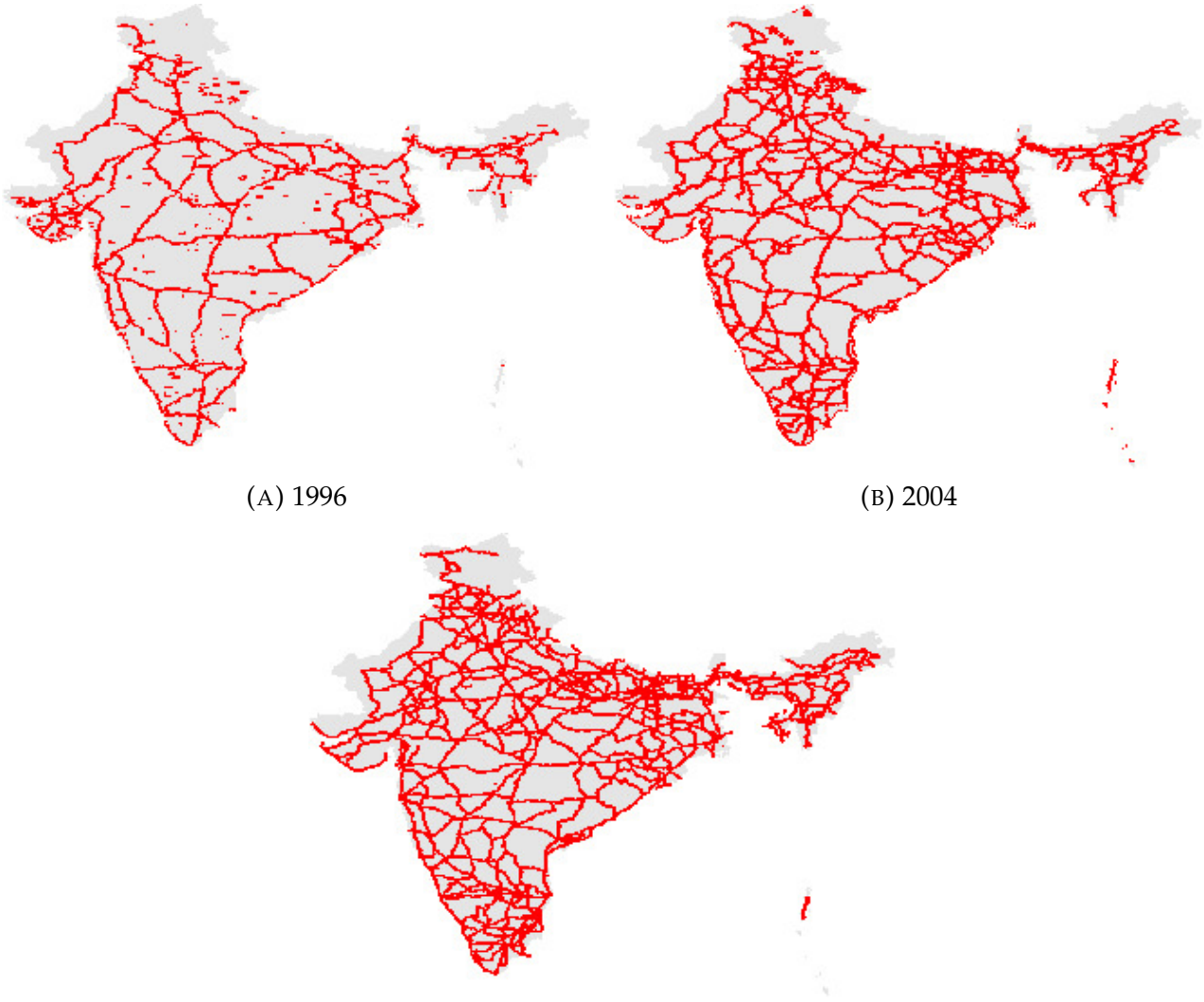
Tombe, Trevor, "The missing food problem: Trade, agriculture, and international productivity differences," *American Economic Journal: Macroeconomics*, 2015, *7* (3), 226–258.

Vogel, Kathryn Baragwanath, Gordon H Hanson, Amit Khandelwal, Chen Liu, and Hogeun Park, "Using Satellite Imagery to Detect the Impacts of New Highways: An Application to India," Technical Report, National Bureau of Economic Research 2024.

World Bank, *India's Transport Sector : The Challenges Ahead*, World Bank Publications, 2022.

A. Appendix Figures and Tables

FIGURE A.1: Road Network in India 1996-2011



(A) 1996

(B) 2004

(C) 2011

Notes: The figure shows the Indian road network in 1996, 2004, and 2011 in Allen and Atkin (2022).

FIGURE A.2: Travel Time By Distance and Across Years

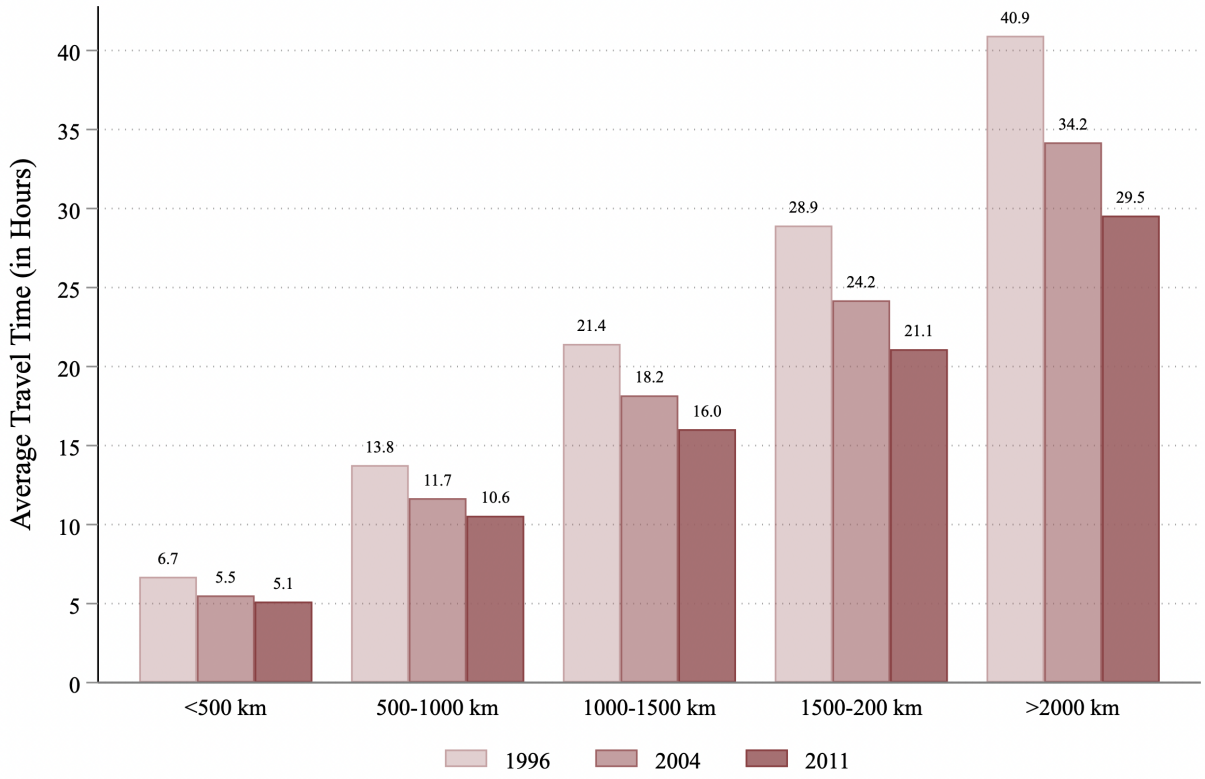
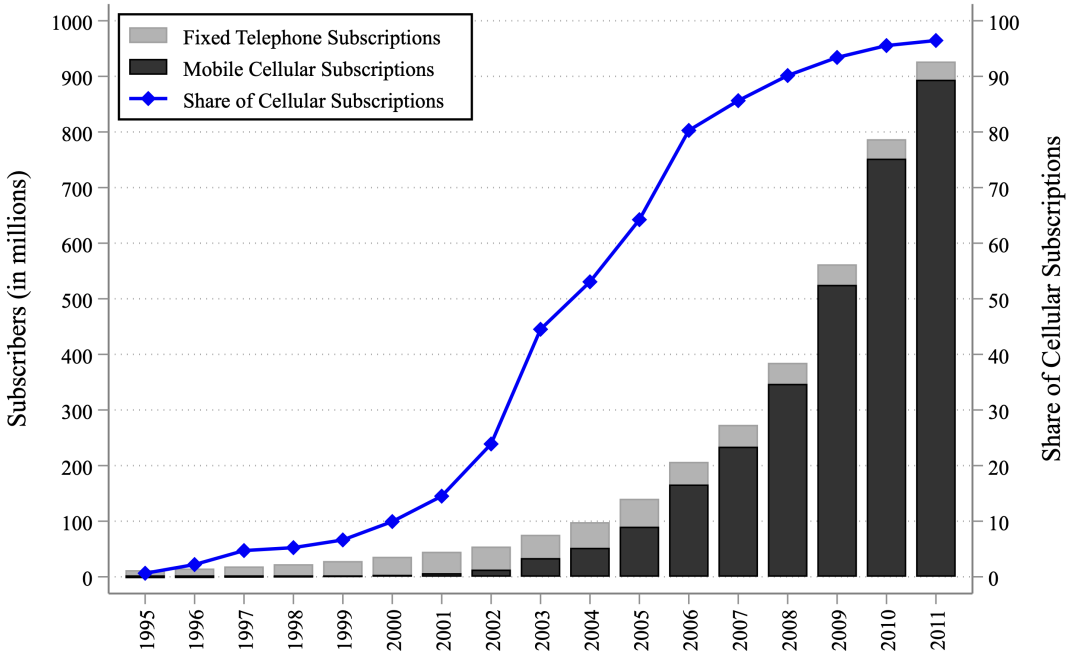


FIGURE A.3: Change in Travel Time Across Decades

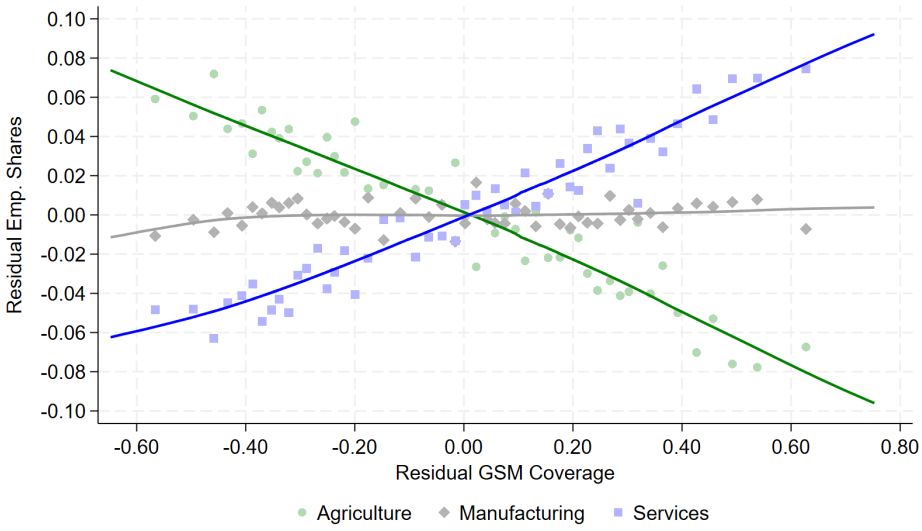
Notes: The above figure shows the change in average travel time based on distance between two districts in 1996, 2004, and 2011.

FIGURE A.4: Demand For Fixed vs Mobile Subscriptions between 1995-2011

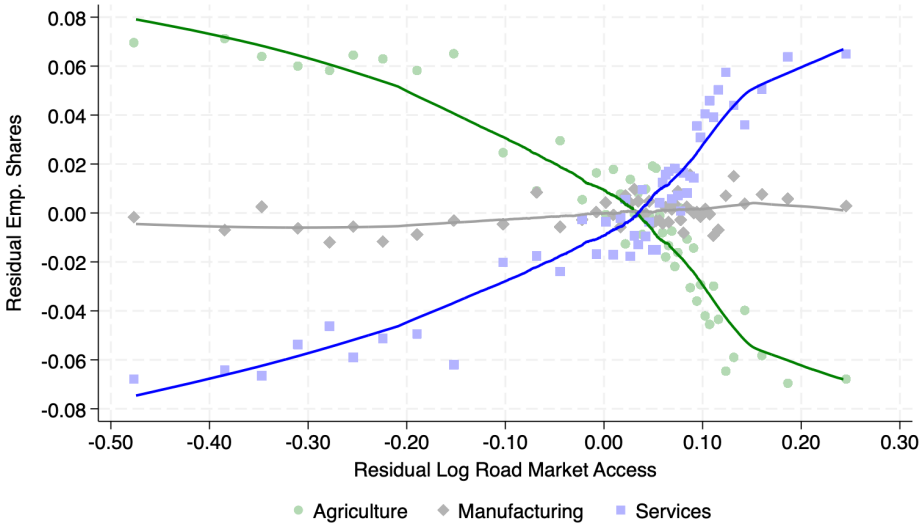


Notes: The figure uses data from the World Bank to examine the growth of telephone subscribers growth in India by fixed telephone subscribers (gray) and mobile subscribers (black) between 1995-2011. The blue line shows the share of cellular subscriptions for each year.

FIGURE A.5: Within-District Correlation Between Infrastructure and Sectoral Employment



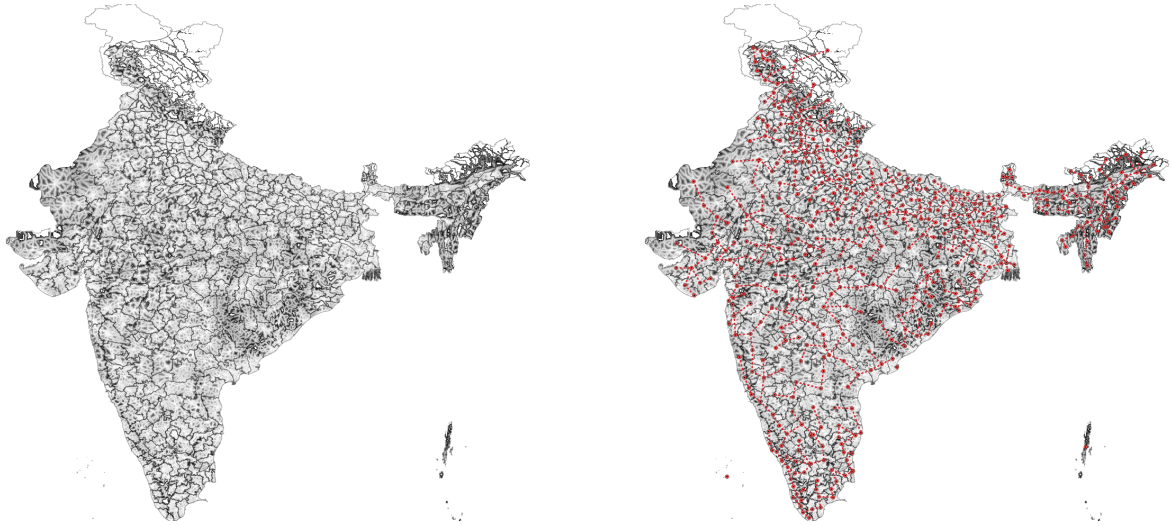
(A) Cellular GSM Coverage



(B) Log Road Market Access

Notes: Figures (a) and (b) plot the non-parametric correlations between the residualized district-level share of people employed in agriculture or manufacturing or services with residualized GSM coverage and market access in them. The residuals are constructed from a regression of the variable of interest on a district fixed effect.

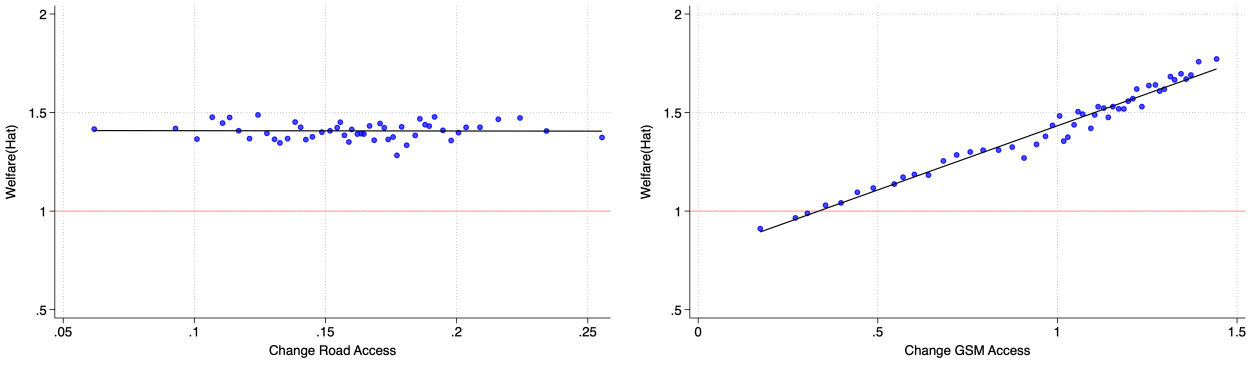
FIGURE A.6: Optimal Transportation Network



(A) Friction-Surface Raster of Accessibility (B) Optimal Transportation Network

Notes: TBD.

FIGURE A.7: Welfare and Market Access



(A) Road Market Access (B) Communication Market Access

Note:

TABLE A.1: **First Stage**

	(1) Road Access	(2) GSM Access
Road IV	-0.023***	0.072**
	(0.006)	(0.034)
𝟙(Elevation Bin = 2) * t	-0.002***	0.016***
	(0.001)	(0.004)
𝟙(Elevation Bin = 3) * t	-0.000	0.001
	(0.001)	(0.004)
𝟙(Elevation Bin = 4) * t	-0.002**	-0.012**
	(0.001)	(0.005)
𝟙(Elevation Bin = 5) * t	-0.002**	-0.004
	(0.001)	(0.005)
𝟙(Terrain Bin = 2) * t	0.000	-0.007*
	(0.001)	(0.004)
𝟙(Terrain Bin = 3) * t	0.001	-0.014***
	(0.001)	(0.004)
𝟙(Terrain Bin = 4) * t	0.000	-0.011***
	(0.001)	(0.004)
𝟙(Terrain Bin = 5) * t	0.000	-0.009*
	(0.001)	(0.005)
Observations	1,480	1,480
R-squared	0.995	0.937
Employment Controls	Y	Y
Socio-Economics Controls	Y	Y
Roads IV F-Stat	15.12	4.40
GSM IV F-Stat	1.96	11.74

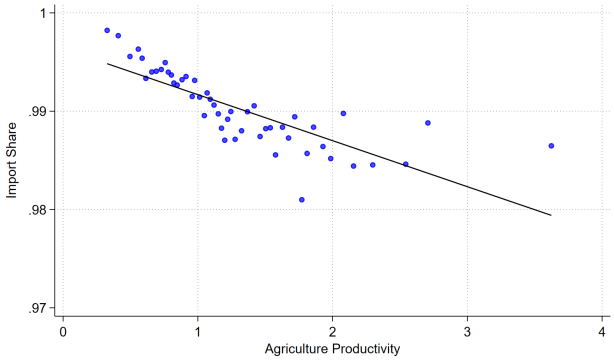
*** p<0.01, ** p<0.05, * p<0.1

TABLE A.2: **Individual Level Structural Change**

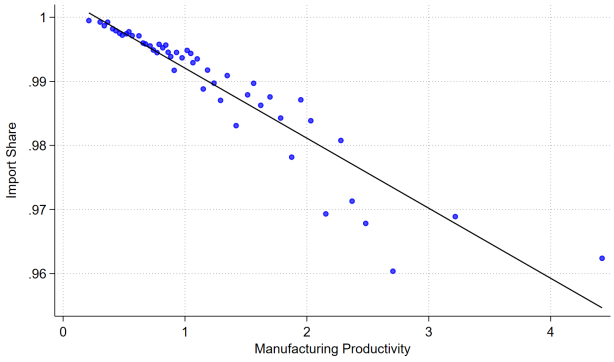
	Self	Wage	Unpaid	Employer
	(1)	(2)	(3)	(4)
Panel A. Agriculture				
Road Access	-0.060	-0.425	-0.317	0.025
	(0.242)	(0.363)	(0.255)	(0.074)
GSM Access	-0.094***	-0.009	-0.064*	0.008
	(0.028)	(0.044)	(0.034)	(0.009)
Panel B. Manufacturing				
Road Access	-0.047	-0.079	0.041	0.000
	(0.123)	(0.214)	(0.084)	(0.016)
GSM Access	0.009	0.044*	-0.010	0.001
	(0.014)	(0.023)	(0.011)	(0.002)
Panel C. Services				
Road Access	0.251	0.411	0.064	-0.038
	(0.179)	(0.291)	(0.091)	(0.028)
GSM Access	0.035	0.102***	-0.015	0.003
	(0.022)	(0.037)	(0.011)	(0.006)
Observations	611,095	611,095	611,095	611,095
Individual Controls	Y	Y	Y	Y
Employment Controls	Y	Y	Y	Y
Socio-Economic Controls	Y	Y	Y	Y

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

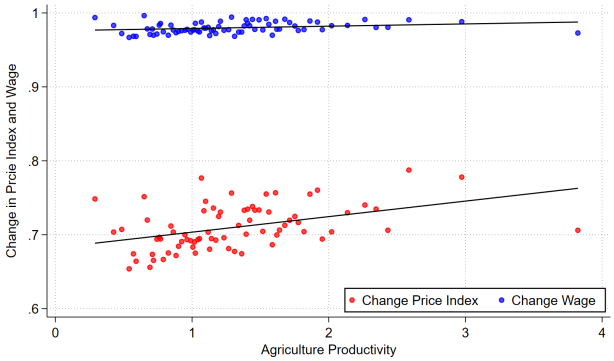
FIGURE A.8: Import Share and Initial Productivity



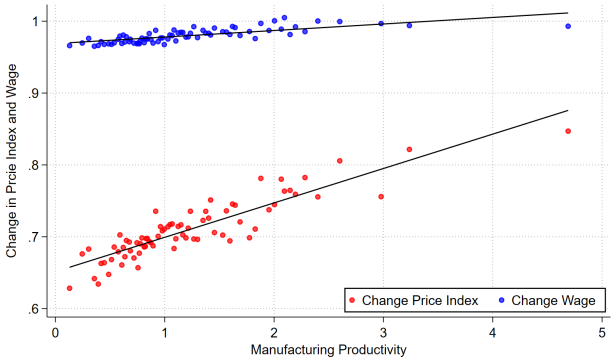
(A) Import Share: Agriculture



(B) Import Share: Manufacturing

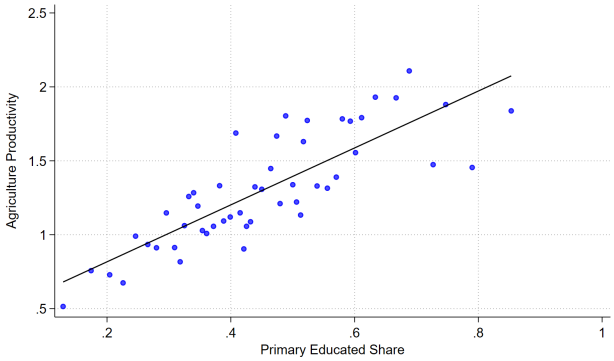


(C) Real Income: Agriculture

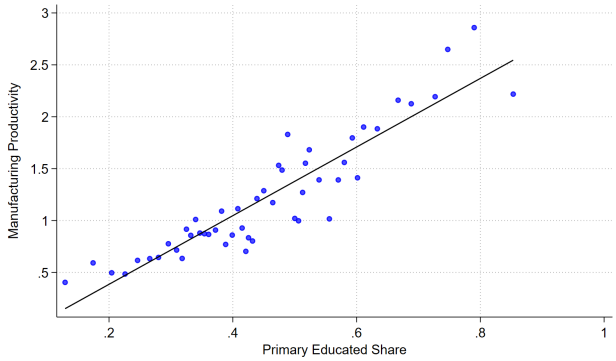


(D) Real Income: Manufacturing

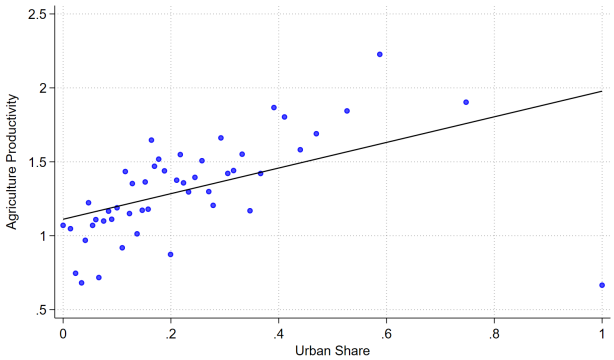
FIGURE A.9: Productivity, Education, and Urbanization



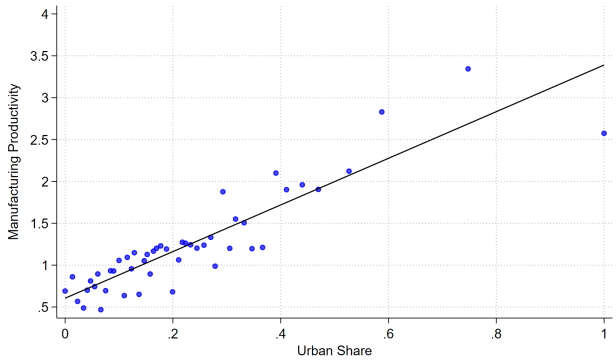
(A) Agriculture



(B) Manufacturing



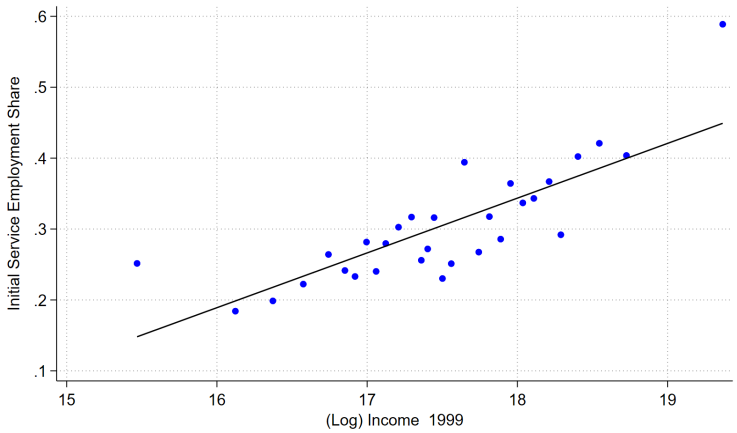
(C) Agriculture



(D) Manufacturing

Note:

FIGURE A.10: Service Employment and Initial Income



Note:

B. Theoretical Appendix

B.1. Non-homothetic CES Across Sectors

For any region i (dropping this subscript for notational clarity), let P_k denote the sectoral prices of sector k in region i and I denote total income. Consumers maximize utility (Equation 1) subject to the budget constraint $\sum_k P_k C_k \leq I$. Cost minimization implies:

$$\mathcal{L} = \sum_k P_k C_k - \gamma \left[\sum_k \eta_k^{1/\sigma} \left(\frac{C_k}{C^{\varepsilon_k}} \right)^{\frac{\sigma-1}{\sigma}} - 1 \right].$$

FOC w.r.t. C_k gives:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_k} &= P_k - \gamma \eta_k^{1/\sigma} \frac{\sigma-1}{\sigma} \left(\frac{C_k}{C^{\varepsilon_k}} \right)^{\frac{\sigma-1}{\sigma}-1} \cdot \frac{1}{C^{\varepsilon_k}} = 0. \\
\Rightarrow P_k &= \gamma \eta_k^{1/\sigma} \frac{\sigma-1}{\sigma} C_k^{-1/\sigma} C^{-\varepsilon_k(1-1/\sigma)} \\
\sum_k P_k C_k &= \gamma \frac{\sigma-1}{\sigma} \underbrace{\sum_k \eta_k^{1/\sigma} \left(\frac{C_k}{C^{\varepsilon_k}} \right)^{\frac{\sigma-1}{\sigma}}}_{=1} \\
\Rightarrow \gamma &= \frac{\sigma}{\sigma-1} I
\end{aligned}$$

Using $C = I/P$ and substituting back in the price expression we get:

$$\begin{aligned}
P_k &= I \;\; \eta_k^{1/\sigma} C_k^{-1/\sigma} \;\; (I/P)^{-\varepsilon_k(1-1/\sigma)} \\
C_k &= \eta_k \left(\frac{P_k}{P} \right)^{-\sigma} \;\; (I/P)^{\sigma+\varepsilon_k(1-\sigma)} \\
x_k \equiv \frac{P_k C_k}{I} &= \eta_k \left(\frac{P_k}{P} \right)^{1-\sigma} \;\; \left(\frac{I}{P} \right)^{(\varepsilon_k-1)(1-\sigma)}
\end{aligned} \tag{B.1}$$

Ratios of Expenditure Shares. For two sectors k and m, the ratios of expenditure shares can be given by:

$$\frac{x_k}{x_{k'}} = \left(\frac{\eta_k}{\eta_{k'}} \right) \left(\frac{P_k}{P_{k'}} \right)^{1-\sigma} \left(\frac{I}{P} \right)^{(\varepsilon_k - \varepsilon_{k'})(1-\sigma)}. \tag{B.2}$$

Aggregate Price Index. From Equation (B.1), the aggregate price index can be calculated

as follows:

$$\sum_k P_k C_k = I \cdot \underbrace{\sum_k \eta_k \left(\frac{P_k}{P}\right)^{1-\sigma} \quad \left(\frac{I}{P}\right)^{(\varepsilon_k - 1)(1-\sigma)}}_{=1}$$

$$\text{Rearranging:} \quad P = \left[\sum_k \eta_k P_k^{1-\sigma} \quad \left(\frac{I}{P}\right)^{(\varepsilon_k - 1)(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \tag{B.3}$$

B.2. General Equilibrium

The general equilibrium of the model solves the following system of equations:

Aggregate price index in district i:

$$P_i = \left[\sum_k \eta_k \, P_{ki}^{1-\sigma} \left(\frac{I_i}{P_i}\right)^{(\varepsilon_k - 1)(1-\sigma)}\right]^{\frac{1}{1-\sigma}} \tag{B.4}$$

Total income in the district:

$$I_i = w_i L_i \tag{B.5}$$

Total labor supply in the district:

$$L_i = \sum_k L_{ki} \tag{B.6}$$

Expenditure share of sector k in the district:

$$x_{ki} = \eta_k \left(\frac{P_{ki}}{P_i}\right)^{1-\sigma} \quad \left(\frac{I_i}{P_i}\right)^{(\varepsilon_k - 1)(1-\sigma)} \tag{B.7}$$

Share of district j in district i's expenditure of on sector k:

$$s_{kij} = \left(\frac{p_{kij}}{P_{ki}}\right)^{1-\sigma_k}, \quad \forall\, k \in \{A, M\} \tag{B.8}$$

Bilateral price of tradables from district j for consumers in district i:

$$p_{kij} = \tau_{kij} \lambda_{ij} p_{kj} = \frac{\tau_{kij} \lambda_{ij} \delta_j w_j}{z_{kj}}, \quad \forall\, k \in \{A, M\} \tag{B.9}$$

Sectoral price index of tradables in district i:

$$P_{ki} = \left[\sum_{j \in \mathcal{D}} p_{kij}^{1-\sigma_k} \right]^{\frac{1}{1-\sigma_k}}, \quad \forall\, k \in \{A, M\} \tag{B.10}$$

Price of services in district i:

$$P_{Si} = \frac{\lambda_{ii} \delta_i w_i}{z_{Si}} \tag{B.11}$$

Labor employed in sector k in district i:

$$L_{ki} = \frac{\sum_j s_{kji} x_{kj} w_j L_j}{w_i}, \quad \forall\, k \in \{A, M\} \quad \text{and} \quad L_{Si} = x_{Si} L_i \tag{B.12}$$

Revenue of sector k in district i:

$$R_{ki} = \delta_i w_i L_{ki} \tag{B.13}$$

B.3. Exact Hat Algebra

Let y denote the value of the variable in the initial equilibrium and y' denote a counterfactual value, then define $\widehat{y} = y'/y$ as the relative change in the variable across the two equilibria. Note that for any variable $y = \sum_i x_i$, define $\alpha_i = x_i / y_i$. Then, $\widehat{y}_i = \sum_i \alpha_i \widehat{x}_i$.[19]

Changes in $\{\tau_{ij}^T, \tau_{ij}^C, \delta_i\}$ can be given by:

$$\widehat{\tau}_{ij}^T = \left(\frac{1 + T_{ij}'}{1 + T_{ij}} \right)^{\theta_T}; \quad \widehat{\tau}_{ij}^C = \left(\frac{1 + C_{ij}'}{1 + C_{ij}} \right)^{\theta_C}; \quad \widehat{\delta}_i = \left(\frac{1 + C_i'}{1 + C_i} \right)^{\theta_L} \tag{B.14}$$

Holding constant productivities (z_{ki}), preference parameters (η_k), and labor supply in each district (L_i), and initial guesses for changes in wages $\widehat{w}_i^0$ for all districts, we can express

[19] To see this, note that $\frac{y'}{y} = \sum_i \frac{x_i'}{y} = \sum_i \frac{x_i}{y} \cdot \frac{x_i'}{x_i} = \sum_i \alpha_i \widehat{x}_i$.

sectoral price changes in general equilibrium as follows:

$$\widehat{I}_i = \widehat{w}_i^0 \tag{B.15}$$

$$\text{From Equation 2:} \quad \widehat{P}_i = \left[\sum_k x_{ki} \, \widehat{P}_{ki}^{1-\sigma} \left(\frac{\widehat{I}_i}{\widehat{P}_i} \right)^{(\varepsilon_k - 1)(1-\sigma)} \right]^{\frac{1}{1-\sigma}} \tag{B.16}$$

$$\text{From Equation 6:} \quad \widehat{P}_{ki} = \left[\sum_{j \in \mathcal{D}} s_{kij} \, \widehat{p}_{kij}^{1-\sigma_k} \right]^{\frac{1}{1-\sigma_k}}, \quad k \in \{A, M\} \tag{B.17}$$

$$\text{From Equation 8:} \quad \widehat{p}_{kij} = \widehat{\tau}_{ij}^T \, \widehat{\tau}_{ij}^C \, \widehat{\delta}_i \, \widehat{w}_j^0, \quad k \in \{A, M\} \tag{B.18}$$

$$\widehat{P}_{Si} = \widehat{\tau}_{ii}^C \, \widehat{\delta}_i \, \widehat{w}_i^0 \tag{B.19}$$

Given these, the changes in labor demand from Equation (10) can be given by:

$$\widehat{L}_{ki} = \frac{1}{\widehat{w}_i^0} \left[\sum_j \Psi_{kji} \;\; \widehat{s}_{kji} \, \widehat{x}_{kj} \, \widehat{I}_j \right], \quad k \in \{A, M\} \tag{B.20}$$

$$\widehat{L}_{Si} = \widehat{x}_{Si} \tag{B.21}$$

$$\text{where:} \quad \Psi_{kji} = \frac{s_{kji} \, x_{kj} \, I_j}{\sum_m s_{kmi} \, x_{km} \, I_m}$$

Given these, the change in expenditure shares can be given by:

$$\text{From Equation 4:} \quad \widehat{x}_{ki} = \left(\frac{\widehat{P}_{ki}}{\widehat{P}_i} \right)^{1-\sigma} \left(\frac{\widehat{I}_i}{\widehat{P}_i} \right)^{(\varepsilon_k - 1)(1-\sigma)} \tag{B.22}$$

$$\text{From Equation 6:} \quad \widehat{s}_{kij} = \left(\frac{\widehat{p}_{kij}}{\widehat{P}_{ki}} \right)^{1-\sigma_k} \tag{B.23}$$

From the above equations, the implied change in wages, $\widehat{w}^0$ for all districts i is:

$$\widehat{w}_i = \widehat{w}_i^0 \sum_k \left(\frac{L_{ki}}{L_i} \right) \cdot \widehat{L}_{ki} \tag{B.24}$$