International Access, Structural Change and The Distributional Effects of Trade Liberalization*

Anmol Agarwal
University of Virginia

Aishwarya Kekre
University of Virginia

March 26, 2026

Abstract

We study how access to international markets shapes the spatial distribution of gains from large-scale external integration. Exploiting India's 1990s trade liberalization against the backdrop of an underdeveloped domestic road network, we show that districts with better port access experienced faster growth in manufacturing output, average firm size, firm entry, worker inflows, and reallocation of labor out of agriculture. We develop a quantitative spatial model to evaluate the distributional consequences of this liberalization. Counterfactual exercises reveal stark spatial inequalities in welfare gains, disproportionately favoring well-connected districts. Hypothetical parallel highway investments targeting port connectivity substantially reduce these spatial disparities.

JEL Classification: F14, F15, F16, F63, R12, R13

Keywords: Structural Transformation, International Trade, Roads

*We thank Kerem Coşar, James Harrigan, John McLaren, Jonathan Colmer, members of the Trade and Development Groups at the University of Virginia, and conference participants at EIIT, Midwest International Trade Conference, STEG Annual Conference, and ISI Delhi for their valuable comments. We are grateful to Treb Allen, Devaki Ghose, and Petia Topalova for sharing data which was instrumental for this project. All errors are our own.

1 Introduction

Domestic transport infrastructure is a critical determinant of how well regions within a country are integrated with international markets. Remoteness insulates regions from foreign competition, while dampening any positive effects from declining international trade barriers. Poor transport infrastructure can amplify these margins of gains and losses. In this paper, we examine how access to international markets shapes the regional effects of external integration. Specifically, what are the distributional effects of a large-scale external integration when internal integration is incomplete? Can concurrent investments in improving road infrastructure mitigate spatial welfare inequalities? We address these questions using one of the largest episodes of trade liberalization– India's trade liberalization during the 1990s, affecting a sixth of the global population. A defining feature of this period was India's weak road network, which created marked disparities in regional access to global markets via ports.

We empirically establish that access to international ports is a key determinant of the spatially uneven effects of India's trade liberalization. We find that trade liberalization led to a growth in manufacturing activity across all districts, with significantly larger gains in districts with better international access. These positive effects are driven by reductions in input tariffs, with no offsetting effects from output tariffs. Districts with better international access experience stronger structural change across the tradable sectors of the economy, with employment shifting out of agriculture and into manufacturing and producer services. We also document substantial migration from low- to high-international access districts in the decade following liberalization.

To account for these empirical findings, we develop a quantitative spatial model of structural change building on Heblich et al. (2024). The model introduces tradable intermediate inputs in manufacturing production and allows for heterogeneity in the time sensitivity of agricultural output, manufacturing intermediates, and manufacturing output, which is the key mechanism driving the spatial effects. We estimate the time sensitivity of these intermediates and find that they are more time-sensitive than manufacturing output but less time-sensitive than agricultural output. Our counterfactual trade liberalization exercise reveals pronounced spatial disparities in welfare gains, with substantially larger gains in districts with high international access. Finally, we construct a hypothetical road network targeting port connectivity and show that parallel improvements in international access substantially attenuate the spatial disparities generated by trade liberalization.

The setting of this paper is India's trade liberalization in the 1990s, which reversed decades of inward-looking policies. Import restrictions were eased, accompanied by a sharp decline in manufacturing tariffs. India witnessed a massive surge in manufacturing imports, particularly in intermediate goods, which grew by 227% between 1987 and 2000 (Goldberg et al., 2010). In agriculture, where India was consistently a net exporter, export restrictions were further relaxed through the gradual removal of export quotas, licenses, and prohibitions (Gulati and Mullen, 2003). This made agriculture more vulnerable to world price movements in a period when world prices in major agriculture products were falling. Turning to internal integration, we document substantial

disparities in regional access to major international gateways. Utilizing the 1996 road network of India, we find that travel time from a district to its nearest major port ranges from less than one hour to over 21 hours. Using these travel times, we construct a measure of international access, expressed as an iceberg trade cost. India in the 1990s, therefore, provides an ideal spatial setting to study how initial access to world markets shapes the gains from trade liberalization.

We first examine the spatial effects of liberalization in the manufacturing sector. While much work emphasizes the adverse local effects of import competition (Kovak, 2013; Autor et al., 2013), in a developing country such as India, where imports of final goods are limited, the negative impact of competition can be outweighed by improved access to cheaper or higher quality foreign intermediates that lower production costs.[1] Using a panel dataset covering 340 districts and 100 industries from 1987 to 2004, combined with our estimated trade costs, we show that in response to input tariff declines, within industry manufacturing output rises, with substantially stronger gains in districts that have better international access. For a 10 percentage points (pp) reduction in input tariffs, reducing trade costs by the median value is associated with a 16.69% larger increase in total output, an 11.65% increase in average firm size, a 14.97% increase in the number of firms, and a 31.27% increase in total employment. We do not find significant effects of output tariffs, either in the aggregate or across space. These results are not driven by a mere reallocation of output from high to low trade cost districts, since the marginal effect of an input tariff reduction remains positive even in districts with high trade costs.

We next examine whether the unequal distribution of manufacturing gains translates into differences in sectoral labor reallocation across districts. Using the Population Census and nationally representative household surveys (NSS), we find that between 1991 and 2004 the lowest trade cost district experienced a 6.25 percentage point larger decline in agricultural employment, a 3.21 percentage point larger increase in manufacturing, and a 1.69 percentage point larger increase in producer services, relative to the median trade cost district. These effects are sizeable – on the conservative side, roughly 1.16 million additional workers moved out of agriculture in below-median trade cost districts relative to above-median districts. Event studies rule out the presence of any differential pre-trends in employment outcomes. Further, using confidential Census data on district-to-district migration, we document significant post-liberalization migration flows from low to high international access districts.

Next, we develop a quantitative spatial model to examine the implications for the distribution of welfare. We consider production across two tradable sectors - agriculture and manufacturing, and model districts as small open economies facing exogenous world prices for these sectors. Crucially, we allow factor usage to differ across sectors – agriculture uses land and labor, while manufacturing uses land, labor, and intermediates.[2] We treat intermediate inputs as tradables with their

[1] According to WITS Data, consumer goods made up only 10.7% of India's imports in 2000 while the corresponding U.S. share was 36%

[2] Calculations from the 1993 I-O Table of India show that intermediate inputs accounted for only 16% of gross value of output in agriculture whereas the same share was over 50% in manufacturing. Further, Peter and Ruane (2023) show an average share of intermediates in output of 81.6% for Indian manufacturing firms in 1989.

price determined exogenously on the world market. Land is owned by geographically immobile landlords and labor is mobile across sectors and districts. Each district has an endowment of a continuum of land plots differing in their productivity for alternative sectoral use.

An important feature of our model is that we allow time sensitivity to differ across manufacturing output, manufacturing intermediates, and agricultural output. Using firm by input import data, we estimate the economies of time for intermediates to be lower than for manufacturing output but higher than for agricultural output. We use household surveys on employment and consumption, satellite data on land cover, manufacturing survey data from ASI, and estimates from the literature to calibrate model parameters to match the observed data to the pre-liberalization Indian economy. We show that our observed data are sufficient to recover the unobserved endogenous variables required for the counterfactuals.

As our first counterfactual exercise, we liberalize the Indian economy. We lower the prices of manufacturing output and intermediates using the exogenous changes in output and input tariffs, respectively, and adjust agricultural prices in line with the observed changes in world commodity prices for agricultural goods.[3] We use exact hat algebra to isolate the effects of trade liberalization while holding constant productivities, amenities, and trade costs. We find that the model rationalizes our empirical findings. In response to the falling prices of tradables, districts with better international access experience a significantly larger increase in manufacturing output, greater structural change from agriculture towards manufacturing, and an increase in population relative to districts with weaker international access.

Both the use of intermediates in manufacturing and the higher time sensitivity of intermediates relative to manufacturing output are central to our results. First, although the relative price of manufacturing to agriculture falls after trade liberalization, the relative cost-adjusted price of manufacturing rises because intermediate input costs decline. Since agriculture does not benefit from a similar cost reduction, this raises rents in manufacturing and generates structural change away from agriculture. Second, because intermediates are more time sensitive, interior districts experience a smaller decline in input costs relative to the decline in output prices. As a result, compared with port districts, they face a lower cost-adjusted price of manufacturing and a lower cost-adjusted relative price of manufacturing to agriculture. This implies that, although labor moves out of agriculture and into manufacturing in all districts, the reallocation is weaker in interior districts.

We document a strong negative relationship between welfare gains and districts' trade costs to ports. Although real income rises in every district, the gains are concentrated in regions with better international access. A reduction in trade cost by the median amount is associated with a 14.42% larger increase in real income. We find similarly negative relationships between district trade costs and changes in real rents and real wages.

Finally, we examine the role of improvements in road infrastructure in mitigating the spa-

[3] We show that our imposed price changes match well with the observed relative price change of agriculture to manufacturing during the period.

tial welfare inequalities from trade liberalization. We construct a hypothetical highway network targeting port connectivity, taking into account the cost for traveling across any 1×1 km pixel in India.[4] Re-running the trade liberalization counterfactual but with simultaneous reduction in trade costs substantially reduces the spatial inequality in welfare gains from India's liberalization. For a given decline in trade costs, the additional welfare gains are cut by over a factor of seven of the baseline effect, indicating a much flatter welfare–international access gradient. However, the spatial inequalities in structural change fall only marginally as a larger decline in the cost of intermediates is counteracted by a smaller fall in the price of agriculture for the interior districts. At the aggregate, we find a 34.3% increase in expected worker utility from trade liberalization with optimal road investments compared to the counterfactual without such investments.

Our paper contributes to the literature on local labor market effects of trade shocks, including Topalova (2010), Kovak (2013), Autor et al. (2013), Costa et al. (2016), Dix-Carneiro and Kovak (2017), and Caliendo et al. (2019).[5] Most of this research emphasizes import-competition effects from declining trade barriers. Because the import basket of developing countries such as India is more intensive in intermediate inputs than that of developed economies like the United States, the input cost saving channel is particularly important. We explicitly incorporate this channel in our analysis and focus on the role of international access, which has received relatively little attention in this literature. We show that the manufacturing sector expands when input tariffs fall and that these gains are larger in districts with better international access. Further, we extend our analysis beyond the manufacturing sector to examine broader labor market outcomes, including regional structural change and welfare gains.

Our paper also contributes to the recent literature on how improvements in domestic transport infrastructure affect economic activity. Donaldson and Hornbeck (2016), Donaldson (2018), Asher and Novosad (2020), Ghani et al. (2016), Asturias et al. (2019), Alder (2025), and Vogel et al. (2024) study these effects in predominantly closed-economy settings.[6] More related to our paper, Coşar and Demir (2016), Baum-Snow et al. (2020), and Xu and Yang (2021) examine how domestic transport improvements reshape regions' access to international markets. In the context of India, Van Leemput (2021) show that international access constitutes a substantial share of the overall barriers to international trade, roughly 40%. We complement this literature by showing that initial access to international markets is crucial for the distribution of gains from a major episode of external integration in the presence of a weak road network. Using counterfactual exercises, we then quantify how improvements in port connectivity, undertaken alongside falling international trade costs, reshape the spatial distribution of welfare and factor incomes at a fine geographic scale, covering 340 districts.

[4] We use the global map for accessibility for 2000 as developed by the European Commission and the World Bank that incorporates several spatial datasets on topology, terrain, environmental, and political factors to calculate cost (Nelson, 2008). With the counterfactual road network, travel time for the farthest district from ports decreases from 21 to 10.8 hours.

[5] Redding (2022) provides a comprehensive review of the literature on the local labor market effects of trade.

[6] Refer to Redding (2022) for an insightful discussion on the effects of transport infrastructure.

Our paper is also related to the growing literature on quantitative spatial models. We build on the insights of Sotelo (2020), Fajgelbaum and Redding (2022), and Heblich et al. (2024). Fajgelbaum and Redding (2022) provide micro-foundations for the spatial Balassa–Samuelson effect and quantify aggregate welfare gains from Argentina's external and internal integration in the late nineteenth century, while Heblich et al. (2024) show that, in response to the Grain Invasion in Britain, regions with higher wheat suitability experience stronger structural change away from agriculture and larger welfare gains. We extend this class of models to study how international access shapes the distribution of welfare gains across districts. We introduce two new channels. First, we allow tradable intermediate inputs to enter the production of manufacturing output but not agricultural output. Second, we estimate that manufacturing intermediates are more sensitive to time in transit than manufacturing output itself. Other notable contributions in this literature include Peters (2022), Eckert and Peters (2022), Farrokhi and Pellegrina (2023), Coşar and Fajgelbaum (2016), Donaldson and Hornbeck (2016), and Allen and Arkolakis (2014).[7]

Our paper also relates to the literature on trade and structural change, including Matsuyama (1992, 2009), Bustos et al. (2016), Cravino and Sotelo (2019), and Teignier (2018). This literature shows that greater openness to trade affects structural change by promoting specialization and by shaping how economy-wide forces, such as productivity changes, translate into structural change.[8] More closely related to our work are Fajgelbaum and Redding (2022) and Heblich et al. (2024), who examine how trade shapes structural change in historical spatial settings. We add to this literature by showing that international access is an important margin through which trade generates spatially uneven structural change *within* the tradable sectors of a large modern economy such as India.

Lastly, our paper contributes to the extensive literature on India's trade liberalization. Goldberg et al. (2010), Topalova and Khandelwal (2011), Goldberg et al. (2009), and Nataraj (2011) study firm-level outcomes such as productivity, product scope, and R&D, while Topalova (2010), Hasan et al. (2012), and Edmonds et al. (2010) examine local labor market effects on poverty, unemployment, and human capital. We add a spatial dimension to this literature by focusing on variation in districts' international access. Empirically, we assemble a *district*$\times$*industry* panel dating back to 1987 and study a novel set of outcomes, including total manufacturing output, average firm size, number of firms, sectoral labor shares, and district to district migration flows. Our theoretical framework rationalizes these empirical patterns and highlights the role of international access in shaping the spatial distribution of welfare gains.

The remainder of this paper is organized as follows: Section 2 provides background about India's trade liberalization and internal geography. Section 3 describes the data used in the analysis. Section 4 presents the empirical evidence. Section 5 outlines the quantitative model and discusses the results of our counterfactuals. Section 6 concludes.

[7] See Redding and Rossi-Hansberg (2017) for a comprehensive review of quantitative spatial models.
[8] Alessandria et al. (2023) provide a comprehensive review of the literature on trade and structural change.

2 Background

2.1 India's Trade Liberalization

Prior to 1991, India's trade regime was highly restrictive and centered on an import substitution led strategy of industrialization. The country had some of the highest tariff and non-tariff barriers in Asia, with an average nominal tariff of 95% and 87% of products subject to quantitative restrictions in 1987-88. In the late 1980s, a combination of external and internal crises paved the way for trade liberalization. The Gulf War in 1990 led to a sharp increase in oil prices and a decline in remittances from Indian workers abroad. At the same time, political instability triggered capital outflows, resulting in a severe balance-of-payments crisis. By June 1991, India's foreign exchange reserves were sufficient to cover only two weeks of imports.

To avoid defaulting on its balance-of-payments obligations, India sought financial assistance from the International Monetary Fund (IMF), which was granted conditional on the implementation of macroeconomic reforms. Central to these reforms was a complete overhaul of trade policy. The reforms, both unexpected and externally mandated, required a substantial reduction in both the level and dispersion of tariffs and the removal of quantitative restrictions. Import licensing for nearly all intermediate inputs and capital goods was abolished. As shown in Figure 1, average tariffs declined by roughly 50 pp within the first five years of the reforms.[9] As an immediate consequence, overall imports rose by 130% between 1987 and 2000. This growth was mostly driven by imports of intermediate inputs in the manufacturing sector, which grew by 227% over the same period (Goldberg et al., 2010).

Figure 1: Average Tariffs in India (1987-2004)



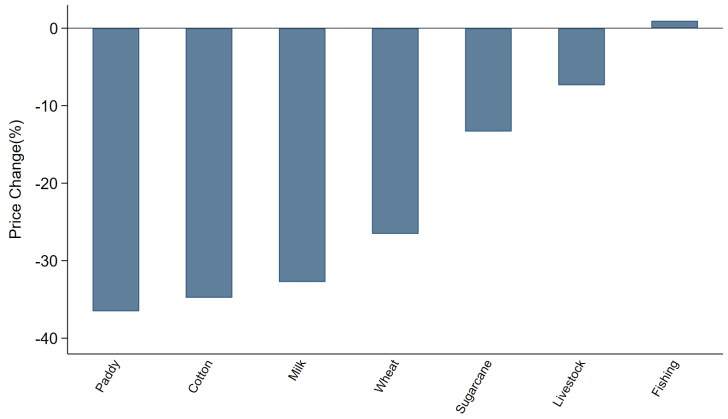
Note: The figure shows the decline in average tariffs 1987-2004.

Trade reforms were implemented in agriculture as well. While quantitative restrictions on agricultural imports were maintained until 1997, tariff barriers adjusted more quickly. With the exception of cereals and oilseeds, agricultural products experienced a sharp decline in average

[9] Refer to Topalova (2010) for a detailed discussion of the liberalization reforms.

tariffs from 90 to 20 percentage points between 1987 and 2000 (Topalova, 2010). On the export side, reforms accelerated substantially after 1994, with the elimination of export quotas, increased credit availability for exporters, and income tax exemptions on export profits (Gulati and Mullen, 2003). Throughout 1987–2004, India remained a net exporter of food. However, greater integration of Indian agriculture into world markets also increased its exposure to the decline in world agricultural prices during the 1990s. Figure 2 shows the change in average world commodity prices for India's major agricultural products between 1975–89 and 1990–2004.[10] The figure illustrates that the liberalization of the Indian economy coincided with increased international competition in agriculture, as world prices for major Indian crops such as paddy and wheat fell by over 25%, potentially affecting the domestic profitability of agriculture.[11] Data from the 1999-00 Economic Survey of India document a steep decline in the growth rate of food grains production from 3.54% over 1979-89 to 1.80% over 1989-99.[12]

Figure 2: Change in World Prices for India's Major Agriculture Commodities



Note: The figure shows the change in average world agriculture prices between 1975-89 to 1990-04 for India's major agriculture commodities.

2.2 Major International Gateways

Despite sharing extensive land borders, India's trade with its neighboring countries has historically remained low. In 2000, trade with neighbors accounted for only 1.8% of India's total trade by value.[13] Instead, India's international trade is predominantly maritime, with seaborne trade accounting for nearly 65% of total trade value. The country's long coastline is served by a network

[10]The commodity price data, except milk, are taken from the publicly available World Bank Pink Sheet on commodity prices. We use real annual price data. Since dairy products are not included in this dataset, we use EU prices for raw milk from the CLAL database. We compute the real average price across Belgium, Germany, Denmark, Ireland, Italy, Spain, and the Netherlands, for which data are available from 1977. The EU is the largest exporter of milk and dairy products, and its prices are therefore representative of world dairy prices.

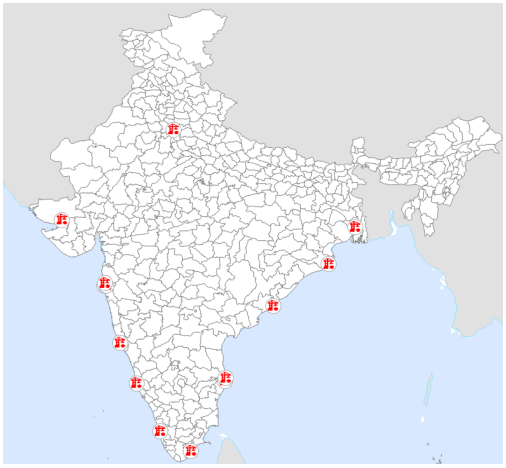
[11]Average unit value of cereal exports, a major Indian agricultural export, fell from 419 $/tonne in 1990-94 to 324 $/tonne in 1995-98 (FAO Report).

[12]Specifically, growth rate of rice slowed from 4.29 to 1.6%, wheat from 4.24 to 3.62%, and pulses from 2.78 to 1.19%. Data available here.

[13]Based on trade data from UNCOMTRADE. Neighbors include Afghanistan, Bangladesh, Bhutan, Myanmar, Nepal, and Pakistan. China is excluded as trade between India-China largely occurs via sea.

of 12 major and 185 minor ports, with the 12 major ports handling 87.5% of all seaborne trade by weight (Basic Ports Statistics, 2001–02). Two major airports – Delhi and Mumbai – are also of particular importance, together accounting for 71% of India's total airborne freight by weight (Airport Authority of India, 1998–99). As shown in Figure 3, there is substantial heterogeneity across regions in access to these 12 major ports and two airports.

Figure 3: Major Sea and Air Ports in India



Note: The figure shows the 12 major seaports and two major airports in India, overlaid on 1987 district boundaries.

2.3 Road Infrastructure

Until the early 2000s, India's road infrastructure was particularly poor. Traveling 1,500 km by truck typically took four to five days, around five times longer than in the United States (Asturias et al., 2019). The economic losses from congestion and poor road quality were substantial, estimated at INR 120–300 billion (approximately USD 2.6–6.5 billion) per year (Mohan et al., 1996). Despite an extensive railway network, roads carried 61.3% of total freight in 2000–01, with an even higher share for non-bulk commodities (GoI). For distances up to 500 km, road transport was more cost-effective than rail (Sahu et al., 2022). Given that 95% of Indian districts lie within 600 km of a major port, the road network plays a crucial role in connecting inland regions to international markets.

Large-scale investment in highway infrastructure began only in late 2001 with the launch of the Golden Quadrilateral (GQ) project. Designed to improve connectivity among the four major Indian cities - Delhi, Mumbai, Chennai, and Kolkata - the GQ became fully operational only by 2006.[14] Consequently, the period from 1987 to 2004 provides an ideal setting to examine how travel costs to ports shape the distributional consequences of trade liberalization.

[14]Other major initiatives, such as the North–South and East–West corridors, had only about 10% of construction completed by 2006.

3 Data

The spatial unit of analysis is the district, an administrative unit below the state level that is responsible for governance, revenue collection, and law and order. Because district boundaries change over time, we harmonize them to their 1987 configuration, yielding a sample of 340 districts.[15] We next describe the primary data sources used in the analysis.

International Access: The key spatial variable in this analysis is the proximity of a district to its nearest major seaport or airport, which we define as *International Access*. We compute international access as an iceberg trade cost and proceed as follows.[16] First, we digitize the complete road network of India for 1996, as published by the Government of India.[17] This allows us to identify five types of roads: National and State Highways, All Weather Motorable Roads, Fair Weather Motorable Roads, Roads Where Delays May Occur, and Other Roads. Panels (A) and (B) of Figure 4 show the placement of highways and all weather motorable roads in 1996, respectively.[18] Next, consistent with the literature and available reports, we assign an average speed to each road type: 30 mph for highways, 20 mph for all weather motorable roads, and 10 mph for all other road types.[19]

Next, using the geocodes of the 12 major seaports and 2 airports from publicly available government websites, we compute the shortest travel time between each district centroid and its nearest port using Dijkstra's algorithm.[20] The resulting travel times are summarized in Table 1. For the 340 districts in our sample, the average and median travel times are 9.25 hours and 8.88 hours, respectively. The inter-quartile range is 7.4 hours, and the most remote district is 21 hours away from the nearest port. Figure 5 shows the distribution of districts by travel time to the nearest international gate. Together, Table 1 and Figure 5 document substantial variation across districts in access to ports. The advantage of using the complete road network over straight line distance is illustrated in Figure A.2. While both *Bharuch* and *Aurangabad* (district centroids) are at roughly the same linear distance from *Mumbai* (port), *Bharuch*, which lies closer to the highway network, is 2.6 hours closer in terms of travel time.

Finally, we use these travel times to compute *International Access* in the form of iceberg trade costs as follows:

$$Trade\,Cost_{od} = \gamma(Travel\,Time_{od})^{\rho}$$

[15] We exclude districts from eight North Eastern states, Jammu & Kashmir, and the Islands due to inadequate sample coverage in the administrative datasets used in this paper.

[16] In the context of international access, iceberg trade cost measures the value of goods lost (in percent) while moving between a district center and its nearest port.

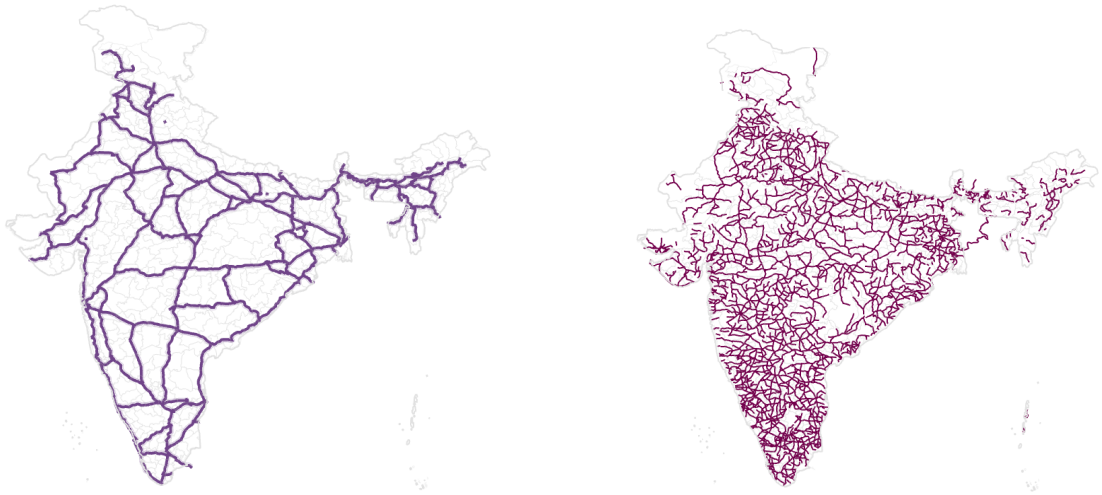
[17] We thank Treb Allen for sharing the older raw road maps which are no longer publicly available. We also thank CyberSwift, a digital mapping company, for their assistance with digitization.

[18] We use the 1991 district shapefile of India available on IPUMS International.

[19] The World Bank Report (2002) estimates an average speed of 22 to 25 mph on highways in India. Allen and Atkin (2022) and Vogel et al. (2024) assign speeds of 40 mph for highways and 20 mph for the lowest quality roads. The results of this paper are robust to changing the average speed on each road by ± 5 mph.

[20] Refer to Appendix A.1 for details on Dijkstra's algorithm.

Figure 4: Digitized Highways and All-Weather Motorable Roads



(a) Highways **(b)** All-Weather Motorable

Note: The figure shows the two major road categories from India's digitized 1996 road network.

Table 1: Summary Statistics - Travel Time and International Access

	Mean	S.D.	25th pctile	Median	75th pctile	Min	Max
Travel Time (in hours)	9.25	4.81	5.39	8.88	12.79	0.11	21.05
Trade Cost	0.15	0.04	0.12	0.15	0.18	0.02	0.22
# Districts	340						

Note: The table reports summary statistics for travel time and trade cost from district's centroid to its nearest major port.

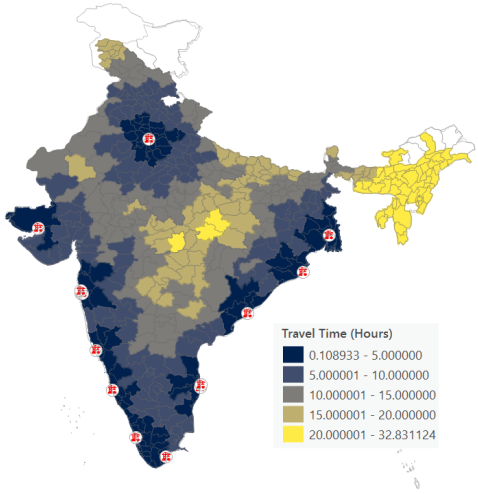
where o is the origin (district centroid) and d is the destination (nearest port). In our empirical analysis, we set $\rho = 0.45$ (Roberts et al., 2012) and calibrate γ such that the median trade cost between any two district centroids is 0.25 in the 1996 road network (Alder, 2025).[21,22] We summarize the resulting trade costs in Table 1. The median trade cost is 0.15, which implies that a travel time of 8.88 hours to a port results in a loss of 15% of the value shipped. The closest district to a port experiences a loss of 2%, whereas the most remote district loses 22%.

Annual Survey of Industries: We construct a novel *district*×*industry* panel on manufacturing outcomes for 1987 to 1997 using establishment level data from the Annual Survey of Industries (ASI), conducted by the Ministry of Statistics and Programme Implementation (MoSPI), Government of India. The ASI is the primary data source on organized manufacturing in India. It covers all large establishments (at least 100 workers) and a representative sample of smaller establish-

[21] ρ is a concavity parameter that captures economies of scale in transport cost over travel time. The value of 0.45 is estimated by Roberts et al. (2012) for the urban sector in the context of China. Results are robust to using $\rho = 0.8$ as in Alder (2025); Baum-Snow et al. (2020); Barnwal et al. (2024). They are also robust to a range of values of γ calibrated such that the median trade cost lies between 0.20 and 0.30.

[22] In our model, we allow ρ to differ across manufacturing sector output and intermediates, and estimate ρ at 0.63 for intermediates. We show that our main empirical findings are robust to using separate measures of trade cost for manufacturing output and inputs based on differential time sensitivity.

Figure 5: Travel Time to a Nearest Port across Districts



Note: The figure depicts the distribution of travel times to the nearest major port across all Indian districts.

ments, drawn with probability of roughly one third.[23] We use annual data on sales at the establishment level to construct measures of total output and average firm size for each available district-industry pair. A key challenge in creating this panel is the lack of consistent geographic identifiers in the ASI. The survey reports a district code for each establishment but withholds district names, which makes it difficult to track districts over time, especially as boundaries change. We therefore draw on all available publications on districts released by MoSPI for its other flagship surveys to augment the ASI with a consistent district identifier from 1987 to 2004.[24] Our final data set consists of 340 districts at 1987 district boundaries and 100 four digit industries under the 1998 National Industry Classification (NIC).

Economic Census: Because a large share of India's labor force works in the unorganized sector, we use the Economic Census (EC) to measure extensive margin outcomes, in particular the total number of firms and total employment in each industry by district pair. The EC covers all establishments in both the organized and unorganized sectors. We use three rounds, 1990, 1998, and 2005, which align with the ASI sample period.

National Sample Survey: We measure structural transformation using the Population Census of India and the National Sample Surveys (NSS). The NSS are nationally representative household surveys, conducted at roughly five year intervals, that provide individual level information on employment status, industry of work, and demographics. We use the 1987-1988, 1999-2000, and 2004-2005 NSS rounds.[25] Using the sample weights provided by the NSS, we aggregate in-

[23] Sampling for smaller firms in the ASI is stratified at the state by industry level. If smaller firms account for a large share of output, this can generate measurement error at the district by industry level. Larger firms, which are surveyed each year and are unaffected by the sampling strategy, account for the bulk of output. For the years prior to 1998 to 1999, large firms (with weight one) account for around 90 percent of total output within a state on average, so sampling is not a significant concern. The share of large firms falls after 1998 to 1999, and to avoid sampling problems we restrict our use of ASI data to 1997 to 1998.

[24] Refer to Appendix A.1 for details of this exercise.

[25] We exclude the 1993-1994 NSS round because it does not report district level identifiers.

dividual records to construct district level employment shares in agriculture, manufacturing, and services. We supplement these data with the 1991 Population Census, which reports district level employment by industry.

Tariffs: We construct a HS6 product level tariff series for 1987 to 2005 by combining two sources. For 1987 to 2001 we use the tariff data from Topalova (2010), and for 2002 to 2005 we use the UNCTAD TRAINS database hosted on the World Integrated Trade Solutions (WITS) portal. The combined dataset contains tariffs for 5,045 HS6 products. Using product concordances from WITS, we first map HS6 product codes to ISIC Revision 3, which maps one to one to India's NIC 1998 industry codes. We then compute four digit NIC 1998 industry level *output tariffs* as simple averages of the corresponding product level tariffs. Finally, we combine these industry tariffs with India's Input Output Matrix for 1993 to 1994 to derive *input tariffs*.[26]

Migration: We use confidential data on district to district migration flows from the 2001 Population Census of India.[27] For each origin and destination district pair, the data report the total number of migrants between 1991 and 2001, disaggregated by gender. We restrict the sample to men, as most inter district female migration is driven by marriage.

4 Empirical Evidence

4.1 Manufacturing Activity

To motivate our analysis, we begin by examining the correlation between manufacturing output growth and international access in Figure 6. In Panel (A), we show that in the pre-liberalization period (1987–1991), when India was largely a closed economy, output growth exhibits no systematic relationship with international access. In Panel (B), however, following trade liberalization (1991–1997), we observe a strong negative correlation between trade costs and manufacturing output growth, indicating that proximity to ports becomes an important determinant of manufacturing growth.

The effects of trade liberalization on overall manufacturing activity in a district are *apriori* ambiguous. Lower output tariffs increase import competition, which can adversely affect smaller or less productive firms, potentially leading to exit (Melitz, 2003). The direct negative competition effects can be counteracted by pro-competitive effects by inducing efficiency gains. On the other hand, access to cheaper or better intermediate inputs might enable firms to expand and raise productivity (Topalova and Khandelwal, 2011). The overall impact on manufacturing reflects the balance of these forces and, crucially, depends on a district's international access. High trade costs to the ports can attenuate both the positive and negative effects of liberalization.

[26] Input tariffs, τ^{ip} are computed as follows: $\tau^{ip}_{jt} = \sum_i \alpha_{ij}\tau^{op}_{it}$. Weights α_i are the share of input i in the total value of industry j's output and τ^{op} is the tariff on input i in year t. Weights α are taken from India's Input-Output Table (1993-94) after mapping inputs to 4-digit NIC 1998 industry codes.

[27] We are grateful to Devaki Ghose and Kerem Coşar for generously sharing the data. The data were originally obtained through a special agreement with the Government of India.

Figure 6: Manufacturing Output Growth and International Access



Note: The figure shows the correlation between international access and manufacturing output growth (%) in the pre- and post-liberalization period. Top and bottom 2.5% sample is trimmed in both periods to remove outliers. Regressions are weighted by the start of the period (log) real output.

4.1.1 Empirical Strategy

To capture these different channels, our empirical strategy exploits the district-level variation in international access to estimate the effects of input and output tariff declines on industry-level manufacturing outcomes:

$$y_{jdt} = \beta_0 + \beta_1 \tau_{jt}^{ip} \times Trade\,Cost_d + \beta_2 \tau_{jt}^{op} \times Trade\,Cost_d + \beta_3 \tau_{jt}^{ip} + \beta_4 \tau_{jt}^{op} + \beta_5 DMA_{dt} + X_{jd,pre} \cdot \delta_t + \lambda_{jd} + \delta_t + \epsilon_{jdt} \quad (1)$$

Here, y is the outcome variable for industry j in district d in year t. We look at four key outcomes – total output, average firm size, total number of firms, and total employment in an industry-district. τ_{jt}^{ip} and τ_{jt}^{op} denote the input and output tariffs faced by industry j in year t, respectively. *Trade Cost* measures the iceberg trade cost of moving goods between the district's centroid and its nearest port, using the 1996 road network. DMA_{dt} captures the domestic market access of each district to account for the possibility that districts closer to ports may also have better access to domestic markets, which can boost manufacturing activity.[28] $X_{jd,pre}$ contains pre-liberalization characteristics of industry j in district d, such as the initial level of y and the share of industry j's output in d's total output. We interact these with year fixed effects to allow each industry-district pair to follow a distinct growth path conditional on its initial characteristics. Year fixed effects, δ_t, control for time-varying unobservables, while industry-district fixed effects, λ_{jd}, absorb time-invariant industry-district characteristics. Standard errors are two-way clustered at the 4-digit industry-year and district levels.

The coefficients of interest, β_1 and β_2, capture the role of international access in explaining cross-district differences in changes in y arising from declining input and output tariffs within an industry, respectively. A positive β_1 (or β_2) implies that, for a given decline in input (or output) tariffs, the resulting change in y within an industry is larger in districts with lower trade costs (i.e.,

[28] We follow Donaldson and Hornbeck (2016) to construct *DMA*. Refer to Appendix A.1 for details on the construction.

better international access).

Identification: Causal identification of β_1 (and β_2) comes from comparing two districts within the same 4-digit industry that have similar pre-liberalization outcome levels and face the same input (or output) tariff changes, but differ in their degree of international access. Since the comparison is made within an industry, endogeneity of tariff reductions does not pose a concern for identification. Irrespective, a large literature on India's trade liberalization provides strong evidence that tariff changes across industries were exogenous, at least through the late 1990s (Varshney, 1998; Topalova and Khandelwal, 2011; Goyal, 1996).

A potential threat to identification could arise if major ports were constructed after 1991, particularly in districts that were already better positioned for future growth. However, this concern is mitigated because almost all major ports in India were established either during British rule to facilitate trade or in the years immediately following India's independence in 1947. The only exception is the Kamarajar Port, built in the Chennai district in 2002. Since it is located only 18 km from the pre-existing Chennai Port (established in 1881), our measure of international access is not meaningfully affected. As discussed below, our results are also robust to restricting the sample to the year 2000.

Since we use the 1996 road network, another potential concern for identification is that road improvements undertaken after 1991 may have targeted districts better positioned to benefit from trade liberalization. Such improvements could bias our results if districts closer to ports experienced gains primarily through improved infrastructure rather than greater international access. However, this does not appear to be the case as road expansion during the 1990s largely focused on upgrading lower-quality rural roads to improve rural and interior connectivity. In contrast, high-standard arterial highways connecting India's four main port cities were largely neglected until 2001 (World Bank, 2005).

The remaining potential threats arise from other changes during the sample period that may simultaneously affect manufacturing outcomes and be correlated with districts' international access. Tariff cuts were only one component of a broader liberalization package. A second major reform allowed automatic approval of Foreign Direct Investment (FDI) of up to 51% of a firm's equity in several manufacturing industries. If FDI inflows were disproportionately larger in districts with better international access, our estimates could be biased. Using data on FDI inflows from CMIE's CapEx database, we find that FDI was highly spatially concentrated: 55% of all FDI projects during 1993–2008 were located in just 11 districts.[29] Unsurprisingly, these were among the most urban districts in the country at the time, and several – such as Mumbai, Delhi, Bangalore, Chennai, and Kolkata – are also port districts. We show that our results remain unchanged when these high FDI-receiving districts are excluded from the sample. This further addresses the concern that the benefits of liberalization may have accrued disproportionately to urban districts that also happen to be closer to ports.

[29] These districts are Delhi, Mumbai, Bangalore, Kolkata, Chennai, Hyderabad, Pune, Gurgaon, Ghaziabad, Chengalpattu, and Thane.

Finally, a potential threat to identification is India's ambitious National Highway Development Program (NHDP), approved in 1998. Under the NHDP, India initiated construction of the Golden Quadrilateral (GQ), a network of high-speed highways connecting the four major cities – Delhi, Mumbai, Chennai, and Kolkata. Prior work has documented significant manufacturing gains for districts located along the GQ network (Asturias et al., 2019; Ghani et al., 2016). As discussed in Section 2, construction of the GQ began in late 2001 and was largely completed only by 2006. Because a substantial portion of the GQ passes through districts located closer to major ports (see Figure A.3), our estimates could be biased if these districts experienced growth primarily due to improved road connectivity. To address this concern, we show that our results remain robust when restricting the sample to the pre-2001 period.

4.1.2 Total Output

Our first outcome is the total manufacturing output for an industry-district pair. We use annual data from the ASI from 1991-92 to 1997-98.[30] Results are presented in Table 2. Column (1) reports estimates from Equation 1. The coefficient β_1 is positive and statistically significant at 1%, indicating that for a given decline in input tariffs, output growth within an industry is larger in districts with better international access (i.e., lower trade cost). The magnitude implies that for a 10 pp decline in τ^{ip}, reducing trade costs by the median difference of 0.13 leads to 16.69% (= 12.842 $\times$ (0.02–0.15) $\times$ –10) higher output. By contrast, the estimate of β_2 is small and statistically insignificant, suggesting that the gains from lower input tariffs far outweigh any potential losses from increased foreign competition.

In Column (2), we exclude the 11 high FDI-receiving districts from the sample to ensure our results are not driven by the FDI reform. In Column (3), we include state-year fixed effects to account for time-varying state-level policies, such as changes in labor regulations, that could influence manufacturing outcomes. In Column (4), we add 4-digit industry-year fixed effects to absorb industry-specific shocks, such as de-licensing. This fully absorbs the main effects of τ^{ip} and τ^{op}, but our coefficients of interest, β_1 and β_2, remain identified. Across all specifications in Table 2, the estimates of β_1 remain positive, statistically significant, and remarkably stable. Economically, for a 10 pp decline in input tariffs, reducing trade costs by the median amount increases output by roughly 17-18%. Between 1990-98, τ^{ip} declined on average by 21 pp. Our estimates imply that, for an industry experiencing the average input tariff reduction, moving to a district whose international access is better by the median amount is associated with 35 to 38% higher manufacturing output growth relative to an industry facing no tariff change. [31]

In Column (5), we examine average firm output within an industry–district pair and find that, in response to declining input tariffs, the average firm becomes larger in districts with better international access. For a 10 pp decline in input tariffs, a median reduction in trade cost leads to an

[30] We deflate the value of output using the industry-level Wholesale Price Index at 1993-94 constant prices.

[31] We also test robustness using the "Effective Rate of Protection" (Corden, 1966) as an alternative measure of trade liberalization, also used in the Indian context by Topalova and Khandelwal (2011). Table A.1 shows that the main findings remain unchanged.

additional 11.65% increase in the average firm output. Lastly, to rule out pre-existing trends, we estimate Column (3) and Column (5) of Table 2 for the pre-liberalization period (1987–1991). The results show no significant pre-trends.

Table 2: The Effect of International Access on Manufacturing Output

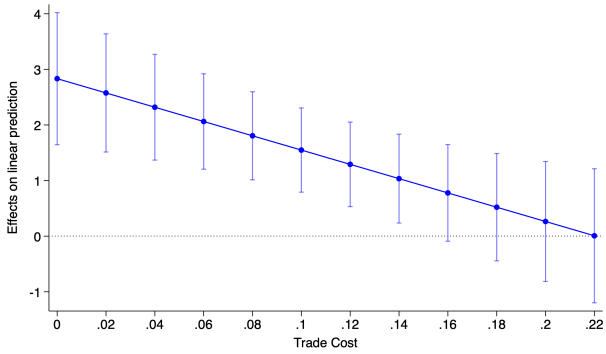
	(log) Output				Avg. Firm Output	Placebo (1987-91)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
τ^{ip}	-2.833***	-2.990***	-2.782***		-2.546***	1.787	-0.887
	(0.606)	(0.731)	(0.666)		(0.745)	(1.733)	(1.266)
τ^{op}	-0.125	-0.072	-0.114		-0.114	-0.455	0.090
	(0.149)	(0.193)	(0.154)		(0.168)	(0.281)	(0.265)
$\tau^{ip} \times$ Trade Cost	12.842***	13.768***	12.867**	12.787***	8.964**	-19.360	10.747
	(4.310)	(5.235)	(5.104)	(4.388)	(4.385)	(12.607)	(8.602)
$\tau^{op} \times$ Trade Cost	1.103	0.890	1.155	1.645	1.078	1.399	-0.992
	(1.193)	(1.432)	(1.207)	(1.242)	(1.248)	(1.917)	(1.879)
Observations	41,332	37,384	41,332	41,327	41,332	23,696	23,696
District-Industry FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	N	Y	Y	Y
State-Year FE	N	N	Y	N	Y	Y	Y
Industry-Year FE	N	N	N	Y	N	N	N

Note: Dependent variables are in logs at the industry–district level. *Trade Cost* is the iceberg trade cost between a district's centroid and its nearest port. The outcome variable in Columns (1)–(4) is total manufacturing output (real sales). Column (2) drops the 11 high-FDI districts. Column (3) adds state-year fixed effects. Column (4) adds 4-digit industry–year fixed effects. The outcome variable in Column (5) is average firm-level output. Columns (6)–(7) report placebo estimates from 1987–91 for total output and average firm-level output, respectively. Standard errors are two-way clustered at the 4-digit industry–year and district levels. The sample covers 97 4-digit industries across 333 districts. ***, **, * denote significance at the 1%, 5%, and 10% levels.

Insofar, we have established that districts with better international access experience relatively larger output gains. However, these relative effects of input tariff declines could simply reflect a reallocation of output, with internationally better-connected districts gaining at the expense of more remote ones. Such a pattern would imply a redistribution of the existing economic pie rather than an expansion of it. To assess the aggregate implications of falling input tariffs, Figure 7 plots the marginal effect of a decline in input tariffs (estimated from equation 1) across the range of trade cost values. The figure shows that the effect on manufacturing output is positive for the entire range of trade costs. Moreover, the effect is statistically significant at the 5% up to a trade cost of 0.14, which encompasses roughly 50% of districts. These results imply that, although unequal, declining input tariffs generate aggregate gains in manufacturing output rather than merely redistributing output toward better-connected districts. Additionally, Appendix A.2 presents semi-parametric estimates showing that the effects increase monotonically with international access. Appendix Section A.2.3 reproduces Table 2 with a different value of economies of time for manufacturing inputs and output. Our results remain robust to this alternative calculation of trade costs.

To further address concerns that our results may be driven by specific years or pre-existing

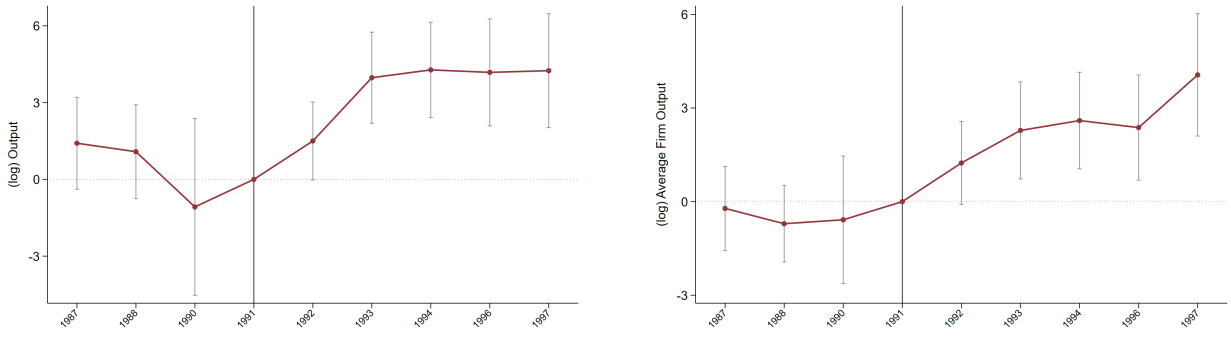
Figure 7: Total Effect of Declining Input Tariffs across Levels of Trade Cost



Note: The figure shows the marginal effect of input-tariff reductions at different levels of trade costs. Each point reports the estimated effect of a 10% decline in input tariffs for an industry–district at a given trade-cost level, relative to an industry experiencing no tariff decline. 95% confidence intervals are shown.

trends, we present the dynamic effects of trade liberalization in Figure 8. Since most industries experienced sharp tariff reductions in 1992, we take 1991 as the base year. Controlling for time trends in pre-liberalization district-industry characteristics, as well as industry-year, state-year, and district-industry fixed effects, Panel (a) shows that districts with lower trade costs experienced additional gains in manufacturing output beginning in 1993, with these effects remaining stable through 1997-98. Panel (b) reports similar patterns for average firm output, with significant positive effects emerging as early as 1993 and persisting throughout the sample period, without evidence of any significant differential pre-trends.

Figure 8: Dynamic Effects



(a) Total Output **(b)** Average Firm Output

Note: The figure reports event-study estimates of the effect of international access. The dependent variable is (log) real industry–district sales in Panel (a) and (log) average firm sales in Panel (b). Each point shows the additional effect in the least trade cost district relative to the median trade cost district, in year t relative to 1991, with 95% confidence intervals. Controls include 1991 industry–district size and the industry's share in district output in 1991, both interacted with year fixed effects. District–industry, state–year and industry–year fixed effects are included. Standard errors are two-way clustered at the industry–year and district levels.

4.1.3 Total Number of Firms and Labor

Next, we turn to manufacturing activity on the extensive margin, specifically the total number of firms and the labor employed in an industry-district pair. For this analysis, we use the Economic Census (EC) of India. Unlike the ASI, which covers only formal-sector firms, the EC surveys both formal and informal establishments. The inclusion of informal firms is crucial, as even though they account for under 20% of total manufacturing output, they employ over 80% of the manufacturing workforce in India (World Bank Blogs, 2015). Since EC data are available only for 1990, 1998, and 2005 during our period of interest, we use 1990 and 2005 to estimate the long-run effects of trade liberalization.[32]

We present the results for (log) number of firms in Panel (A) of Table 3. Column (1) shows the results from Equation 1. The interaction between input tariffs and trade costs is positive and significant, indicating that falling input tariffs increase the number of firms differentially across jd pairs, whereas output tariffs have no significant effect. Specifically, for a 10 pp decline in input tariffs, a median reduction in trade costs leads to a 14.97% increase in the number of firms within an industry. In Column (2), we drop the 11 high FDI-receiving districts. In Column (3), we exclude year 2005 from our sample to avoid confounding from the GQ and instead use the 1998 EC round. In Column (4), we include state–year fixed effects, and in Column (5), 4-digit industry–year fixed effects to account for time-varying state-level and industry-level policies that may be correlated with a district's international access. International access remains important for the distribution of extensive-margin gains from lower input tariffs. In particular, given an average decline in τ^{ip} of 24 pp between 1990-2005, our estimates imply a 25-43% relative increase in the total number of firms in districts with lower trade costs by the median value.

Panel (B) of Table 3 documents very similar patterns for total manufacturing employment. For a given decline in input tariffs, the increase in total employment within an industry is larger in districts with better international access. For a 10 pp reduction in input tariffs, the point estimate in Column (1) implies a 31.27% additional increase in manufacturing employment for a median decline in trade costs. Columns (2)–(5) show that these results are robust to the alternative specifications discussed for total firms. These findings indicate that the reallocation of labor within manufacturing toward industries experiencing larger input tariff reductions is faster in districts with lower trade costs. As before, output tariffs do not have any significant effects on total employment, either overall or differentially. Thus, in contrast to a lot of the literature that emphasizes the negative effects of import competition on local labor markets, we find that positive effects of input tariff declines dominate. Overall, given an average decline of 24 pp in τ^{ip} during our sample period, reducing trade costs by the median value leads to a 65-81% increase in total employment in an industry facing this average tariff decline, relative to an industry with no change in τ^{ip}.[33,34]

[32] We make one modification to Equation 1, in $X_{jd,pre}$ we include industry j's share in district d's manufacturing employment instead of output, to align with the outcomes available in the EC.

[33] Refer to Appendix A.2 for semi-parametric estimates on the effects of international access on number of firms and employment.

[34] See Table A.1 for results using Effective Rate of Protection as a measure of trade liberalization.

Table 3: The Effect of International Access on Number of Firms and Employment

	Baseline	Drop FDI	Pre-GQ	State-Year FE	Industry-Year FE
	(1)	(2)	(3)	(4)	(5)
Panel A: (log) Number of firms					
$\tau^{ip} \times$ Trade Cost	11.513***	10.677**	8.102**	10.302**	13.792***
	(4.020)	(4.246)	(3.556)	(4.540)	(3.963)
$\tau^{op} \times$ Trade Cost	1.399	0.857	0.954	1.425	1.710
	(1.263)	(1.379)	(1.103)	(1.262)	(1.151)
Panel B: (log) Employment					
$\tau^{ip} \times$ Trade Cost	24.052***	20.943***	21.281***	21.467***	26.075***
	(4.371)	(4.630)	(4.700)	(4.956)	(4.482)
$\tau^{op} \times$ Trade Cost	1.468	1.107	1.613	1.258	1.869
	(1.199)	(1.341)	(1.308)	(1.207)	(1.136)
Observations	44,278	42,774	41,936	44,278	44,278
District-Industry FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	N	N

Note: The dependent variable in Panel (A) is the (log) total number of firms in an industry–district and (log) total manufacturing employment in Panel (B). *Trade Cost* is the iceberg trade cost between the district centroid and its nearest port. Column (2) drops the 11 high-FDI districts. Column (3) restricts the sample to years up to 1998 (pre-GQ highway). Column (4) adds state–year fixed effects. Column (5) adds 4-digit industry–year fixed effects. Standard errors are two-way clustered by 4-digit industry–year and district. The sample includes 98 4-digit industries and 339 districts. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

While the EC offers several advantages over the ASI in terms of coverage, one caveat is that we cannot perform a pre-trend test, as EC data are not digitized prior to 1990. Causal interpretations of the results in this section should therefore be viewed with appropriate caution. However, the inclusion of 1990-level district–industry outcomes interacted with year fixed effects strengthens the assumption that pre-existing trends in the growth of the number of firms or total labor do not drive these results.

4.2 Structural Transformation

In this section, we examine whether international access shapes structural transformation in response to trade liberalization. Using multiple rounds of the NSS and the Population Census, we assemble data on district-level employment shares in agriculture, manufacturing, and services for the years 1987–88, 1991, 1999–2000, and 2004–05.[35] The labor force in our sample includes all individuals aged 15–59. Table 4 summarizes our key outcome variables. Over this period, the share of employment in agriculture declines, accompanied by increases in the shares of manufacturing and services.

To motivate our analysis, we present the correlation between a district's (%) change in the agricultural employment share and its trade cost in Figure 9. Panel (A) shows no significant relationship between the two in the pre-liberalization period (1987–1991). By contrast, Panel (B) shows that in the post-liberalization period (1991–2004), districts with lower trade costs (i.e., better

[35] We include non-tradable services, such as construction, in our classification of the service sector.

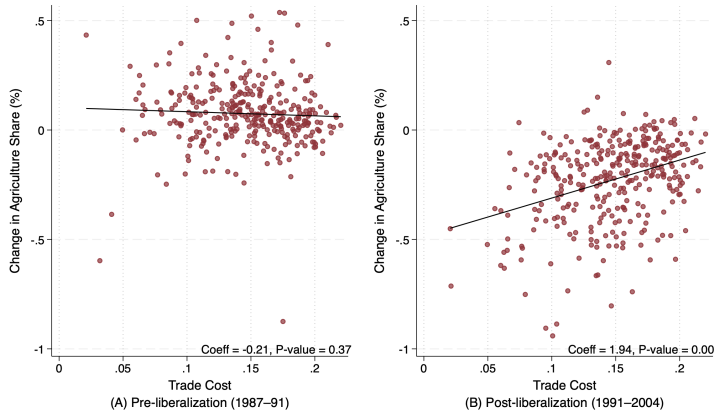
Table 4: Employment Shares Across Sectors in India

	1991	1999 & 2004
	(Pre)	(Post)
Agriculture	0.67	0.56
Manufacturing	0.11	0.13
Services	0.22	0.31
– Producer Services	0.02	0.04
– Consumer Services	0.20	0.27
#Districts	340	

Note: The table reports India's employment shares in the three broad sectors for 1991 (pre-liberalization) and for the post-liberalization period, the average of 1999–2000 and 2004–2005. Calculations based on a sample of 340 districts.

international access) experience significantly larger declines in agricultural employment shares. Most points lie below 0, indicating that the movement from agriculture to non-agricultural sectors is a nationwide phenomenon. This pattern is consistent with structural transformation being more pronounced in internationally better-connected districts, driven by greater expansion in non-agricultural sectors following trade liberalization, as discussed previously.

Figure 9: Changes in Agriculture Employment Share and International Access



Note: The figure shows the correlation between districts' international access and (%) change in agriculture employment share in the pre- and post-liberalization period. Top and bottom 1% sample is trimmed in both periods to remove outliers.

4.2.1 Empirical Strategy

To evaluate the impact of international access in facilitating structural change following India's trade liberalization, we estimate the following equation:

$$y^k_{dst} = \beta_0 + \beta^k \, Trade\, Cost_d \times Post_t + X_{d,pre}\, \delta_t + \Gamma' X_{dt} + \gamma_d + \delta_t + \epsilon_{dt} \tag{2}$$

Here, y^k is the share of employment in district d and year t in sector k, where $k \in \{agriculture, manufacturing, services\}$. *Trade Cost* measures the iceberg trade cost between a district d's centroid and its nearest port. *Post* is a dummy variable for the post-liberalization period and takes the value 1 for 1999–2000 and 2004–05, and 0 for 1991. In the baseline specification, $X_{d,pre}$ controls for the shares of employment in manufacturing and services in district d in 1991 (agriculture is

the omitted category). We interact $X_{d,pre}$ with year fixed effects to allow districts with larger pre-liberalization non-agricultural employment shares to follow different post-liberalization trends. Later, we expand $X_{d,pre}$ to include pre-liberalization district characteristics such as total population, the share of females, the share of the primary-educated population, the working-age population, and average elevation. Interacting these with year fixed effects accounts for differential trends that may arise post-liberalization due to a district's baseline economic, geographic, and demographic characteristics.[36]

X_{dt} controls for time-varying district characteristics such as tariff exposure and domestic market access. If districts near ports systematically have a larger presence of industries that experienced greater tariff cuts during trade liberalization, we might overestimate the effect of international access on structural change. To address this concern, we control for a district's tariff exposure based on its pre-liberalization industrial composition, following Topalova (2010).[37] Controlling for a district's domestic market access captures the effect of domestic connectivity on local labor demand. γ_d denotes district fixed effects and δ_t year fixed effects. We cluster standard errors at the district level.

The coefficient of interest, β^k, measures the effect of proximity to an international port on a district's employment share in sector k, on average, after trade liberalization. A negative β^k implies that trade liberalization *increases* the employment share of sector k as international access *improves* (i.e., as trade costs fall). Since manufacturing activity increased more in districts closer to international gateways than in those further away, we expect a $\beta^M < 0$ for manufacturing and $\beta^A > 0$ for agriculture. Moreover, producer services could rise as stronger growth in manufacturing activity in districts closer to ports could lead to a larger demand for complementary producer services, such as cargo handling, financial, and accounting services.

Identification: In this setup, β^k is identified by comparing the effects of trade liberalization across districts with similar observable characteristics but different international access. Threats to identification arise if factors that affect a district's sectoral employment shares are correlated with its international access. We identify three such potential threats. First, β^k would be biased if a district's exposure to trade liberalization is correlated with its international access. As discussed above, we control for districts' tariff exposure in the baseline specification. Other threats include increased FDI in districts with better international access and the launch of GQ construction in 2001, discussed in detail in Section 4.1. Both channels can lead to upward bias in the estimates. We show that the results are robust to dropping high-FDI districts from the sample (which also removes the most urban districts in India) and to restricting the post period to 1999–2000 (pre-GQ) only.

4.2.2 Results

We summarize the results from Equation 2 in Table 5. We begin with agriculture in Panel (A). Column (1) reports estimates from the baseline specification. We find that districts with better

[36] Demographic characteristics are computed from the Population Census of 1991.
[37] Construction of this measure is discussed in Appendix A.1.

international access experience a larger decline in the agricultural employment share following trade liberalization. Specifically, comparing a district at the median of the trade cost distribution (0.15) to a district with the lowest trade cost (0.02), we find a 6.12 pp (= 0.471×(0.15–0.02)×100) greater decline in the agricultural employment share in the lower cost district. In Column (2), our preferred specification, we control for the vector of pre-liberalization district characteristics interacted with year fixed effects and find an additional decline of 6.25 pp in the least-cost district relative to the median district. In Columns (3) and (4), we address threats to identification from the FDI reform and the Golden Quadrilateral construction, respectively. In Column (3), we drop the 11 high FDI-receiving districts from the sample, and in Column (4), we restrict the post period to 1999–00. The effects remain stable and statistically significant across these specifications.

Table 5: The Effect of International Access on Structural Transformation

	Baseline	District$_{91} \times t$	Drop FDI	Pre-GQ
	(1)	(2)	(3)	(4)
Panel A: Agriculture				
Trade Cost × Post	0.471***	0.481***	0.464***	0.344**
	(0.123)	(0.135)	(0.134)	(0.165)
Panel B: Manufacturing				
Trade Cost × Post	-0.236***	-0.247***	-0.217***	-0.228***
	(0.064)	(0.067)	(0.065)	(0.081)
Panel C: Services				
Trade Cost × Post	-0.234**	-0.234**	-0.247**	-0.117
	(0.097)	(0.106)	(0.106)	(0.141)
Panel D: Producer Services				
Trade Cost × Post	-0.121***	-0.130***	-0.145***	-0.135***
	(0.025)	(0.026)	(0.024)	(0.035)
Panel E: Consumer Services				
Trade Cost × Post	-0.113	-0.104	-0.102	0.019
	(0.096)	(0.103)	(0.100)	(0.138)
Observations	1,020	1,020	987	680
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y

Note: The dependent variable is the share of a district's employment in sector k. *Trade Cost* is the iceberg trade cost between the district centroid and its nearest port. *Post* equals 1 in 1999–00 and 2004–05, and 0 in 1991. Column (2) adds pre-liberalization district characteristics interacted with year fixed effects. Column (3) drops the 11 high-FDI districts. Column (4) excludes 2004–05 from the post period to remove potential GQ highway effects. The sample includes 340 districts. Standard errors are clustered at the district level. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel (B) summarizes the results for the manufacturing sector. The point estimates of β^M are negative and highly significant. Our preferred specification in Column (2) indicates that a median reduction in trade costs across districts leads to a 3.21 pp increase in the manufacturing employment share on average due to trade liberalization. Consistent with our previous estimates, stronger gains in manufacturing activity in districts closer to international gateways translate into

a greater increase in the manufacturing employment share in these districts. Across all specifications, the effect size ranges between 2.82 and 3.21 pp for a median reduction in trade costs. These effects are smaller than the decline in agricultural shares, indicating some reallocation toward services.

In Panel (C), we examine the service sector. Our preferred specification in Column (2) shows a 3.04 pp increase in the service employment share due to trade liberalization for a median reduction in trade cost across districts. While the coefficient is robust across specifications, it is insignificant in the pre-GQ period. To explore this further, we split services into producer and consumer services in Panels (D) and (E), respectively. We find that the reallocation into services is entirely driven by producer services, with effects in the range of 1.6 to 1.9 pp for a median decline in trade cost.

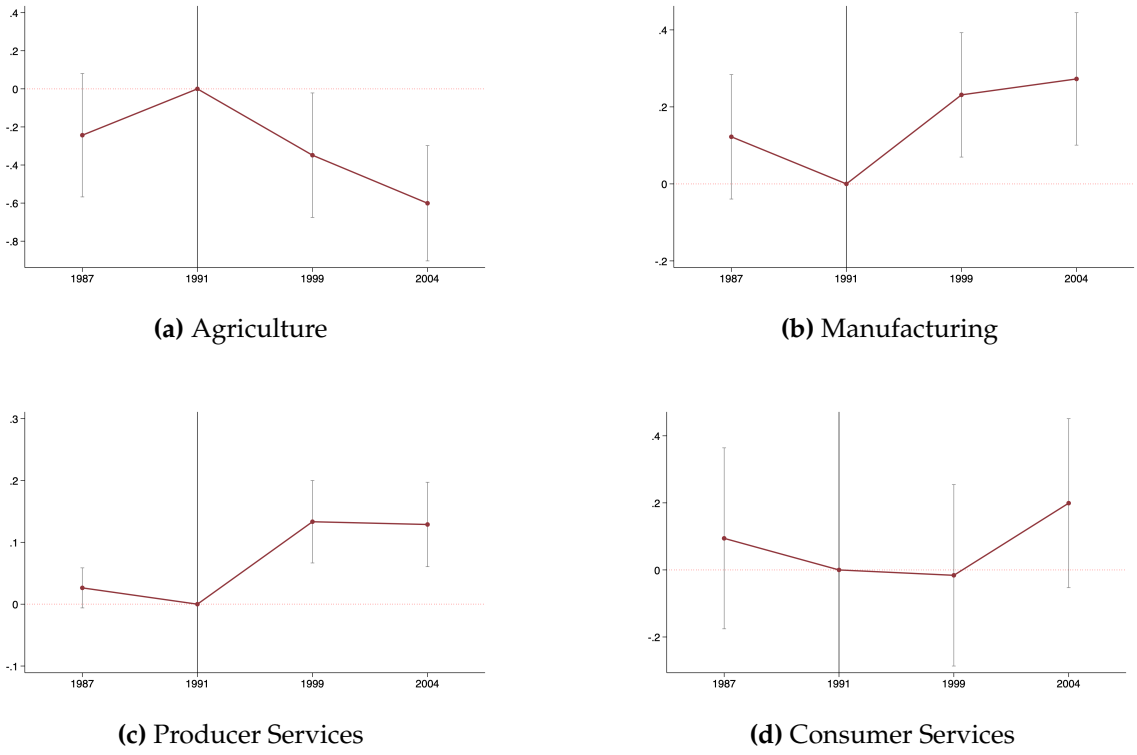
To confirm that these effects are not driven by pre-trends in employment shares, we use data from 1987–88 and estimate the specification in Column (2) of Table 5 in an event-study framework. The event-study plots in Figure 10 show that the results in Table 5 are not driven by pre-existing trends in employment shares. They also indicate that, for both agriculture and manufacturing, the effects are not concentrated in a single year but instead evolve over time. For services overall, the differential effects are driven by producer services, while consumer services do not evolve differentially with a district's international access. Overall, we find strong evidence that, following a major trade liberalization episode, better international access is an important determinant of structural transformation across the tradable sectors of the economy.

We conduct several additional robustness checks. Since many districts with better port access, such as Mumbai, Kolkata, and Delhi, were already large urban centers before trade liberalization, one could argue that these differential effects might have arisen even in the absence of trade reforms, as urban centers generally have a larger share of employment in non-agricultural sectors. To ensure our results are not driven by urbanization, we exclude districts where over 50% of the population resides in urban areas and confirm that the findings hold for predominantly rural districts. Additionally, since labor laws in India are set at the state level, β^k might be biased if states with better international access enacted more pro-manufacturing labor laws to drive growth. To address this, we include state-year fixed effects and find consistent results. Lastly, if rainfall shocks are correlated across districts closer to ports versus inland, supply-side shocks to agriculture due to rainfall could differentially affect sectoral employment shares across districts in any given year. We show that our results are robust to controlling for annual rainfall and its square.[38] Results for these additional robustness checks are shown in Table A.3.

Benchmarking the importance of international access: To understand the economic significance of our estimates for structural transformation, we conduct a simple benchmarking exercise. We categorize each district as below or above-median international access based on its trade cost to the nearest port. We then estimate Equation 2 for the agriculture sector, replacing the continuous

[38] We use Indian Meteorological Department's grid-level data on daily rainfall to compute total rainfall received by a district in a given year.

Figure 10: Event Study for the Effect of International Access on Structural Transformation



(a) Agriculture

(b) Manufacturing

(c) Producer Services

(d) Consumer Services

Note: The figure reports event-study estimates of the effects of international access on sectoral employment shares. Outcome variables are multiplied by (-1) for ease of visualization. Positive coefficients indicate an increase in the employment share of sector k in year t relative to 1991 as trade costs to access ports fall. Services are classified as producer and consumer services at the 4-digit industry level. 95% confidence intervals are shown.

trade cost measure with a binary indicator:

$$y_{dst}^{A} = \beta_0 + \beta^{A} Above\ Median_d \times Post_t + X^{d,pre}\,\delta_t + \Gamma' X_{dt} + \gamma_d + \eta_{st} + \epsilon_{dt} \tag{3}$$

Here, *Above Median* takes the value 1 if the district is above the median in the distribution of international access (i.e., has below-median trade cost), and 0 otherwise. We include the full vector of pre-liberalization district-level geographic, demographic, and economic controls, as well as the pre-liberalization level of the outcome, each interacted with year fixed effects. We also control for tariff exposure, domestic market access, and state–year fixed effects. The estimate of β^A captures the average additional change in the agricultural employment share in above-median international access districts relative to below-median districts due to trade liberalization. Since employment shares in the three sectors add up to 1, $-\beta^A$ gives the corresponding change in the non-agricultural sector.

The additional effect on the agricultural employment share is negative and statistically significant. In the data, the *additional* decline in the share of employment in agriculture from the pre- to the post-period in the above-median districts relative to the below-median districts is 2.02 pp. For a conservative assessment of economic significance, we focus on the lower bound (in absolute terms) of the 95% confidence interval for β^A, equal to 0.7 pp ($\approx$ 0.68 pp). This lower-bound ef-

Table 6: Benchmarking Exercise for Structural Change Estimates

	Agriculture
	(1)
Above Median × Post	-0.027***
	(0.010)
Observations	1,011
95% C.I.	[-0.048, -0.007]
District FE	Y
State-Year FE	Y

Note: The dependent variable is the share of a district's employment in agriculture. *Above Median* equals 1 for districts with above-median *International Access* (i.e., below-median trade cost). *Post* equals 1 in 1999–00 and 2004–05, and 0 in 1991. The sample includes 337 districts. Standard errors are clustered at the district level. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

fect explains 34.15% of the overall additional decline. It implies *at least* an additional 1.16 million people transitioning out of agriculture into non-agricultural sectors in the above-median districts. International access is therefore quantitatively important for understanding the magnitude of spatial structural change in India after the 1991 reforms.

4.3 Migration

Lastly, we examine whether international access shapes migration flows across districts. We use detailed confidential district-to-district migration data from the 2001 Census, which allows us to identify, for each destination district d and origin district o, the total number of migrants who moved between 1991 and 2001. We estimate the following equation:

$$MigFlow_{ds,os'} = \beta_0 + \beta_1 Trade\ Cost_d + \beta_2 Trade\ Cost_o + \beta_3 Travel\ Time_{do} + \alpha_1 L_{do} + X_d + X_o + \eta_{ss'} + \epsilon_{ds,os'} \quad (4)$$

where *MigFlow* is the number of migrants in destination district d in state s from origin district o in state s' who migrated between 1991 and 2001. *Trade Cost*$_d$ and *Trade Cost*$_o$ are the trade costs of the destination and origin districts, respectively. *Travel Time*$_{do}$ is the travel time (in hours) between the two districts, and L_{do} is the probability that people in districts d and o speak the same mother tongue, and thus can communicate more easily. X contains pre-liberalization district characteristics, such as (log) population and, as a robustness check, the share of the population residing in urban areas. The latter accounts for the fact that migration flows may be larger from less urban to more urban districts. $\eta_{ss'}$ is a destination state-origin state fixed effect that captures any state-pair-specific migration costs.[39] Standard errors are two-way clustered at the destination and origin district levels. We use the Poisson Pseudo Maximum Likelihood (PPML) estimator to account for zero migration flows.

The coefficients of interest are β_1 and β_2. A positive β_2 indicates that the outflow of migrants is larger from districts with high trade costs (i.e., poor international access). A negative β_1, in

[39] We drop the 6 major urban centers – Delhi, Mumbai, Kolkata, Bangalore, Hyderabad, and Chennai – as migration inflows are expected to be much larger in these districts and could bias the results. Removing these districts did not affect any of our previous results in the paper. Additionally, we restrict the sample to males because women in India predominantly migrate for reasons of marriage.

contrast, indicates that the inflow of migrants is larger in districts with low trade costs (i.e., high international access).

Table 7 presents the results. Column (1) shows the estimates from the baseline specification, and Column (2) adds a control for the urbanity of the origin and destination districts. β_2 is positive and significant at the 1% level across specifications. Similarly, β_1 is negative and significant at the 1% level. We therefore find strong evidence that migration flows are significantly larger from districts with low international access to those with high international access.

Table 7: The Effect of International Access on Inter-District Migration

	(1)	(2)
$Trade\,Cost_d$	-11.843***	-8.168***
	(2.785)	(1.869)
$Trade\,Cost_o$	7.108***	6.149***
	(1.445)	(1.296)
Observations	114,554	111,884
Origin-destination State FE	Y	Y

Note: The dependent variable is the total number of migrants from origin district o to destination district d between 1991 and 2001. Column (2) additionally controls for the share of the population residing in urban areas in both origin and destination districts. Both columns include destination–origin state fixed effects. Standard errors are two-way clustered at the destination and origin district levels. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

5 Quantitative Model

We develop a quantitative spatial model to rationalize our empirical findings on how international access shapes manufacturing gains and structural change. We use the model to study the distribution of welfare gains from India's trade liberalization and assess the role of road investments targeting port connectivity in reversing spatial inequalities.

Following Fajgelbaum and Redding (2022) and Heblich et al. (2024), we model districts, $d \in \mathcal{D}$, as a set of small open economies. The economy consists of two tradable sectors, Agriculture (A) and Manufacturing (M).[40] We make a crucial distinction between the two sectors by including tradable materials (Q_M) as inputs in the production of the manufacturing sector. Importantly, the travel time sensitivity differs across agriculture, manufacturing output and manufacturing intermediates.[41] Land is fixed across districts and workers are mobile. Each district d is endowed with a continuum of land plots, (ζ) owned by immobile landlords who decide the sector to which the land plot is put to use. Workers are endowed with one unit of labor which is supplied inelastically. Districts are net exporters of agriculture and net importers of manufacturing output and materials. They differ in sectoral productivities, amenities, trade costs by sector, and total land area.

5.1 Consumer Preferences

The utility of a worker ψ who chooses to live in district d depends on the district's amenities B_d, consumption of goods from each sector k, C_{kd}, and an idiosyncratic preference shock $b_d(\psi)$.

[40] We include producer services as part of the manufacturing sector in our estimation.

[41] Hummels and Schaur (2013) show that the cost of time is higher in perishable and intermediates.

Preferences take the following Constant Elasticity of Substitution (CES) form:

$$u_d(\psi) = B_d b_d(\psi) \left[\sum_{k \in \{A,M\}} \eta_k^{\frac{1}{\sigma}} C_{kd}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \quad , \quad \sigma \in (0,1) \tag{5}$$

where η_k governs the relative weight of sector k in the worker's utility. Location preferences for each individual are drawn independently from a Fréchet distribution given by,

$$F(b) = e^{-b^{-\chi}} \tag{6}$$

where $\chi > 1$. The scale parameter enters the model isomorphically to amenities (B_d), and hence is normalized to 1.

Given the CES preferences, the price index (P_{Cd}) in district d takes the following form:

$$P_{Cd} = \left[\sum_{k \in \{A,M\}} \eta_k P_{kd}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{7}$$

where P_{kd} is the price of the good produced by sector k.

Since the output of both sectors is traded on the world markets, the no arbitrage condition holds. Domestic price of each sector is then a function of sectoral world prices (P_k^*), tariffs on the world output (τ_k), and the sectoral trade cost of moving the goods between the world markets (i.e., ports) and the district (t_{kd}):

$$\begin{aligned} P_{kd} &= P_k^*(1 + \tau_k + t_{kd}), \quad \text{if} \quad C_{kd} > Y_{kd} \quad , \quad t_{kd} > 0 \\ P_{kd} &= \frac{P_k^*}{(1 + t_{kd})}, \quad \text{if} \quad Y_{kd} > C_{kd} \end{aligned} \tag{8}$$

where Y_{kd} is the production of sector k's good in district d.[42] Trade cost t_{kd} is given by:

$$t_{kd} = \gamma (Time_d)^{\rho_k}, \qquad \rho_M < \rho_A \tag{9}$$

where $Time$ is the hourly travel time between a district and its nearest international port, and ρ_k governs the economies of time. We allow economies of time to be stronger in manufacturing as compared to agriculture, representing the higher perishability of agriculture output. Consistent with India's production patterns in the 1990's, we model all districts as net exporters in agriculture and net importers in manufacturing output. Utility maximization gives a sector's expenditure share in total expenditure as:

$$x_{kd} = \frac{\eta_k P_{kd}^{1-\sigma}}{\sum_{f \in \{A,M\}} \eta_f P_{fd}^{1-\sigma}} \tag{10}$$

5.2 Production

We assume perfectly competitive markets and constant returns to scale in sectoral production. A key feature of the environment is that factor use differs across sectors. Agriculture combines land and labor, while manufacturing combines land, labor, and material inputs. Output per unit of

[42] We abstract away from rest of the world tariffs on Indian exports.

land, $y_{kd}(\zeta)$, can be written as:

$$\begin{aligned} y_{Ad}(\zeta) &= \nu_{Ad}\, n_{Ad}(\zeta)^{1-\alpha_A}\, z_{Ad}(\zeta)^{\alpha_A}, \quad 0 < \alpha_A < 1 \\ y_{Md}(\zeta) &= \nu_{Md}\, n_{Md}(\zeta)^{1-\alpha_M-\beta_M}\, q_{Md}^{\beta_M}\, z_{Md}(\zeta)^{\alpha_M}, \quad 0 < \alpha_M, \beta_M < 1 \end{aligned} \tag{11}$$

where $n_{kd}(\zeta)$ denotes employment per unit of land in sector $k \in \{A, M\}$, q_{Md} denotes materials per unit of land in manufacturing, and $z_{kd}(\zeta)$ captures the productivity differences across sectors.[43] α_k captures the land intensity in sector k, and we assume that agriculture is more land intensive, $\alpha_A > \alpha_M$.

Materials (Q_M) in manufacturing are tradable on the world markets. For simplicity, we assume that materials are completely imported and face a perfectly elastic foreign supply at the world price, P_q^*. We model districts as net importers of materials so their price is determined by the no arbitrage condition as:

$$P_{qd} = P_q^*(1 + \tau_q + t_{qd}), \tag{12}$$

where P_q^* is the world price of materials and trade cost of materials between a district and its nearest port (t_{qd}) is given by,

$$t_{qd} = \gamma(Time_d)^{\rho_q} \tag{13}$$

where ρ_q denotes the economies of time in manufacturing intermediates. We assume that materials are more time-sensitive than manufacturing output, but less sensitive than the perishable agriculture sector output ($\rho_M < \rho_q < \rho_A$).

Sectoral productivity (z_{kd}) for each land plot (ζ) is independently drawn from the following Fréchet distribution given by,

$$F_{kd}(z) = e^{-\mathcal{T}_{kd} z^{-\theta}} \tag{14}$$

where $\mathcal{T}_{kd}$ governs the absolute advantage and θ regulates the comparative advantage.

5.2.1 Sectoral Land Allocation

Landlords observe the sectoral productivity draw and allocate the land to the sector that pays the highest rent. The share of total land allocated to a sector in a district is given by the probability that a land plot is allocated to that sector. Labor mobility across sectors implies wages (w_d) are equalized across sectors. Given realized productivities and the zero profit condition, profit maximization pins down the equilibrium rental rate of land in sector k in district d as:

$$\begin{aligned} r_{Ad}(\zeta) = z_{Ad}(\zeta)\mathcal{P}_{Ad}, \quad r_{Md}(\zeta) = z_{Md}(\zeta)\mathcal{P}_{Md}, \\ \text{where} \quad \mathcal{P}_{Ad} = \left[\frac{P_{Ad}}{w_d^{1-\alpha_A}}\right]^{\frac{1}{\alpha_A}} \quad \text{and} \quad \mathcal{P}_{Md} = \left[\frac{P_{Md}}{P_{qd}^{\beta_M} w_d^{1-\alpha_M-\beta_M}}\right]^{\frac{1}{\alpha_M}} \end{aligned} \tag{15}$$

where $\mathcal{P}_{kd}$ refers to the sectoral prices adjusted by wages, cost of materials and factor intensity. We refer to $\mathcal{P}_{kd}$ as cost-adjusted price from hereon. Using the Fréchet distribution in 14, we get the

[43] We use the terms materials in manufacturing and manufacturing intermediates interchangeably from hereon

following probability of a land plot being allotted to sector k,

$$l_{kd} = \frac{L_{kd}}{L_d} = \left[\frac{\mathcal{T}_{kd}\mathcal{P}_{kd}^{\theta}}{\sum_{f\in\{A,M\}}\mathcal{T}_{fd}\mathcal{P}_{fd}^{\theta}}\right] \tag{16}$$

where l_{kd} is the share of land in a district allocated to sector $k \in \{A, M\}$. Intuitively, the share of land in agriculture is increasing in the average productivity of agriculture and cost-adjusted price of agriculture, and decreasing in average productivity and cost-adjusted price of manufacturing. Using the properties of Fréchet distribution, the expected rent is equalized across sectors and given by,

$$\mathbb{E}_z(r_d) = \gamma_\theta \left[\sum_{k\in\{A,M\}} \mathcal{T}_{kd}\mathcal{P}_{kd}^{\theta}\right]^{\frac{1}{\theta}} \quad , \quad \gamma_\theta = \Gamma\left(\frac{\theta-1}{\theta}\right) \tag{17}$$

where $\mathbb{E}_z$ is the expectation over the distribution of sectoral productivity and Γ is the Gamma function.

Using the first order conditions from sectoral profit maximization, labor per unit of land across sectors and materials per unit of land in manufacturing can be written as:

$$n_{Md} = \frac{1-\alpha_M-\beta_M}{\alpha_M}\left[\frac{\mathbb{E}_z(r_d)}{w_d}\right], \quad n_{Ad} = \frac{1-\alpha_A}{\alpha_A}\left[\frac{\mathbb{E}_z(r_d)}{w_d}\right], \quad q_{Md} = \frac{\beta_M}{\alpha_M}\left[\frac{\mathbb{E}_z(r_d)}{P_{qd}}\right] \tag{18}$$

5.3 Population Mobility

Workers draw idiosyncratic preferences over districts and choose the location that offers them the highest utility. Using the Fréchet distribution in 6, the probability that worker ψ chooses to reside in district d is given by:

$$\lambda_d = \frac{(B_d w_d/P_{cd})^{\chi}}{\sum_{f\in\mathcal{D}}(B_f w_f/P_{cf})^{\chi}} \tag{19}$$

where λ_d is the share of total population in district d. Intuitively, the probability of a worker choosing a specific district is increasing in the amenity-adjusted real wage of that district and decreasing in the amenity-adjusted real wage of all other districts.

The expected utility of workers, conditional on choosing a specific district is equalized and written as,

$$\mathbb{E}_u = \gamma_\chi \left[\sum_{f\in\mathcal{D}}(B_f w_f/P_{cf})^{\chi}\right]^{\frac{1}{\chi}}, \quad \gamma_\chi = \Gamma\left(\frac{\chi-1}{\chi}\right) \tag{20}$$

where Γ is, again, the Gamma function.

5.4 Market Clearing

We can now lay out the conditions for market clearing. First, the total income in a district, composed of workers and landlords, is equal to the total expenditure in the district.

$$X_d = w_d\, N_d + r_d\, L_d \tag{21}$$

Second, labor market clearing within a district requires that the total labor in a district is the sum total of labor employed across agriculture and manufacturing.

$$N_d = N_A + N_M \tag{22}$$

Third, labor market clearing across districts implies that total labor in a district equals the share (λ_d) of total workers in the economy (N) choosing to stay in that district,

$$N_d = \lambda_d N \tag{23}$$

Fourth, land market clearing implies that total rent for the landlords equals the total payments for land use across sectors,

$$r_d L_d = w_d \left[\frac{\alpha_M N_{Md}}{1 - \alpha_M - \beta_M} + \frac{\alpha_A N_{Ad}}{1 - \alpha_A} \right] \tag{24}$$

Finally, materials market clearing implies the total income of foreign material owners equals the total payments for materials by manufacturing sector,

$$P_{qd} Q_{Md} = r_d L_{Md} \frac{\beta_M}{\alpha_M} \tag{25}$$

5.5 General Equilibrium

Given a set of exogenous parameters $(\eta_A, \eta_M, \sigma, \gamma, \rho_A, \rho_M, \rho_q, \alpha_A, \alpha_M, \beta_M, \theta, \chi)$, productivities ($\mathcal{T}_{Ad}$, $\mathcal{T}_{Md}$), amenities (B_d), world prices and tariffs ($P_A^*, P_M^*, P_q^*, \tau_M, \tau_q$), trade costs ($t_{Ad}, t_{Md}, t_{qd}$), land supply ($\bar{L}$) and labor supply ($\bar{N}$), the general equilibrium of the model is characterized by the following endogenous variables, (i) Domestic price of agriculture, manufacturing and materials (P_{Ad}, P_{Md}, P_{qd}); (ii) Wages (w_d); (iii) Expected Rent (r_d); (iv) Share of land in manufacturing and agriculture (l_{Ad}, l_{Md}); and (v) Labor Share (λ_d)

All remaining endogenous variables ($L_{Md}, L_{Ad}, N_{Md}, N_{Ad}, Q_{Md}, x_{Ad}, x_{Md}, P_{cd}$) are functions of these equilibrium objects.[44]

5.6 Calibration

We study the distributional effects of trade liberalization in an environment with high internal trade costs. Our counterfactual exercises replicate the reforms in India since 1991. We therefore calibrate the baseline equilibrium to match key features of the Indian economy before these reforms.

5.6.1 Parameter Estimation

The model parameters are given by the vector $(\eta_A, \eta_M, \sigma, \gamma, \rho_A, \rho_M, \rho_q, \alpha_A, \alpha_M, \beta_M, \theta, \chi)$. Our quantitative exercises rely on the exact hat algebra approach and do not require values for η_A and η_M. Following the macroeconomic literature on structural transformation, we set the elasticity of substitution across sectors to $\sigma = 0.5$.[45]

Next, we estimate factor intensities. For α_A and α_M we follow Dholakia (1996), who estimates

[44] For a full set of equilibrium conditions, refer to Appendix B.2

[45] Several papers using the CES assumption estimate this elasticity between 0.4 and 0.5, including Fajgelbaum and Redding (2022); Bah (2007); Rogerson (2008); Duarte and Restuccia (2010).

the average share of land in the primary and secondary sectors in India during 1961-91 as 0.3 and 0.035, respectively. To determine β_M, we use ASI data for 1990-91, which report total material expenditure and total labor compensation for each firm in the sample.[46] We compute the ratio of total material expenditure to total labor compensation, which corresponds to $\frac{\beta_M}{1-\alpha_M-\beta_M}$, and obtain a value of 5.1. Using the estimate of α_M and this ratio, we back out $\beta_M = 0.8$.

We infer the values of θ and χ from the existing literature. Following Heblich et al. (2024), we set $\theta = 1.2$. For χ, Heblich et al. (2024) use a value of 2, which lies near the center of the range of estimates reported in Fajgelbaum and Redding (2022), Gallé et al. (2022), and Bryan and Morten (2019). We adopt $\chi = 2$ in our baseline model.

Finally, ρ_A and ρ_M govern the economies of time in agriculture and manufacturing output, respectively. We set $\rho_A = 0.75$ and $\rho_M = 0.45$ following Roberts et al. (2012), who estimate these parameters for rural and urban goods in China. These values are consistent with our assumption that $\rho_M < \rho_A$. To estimate the economies of time in materials, ρ_q, we use ASI data on import of intermediate inputs from 2001. For all manufacturing firms, we observe import values and quantities for narrowly defined intermediate inputs and construct unit values as proxies for import prices. For each firm, we assume that all imports arrive through its nearest major port. We can then express the firm-input import price as a function of the world price and trade cost from the firm's district to that port as in Equation 12. A nonlinear least squares estimation yields $\rho_q = 0.63$, which is in line with our assumption that $\rho_M < \rho_q < \rho_A$. As a validation for our choice of $\rho_M = 0.45$, we re-estimate the above specification using only goods that are closer to final consumption and less likely to be intermediates, based on the Broad Economic Categories (BEC) classification. Although the estimate is imprecise due to the smaller sample size, we obtain a time-sensitivity estimate of 0.457, which matches the value used in the calibration. The estimation procedure is described in detail in Appendix B.3.1. As in the empirical section, γ is a calibration parameter and set to 0.0559 so that the median trade cost between a district pair in the manufacturing sector is 25%.

5.6.2 Model Inversion

In this section, we show that, given the calibrated parameters, the observed data allow us to recover unique values of all unobserved endogenous variables that rationalize the data as an equilibrium of the model.[47] In particular, we require values for the following objects, $(w_d, r_d, Q_{MD}, L_d, L_{Ad}, L_{Md}, l_{Ad}, l_{Md}, x_{Ad}, x_{Md}, N_{Md}, N_{Ad})$.

Observed Data: We take wages, expenditure shares, sectoral employment, and land area directly from the data, as described below.[48]

1. Wages (w_d): We obtain data on district level wages from the 1987-1988 round of the NSS. For each individual, the NSS reports the number of days worked and total earnings in the

[46] Labor compensation includes wages, salary, and bonus. Materials refer to materials consumed and exclude fuel expenditure.

[47] For a discussion of the identification and uniqueness (up to scale) of productivity-adjusted sectoral prices and amenity-adjusted aggregate prices, we refer the reader to the online Appendix of Heblich et al. (2024).

[48] See Appendix B for details.

reference week, along with the industry of employment. We restrict the sample to workers employed in agriculture, manufacturing, and producer services and use the sample weights to compute the average daily wage in each district.

2. Expenditure Shares (x_{kd}): We use household-level consumption data from 1987-88 NSS round. The survey records item-wise household expenditure over the previous 30 days on a broad set of goods in agriculture, manufacturing, and services. We focus on expenditure on agricultural and manufacturing goods and use sample weights to aggregate these to district-level monthly sectoral expenditures, from which we compute expenditure shares.
3. Sectoral Employment (N_{kd}): We use the Population Census of 1991 to obtain district-level employment in agriculture and manufacturing. We also include employment in producer services in the manufacturing employment.
4. Land (L_d): To measure usable land in each district, we combine the 1987 district boundary shapefile with Global Land Cover data in ArcGIS to identify areas such as forests, mountains, and lakes that are not available for production. We overlay these classifications on total district area to compute usable land L_d for each district.[49]

Inversion: Using observed data and conditions from the model, we recover rent, land shares, and total materials as follows:

1. Rent (r_d): The land market clearing condition implies,
$$r_d L_d = w_d \left[\frac{\alpha_M N_{Md}}{1-\alpha_M-\beta_M} + \frac{\alpha_A N_{Ad}}{1-\alpha_A} \right]$$
Using observed data on land, wages, sectoral employment, and production parameters, we pin down district-level rents.
2. Sectoral land use (L_{kd}) and land shares (l_{kd}): Using the first-order conditions, we recover land used in agriculture and manufacturing as,
$$L_{Ad} = \frac{w_d}{r_d} \frac{\alpha_A}{1-\alpha_A} N_{Ad} \quad , \quad L_{Md} = \frac{w_d}{r_d} \frac{\alpha_M}{1-\alpha_M-\beta_M} N_{Md}$$
As L_d is observed in the data, sectoral land shares follow as $l_{Ad} = L_{Ad}/L_d$ and $l_{Md} = L_{Md}/L_d$.
3. World Price of Materials (P_q^*): Since the world price of materials (P_q^*) is exogenous, we normalize it to 1. The delivered materials price for districts (P_{qd}) can then be written as $(1+\tau_q+t_{qd})$.
4. Material Use (Q_{Md}): We use the first order condition on materials to obtain the total material usage in a district as,
$$Q_{Md} = \frac{r_d}{P_{qd}} \frac{\beta_M}{\alpha_M} L_{Md}$$

[49] See Appendix B.3.2 for details.

5.7 Counterfactuals

5.7.1 Trade Liberalization

As our first counterfactual, we liberalize the Indian economy. In the manufacturing sector, we reduce tariffs on both output and intermediates in line with the observed tariff cuts between 1989 and 2004. For agriculture, we lower the price of agricultural goods to match the average decline in world commodity prices between 1975–89 and 1990–2004.[50] We first show that the model reproduces our empirical findings of stronger manufacturing growth and structural change toward manufacturing in districts with better international access. We then use the model to examine how differences in international access shape the distribution of welfare gains from trade liberalization across districts.

Between 1989 and 2004, large tariff reductions in manufacturing lowered the price of manufacturing output relative to agriculture. At the same time, input tariffs fell sharply, so the prices of manufacturing intermediates declined, while agriculture experienced no additional comparable cost reduction.[51] As a result, the cost-adjusted relative price of manufacturing to agriculture ($\mathcal{P}_\mathcal{M}/\mathcal{P}_\mathcal{A}$) rises and rent in manufacturing increases even though the relative price of manufacturing to agriculture output (P_M/P_A) falls. Equation 16 then implies that more land is allocated to manufacturing and all districts experience structural change away from agriculture.

The magnitude of manufacturing growth and structural change depends crucially on heterogeneity in trade costs to ports across sectors and districts. Figure 11 plots trade costs as a function of travel time for manufacturing output, manufacturing intermediates, and agriculture. Relative to a port district, interior districts face higher trade costs and therefore experience both a smaller decline in manufacturing output prices when output tariffs fall and a smaller decline in intermediate input costs when input tariffs fall. Importantly, because the economies of time are stronger for manufacturing output than for intermediates, within interior districts, decline in input costs are smaller relative to the decline in output prices. As a result, relative to port districts, they have a lower cost-adjusted price of manufacturing and a lower cost-adjusted relative price of manufacturing to agriculture. This implies that, although labor moves out of agriculture and into manufacturing in all districts, the reallocation is weaker in interior districts.

International access therefore shapes the spatial distribution of economic activity across sectors in response to changing relative prices. Although agriculture prices fall uniformly across all districts, districts with lower trade costs experience larger proportional declines in manufacturing prices and, in turn, a sharper fall in the consumption price index. As a result, real wage rises more in these districts, which induces workers to migrate from low to high international access locations, consistent with our empirical findings.

Finally, because agriculture is more land-intensive, a sharper decline in agriculture in port districts would, in a standard Heckscher-Ohlin framework, put downward pressure on land rents.

[50] We show that the agriculture to manufacturing relative price changes we feed in the model track closely the relative price changes in the data using the official terms of trade data.

[51] Wages are endogenous and common across sectors so they affect the cost-adjusted price in both sectors similarly.

However, labor mobility implies that inflow of workers into port districts would bid up the return to the fixed factor as the demand for land would rise. This could potentially overturn the Heckscher-Ohlin prediction. As a result, the effect on the spatial distribution of real rents is ambiguous.

We use y_d to denote the value of a variable in the observed pre-liberalization equilibrium and y'_d for its value in a counterfactual equilibrium. Relative changes between counterfactual and observed values are written with hats, $\widehat{y_d}(= \frac{y'_d}{y_d})$. In Appendix B.3.3, we apply exact hat algebra to express the full set of general equilibrium conditions in changes, using the observed initial values of endogenous variables and the specified price changes.

We use the following price changes between 1989 and 2004. We first construct average output and input tariffs by aggregating industry-level tariffs using weights based on each industry's share in total output and total input use, as reported in the 1993 to 1994 Input Output Table. Over this period, average output tariffs fall from 93.28% in 1989 to 34.20% in 2004, and average input tariffs fall from 98.1% to 29.84%.[52] Using these values, we can write the relative changes in manufacturing and materials prices as,

$$\widehat{P_{Md}} = \frac{1 + 0.3420 + t_{Md}}{1 + 0.9328 + t_{Md}} \quad , \quad \widehat{P_{qd}} = \frac{1 + 0.2984 + t_{qd}}{1 + 0.981 + t_{qd}} \tag{26}$$

To obtain changes in agricultural prices, we use the World Bank Pink Sheet data on commodity prices. Given the volatility of commodity prices, we construct long-run averages to capture pre- and post-liberalization levels. Specifically, we compute an industry-size-weighted average of world agricultural prices over 1975-1989 for the pre-liberalization period and over 1990-2004 for the post-liberalization period. This yields a decline in the world agricultural price, $\widehat{P}^*_A$, of 0.78.[53] Assuming zero global tariffs on Indian agricultural exports, we set the corresponding change in the domestic agricultural price, $\widehat{P}_{Ad}$, equal to 0.78.[54]

An accurate measure of the change in relative prices between agriculture and manufacturing is crucial for our model predictions. As a validation exercise, we use official data on India's agriculture to manufacturing terms of trade from the Economic Survey of India for 1981 to 2003 and compare the change from the pre-liberalization period to the post-liberalization period with the change we feed into the model. According to the Economic Survey, the agriculture terms of trade improved by 11.27% on average from 1981-89 to 1990-2003, while the price changes we impose in the model, based on manufacturing tariff reductions and world agriculture commodity price movements, imply an improvement in relative agriculture prices of 12.3%. This close correspondence indicates that the price changes we feed into the model provide an empirically grounded representation of the evolution of relative prices in the Indian economy.[55]

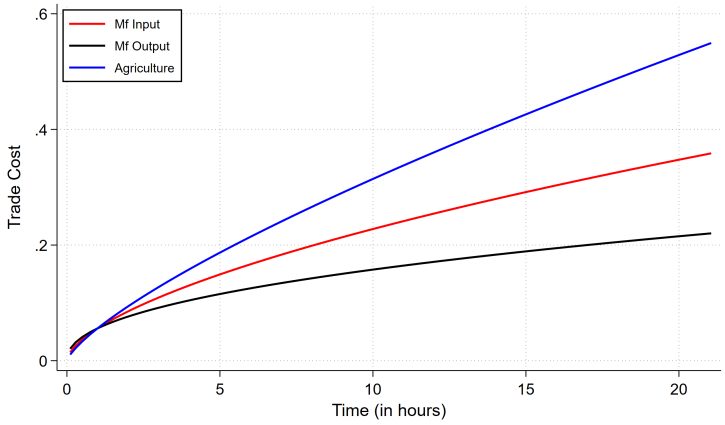
[52] The larger decline in input tariffs is consistent with the emphasis of India's liberalization on inputs, especially in the early years.

[53] Our results are robust to using a shorter pre-liberalization window, such as 1979-1989, which gives a $\widehat{P}_{Ad} = 0.81$.

[54] Refer to Appendix B.3.4 for details on the construction of changes in average input and output tariffs and world agricultural prices.

[55] For details on the calculation, see Appendix Section B.3.5

Figure 11: Trade Costs



Note: The figure shows the increase in trade cost across manufacturing output, manufacturing input and agriculture.

Results: We first show that the model rationalizes our empirical findings.[56] In Panel (a) of Figure 12, we report results for structural change, measured in percentage points. The model again matches our empirical finding that districts with better international access experience a stronger shift of employment out of agriculture and into manufacturing and producer services. Additionally, consistent with Figure 9(b), all districts experience a decline in the share of agriculture employment, with the decline for the average district being around 5 pp. The slope of the fitted line is 0.9, which is close to the corresponding slope of 0.79 in the raw data between 1991 and 2004. The model, therefore, rationalizes the pronounced spatial inequality in structural change in India after liberalization.

Recent work on structural change in India (Fan et al., 2023), together with the evidence in Table 4, highlights the importance of growth in consumer services in driving the decline in agricultural employment. Before turning to the welfare analysis, it is therefore important to verify that our model does not overstate the role of manufacturing and producer services in shaping the broad patterns of structural change observed in India over our sample period. Restricting attention to agriculture, manufacturing, and producer services for the districts in our sample, the aggregate share of employment in agriculture declines by 6.81 pp in the data. The corresponding model predicted decline is 4.45 pp. The model therefore captures a substantial share of the observed reallocation out of agriculture, while leaving room for other forces, including the expansion of consumer services, that lie outside our framework.

Panel (b) plots the percent growth in manufacturing output in each district against the manufacturing output trade cost, together with the fitted line.[57] We find a strong negative relationship between output gains and trade cost and, consistent with our empirical evidence, even districts

[56] We drop four districts from the model with agriculture employment share of less than 1 percentage point. Including them does not affect our results.

[57] For the discussion of our model results, we always use the trade cost in the manufacturing sector and refer to it simply as trade cost for brevity, to maintain comparability with the empirical section. The core mechanisms and results are unchanged if we use any other trade cost on the x-axis.

with very weak international access experience net gains in manufacturing output. Thus, trade liberalization raises manufacturing output in all districts, but the gains are larger in districts with better international access. Again the negative slope of 3.88 for the fitted line matches well with the slope of -3.95 in the raw data.[58]

In Panel (c), we show that the model is consistent with our empirical finding of increased worker outflows from high to low trade cost districts. The coefficient on the best linear fit between the change in total workers and trade cost is negative and statistically significant. Districts with low trade costs typically experience an increase in total workers, while most high trade cost districts have values below one, indicating a decline in their worker population. To get the corresponding slope in the data, we use the migration data between 1991 to 2001 to measure the net migration flows in each district and add that to the total workers in 1991. This yields migration-induced population changes that are comparable to those in the model. The slope of the fitted line in the model, -0.82, is close to its counterpart in the data, -0.87.

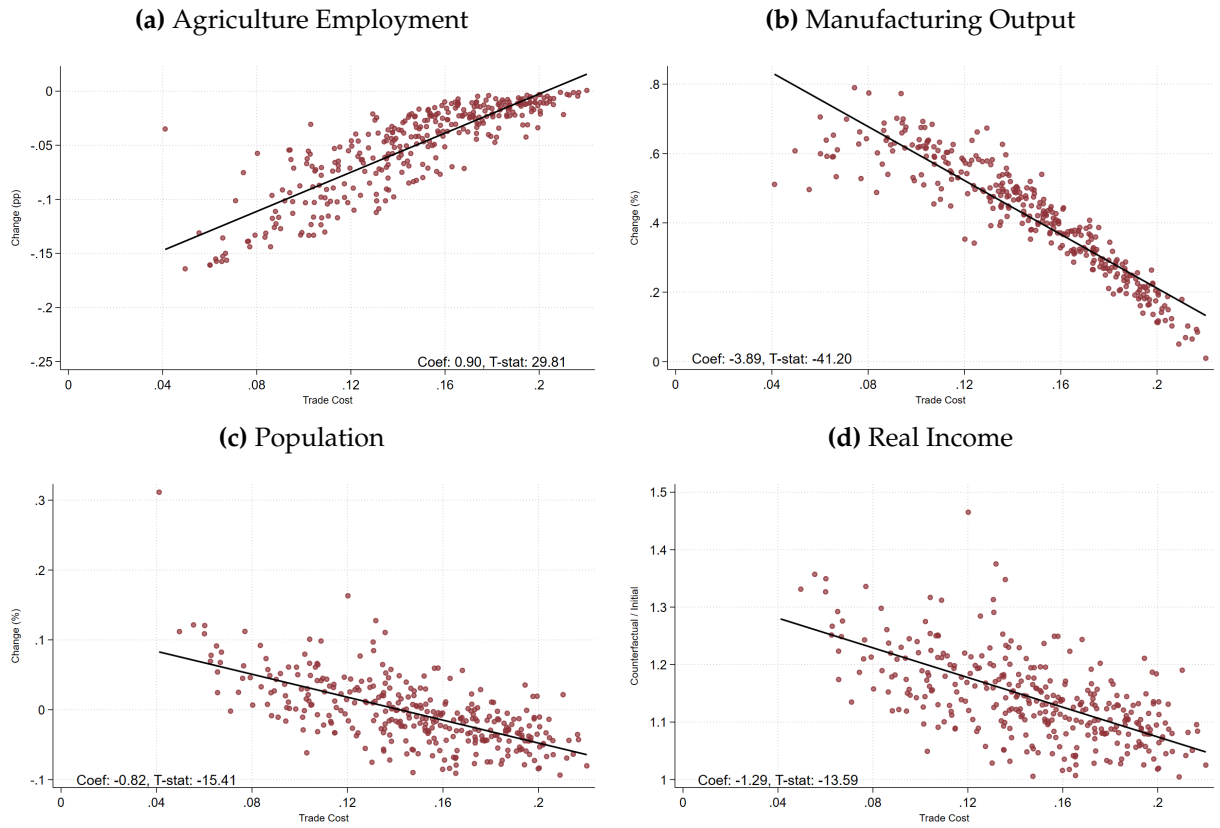
We next use the model to study how trade costs shape the distribution of real income.[59] We find substantial heterogeneity in welfare gains across districts. The average district gains 14.3%, but the smallest gainer sees an increase of only 0.46%. Panel (d) of Figure 12 plots the change in real income against trade cost for districts and shows a strong negative and statistically significant linear relationship. A reduction in trade cost by the median amount is associated with a 14.19% higher real income gain. This highlights international access as a key margin shaping the distribution of welfare gains from external integration.

Finally, we examine how international access shapes the distribution of factor payments. Figure 13 plots changes in real rents and real wages against trade costs. Districts with better international access experience larger increases in payments to both factors. On average, real rents to landlords rise by 11.8%, but landlords in the worst affected district lose by 0.58% increase. Real wages to workers increase on average by 15.4% with a minimum gain of 10.3%. Since the decline in agriculture is larger in districts with better international access, and agriculture is more land-intensive, a standard Heckscher–Ohlin framework predicts a stronger fall in real rents in these districts. However, labor mobility in our model alters this logic – the inflow of workers from high to low trade cost districts increases pressure on the fixed land endowment in low trade cost districts, raising land rents and generating the positive relationship between international access and real rents.

[58] As discussed above, the ASI manufacturing data are not representative at the district level after 1998. We therefore use changes in real manufacturing output between 1990-91 and 1997-98 to compute the slope in the raw data. In addition, because of missing data, the number of districts in the data, 309, is smaller than in the model, 303.

[59] For clarity of presentation, we omit Delhi from the scatter plots because it is an outlier with real income gains above 80%. We also omit Delhi when discussing results for rents and wages. The fitted lines, however, are always estimated using the full sample, including Delhi.

Figure 12: Model Results from Trade Liberalization Counterfactual



Note: The figure shows the role of international access in the effects from trade liberalization on manufacturing output, agriculture employment, total worker population, and real income.

5.7.2 Trade Liberalization with Optimal Road Network

Having established the role of India's poor road infrastructure in shaping the distribution of gains from trade liberalization, we now ask whether simultaneous road investments could reduce this inequality. In particular, we ask how welfare gains would be distributed across districts if, in the 1990s, India had invested in a high-speed road network targeting port connectivity. To answer this question, we construct a counterfactual optimal highway network that connects each district centroid to its nearest major port, taking into account spatial variation in terrain that affects the cost of building roads.

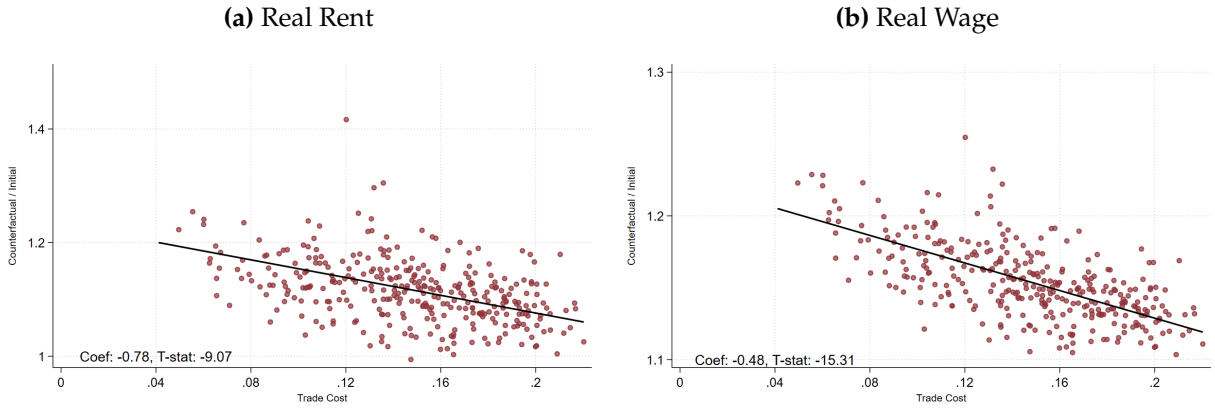
We begin by measuring the cost of traveling across space in India using the Global Accessibility Map for the year 2000, developed by the European Commission and the World Bank (Nelson, 2008). The map assigns a "cost" or friction to travel across each 1×1 km pixel, based on an algorithm that combines information on topology, terrain, environmental conditions, and political factors.[60] Figure A.5 shows the distribution of this cost across India.

We next use a Minimum Spanning Tree (MST) algorithm to connect each district centroid to an international port.[61] This delivers a cost-efficient network in which every district is connected

[60] Details on the construction of the index are available here.

[61] The MST algorithm takes the cost raster as the weight on each possible link and selects the set of connections that

Figure 13: Trade Costs and Distribution of Factor Payments from Trade Liberalization



Note: The figure shows the role of international access in the distribution of factor payments from trade liberalization.

to an international gate and aggregate travel cost is minimized. Figure 14 shows the resulting road network. We assign a speed of 45 mph to this optimal highway network.[62] Using these new links, we recompute travel times between each district center and its closest port and use the resulting travel times to construct counterfactual trade costs for manufacturing output, materials, and agriculture.

Figure 14: Counterfactual Optimal Highway Network



Note: The figure shows the counterfactual highway network targeting port connectivity using the Optimal Region Connections tool in ArcGIS.

We repeat the trade liberalization counterfactual with the previously discussed changes in tariffs and the world price of agriculture while changing the trade costs to examine whether spatial welfare inequalities can be dampened when a country invests in internal integration as it becomes

link all district centroids to the port system at minimum total cost. More details on the tool are available here.

[62] This is close to the speed of 75 kmph on the GQ completed in the late 2000s (Alder, 2025). We treat this as an upper bound for speeds on Indian highways.

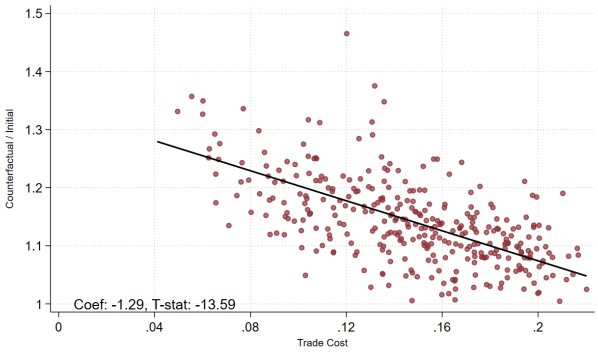
more exposed to world price changes. Our relative price changes can now be expressed as:

$$\widehat{P_{Md}} = \frac{1 + 0.3420 + t'_{Md}}{1 + 0.9328 + t_{Md}} \quad , \quad \widehat{P_{qd}} = \frac{1 + 0.2984 + t'_{qd}}{1 + 0.981 + t_{qd}} \quad , \quad \widehat{P_{Ad}} = 0.78 \cdot \frac{1 + t_{Ad}}{1 + t'_{Ad}}$$

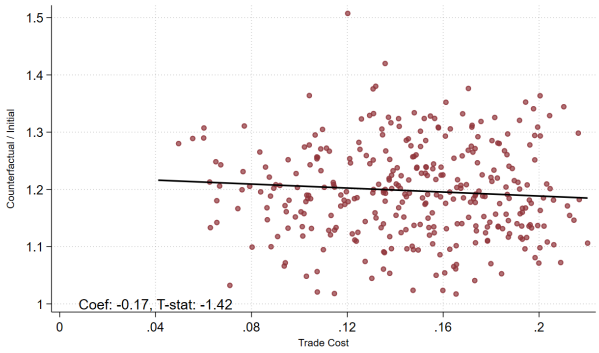
Results: We present the distribution of real income changes in Figure 15. Panel (a) reproduces the results from the first counterfactual with trade liberalization alone. Panel (b) reports the results from the current counterfactual, where along with trade liberalization, trade costs fall under the optimal road network, and the x-axis continues to show initial trade costs. Initial international access is now substantially less important for the distribution of welfare gains. The slope of the linear fit remains negative, but its magnitude is seven times smaller.

Figure 15: Counterfactual Road Network and Distribution of Welfare from Trade Liberalization

(a) Without Optimal Roads



(b) With Optimal Roads

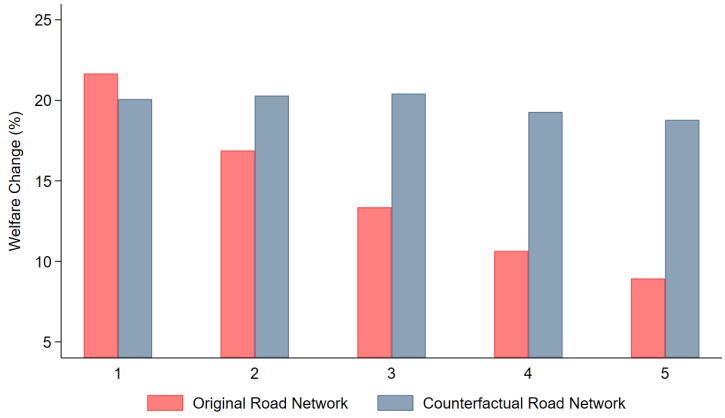


Note: Panel(a) plots the relative change in real income from the trade liberalization counterfactual without changes in the road network. Panel(b) plots the relative change in real income from trade liberalization with the counterfactual road network.

In Figure 16, we group districts into quintiles of initial trade cost and plot the average change in welfare for each quintile. In the first counterfactual, average welfare gains decline sharply with higher trade costs. Under the counterfactual with improved international access, the gains are much more evenly distributed. For the fifth quintile, the average increase in real income doubles, from 8.93% to 18.78%. Thus, the model predicts a markedly more equal distribution of welfare gains from external integration when the economy simultaneously invests in targeted road infrastructure that improves international access for initially disadvantaged locations. At the aggregate level, expected worker utility rises from 15.63% without road investment to 20.5% with road investment, an increase of about 31.15%.

We next examine whether road investments change how districts experience structural change from agriculture to manufacturing. Figure 17 shows that the slope of structural change flattens only slightly. With targeted road investments, initially high trade cost districts benefit more from input cost savings due to lower tariffs on materials. At the same time, these districts experience a smaller decline in agricultural prices, because lower trade costs to international ports partially offset the fall in the world price. This cushioning of agricultural prices counteracts the stronger input cost savings in high trade cost districts. As a result, the cost adjusted price of agriculture relative

Figure 16: Average Welfare across Quintiles of Trade Cost

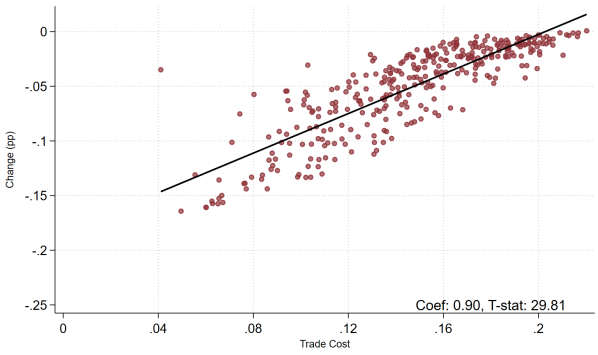


Note: The figure plots the average welfare gain across quintiles of trade cost for the trade liberalization counterfactual without and with the optimal road network.

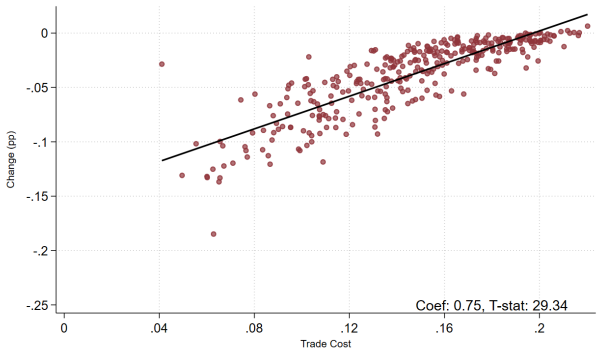
to manufacturing remains similar in these locations, and the decline in agricultural employment is close to that in the earlier counterfactual without optimal roads. Notably, the unevenness of structural change could be much stronger in an economy that is a net importer of agricultural goods.

Figure 17: Counterfactual Road Network and Structural Change

(a) Without Optimal Roads



(b) With Optimal Roads



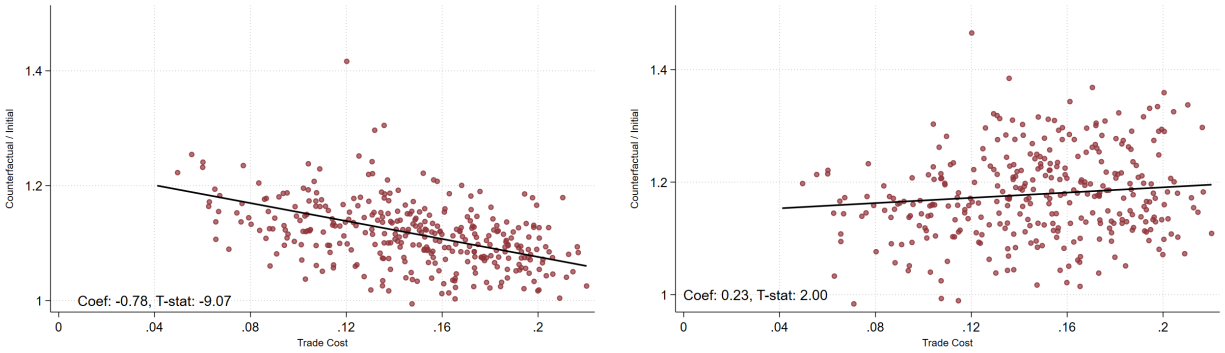
Note: Panel(a) plots the change in share of agriculture employment from the trade liberalization counterfactual without changes in the road network. Panel(b) plots the change in share of agriculture employment from trade liberalization with the counterfactual road network.

In Figure A.6, we show that highway investments substantially reduce the migration flows from districts with initially low to those with high international access. This alters the distribution of real rents as shown in Figure 18. With the migration channel muted, the standard Heckscher-Ohlin channel dominates, with smaller increases in real rent in the better international access districts, reversing the trend from the earlier counterfactual without infrastructure investments. Finally, Figure A.7 shows that simultaneous highway investments also substantially dampen real wage inequalities across districts.

Figure 18: Counterfactual Road Network and Real rents

(a) Without Optimal Roads

(b) With Optimal Roads



Note: Panel(a) plots the relative change in real rent from the trade liberalization counterfactual without changes in the road network. Panel(b) plots the relative change in real rent from trade liberalization with the counterfactual road network.

6 Conclusion

We look at the distributional impact of a major trade liberalization episode in the presence of poor domestic infrastructure. Exploiting the heterogeneity in Indian districts' connectivity to major international ports, we show how a common national trade shock produces diverse regional implications for structural transformation and inequality.

To answer these questions, we assemble an array of rich data. We digitize India's complete road network from 1996 and construct a novel spatially disaggregated industry-level panel on manufacturing activity from 1987–2004. We show that districts with better international access experienced significantly larger gains in manufacturing activity along several dimensions. In contrast to much of the existing literature that emphasizes negative effects from import competition, our results highlight the central role of input tariffs in driving manufacturing growth. At the local labor markets, we find that these gains induce faster structural transformation, with labor reallocating away from agriculture and into manufacturing and producer services in districts better connected to the world markets.

Using a quantitative spatial model we show stark spatial inequalities in distribution of gains from India's liberalization. Districts with better international access experience larger increase in real income, real wages, and real rent. Through a policy counterfactual, our model shows that as countries open to trade, simultaneous investments in road networks targeting port connectivity can substantially reduce the spatial welfare inequalities from trade.

References

Alder, S. (2025). Chinese roads in india: The effect of transport infrastructure on economic development. *Journal of International Economics*, 104140.

Alessandria, G., R. C. Johnson, and K.-M. Yi (2023). Perspectives on trade and structural transformation. *Oxford Development Studies 51*(4), 455–475.

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

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

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

Asturias, J., M. García-Santana, and R. Ramos (2019). Competition and the welfare gains from transportation infrastructure: Evidence from the golden quadrilateral of india. *Journal of the European Economic Association 17*(6), 1881–1940.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *American economic review 103*(6), 2121–2168.

Bah, E.-h. M. (2007). A three-sector model of structural transformation and economic development.

Barnwal, P., J. I. Dingel, D. Iurchenko, P. Krishna, and E. Van Leemput (2024). Internal trade barriers in india.

Baum-Snow, N., J. V. Henderson, M. A. Turner, Q. Zhang, and L. Brandt (2020). Does investment in national highways help or hurt hinterland city growth? *Journal of Urban Economics 115*, 103124.

Bryan, G. and M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Bustos, P., B. Caprettini, and J. Ponticelli (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review 106*(6), 1320–1365.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica 87*(3), 741–835.

Corden, W. M. (1966). The structure of a tariff system and the effective protective rate. *Journal of Political Economy 74*(3), 221–237.

Coşar, A. K. and B. Demir (2016). Domestic road infrastructure and international trade: Evidence from turkey. *Journal of Development Economics 118*, 232–244.

Coşar, A. K. and P. D. Fajgelbaum (2016). Internal geography, international trade, and regional specialization. *American Economic Journal: Microeconomics 8*(1), 24–56.

Costa, F., J. Garred, and J. P. Pessoa (2016). Winners and losers from a commodities-for-manufactures trade boom. *Journal of International Economics 102*, 50–69.

Cravino, J. and S. Sotelo (2019). Trade-induced structural change and the skill premium. *American Economic Journal: Macroeconomics 11*(3), 289–326.

Dholakia, B. H. (1996). Functional distribution of national income in india. *Economic and Political Weekly*, 229–236.

Dix-Carneiro, R. and B. K. Kovak (2017). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–2946.

Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review 108*(4-5), 899–934.

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

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

Eckert, F. and M. Peters (2022). Spatial structural change. Technical report, National Bureau of Economic Research.

Edmonds, E. V., N. Pavcnik, and P. Topalova (2010). Trade adjustment and human capital investments: Evidence from indian tariff reform. *American Economic Journal: Applied Economics 2*(4), 42–75.

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

Fan, T., M. Peters, and F. Zilibotti (2023). Growing like india—the unequal effects of service-led growth. *Econometrica 91*(4), 1457–1494.

Farrokhi, F. and H. S. Pellegrina (2023). Trade, technology, and agricultural productivity. *Journal of Political Economy 131*(9), 2509–2555.

Gallé, J., D. Overbeck, N. Riedel, and T. Seidel (2022). Place-based policies, structural change and female labor: Evidence from india's special economic zones. *Available at SSRN 4210059*.

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

Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova (2009). Trade liberalization and new imported inputs. *American economic review 99*(2), 494–500.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010). Imported intermediate inputs and domestic product growth: Evidence from india. *The Quarterly journal of economics 125*(4), 1727–1767.

Goyal, S. K. (1996). Political economy of india's economic reforms. *Institute for Studies in Industrial Development (ISID) Working Paper 4*.

Gulati, A. and K. Mullen (2003). Responding to policy reform: Indian agriculture during the 1990s and after. In *Fourth Annual Conference on Indian Economic Policy Reform, Stanford University, California*. Citeseer.

Hasan, R., D. Mitra, P. Ranjan, and R. N. Ahsan (2012). Trade liberalization and unemployment: Theory and evidence from india. *Journal of Development Economics 97*(2), 269–280.

Heblich, S., S. J. Redding, and Y. Zylberberg (2024). The distributional consequences of trade: Evidence from the grain invasion. Technical report, National Bureau of Economic Research.

Hummels, D. L. and G. Schaur (2013). Time as a trade barrier. *American Economic Review 103*(7), 2935–2959.

Kovak, B. K. (2013). Regional effects of trade reform: What is the correct measure of liberalization? *American Economic Review 103*(5), 1960–1976.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of economic theory 58*(2), 317–334.

Matsuyama, K. (2009). Structural change in an interdependent world: A global view of manufacturing decline. *Journal of the European Economic Association 7*(2-3), 478–486.

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

Mohan, R. et al. (1996). The india infrastructure report: policy imperatives for growth and welfare. *(No Title)*.

Nataraj, S. (2011). The impact of trade liberalization on productivity: Evidence from india's formal and informal manufacturing sectors. *Journal of International Economics 85*(2), 292–301.

Nelson, A. (2008). 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)*.

Peter, A. and C. Ruane (2023). The aggregate importance of intermediate input substitutability. Technical report, National Bureau of Economic Research.

Peters, M. (2022). Market size and spatial growth—evidence from germany's post-war population expulsions. *Econometrica 90*(5), 2357–2396.

Redding, S. J. (2022). Trade and geography. *Handbook of International Economics 5*, 147–217.

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

Roberts, M., U. Deichmann, B. Fingleton, and T. Shi (2012). Evaluating china's road to prosperity: A new economic geography approach. *Regional Science and Urban Economics 42*(4), 580–594.

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

Sahu, P. K., A. Pani, and G. Santos (2022). Freight traffic impacts and logistics inefficiencies in india: Policy interventions and solution concepts for sustainable city logistics. *Transportation in Developing Economies 8*(2), 31.

Sotelo, S. (2020). Domestic trade frictions and agriculture. *Journal of Political Economy 128*(7), 2690–2738.

Teignier, M. (2018). The role of trade in structural transformation. *Journal of Development Economics 130*, 45–65.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from india. *American Economic Journal: Applied Economics 2*(4), 1–41.

Topalova, P. and A. Khandelwal (2011). Trade liberalization and firm productivity: The case of india. *Review of economics and statistics 93*(3), 995–1009.

Van Leemput, E. (2021). A passage to india: Quantifying internal and external barriers to trade. *Journal of International Economics 131*, 103473.

Varshney, A. (1998). Mass politics or elite politics? india's economic reforms in comparative perspective. *The Journal of Policy Reform 2*(4), 301–335.

Vogel, K. B., G. H. Hanson, A. Khandelwal, C. Liu, and H. Park (2024). Using satellite imagery to detect the impacts of new highways: An application to india. Technical report, National Bureau of Economic Research.

Xu, Y. and X. Yang (2021). Access to ports and the welfare gains from domestic transportation infrastructure. *Journal of Urban Economics 126*, 103392.

A Empirical Appendix

A.1 Data Construction

A.1.1 Dijkstra's Algorithm

Dijkstra's algorithm computes the shortest travel time between a district's centroid and a port by breaking the road network into nodes and edges. The nodes represent intersections and the edges represent road segments. It assigns a weight to each edge, calculated using road length and speed. It iteratively explores neighboring nodes, always choosing the path with the lowest cumulative travel time, until it reaches a port. Since both district centroids and ports may not be directly connected to the road network, we assume that a costless connection exists between these points and the nearest road segment.

A.1.2 Methodology to obtain district identifiers in ASI data

This section provides details of the matching exercise undertaken to augment district identifiers to the ASI data from 1987-88 to 2009-10. We begin by collecting all available publications of MoSPI on district codebooks for the period of our ASI sample. This includes the district codebooks released for ASI 2009-10, the Economic Census (EC) of 2005, 1998, and 1990, and the National Sample Surveys (NSS) of 2004-05, 1999-2000, 1993-94 and 1987-88.

The first step involves identifying which publications out of the above-mentioned can be used as a valid codebook for the ASI rounds in our sample. To do this, we perform string matching algorithms over successive codebooks going backwards in time, i.e., we match ASI 2009-10 to EC 2005, EC 2005 to NSS 2004-05, and so on. This helps in identifying the years in which MoSPI changed the codebooks. We find that codebooks are identical for ASI 2009-10, EC 2005, and NSS 2004-05. The district nomenclature changes in NSS 1999-2000 and is identical across EC 1998 (except for the 3 states) and NSS 1993-94. Next, changes occur for EC 1990 and NSS 1987-88. To sum up, district nomenclatures in ASI are expected to be different for the periods 1987-90, 1990-93, 1993-2000, and 2000-04, 2004-10. This is further strengthened by examining the changes in the state codebooks available for each round of ASI. Inspection of ASI state codebooks reveals that state codes were similar from 1987-90, 1990-2000, and 2000-10. Combining the two, we obtain the following bands: (i) ASI 2009-10 codebook applies to ASI 2000-10; (ii) a combined codebook of NSS 1999-00 and EC 1998 applies to ASI 1993-2000; (iii) EC 1990 codebook applies to ASI 1990-93; and (iv) NSS 1987-88 codebook applies ASI 1987-90.

Next, as district boundaries change over time, we need to create a concordance across these codebooks so that district boundaries are consistent over the period 1987-2010. To do so, we use the codebook of EC 1990 as the basefile and construct a concordance of all district codebooks from 1990-2009 with district EC 1990. We use a string-matching algorithm that matches two codebooks on common district names. We manually inspect and match all districts that are matched with a similarity score ¡ 90%. As district boundaries change over time, to ensure that the district codes in the concordance consistently represent the district boundaries that existed in 1990, we undertake the following process:

- If a new district was carved out of one parent district, we map the new district to the code of the EC 1990 parent district. For example, in ASI 2009-10 [2723] Mumbai Suburban was carved out of the EC 1990 district [1401] Greater Mumbai. We map [2723] Mumbai Suburban in ASI 2009-10 to [1401] Greater Mumbai in EC 1990.
- If a new district is carved out of multiple parent districts, we aggregate all parent districts and assign them the code of the parent with the smallest code in EC 1990. For example, in ASI 2009-10 [0318] Mohali was formed by combining parts of EC 1990 districts [2008] Roopnagar and [2009] Patiala. We aggregate Roopnagar and Patiala in EC 1990 and code the aggregated district as [2008]. We map [0318] Mohali in ASI 2009-10 to [2008] aggregated Roopnagar in EC 1990.

Lastly, we map the codebooks from 1987-90 to EC 1990 such that the district boundaries are representative of the 1987 level. This exercise yields 345 unique districts at 1987 administrative boundaries. We then apply these concordances to the ASI rounds so that each district in an ASI round gets mapped to a district in 1987. As a result, we are able to consistently identify districts across all rounds of ASI from 1987-88 to 2009-10.

A.1.3 Domestic Market Access

Following Donaldson and Hornbeck (2016), we compute a first-order approximation of domestic market access of a district as follows:

$$DMA_{dt} = \sum_{d' \neq d} Trade\,Cost_{dd'}^{-\theta} N_{d't}$$

where, d is the origin district and d' is the destination district. *Trade Cost* is the iceberg trade cost of shipping a good from d to d' and N is the population of the destination district d'. As annual population data is not available, we use district populations from the 1991 and 2001 Population Censuses and derive the compounding annual growth rate to predict populations for 1987-2005. θ is the trade elasticity which is assumed to be 1.5 following Vogel et al. (2024). During the period of our analysis, there are no major changes to the Indian road network, so we estimate trade costs using the road network of 1996. Thus, the only source of variation in *DMA* comes from population changes. In all specifications, we control for the log of *DMA*.

A.1.4 District-level Trade Liberalization Exposure

Following Topalova (2010), we construct district-level measures of exposure to input and output tariff changes as

$$Exposure_{dt}^{ip} = \sum_{j} \omega_{jd,1991} \times \tau_{jt}^{ip} \quad , \quad Exposure_{dt}^{op} = \sum_{j} \omega_{jd,1991} \times \tau_{jt}^{op}.$$

Here, $\omega_{jd,1991}$ denotes the share of manufacturing workers in district d who are employed in industry j in 1991 (i.e., before trade liberalization). τ_{jt}^{ip} and τ_{jt}^{op} are the input and output tariffs for industry j in year t, respectively. Intuitively, these measures summarize the tariff changes faced by a district, given its pre-liberalization industrial composition, and allow us to control for differential effects of trade liberalization arising from cross-district differences in industrial structure

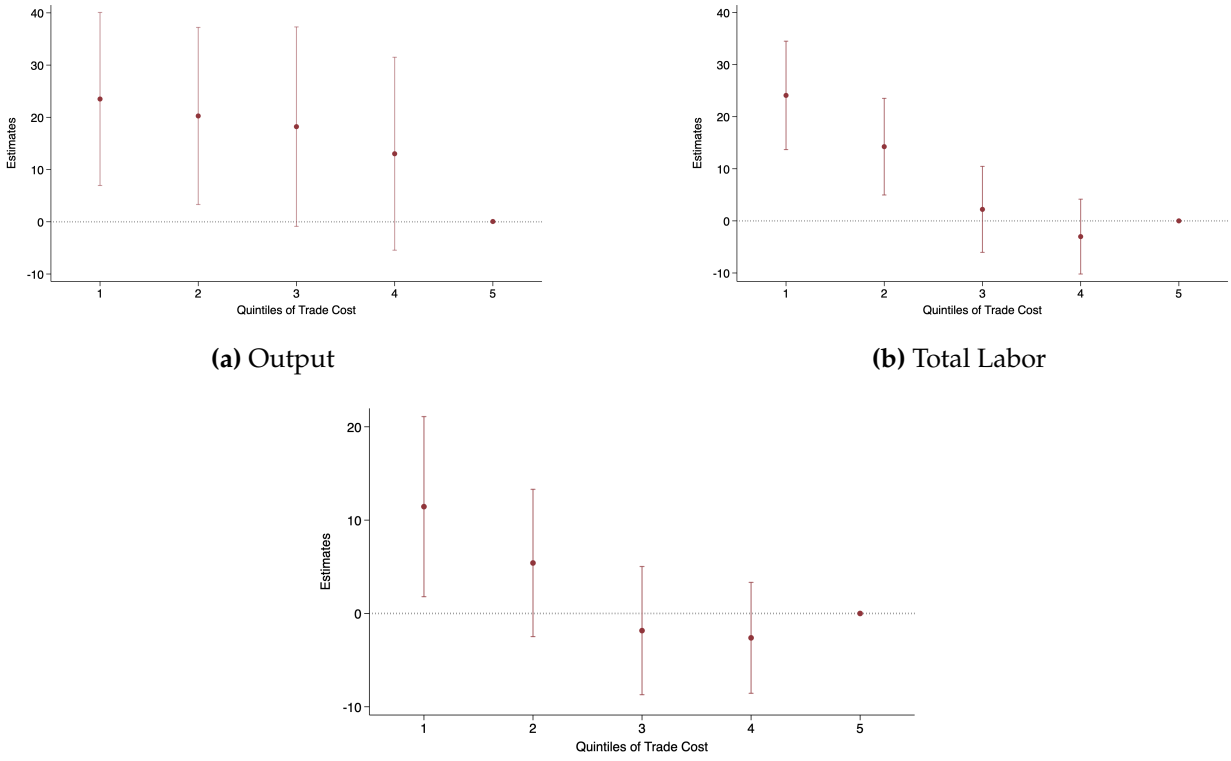
that may be correlated with a district's trade cost.

A.2 Additional Results

A.2.1 Semi-Parametric Estimation

We also estimate the effect of trade liberalization on manufacturing-sector outcomes using a semi-parametric specification in international access. We divide trade costs into quintiles, where the 1st quintile represents districts with highest international access (i.e., lowest trade cost) and the 5th quintile represents districts with the lowest international access. This approach allows us to examine whether the effects are concentrated in the most internationally integrated districts or evenly distributed across the spectrum. Results are presented in Figure A.1. For total output (Panel a), relative to the 5th quintile, all other quintiles experience larger gains, with the effects decreasing monotonically with trade cost.

Figure A.1: Semi-Parametric Estimation of Trade Cost



(a) Output

(b) Total Labor

(c) Number of Firms

Note: This figure reports results from a semi-parametric estimation of equation 1 for (log) total output, number of firms, and labor. Each point estimate represents the additional effect for the i'th quintile relative to the 5th quintile for a 10 pp decline in input tariffs. Confidence intervals at 95%.

A.2.2 Effective Rate of Protection

Instead of using input and output tariffs separately, we combine them to create a measure of effective rate of protection (erp) of an industry following Corden (1966).

$$erp_{jt} = \frac{Output\ tariff_{jt} - Input\ tariff_{jt}}{1 - \sum_s \alpha_{js}}$$

where α_{js} is the share of input from industry s in the value of output of industry j. This measure has been widely used to measure the change in net protection in industries, especially in the Indian context by Topalova and Khandelwal (2011). The results shown in Table A.1 confirm that declining erp has a positive effect on the manufacturing output with the effects being relatively higher for districts with better international access.

Table A.1: The Effect of International Access on Manufacturing Activity via Effective Rate of Protection

	Output	Number of Firms	Employment
	(1)	(2)	(3)
erp	-0.450***	-0.645***	-1.060***
	(0.109)	(0.193)	(0.200)
erp×Trade Cost	3.377***	3.711***	6.572***
	(0.839)	(0.777)	(0.972)
Observations	41,322	44,278	44,276
District-Industry FE	Y	Y	Y
Year FE	Y	Y	Y

Note: Dependent variables are (log) total output (Column 1), number of firms (Column 2), and employment (Column 3) in an industry-district. erp is the effective rate of protection for an industry. *Trade Cost* measures the iceberg trade cost between a district's centroid and its nearest port. Standard errors are two-way clustered at the 4-digit industry-year and district level. The sample consists of 98 4-digit industries × 339 districts. ***, **, * show significance at 1%, 5%, and 10%, respectively.

A.2.3 Alternate Calculations of Trade cost

In our theoretical framework, we allow the time sensitivity of manufacturing intermediates to differ from that of manufacturing output. In Section B.3.1, we estimate this parameter to be 0.63 for inputs, compared to 0.45 for output. In the empirical analysis, however, the baseline results are obtained using a common value of 0.45 for both. We now examine whether using differential trade cost values for manufacturing inputs and output, based on these estimates, affects our main results. Table A.2 reports the corresponding regressions, where $Trade\ Cost_M$ denotes the trade cost for output using the time sensitivity parameter $\rho = 0.45$, and $Trade\ Cost_q$ denotes the trade cost for intermediates using $\rho = 0.63$. The median change in input trade cost under this definition is 0.192. The results are very similar to those presented in the main text. In the baseline regression in Column (1), in response to a 10 pp decline in input trade cost, moving to a district with a lower input trade cost by the median value leads to a 16.31% additional increase in output. The corresponding increase in Table 2 is 16.69%. The effects of output tariff remain small and statistically insignificant.

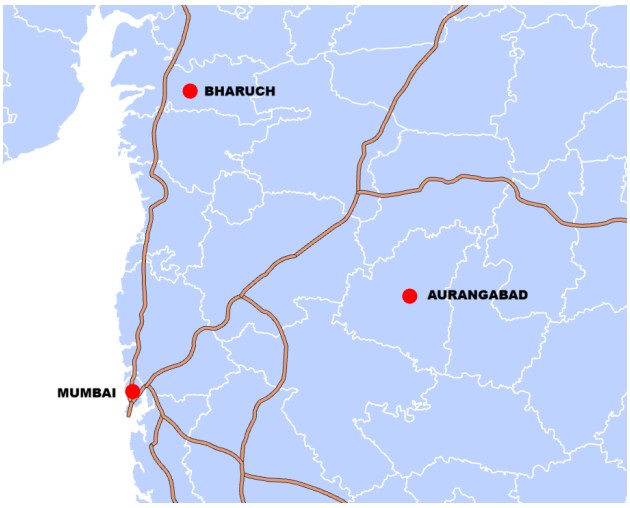
Table A.2: The Effect of International Access on Manufacturing Output

	(log) Output				Avg. Firm Output	Placebo (1987-91)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
τ^{ip}	-2.653***	-2.689***	-2.512***		-2.284***	1.224	-0.582
	(0.544)	(0.628)	(0.559)		(0.693)	(1.339)	(1.064)
τ^{op}	-0.065	-0.026	-0.094		-0.125	-0.445	0.086
	(0.153)	(0.203)	(0.155)		(0.169)	(0.270)	(0.260)
$\tau^{ip} \times TradeCost_q$	8.496***	8.487***	7.977***	8.225***	5.115*	-11.085	6.197
	(2.673)	(3.087)	(2.978)	(2.734)	(2.718)	(6.871)	(4.897)
$\tau^{op} \times TradeCost_M$	0.623	0.539	0.992	1.287	1.166	1.325	-0.961
	(1.220)	(1.510)	(1.223)	(1.267)	(1.253)	(1.827)	(1.825)
Observations	41,332	37,384	41,332	41,327	41,332	23,696	23,696
District-Industry FE	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	N	Y	Y	Y
State-Year FE	N	N	Y	N	Y	Y	Y
Industry-Year FE	N	N	N	Y	N	N	N

Note: Dependent variables are in logs at the industry–district level. $TradeCost_M$ and $TradeCost_q$ are the iceberg trade cost in manufacturing output and intermediates, respectively between a district's centroid and its nearest port. The outcome variable in Columns (1)–(4) is total manufacturing output (real sales). Column (2) drops the 11 high-FDI districts. Column (3) adds state-year fixed effects. Column (4) adds 4-digit industry–year fixed effects. The outcome variable in Column (5) is average firm-level output. Columns (6)–(7) report placebo estimates from 1987–91 for total output and average firm-level output, respectively. Standard errors are two-way clustered at the 4-digit industry–year and district levels. The sample covers 97 4-digit industries across 333 districts. ***, **, * denote significance at the 1%, 5%, and 10% levels.

A.2.4 Additional Tables and Figures

Figure A.2: Importance of Road Network in Measuring Travel Time



Note: Both Bharuch and Aurangabad are district centers, roughly 170-175 miles away from Mumbai, a port. Although equidistant, Bharuch's travel time to Mumbai is 2.6 hrs less than Aurangabad's when accounting for the entire road network.

Figure A.3: Golden Quadrilateral Highway



Note: This figure shows the placement of the Golden Quadrilateral (GQ) highway. Large parts of the GQ pass through coasts and connect to Delhi.

Table A.3: Robustness – International Access and Structural Transformation

	Rural Sample	Rainfall Shocks	State-Year
	(1)	(2)	(3)
Panel A: Agriculture			
Trade Cost × Post	0.511***	0.523***	0.256*
	(0.150)	(0.136)	(0.146)
Panel B: Manufacturing			
Trade Cost × Post	-0.263***	-0.265***	-0.204***
	(0.071)	(0.069)	(0.075)
Panel C: Services			
Trade Cost × Post	-0.248**	-0.257**	-0.052
	(0.117)	(0.107)	(0.123)
Panel D: Producer Services			
Trade Cost × Post	-0.142***	-0.129***	-0.123***
	(0.024)	(0.025)	(0.031)
Panel C: Consumer Services			
Trade Cost × Post	-0.106	-0.129	0.071
	(0.109)	(0.103)	(0.117)
Observations	945	1,020	1,011
District FE	Y	Y	Y
Year FE	Y	Y	N
State-Year FE	N	N	Y

Note: The dependent variable is the share of a district's employment in sector k. *Trade Cost* measures the iceberg trade cost between a district's centroid and its nearest port. *Post* is an indicator variable equal to 1 if the year is 1999-00 or 2004-05 and 0 if the year is 1991. Column (1) excludes districts with high baseline share of urban population ($> 50\%$). Column (2) controls for average annual rainfall and its square in district d in year t. Column (3) includes state-year fixed effects. The sample consists of 340 districts. Standard errors are clustered at the district level. ***, **, * show significance at 1%, 5%, and 10%, respectively.

B Theory Appendix

B.1 Derivations

B.1.1 Production in Manufacturing

Firms choose labor (n_M) and material (q_M) to maximize profits per unit land:

$$\underset{n,q}{max}\left\{\pi_{Md}(\zeta) = P_{Md}y_{Md}(\zeta) - w_d n_{Md}(\zeta) - r_{Md}(\zeta) - P_{qd}q_{Md}(\zeta)\right\}$$

Omitting land plot (ζ) and district (d) subscripts for brevity, the first order conditions produce the following input demands:

$$n_M^* = \frac{(1-\alpha_M-\beta_M)z_M}{\alpha_M}\left[\frac{P_M}{w^{1-\beta_M}P_q^{\beta_M}}\right]^{\frac{1}{\alpha_M}}$$

$$q_M^* = \frac{\beta_M z_M}{\alpha_M}\left[\frac{P_M}{w^{1-\alpha_M-\beta_M}P_q^{\alpha_M+\beta_M}}\right]^{\frac{1}{\alpha_M}}$$

Input demands and zero profit condition give the equilibrium manufacturing rent:

$$r_M = z_M\left[\frac{P_M}{w^{1-\alpha_M-\beta_M}P_q^{\beta_M}}\right]^{\frac{1}{\alpha_M}}$$

B.1.2 Production in Agriculture

Firms choose labor (n_M) to maximize profits per unit land in agriculture:

$$\underset{n}{max}\left\{\pi_{Ad}(\zeta) = P_{Ad}y_{Ad}(\zeta) - w_d n_{Ad}(\zeta) - r_{Ad}(\zeta)\right\}$$

Omitting subscripts for brevity, equilibrium labor demand is given by:

$$n_A^* = \frac{(1-\alpha_A)z_A}{\alpha_A}\left[\frac{P_A}{w}\right]^{\frac{1}{\alpha_A}}$$

Equilibrium labor and zero profit condition implies:

$$r_A = z_A\left[\frac{P_A}{w^{1-\alpha_A}}\right]^{\frac{1}{\alpha_A}}$$

B.1.3 Land Allocation Decision

Landlords allocate land to the sector that maximizes rent:

$$r_d(\zeta) = max\left\{r_{Ad}(\zeta), r_{Md}(\zeta)\right\}$$

The probability that each land plot ζ is allocated to manufacturing is given by:

$$l_{Md} = Pr\left[r_{Ad}(\zeta) \leq r_d(\zeta)\right] \cdot Pr\left[r_{Md}(\zeta) = r_d(\zeta)\right] = \int_r F_{r_{Ad}}(r)\, dF_{r_{Md}}(r)$$

Given z_{kd} follows a Fréchet distribution given by $F_{z_{kd}}(z) = e^{-\mathcal{T}_{kd}z^{-\theta}}$, the distribution of sectoral rents can be written as:

$$F_{r_{kd}}(r) = e^{-\Phi_{kd}r^{-\theta}}$$

where

$$\Phi_{Ad} = \mathcal{T}_{Ad}\left(\frac{P_{Ad}}{w_d^{1-\alpha_A}}\right)^{\frac{\theta}{\alpha_A}} \quad , \quad \Phi_{Md} = \mathcal{T}_{Md}\left(\frac{P_{Md}}{w_d^{1-\alpha_M-\beta_M}P_{qd}^{\beta_M}}\right)^{\frac{\theta}{\alpha_M}}$$

Using these distributions of sectoral rents, we can solve for sectoral land shares:

$$\begin{aligned} l_{Md} &= \int_0^\infty e^{-\Phi_{Ad}r^{-\theta}} e^{-\Phi_{Md}r^{-\theta}}\, \Phi_{Md}\theta r^{-\theta-1}\, dr \\ &= \frac{\Phi_{Md}}{\Phi_{Ad}+\Phi_{Md}}\left[e^{-(\Phi_{Ad}+\Phi_{Md})r^{-\theta}}\right]_0^\infty \\ &= \frac{\Phi_{Md}}{\Phi_{Ad}+\Phi_{Md}} \end{aligned}$$

Note that:

$$\frac{d}{dr}\left(e^{-(\Phi_{Ad}+\Phi_{Md})r^{-\theta}}\right) = \theta(\Phi_{Ad}+\Phi_{Md})e^{-(\Phi_{Ad}+\Phi_{Md})r^{-\theta}}r^{-\theta-1}$$

$$\begin{aligned} \implies l_{Md} &= \frac{\Phi_{Md}}{\Phi_{Ad}+\Phi_{Md}}\left[e^{-(\Phi_{Ad}+\Phi_{Md})r^{-\theta}}\right]_0^\infty \\ &= \frac{\Phi_{Md}}{\Phi_{Ad}+\Phi_{Md}} \end{aligned}$$

Substituting for the expression of Φ_{kd}, we get:

$$l_{Md} = \left[\frac{\mathcal{T}_{Md}\mathcal{P}_{Md}^\theta}{\sum_{f\in\{A,M\}}\mathcal{T}_{fd}\mathcal{P}_{fd}^\theta}\right]$$

where

$$\mathcal{P}_{Ad} = \left[\frac{P_{Ad}}{w_d^{1-\alpha_A}}\right]^{\frac{1}{\alpha_A}} \quad , \quad \mathcal{P}_{Md} = \left[\frac{P_{Md}}{P_{qd}^{\beta_M} \cdot w_d^{1-\alpha_M-\beta_M}}\right]^{\frac{1}{\alpha_M}}$$

Share of land allocated to agriculture (l_{Ad}) can be derived symmetrically.

B.1.4 Expected Rent

Let $r_d(\zeta) = max\{r_{Ad}(\zeta), r_{Md}(\zeta)\}$ denote the rent for a land plot. Then, the expected rent across sectors for each plot can be expressed as:

$$\mathbb{E}_z[r_d] = \int_0^\infty r\, dF_{r_d}(r)$$

where the distribution of the maximum rent r_d can be derived as:

$$F_{r_d}(r) = Pr[r_d \le r] = Pr[r_{Ad} \le r \wedge r_{Md} \le r] = F_{r_{Ad}}(r)\cdot F_{r_{Md}}(r) = e^{-(\Phi_A+\Phi_M)r^{-\theta}}$$

Solving for expected rent:

$$\mathbb{E}_z[r_d] = \int_0^\infty r(\Phi_{Ad}+\Phi_{Md})\theta e^{-(\Phi_A+\Phi_M)r^{-\theta}}r^{-\theta-1}\, dr$$

Define $y = (\Phi_{Ad} + \Phi_{Md})r^{-\theta}$, which gives $dy = -(\Phi_{Ad} + \Phi_{Md})\theta r^{-\theta-1}dr$. Using this change of variable:

$$\begin{aligned}
\mathbb{E}_z[r_d] &= \int_0^\infty (\Phi_{Ad} + \Phi_{Md})^{\frac{1}{\theta}} y^{\frac{-1}{\theta}} e^{-y}\, dy \\
&= (\Phi_{Ad} + \Phi_{Md})\, \Gamma\left(\frac{\theta-1}{\theta}\right) \\
&= \gamma_\theta \Bigg[\sum_{k\in\{A,M\}} \mathcal{T}_{kd}\mathcal{P}_{kd}^{\theta} \Bigg]^{\frac{1}{\theta}} \quad , \quad \gamma_\theta \equiv \Gamma\left(\frac{\theta-1}{\theta}\right)
\end{aligned}$$

B.1.5 Population Mobility

Given that worker ψ's utility is monotonic in amenities (b_d), we can derive the distribution of a worker's utility from living in d as:

$$F_{u_d}(u) = e^{-\Theta u^{-\chi}} \quad , \quad \Theta \equiv \left(\frac{P_{Cd}}{B_d w_d}\right)^{-\chi}$$

A worker will choose to live in district d if it maximizes her utility:

$$u_d(\psi) = \underset{f\in\mathcal{D}}{max}\, \{u_f(\psi)\}$$

The probability that the worker chooses to live in d is given by:

$$\begin{aligned}
\lambda_d &= Pr[u_d \geq u_f\ \forall f] \\
&= \int_0^\infty \prod_{f\neq d} F_{u_f}(u) dF_{u_d}(u) \\
&= \int_0^\infty \prod_{f\neq d} e^{-\Theta_f u^{-\chi}} e^{-\Theta_d u^{-\chi}}\, \Theta_d \chi u^{-\chi-1}\, du \\
&= \int_0^\infty e^{-\sum_f \Theta_f u^{-\chi}}\, \Theta_d \chi u^{-\chi-1}\, du
\end{aligned}$$

Note that:

$$\frac{d}{du}\left(e^{-\sum_f \Theta_f u^{-\chi}}\right) = \sum_f \Theta_f \chi e^{-\sum_f \Theta_f u^{-\chi}} u^{-\chi-1}$$

$$\begin{aligned}
\implies \lambda_d &= \frac{\Theta_d}{\sum_f \Theta_f} \left[e^{-\sum_f \Theta_f u^{-\chi}} \right]_0^\infty \\
&= \frac{\Theta_d}{\sum_f \Theta_f}
\end{aligned}$$

Substituting for Θ_d, we get:

$$\lambda_d = \frac{(B_d w_d / P_{cd})^\chi}{\sum_{f\in\mathcal{D}} (B_f w_f / P_{cf})^\chi}$$

B.1.6 Expected Utility

Let $u_d = \underset{f \in \mathcal{D}}{max} \{u_f\}$ denote the maximum of worker's utilities across all districts. The distribution of u_d can be derived as:

$$F_{u_d}(u) = Pr[u_d \leq u] = Pr[u_f \leq u \, \forall f] = \prod_{f \in \mathcal{D}} F_{u_f}(u) = e^{-\sum\limits_f \Theta_f u^{-\chi}}$$

Therefore, the expected utility of a worker across districts can be expressed as:

$$\begin{aligned}
\mathbb{E}_b[u_d] &= \int_0^\infty u \, dF_{u_d}(u) \\
&= \int_0^\infty u \sum_f \Theta_f \chi e^{-\sum\limits_f \Theta_f u^{-\chi-1}} \, du
\end{aligned}$$

Define $y = \sum\limits_f \Theta_f u^{-\chi}$, which gives $dy = -\sum\limits_f \Theta_f \chi u^{\chi-1} du$. Using this change of variable:

$$\begin{aligned}
\mathbb{E}[u] &= \int_0^\infty \left(\sum_f \Theta_f \right)^{\frac{1}{\chi}} y^{\frac{-1}{\chi}} e^{-y} \, dy \\
&= \left(\sum_f \Theta_f \right)^{\frac{1}{\chi}} \Gamma\left(\frac{\chi - 1}{\chi} \right) \\
&= \gamma_\chi \left[\sum_{f \in \mathcal{D}} \left(\frac{B_f w_f}{P_{cf}} \right)^\chi \right]^{\frac{1}{\chi}} \quad , \quad \gamma_\chi \equiv \Gamma\left(\frac{\chi - 1}{\chi} \right)
\end{aligned}$$

B.2 General Equilibrium

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

$$\begin{aligned}
l_{Ad} &= \left[\frac{\mathcal{T}_{Ad} \mathcal{P}_{Ad}^\theta}{\sum_{f \in \{A,M\}} \mathcal{T}_{fd} \mathcal{P}_{fd}^\theta} \right] \\
l_{Md} &= \left[\frac{\mathcal{T}_{Md} \mathcal{P}_{Md}^\theta}{\sum_{f \in \{A,M\}} \mathcal{T}_{Md} \mathcal{P}_{fd}^\theta} \right] = 1 - l_{Ad}
\end{aligned} \tag{27}$$

$$L_{Ad} = l_{Ad} \cdot L_d \quad , \quad L_{Md} = l_{Md} \cdot L_d \tag{28}$$

$$\begin{aligned}
w_{dt} N_{Adt} &= \frac{1 - \alpha_A}{\alpha_A} r_{dt} L_{Adt} \\
w_{dt} N_{Mdt} &= \frac{1 - \alpha_M - \beta_M}{\alpha_A} r_{dt} L_{Mdt}
\end{aligned} \tag{29}$$

$$P_{qdt} \, Q_{Mdt} = \frac{\beta_M}{\alpha_A} r_{dt} L_{Mdt} \tag{30}$$

$$P_{cdt} = \left[\sum_{k \in \{A,M\}} \eta_k P_{kdt}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{31}$$

$$\lambda_d = \frac{(B_d w_d / P_{cd})^{\chi}}{\sum_{f \in \mathcal{D}} (B_f w_f / P_{cf})^{\chi}} \tag{32}$$

$$N_{dt} = \lambda_{dt} \bar{N} \tag{33}$$

$$w_{dt} N_{dt} = \sum_{k \in \{A,M\}} w_{dt} N_{kdt} \tag{34}$$

$$r_d = \gamma_\theta \left[\sum_{k \in \{A,M\}} \mathcal{T}_{kd} \mathcal{P}_{kd}^{\theta} \right]^{\frac{1}{\theta}} \quad , \quad \gamma_\theta = \Gamma\left(\frac{\theta - 1}{\theta}\right) \tag{35}$$

B.3 Estimation

B.3.1 Parameter Estimation: ρ_q

To estimate the economies of time in manufacturing intermediates, we use firm by input import data from the 2001 round of the ASI.[63] A distinctive feature of the ASI is that firms report both the total value and the quantity imported for each intermediate input used in production.[64] We construct firm by input unit values, which we use as proxies for import prices. In the model, the price paid for each intermediate depends on its world price, the tariff on that input, and the trade cost from the firm's district to the delivery port (Equation 12). We assume that all imports of a firm arrive through the nearest identified port, so we can write the firm by input import price as a function of the world price and travel time from the firm's district to that port as,

$$\ln P_{fkdt} = \ln P^*_{q,kt} + \ln\left(1 + \tau_{kt} + t_{kd}\right),$$

where P_{fkdt} is the price paid by firm f for importing input k in district d in year t, $P^*_{q,kt}$ is the world price, τ_{kt} is the tariff on input k, and t_{kd} is the trade cost between district d and its nearest port.[65] Using a first order approximation of $\ln(1+\tau_{kt}+t_{kd})$ around zero, we write $\ln(1+\tau_{kt}+t_{kd}) \approx \tau_{kt}+t_{kd}$. Substituting the trade cost formula $t_{kd} = \gamma(Time_d)^{\rho_q}$ then yields the estimation equation

$$\ln P_{fkdt} = \gamma(Time_d)^{\rho_q} + \eta_k + \eta_f + \epsilon_{fkdt}, \tag{36}$$

where η_k is an input fixed effect that absorbs the common world price and input specific tariffs, and η_f captures firm characteristics such as the three-digit industry. Consistent with our empirical analysis, we fix γ at 0.0559 and estimate ρ_q by nonlinear least squares. The results are reported in Table A.4.

Column (1) reports the baseline specification, using 2001 ASI data with input and industry fixed effects. The estimate of ρ_q is 0.63, consistent with our assumption that the economy of time

[63] We do not use years after 2001 because construction on parts of the Golden Quadrilateral began after 2001 and may change actual travel times for some districts relative to our 1996 estimates. At the same time, the quality and coverage of the import round are substantially better in later ASI waves, so the 2001 data offer maximal variation without introducing measurement error in travel times. We show that our estimates are robust to including the 2000 round.

[64] Inputs are narrowly defined. They are classified using the ASICC system, which can be matched to seven digit NPCMS product codes. The first five digits of NPCMS can be roughly matched to six digit HS codes, so the data are more granular than the standard six digit HS classification.

[65] We assume that, for a given input, the world price is the same across Indian ports.

for intermediates lies between that for agriculture and manufacturing output. The baseline sample includes 1,745 inputs. In column (2), we restrict the sample to inputs with at least five distinct importing firms. The number of inputs falls to 391, but the estimate remains close to the baseline. In column (3), we restrict the sample to inputs that can be classified as pure intermediate or capital goods.[66] This ensures that the estimates are not driven by imported final goods. In column (4), we add the destination state fixed effect to control for state specific factors such as higher time spent at the border for entry into certain states. Our estimates remain stable around the baseline value. In column (5), we re-estimate the baseline specification using the 2000 ASI round, and in column (6) we pool the 2000 and 2001 rounds. Across all specifications, the estimates of ρ_q are consistently larger than ρ_m (0.45) and smaller than ρ_A (0.75). We therefore use $\rho_q = 0.63$ in the model.

In column (7), we provide a validation for our choice of $\rho_m = 0.45$ in the model. Using the BEC classification, we restrict the sample to goods classified as consumption goods. These goods are less likely to be used as production intermediates and are closer to final demand. Their price response to travel time therefore provides a useful check on the calibration of ρ_m. Although the estimate is imprecise because most firm imports in our data are intended for intermediate use, the estimated time sensitivity is 0.457, which is identical to the value of 0.45 used in the model.

Table A.4: Estimation of ρ_q

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\gamma(Time_d)$	0.614***	0.631***	0.605***	0.630***	0.593***	0.606***	0.457
	(0.107)	(0.113)	(0.121)	(0.148)	(0.156)	(0.089)	(0.593)
Observations	7060	4599	5445	7060	4527	11587	763
R^2	0.857	0.829	0.860	0.858	0.846	0.853	0.862
Number of products	1745	387	1318	1745	1501	3246	224

Note: The table reports the estimates of ρ_q form the nonlinear estimation of Equation 36. Columns (1) to (4) use that from 2001 round of the ASI data. Column (5) uses 2000 round of ASI data and column (6) pools 2000 and 2001 rounds. Column (7) restricts the sample to consumption goods to estimate ρ_m, instead. Robust standard errors are reported in parenthesis. ***, **, * show significance at 1%, 5%, and 10%, respectively.

B.3.2 Model Inversion

In this section, we discuss in detail the construction of endogenous variables from observed data:

Wages (w_d): We obtain district-level wage data from the 1987–88 round of the National Sample Survey (NSS), the latest available dataset on wages before liberalization. The NSS is a household survey that records, for each individual, the number of days worked in a reference week and total weekly earnings, including both cash income and the rupee value of in-kind payments. The reference week is the week in which the individual is interviewed. If an individual holds multiple jobs within a week, the survey reports earnings, days worked, and the one-digit industry code for each job separately.[67]

[66] We use publicly available correspondences to match five digit ASICC codes to seven digit NPCMS codes, where the first five digits correspond to the Central Product Classification (CPC). We then map five digit CPC codes to six digit HS12 codes and apply the BEC classification, which assigns HS products to final use categories. We classify a good as a consumption good if its final use is not purely intermediate or capital.

[67] The NSS records wages only for individuals working in a wage/salary earning job and excludes self-employed individuals.

Since we model only two sectors in our framework – agriculture and manufacturing (inc. producer services), we exclude wages earned from consumer services from the data. Consumer services, on average, account for 92% of service-sector jobs. As service wages are generally higher than other sectors and more concentrated in urban districts, their inclusion would inflate the estimated average wage of the highly urban districts.

To distinguish consumer services from producer services, we require industry classification at the three-digit level, which is available only for an individual's primary activity, defined as the job providing the largest share of income in that week. We therefore use the industry code of the primary activity to exclude individuals working in consumer services. This approach is valid, as 97.82% of individuals in the sample report working only one job during the survey week.

Finally, we compute the district-level average daily wage by first calculating each individual's daily wage as total weekly earnings (across all jobs) divided by total days worked. We then aggregate to the district level using the sample weights in NSS to obtain the district-level average daily wage rate.

Expenditure Shares (x_{kd}): The 1987-88 NSS also records household-level consumption expenditure. The survey records item-wise household expenditure over the past 30 days for 211 items spanning agriculture, manufacturing, and services. Following the National Industry Classification (1987), we classify each item by sector and only keep items that fall under agricultural and manufacturing. Table A.5 shows a subset of items classified by sectors.

Table A.5: Classification of Consumption Items by Sectors

Agriculture	Manufacturing	Services
Cereals (paddy, wheat, etc.)	Food products (rice, flour, bread)	Entertainment (cinema theatre, fairs, etc.)
Pulses	Edible oils (mustard, coconut, etc.)	Consumer services (laundry, cleaning, etc.)
Eggs, meat, and fish (fresh)	Milk products (butter, ice cream, etc.)	Medical (hospitals, veterinary, etc.)
Vegetables (potato, onion, etc.)	Sugar, salt, etc.	Transport (rail, bus fares, etc.)
Fruits (mango, apple, etc.)	Clothing (shirt, trouser, etc.)	Education expenses
Nuts (cashew, walnut, etc.)	Medicines and Footwear	

Note: The NSS 1987–88 survey records household-level consumption over the last 30 days for a set of 211 items. We classify all items into agriculture, manufacturing, and services following the National Industrial Classification (1987). This table illustrates a representative subset of items.

Using sample weights, we aggregate household-level data to district-level. We obtain monthly comsumption expenditure in a district by sector, which we then use to compute expenditure shares as follows:

$$x_{kd} = \frac{\text{expenditure}_{kd}}{\text{expenditure}_{Ad} + \text{expenditure}_{Md}} \quad , \quad k \in \{A, M\}$$

Sectoral Employment (N_{kd}): We obtain district-level employment data from the 1991 Population Census. The Census reports aggregate employment by district for three broad sectors – agriculture, manufacturing, and services. Since our analysis focuses only producer services within services, we use an auxiliary Census dataset that provides employment disaggregated by a four-digit

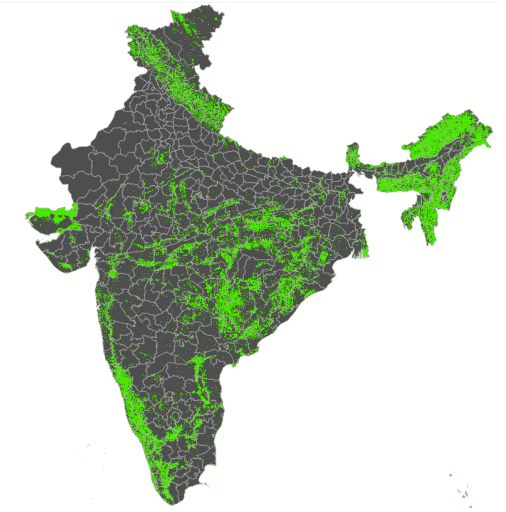
industry classification. However, the sum of producer and consumer service employment in the disaggregated dataset does not match the aggregate service-sector totals for all districts. To address this, we compute the ratio of producer to consumer services from the disaggregated data for each district and use this ratio to impute producer service employment from the aggregate service totals. Our final measure includes employment in agriculture, manufacturing, and producer services, with producer-service employment attributed to the manufacturing sector for consistency with our modeling framework.

Table A.6: Classification of Unusable Land Types

Code	Description
50	Tree Cover Broadleaf Evergreen
60	Tree Cover Broadleaf Deciduous
70	Tree Cover Needleleaved Evergreen
80	Tree Cover Needleleaved Deciduous
90	Mixed Tree Cover
160	Tree Cover Flooded (fresh/brackish)
170	Tree Cover Flooded (saline)
180	Shrub/Herb Flooded
210	Water Bodies
220	Permanent Snow & Ice

Note: The table provides details on the land type from the Global Land Cover data that is classified as unusable for the purpose of out model. All other land types are classified as usable.

Figure A.4: Usable and Unusable Land



Note: The figure shows the classification of land in usable (gray) and unusable (green) land type across the 1987 district boundaries in India.

Total Usable Land Area (L_d): To obtain the usable land area of each district (L_d), we begin by using the Global Land Cover data in ArcGIS. The data provides a time-series map, beginning 1992, of the surface of Earth from the ESA Climate Change Initiative, classified into 36 land cover types. We use these categories to classify land as usable or unusable. The basic intuition of classifying land as unusable is to flag land that is ecologically protected (natural forests, mangroves etc.), physically unusable for building structures or growing crops (open water, permanent snow/ice, permanent

wetlands/flooded vegetation), and very high canopy / swamp forests, salt pans/tidal flats as these are land plots unsuitable for production purposes in our model. Table A.6 summarizes the land cover classes we identify as unusable.

We next use the 1987 Indian districts' shape file to get the boundaries of our spatial units. We project the usable and unusable land on the district boundaries. Figure A.4 shows the results. We then use the Calculate Geometry tool in ArcGIS to calculate the total usuable area of the district. This serves as are measure of L_d.

B.3.3 Exact Hat Algebra

We denote the value of a variable in observed equilibrium in the pre-liberalization period in India without a prime (y_d) and the value of a variable in the counterfactual equilibrium with a prime (y'_d). The relative change between counterfactual and observed equilibrium values of a variable are denoted in hat form ($\hat{y}_d$). Below, we write the full set of general equilibrium conditions in changes, using the observed values of endogenous variables in the initial equilibrium and price changes:

$$\widehat{\mathcal{P}}_{Adt} = \left[\frac{\widehat{P}_{Adt}}{\widehat{w}_{dt}^{1-\alpha_A}}\right]^{\frac{1}{\alpha_A}} \tag{37}$$

$$\widehat{\mathcal{P}}_{Mdt} = \left[\frac{\widehat{P}_{Mdt}}{\widehat{P}_{qdt}^{\beta_M} \cdot \widehat{w}_{dt}^{1-\alpha_M-\beta_M}}\right]^{\frac{1}{\alpha_M}} \tag{38}$$

$$\widehat{l}_{Adt} = \left[\frac{\widehat{\mathcal{P}}_{Adt}^{\theta}}{\sum_{f\in\{A,M\}} l_{fdt}\widehat{\mathcal{P}}_{fdt}^{\theta}}\right] \tag{39}$$

$$\widehat{l}_{Mdt} = \frac{1 - l_{Adt}\widehat{l}_{Adt}}{l_{Mdt}} \tag{40}$$

$$\widehat{L}_{Adt} = \widehat{l}_{Adt} \tag{41}$$

$$\widehat{L}_{Mdt} = \widehat{l}_{Mdt} \tag{42}$$

$$\widehat{r}_{dt} = \left[\sum_{k\in\{A,M\}} l_{kdt}\widehat{\mathcal{P}}_{kdt}^{\theta}\right]^{\frac{1}{\theta}} \tag{43}$$

$$\widehat{w}_{dt}^{o}\widehat{N}_{Adt} = \widehat{r}_{dt}\widehat{L}_{Adt} \tag{44}$$

$$\widehat{w}_{dt}^{o}\widehat{N}_{Mdt} = \widehat{r}_{dt}\widehat{L}_{Mdt} \tag{45}$$

$$\widehat{P}_{qdt}\widehat{Q}_{Mdt} = \widehat{r}_{dt}\widehat{L}_{Mdt} \tag{46}$$

$$\widehat{P}_{cdt} = \left[\sum_{k\in\{A,M\}} x_k\widehat{P}_{kdt}^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{47}$$

$$\widehat{\lambda_{dt}} = \frac{(\widehat{w}_d/\widehat{P}_{cd})^{\chi}}{\sum_{f\in\mathcal{D}} \lambda_f (\widehat{w}_f/\widehat{P}_{cf})^{\chi}} \tag{48}$$

$$\widehat{N}_{dt} = \widehat{\lambda}_{dt} \tag{49}$$

$$\widehat{w}_{dt} = \frac{\widehat{w}^o_{dt}}{\widehat{N}_{dt}} \left[\sum_{k\in\{A,M\}} \frac{N_{kdt}}{N_{dt}} \widehat{N}_{kdt} \right] \tag{50}$$

where $\widehat{w}^o_{dt}$ is the initial guess value for wages. We update this value using (50) until it converges.

B.3.4 Counterfactual

In this section, we discuss in detail the construction of data used in the counterfactual exercises:

Tariffs: We have tariffs at HS 6-digit product-level from Topalova (2010) and WITS portal. We construct the average input and output tariffs for the manufacturing sector using the 1993–94 Input-Output (IO) Table for India, which contains 115 sectors across agriculture, manufacturing, and services. First, we match each HS6 product to an IO sector and compute tariffs at the IO sector-level as simple average of all HS6 products in that sector. If an HS6 product maps to more than one IO sector, we combine the overlapping IO sectors into one group. This gives us tariffs at the IO sector-level.

To aggregate IO sector-level tariffs into a single average input and output tariff for each year, we make use of the Absorption Matrix of the IO table. The Absorption Matrix provides information on the total value of sector j's inputs used by sector i, as well as the total output produced of sector j. We use a sector's input and output usage to construct weights that aggregate the sector-level tariffs into a single average tariff as follows:

$$\tau_{qt} = \sum_j \omega^{ip}_j \times \tau_{jt} \quad , \quad \tau_{Mt} = \sum_j \omega^{op}_j \times \tau_{jt}$$

where ω^{ip} is the share of sector j in the total input usage by the manufacturing sector, ω^{op} is the share of sector j in the total output of the manufacturing sector for final use (i.e., net of input usage).

We first discuss the construction of the input weights. To compute the input tariff weights, we focus on input usage across all manufacturing sectors. We restrict to intermediate inputs produced by the manufacturing sector.[68] For output weights, we construct weights based on the net value of output for each manufacturing IO sector, calculated as the sector's total output minus its intermediate use as inputs in other sectors. This net output measure reflects the sector's value of final use and ensures that sectors producing primarily intermediate goods do not disproportionately influence average output tariffs.

[68] Since we do not model the use of intermediates in agriculture, and most intermediates used in agriculture are themselves agricultural goods, omitting agricultural inputs from manufacturing effectively removes the same set of inputs from both sectors. In any case, once we exclude the food sector, agricultural inputs account for less than 10 percent of total intermediate use in manufacturing, and our welfare results are not meaningfully affected by whether these inputs are included. Finally, the time sensitivity of agriculture intermediates might be higher than the estimated time sensitivity of the broader class of manufacturing intermediates.

Agriculture prices: We use data from the World Bank Pink Sheet database on commodity prices, with the exception of milk, to calculate the average agriculture prices in two periods, 1975-89, and 1990-04. Since commodity prices fluctuate from year to year, we choose pre and post liberalization windows that each span 15 years to ensure a consistent basis for averaging.[69] We use the real annual prices for each commodity.[70] We consider all the agriculture sectors from the 1993–94 Input-Output (IO) of India, and match each sector with the corresponding commodity in the World bank database. For a sector where more than one commodity price index is available (for eg. see coffee below), we first calculate the change in average price across the two periods for each commodity, and then take the average of the price changes across all commodities in the sector. Table A.7 shows the details on agriculture sectors, corresponding commodity matches, and the share of each sector in total final agriculture consumption in India.

We are unable to match any commodities with six sectors – Bajra, Gram, Pulses, Jute, Coconut, and Animal Services. Fortunately, the combined total of these sectors in total agriculture output is only 8.7%, so we can safely exclude them for the purpose of calculating the average change in world price of agriculture.

Using the above matching, for each sector (s) we calculate the change in world agriculture prices, $\widehat{P_s} = \frac{Price_s^{1990-04}}{Price_s^{1975-89}}$. We then take the average of price changes across sectors weighted by the corresponding share of each sector in India's final agriculture consumption to get the average change in world agriculture price as,

$$\widehat{P_A^*} = \sum_s \widehat{P_s} * Share_s$$

This gives us a value of 0.78 for the average change in world agriculture price.

B.3.5 Validating Relative Price Changes

We compare the change in the relative price of agricultural to manufacturing output that we feed into the model with official data on the terms of trade between these two sectors over 1981–2003. To obtain these data, we use the Economic Survey of India reports for FY 2001–02 and FY 2004–05, which provide annual wholesale prices index of agricultural commodities relative to manufactured products from 1981–82 to 2004–05. In these reports, the agricultural-to-manufacturing wholesale price index averages 107.94 over the pre-liberalization period (1981–82 to 1989–90) and 120.10 over the post-liberalization period (1990–91 to 2003–04). This corresponds to an increase of about 11.27% in the relative price of agricultural goods, indicating that agriculture becomes more expensive relative to manufacturing after liberalization.[71]

We now compare this change to the relative price movements that we impose in the model.

[69] Results of the model are robust to using a smaller pre-period, for instance a period of 1979-89

[70] For the Milk sector, we use data on raw cow milk price from the CLAL data base for 6 EU countries – Belgium, Germany, Ireland, Netherlands, Denmark and Spain. The nominal prices are deflated by the CPI of each country to get the real prices and then averaged across countries. The data is available since 1977.

[71] All available publications of the Economic Surveys can be found here.

Table A.7: Average Price across Sectors in Agriculture

Sector	Mapped Commodities	Sector Description	Sector Share
1	ricethai5	Paddy	0.161588
2	wheatushrw, wheatussrw	Wheat	0.085327
3	barley	Jowar	0.010885
5	maize	Maize	0.009310
8	sugarworld	Sugarcane	0.045419
9	groundnuts	Groundnut	0.023274
11	cottonaindex	Cotton	0.025542
12	teaavg3auctions, teacolombo, teamombasa	Tea	0.005460
13	cocoa, coffeearabica, coffeerobusta	Coffee	0.002725
14	rubberrss3	Rubber	0.002413
16	tobaccousimportuv	Tobacco	0.003193
17	bananaus, orange, sorghum, soybeans	Other crops	0.260264
18	raw cow milk EU	Milk and Milk Products	0.134527
20	beef, chicken, lamb	Other livestock products	0.074651
21	logscameroon, logsmalaysian, sawnwoodmalaysian	Forestry and logging	0.035179
22	fishmeal, shrimpsmexican	Fishing	0.032793
4	.	Bajra	0.005122
6	.	Gram	0.0161
7	.	Pulses	0.0276
10	.	Jute	0.0022
15	.	Coconut	0.0114
19	.	Animal Services	0.024

Note: The table provides details on the agriculture sectors in the IO table, along with the commodities that are matched with each sector from the World Bank Pink Sheet Data, with the exception of Milk that uses the CLAL database. The share of each sector in total agriculture final consumption is shown.

First, from the exercise described in the main text, the change in the agricultural price level is $\widehat{P}_{Ad} = 0.78$. Second, for district d, the change in the manufacturing price from Equation 26 is

$$\widehat{P}_{Md} = \frac{1.3420 + t_{Md}}{1.9328 + t_{Md}}.$$

The district level change in the relative price of agriculture to manufacturing is therefore

$$\widehat{\left(\frac{P_{Ad}}{P_{Md}}\right)} = \frac{\hat{P_{Ad}}}{\hat{P_{Md}}} = \frac{0.78}{\left(\frac{1.3420+t_{Md}}{1.9328+t_{Md}}\right)}.$$

To obtain a national average change in the relative price, we consider two approaches. First, we abstract from trade costs and set $t_{Md} = 0$ for all districts, which yields an average relative price change $\widehat{\left(\frac{P_{Ad}}{P_{Md}}\right)} = 1.123$, corresponding to a 12.3% increase in the price of agriculture relative to manufacturing. Second, we incorporate trade costs and compute the average manufacturing price change across districts as

$$\tilde{\hat{P_{Md}}} = \frac{1}{\mathcal{D}} \sum_{d \in \mathcal{D}} \hat{P_{Md}},$$

and then construct the aggregate average relative price change as $\frac{\hat{P_{Ad}}}{\tilde{\hat{P_{Md}}}}$. We obtain $\tilde{\hat{P_{Md}}} = 0.71$, which implies a 9.9% increase in the average relative price of agriculture to manufacturing between the pre and post liberalization periods.

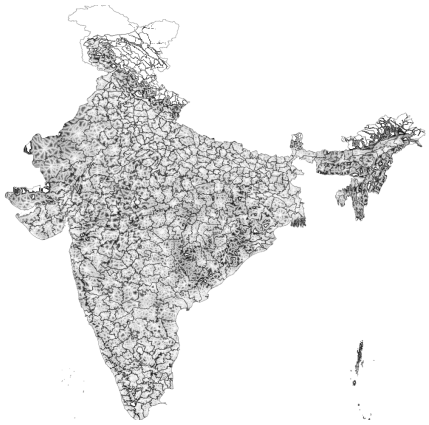
Thus, across both calculations, the relative price changes in the model closely track the official

terms of trade series, giving us confidence that the price changes we feed into the model provide an empirically grounded representation of the evolution of agriculture to manufacturing relative prices in India.

B.3.6 Global Accessibility Map

We use the Global Accessibility Map developed by the European Commission and the World Bank for the year 2000 to assign a cost or friction of traveling across any 1x1 km pixel based on an algorithm that includes transport network, political factors and environmental factors. Figure A.5 shows the distribution of this cost across India.

Figure A.5: Accessibility Map of India



Note: The figure shows the cost or friction of traveling across any 1x1 km pixel across India developed by the European Commission and the World Bank.

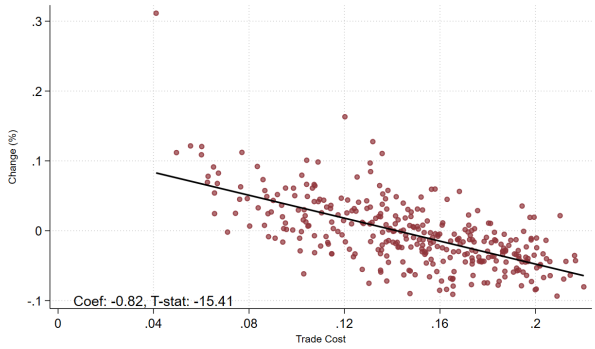
B.4 Additional Model Results

Figure A.6 plots the change in total population against initial trade cost in response to trade liberalization without the counterfactual road network (Panel (a)) and with the counterfactual road network (Panel (b)). The slope of population change becomes flatter by a factor of more than four, indicating that targeted road investments substantially reduce worker outflows from initially low to high international access regions.

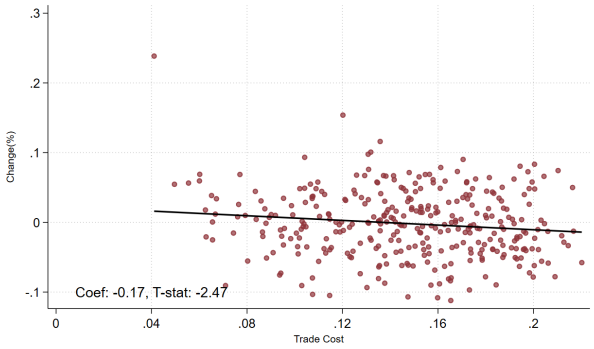
Figure A.7 shows the distribution of real wage changes. Comparing Panel (a) and Panel (b), we see a marked reduction in the unevenness of real wage gains, with the magnitude of the negative slope falling by a factor of 4.45.

Figure A.6: Trade Liberalization and Population

(a) Without Optimal Roads



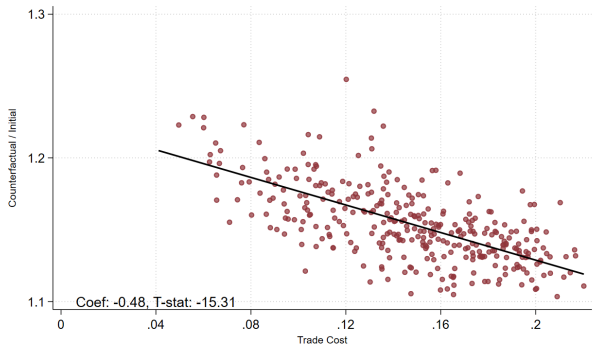
(b) With Optimal Roads



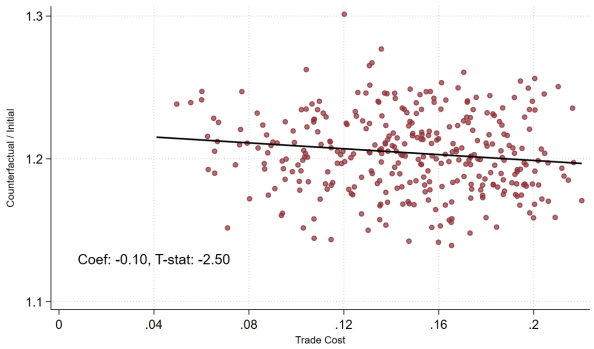
Note: The figure shows the role of international access in the effects from trade liberalization on population, with and without the simultaneous investments in a counterfactual road network.

Figure A.7: Trade Liberalization and Real Wage

(a) Without Optimal Roads



(b) With Optimal Roads



Note: The figure shows the role of international access in the effects from trade liberalization on real wage to workers with and without the simultaneous investments in a counterfactual road network.