Domestic infrastructure and the regional effects of trade liberalization*

Maarten Bosker† Megan Haasbroek‡

Working paper: April 28, 2025

Abstract

We use detailed historical data on India's domestic infrastructure to show how its high domestic transport costs have conditioned the local labor market consequences of its drastic import tariff liberalization in the early 1990s. We find that districts located farther away from the country's main international gateways are better shielded from the resulting increased foreign import competition: their non-agricultural employment falls less than in otherwise similarly exposed districts located closer to India's major ports. At the same time, they also benefit less from improved access to foreign intermediates: non-agricultural employment increases less than in districts with a similar input-output structure, but located closer to the country's main ports. These employment responses also vary across firms of different sizes: employment in small to medium sized firms is hit hardest by increased import competition, whereas employment in medium to large firms benefits most from better access to foreign intermediates. This difference between small and large firms is also most pronounced in districts best-connected to India's major ports.

JEL classification: F14, F15, R11
Keywords: gains from trade, domestic infrastructure, local labor demand, India.

*We would like to extend special thanks to Mark Roberts for his invaluable role in organizing the data collection process, and to Katie L. McWilliams and Arzoo Malhotra for digitizing and operationalizing the historical infrastructure network. Furthermore, we would like to thank Banu Demir, Guido Imbens and Sacha Kapoor, and seminar participants at European Trade Study Group (2022), Erasmus University Rotterdam and University of Cambridge for valuable comments and discussions. Megan Haasbroek gratefully acknowledges the support of the Janeway Institute at Cambridge.

†Erasmus University Rotterdam, The Netherlands, Tinbergen Institute and CEPR. bosker@ese.eur.nl
‡Janeway Institute, University of Cambridge. mh2310@cam.ac.uk

I. Introduction

For many low- and middle-income countries, the costs involved in shipping goods domestically are substantial and often even larger than those paid for the international shipment of goods in and out of these countries [Donaldson, 2015, Atkin and Donaldson, 2015]. Early contributions, by e.g. Krugman and Elizondo [1996] and Hanson [1998], already showed that such domestic trade frictions can have a profound impact on regional economic development patterns following a country's opening up to international trade. Coşar and Fajgelbaum [2016], Sotelo [2020], Xu and Yang [2021], Fan [2019], and Fajgelbaum and Redding [2022], among others, extend these early models to a full-fledged multi-region spatial general equilibrium model of a country's economic geography. Their main message is very much the same: the effects of international market integration depend on a region's access to the country's main international gateways. For example, regions closer to these gateways will see their export opportunities increase by more than those located farther inland. At the same time, the high costs of domestic transportation also shield these inland regions from foreign competition.

The contribution of our paper is to provide quasi-experimental evidence on the role of domestic infrastructure in conditioning the labor market effects of import tariff liberalization. More specifically, we follow e.g. Topalova [2007] and De Loecker et al. [2016], and make use of the natural experiment of India's rapid and unexpected opening up to world markets in the early 1990s. These earlier papers show that this episode affected India's districts differentially depending on their industrial structure prior to this shock. We expand on these findings by bringing novel detailed historical information on India's road and rail network, the exact location of its major international ports, as well as that of its main inland intermodal transshipment hubs, into the picture.

Using this data we assess whether, and if so how, exposure to the reductions in India's import tariffs affected districts differentially depending on their access to India's main international gateways. India's rapid import tariff liberalization provides an ideal setting to study this for two reasons. First, transporting goods domestically was – and still is – very costly at the time of the 1991 trade reforms, and varied substantially across the country both because of differences in the availability of (good quality) infrastructure as well as internal policy barriers to trade. Second, India's extremely high import tariffs in place prior to its trade reforms made the domestic transport costs involved in shipping to one of India's main international ports an unlikely input in firms' location decisions. Moreover, most of its domestic infrastructure investments during the pre-reform period were geared toward improving intra-India connectivity.[1] Combined with the rapid, unexpected, and substantial trade reforms, this alleviates identification concerns related

[1] In Appendix II we show that road construction and improvements in the decade before the trade reforms was not particularly geared to connecting districts to India's main ports. Instead, it was aimed at connecting the larger cities and industrial centers in the country, as well as further integrating its less-connected districts into the network.

to the endogenous placement of infrastructure in anticipation of future trade flows or as a response to already-existing bottlenecks in these flows.

To guide our empirical analysis, we incorporate domestic trade frictions in the simplest possible way in the model developed by Kovak [2013] and Dix-Carneiro and Kovak [2015], while at the same time also extending the scope of the model to the use of foreign intermediates in production.[2] Falling import tariffs in this setting have *two opposing effects* on a district's local labor market, where both are conditioned on the size of the domestic trade costs involved in shipping imported goods from the port to the district.

First, lower import tariffs lead to fiercer import competition, which depresses labor demand. This effect is less pronounced in the more inland destinations, as the domestic transport costs involved in shipping foreign goods to these districts are higher. Second, Indian firms importing foreign inputs benefit from a decrease in tariffs: cheaper intermediate goods allow them to produce more, but to do so they also need more labor, thus increasing labor demand. This effect is also weaker in more inland destinations. The higher domestic transport costs involved in shipping foreign intermediates to these districts imply that firms benefit to a lesser extent from a reduction in tariffs on their intermediate inputs than their counterparts located closer to the port.

Our empirical findings are mostly in line with these predictions. They show a persistent impact of India's rapid trade liberalization that becomes more pronounced over the years, and that depends non-trivially on a district's access to its nearest main international port. Non-agricultural employment falls in regions more exposed to output tariff reductions, and this effect is significantly stronger in regions closer to India's main international ports. In contrast, regions benefiting more from the reduction in input tariffs show faster non-agricultural employment growth, and this effect is also more pronounced in regions closer to India's main ports. Conditional on industrial structure, the effect of import competition (access to foreign intermediates) on local employment is about 50% (17.5%) smaller in the median district compared to a district right next to one of India's main ports. Overall, the positive effect of improved access to foreign intermediates dominates, resulting in overall positive labor demand effects throughout India, but districts closest to the port reap the smallest gains due to the larger negative effects of import competition in those districts.

These findings are robust to using different mappings of our information on India's road, rail, port and inland transshipment hubs to proxy actual (unobserved) transport costs. Moreover, and interestingly so, we find that the effect of import competition is strongest for small- and medium-sized firms, whereas the benefit from cheaper intermediates is strongest for medium-sized and large firms. The former corroborates predictions by e.g. Melitz and Redding [2014b], where smaller, less productive firms are less able to withstand the increase in foreign competition. The latter is in line with Nataraj [2011]:

[2] In the Online Appendix I, we also present a full-fledged multi–district, multi–sector Ricardian model in the spirit of Caliendo and Parro [2015], that – importantly for the case of India – features limited labor mobility [Topalova, 2010] and foreign intermediate use in production.

smaller firms are less likely to use foreign intermediates in their production processes in the first place, and can thus be expected to benefit less from the improved access to cheaper foreign inputs. As a consequence of India's high domestic transport costs, these differential effects on firms of different size are also most pronounced in districts closest to India's main international ports.

Our paper speaks to a number of different strands of research. First, we add to a, by now well-established, empirical literature that uses the rapid opening up to trade of developing countries as a "natural experiment" to identify the differential effects of trade liberalization across regions within these countries depending on their industrial structure ex ante. Important examples concern liberalization episodes in Indonesia [Kis-Katos and Sparrow, 2015], Vietnam [McCaig, 2011, McCaig and Pavcnik, 2013], India [Topalova, 2010], or Brazil [Kovak, 2013, Dix-Carneiro and Kovak, 2015, 2017]. What we add to this literature is a careful analysis of the role of a country's domestic infrastructure in conditioning these effects. We do this using our newly collected, detailed, historical data on India's rail, road, port and inland transshipment infrastructure, combined with other historical sources that contextualize the broader transport sector in India at the time of its trade reforms.

This data allows us to create proxies of the domestic transport costs involved in shipping goods to/from each district from/to the country's main international gateways that go beyond a simple coastal versus non-coastal classification, or using a district's geodesic distance to the nearest coast, as in e.g. McCaig [2011], McCaig and Pavcnik [2013], and Coşar and Fajgelbaum [2016]. Also, we use our historical infrastructure data to confirm that there is no evidence of endogenous infrastructure placement in anticipation of the tariff reforms: road construction and upgrading in the pre-reform period was not geared toward districts most (or least) exposed to India's import tariff reductions.

Second, our paper also speaks to the literature that identifies the role of access to a country's international gateways in conditioning the regional effects of a country's exposure to international markets as in Coşar and Fajgelbaum [2016], Coşar and Demir [2016], Storeygard [2016], Sotelo [2020], van Leemput [2021], Volpe Martincus et al. [2017], Brülhart et al. [2012], Fajgelbaum and Redding [2022] and Xu and Yang [2021]. Our contribution to this literature is twofold. First, we are able to isolate how domestic transport costs condition the effects of *import tariff liberalization* rather than a shock to international trade costs more generally. India's trade reforms in the early 1990s mainly focused on facilitating imports rather than exports, and were not accompanied by large-scale domestic infrastructure improvements.[3] Compared to cases where countries both lowered their import and export barriers as in Brülhart et al. [2012], this allows for a much clearer interpretation of our findings, as we do not have to worry about domestic transport costs also conditioning the effect of improved export opportunities on aggregate

[3] As we will discuss in Section III, export restrictions were also lifted, albeit very gradually. Moreover, the increase in exports after 1991 was mainly driven by services rather than manufacturing.

labor outcomes (as in McCaig [2011] or Coşar and Fajgelbaum [2016]).

Second, and very much related, India's import tariff reductions increased both firms' exposure to import competition, as well as providing much improved access to foreign intermediate inputs [De Loecker et al., 2016]. These two effects have the opposite predicted effect on local labor demand, with domestic transport costs affecting both. Using India's 1993-94 Input-Output Tables, we are able to separately identify these two channels and compare their relative importance. Earlier papers considering the local (labor market) consequences of a country's import tariff reductions have primarily focused on the effect of the resulting import competition (see e.g. Topalova [2010] or Dix-Carneiro and Kovak [2017]), hereby also running the risk of biased inference in case regions' exposure to increased import competition is correlated with their exposure to having improved access to foreign intermediates.

Third, we complement papers that identify the role of changes in domestic transport costs in determining regions' participation in international markets and its consequences using a plausibly exogenous source of variation. Martincus and Blyde [2013] use geographically localized earthquake damage to Chile's domestic infrastructure as a shock to regions' access to the country's main international gateways to identify its effect on a region's import and export performance. Coşar and Demir [2016] show that public investments in road infrastructure in Turkey significantly increased imports and exports of those interior regions now better connected to Turkey's main international gateways; Volpe Martincus et al. [2017] show similar results for Peru. We instead focus on identifying how a country's domestic infrastructure conditions the effect(s) of changes in a country's barriers to *international* trade. Hereby we provide important complementary evidence into the role of domestic transport in shaping regions' participation in international markets, especially as theory predicts that shocks to a country's domestic transport costs can have very different effects than shocks to a country's international trade costs [Coşar and Fajgelbaum, 2016, Fajgelbaum and Redding, 2022].

Fourth, and finally, in extensions to our main findings, we provide evidence that our main focus on district-level aggregate labor demand obscures the heterogeneous impact of import tariff liberalization on firms of different sizes, relating to the literature on heterogeneous firms and trade [Melitz and Redding, 2014a]. In line with predictions from that literature, we find that the effects of decreases in output tariffs are concentrated among small- and medium-sized firms, as the resulting fiercer import competition pushes them to exit the market first [Melitz, 2003]. At the same time, we find that the employment gains from a fall in input tariffs are concentrated in medium to large firms, since these firms are much more likely to actually use these imported intermediates [Amiti and Konings, 2007, Nataraj, 2011, Topalova and Khandelwal, 2011].

The remainder of the paper is organized as follows. The next section presents our conceptual framework. Section III discusses India's opening up to trade and its domestic transportation network. Then, section IV describes our data, and section V our empirical methodology and identification strategy. Finally, section VI presents our findings, and section VII concludes.

II. Conceptual framework

The simplest way to show how domestic transport costs condition the local labor market effects of opening up to trade is to draw on the specific-factors model of regional economies introduced in Kovak [2013]. In this model, a region's local labor demand response to trade liberalization is captured by the following weighted sum:

$$\sum_{i \in S} \beta_{ri} \hat{P}_i, \quad \text{with } \beta_{ri} = \frac{\lambda_{ri} \frac{1}{\phi_i}}{\sum_{j \in S} \lambda_{rj} \frac{1}{\phi_j}}, \tag{1}$$

where $\hat{P}_i$ is the trade liberalization-induced price change in industry $i \in S$, the set of all tradable industries. The impact of this specific industry's price change on local labor market outcomes in region r is captured by β_{ri}, where λ_{ri} depicts the share of employment in industry i over total tradable employment in region r and ϕ_i is 1 minus the labor cost share in sector i. The latter reflects differences in labor demand elasticity across different industries. Intuitively, if an industry is more prominently present in region r and/or has a more elastic labor demand, a reduction in tariffs for this industry will have more weight in determining equilibrium labor demand changes. Thus, while all regions face the same tariff liberalization-induced price changes, the impact of these price changes on local labor demand differs depending on a region's industrial composition.

In the presence of domestic trade frictions, regions will however typically not face a similar price response when a country lowers its import tariffs [Atkin and Donaldson, 2015, Coşar and Fajgelbaum, 2016]. The simplest way to incorporate this in the above setting is to introduce additive domestic transport costs.[4,5] In order to buy a good from industry i from the Rest of the World, one does not only have to pay import tariffs, but also the costs of shipping the good from the port to region r, as formulated below:

$$P_{ri}^W = P_i^W (1 + \tau_i) + d_r, \tag{2}$$

where P_i^W is the world price of good i, τ_i the tariff in industry i and $d_r \geq 0$ refers to the specific transport costs to ship good i from the port to region r.[6] Here, we assume that it is always cheapest for any region to directly ship the good from the port to the region, instead of trading via another region (i.e. the triangular inequality holds).

[4] Additive transport costs make up a large part of total transport costs: e.g. Daudin et al. [2022] show that additive transport costs equal between 30 to 45 percent of total transport costs.

[5] The assumption of additive transport costs is not necessary for tariff liberalization to have a differential impact across space; in Appendix I we show that the same predictions follow from a more elaborate multi-region multi-sector Ricardian trade model featuring multiplicative domestic trade costs and limited labor mobility.

[6] Without loss of generality, we take these transport costs to be the same for each sector in a district.

Now, assuming as in Kovak [2013] that all regions face the same production technologies, we can simply follow Dix-Carneiro and Kovak [2017] and derive the effect of a tariff shock on local regional labor demand by plugging Equation (2) into (1). We call this a region's "output tariff exposure" or *OTE* [7]:

$$OTE_r = \sum_{i \in S} \beta_{ri} \ \mathrm{d} \ln \left(P_i^W (1 + \tau_i) + d_r \right). \tag{3}$$

Without any domestic transport frictions, this would simply collapse into the *RTR* measure of Dix-Carneiro and Kovak [2017]: $RTR_r = \sum_{i \in S} \beta_{ri} \ \mathrm{d} \ln(1 + \tau_i)$.

How does the addition of specific transport costs change the passthrough of the lower import tariffs to local labor markets? Taking the derivative of Equation (3) with respect to d_r yields:

$$\frac{\partial OTE_r}{\partial d_r} = \sum_{i \in S} \beta_{ri} \left[\frac{1}{P_i^W (1 + \tau_i') + d_r} - \frac{1}{P_i^W (1 + \tau_i) + d_r} \right], \tag{4}$$

which is positive when the new tariff $\tau_i' < \tau_i$. As a tariff decline implies a negative shock to labor demand, this means that regions facing higher costs in transporting imported goods from the port are less exposed to the labor demand effects of the tariff reductions. More precisely: consider two regions with the exact same industrial composition, one next to the port and the other more inland. Then, the negative local labor market demand shock to the coastal region is larger than the corresponding shock to the inland region.

This negative local labor demand shock arises due to the increased competition from abroad that firms face following trade liberalization. However, this competition channel is not the sole consequence of trade liberalization; import tariff reductions also provide firms with access to cheaper intermediates [Caliendo and Parro, 2015]. Recent research has shown that this not only results in increased profitability, but also induces firms to take productivity-enhancing decisions [Amiti and Konings, 2007, Topalova and Khandelwal, 2011, Fiorini et al., 2021]. Not accounting for input tariff liberalization can thus underestimate the gains from trade liberalization, and might even bias our inference regarding the effect of import competition when firms facing the strongest/smallest increase in import competition also benefit the least/most from the now-cheaper foreign intermediates [Nataraj, 2011].

The setup in Dix-Carneiro and Kovak [2017] does not incorporate this intermediate input channel. To illustrate its effect, we therefore follow Amiti and Konings [2007] and Kis-Katos and Sparrow [2015], and construct an industry's "input tariff exposure" (*ITE*) that captures how the now cheaper access to foreign intermediates affects regions' local

[7] Strictly speaking, Dix-Carneiro and Kovak [2017] define their regional tariff reduction to be the negative of this equation to facilitate interpretation.

labor demand[8]:

$$ITE_r = -\sum_{i \in S} \beta_{ri} \left[\sum_{j \in \mathcal{J}} \alpha_i^j * \, \mathrm{d} \ln P_{rj}^W \right]. \tag{5}$$

where α_i^j is the cost share of the good produced by industry j in the total material costs of industry i, and with β_{ri} as defined in (1). Note that our measure of *ITE* is negative to highlight that a decline in input tariffs is a positive labor market demand shock.

As firms and consumers in region r face the same world price for each imported good i, the same tariff reductions, and the same domestic transport costs, the price of imported goods is simply the same as in (2). Thus, we can write *ITE* as:

$$ITE_r = -\sum_{i \in S} \beta_{ri} \left[\sum_{j \in \mathcal{J}} \alpha_i^j * \, \mathrm{d} \ln(P_j^W(1+\tau_j) + d_r) \right]. \tag{6}$$

Without domestic trade frictions this would collapse to $-\sum_{i \in S} \beta_{ri} \left[\sum_{j \in \mathcal{J}} \alpha_i^j * \, \mathrm{d} \ln(1+\tau_j) \right]$, and the impact of import tariff reductions on a region's local labor demand would again only depend on its industrial composition: the larger the share of its labor force employed in firms making more intensive use of foreign intermediates whose import tariffs declined most, the larger this impact will be.

The presence of domestic transport costs changes the pass-through of this effect depending on a region's access to the country's main international gateway. To see this, take the derivative of (6) with respect to d_r:

$$\frac{\partial ITE_r}{\partial d_r} = \sum_{i \in S} \beta_{ri} \sum_{j \in \mathcal{J}} \alpha_i^j \left[\frac{1}{P_j^W(1+\tau_j) + d_r} - \frac{1}{P_j^W(1+\tau_j') + d_r} \right], \tag{7}$$

which is negative when the new tariff $\tau_j' < \tau_j$. This shows that regions facing higher costs in transporting imported goods from the port are less exposed to this positive intermediate input effect of trade liberalization. More precisely: consider two regions with the exact same industrial composition, one next to the port and the other more inland. Then, the positive local labor demand shock to the coastal region, as a result of its better access to foreign intermediates, is larger than the corresponding shock to the inland region.

Summing up, the predicted local labor demand effects of import tariff liberalization in the presence of domestic trade frictions are:

1. **Import competition pass-through:**

   (a) Exposure to import competition following output tariff liberalization depends on a region's industrial composition. Higher exposure implies larger negative local labor market demand effects.

[8]The more elaborate model in Appendix I shows this more formally, yielding similar predictions as the ones we outline here.

(b) This effect decreases in a region's domestic transport costs to its main international gateway.

2. **Cheaper intermediates pass-through:**

   (a) Exposure to cheaper intermediates following input tariff liberalization depends on industrial composition. Higher exposure implies larger positive local labor market demand effects.

   (b) This effect decreases in a region's domestic transport costs to its main international gateway.

In the rest of the paper, we test these predictions focusing on India's rapid opening up to world markets in the early 1990s.

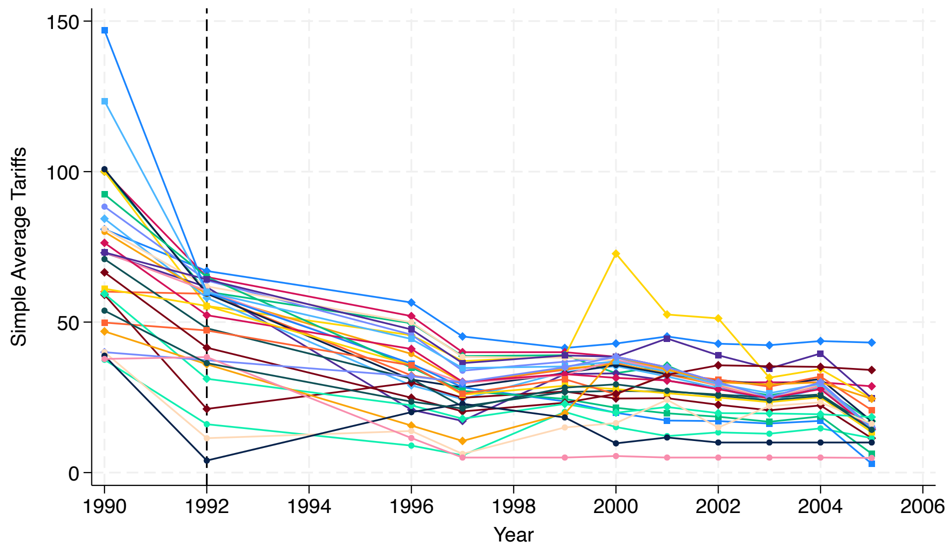
III. India's trade reforms and domestic infrastructure

I. India's trade reforms

After its independence, India decided on a highly protectionist regime, utilizing both tariff and non-tariff barriers to restrict imports [Pandey, 2004]. This came to an abrupt end in 1990, when India, faced with a severe balance-of-payments crisis, was forced to apply to the International Monetary Fund (IMF) and the World Bank for a bailout to fund its obligations [Topalova, 2007]. This bailout was conditional on several structural reforms in the Indian economy. Enacted in July 1991, these reforms entailed substantial reductions in tariffs, increased allowances for foreign direct investment (FDI), all but the abolition of the licensing system, removing exchange rate controls and industrial policies aimed at reducing barriers to entry.

The tariff liberalization was anchored on two principles: harmonization, i.e. reducing the dispersion across tariffs, and rationalization, which involved combining tariff lines to simplify the existing system [Topalova, 2010]. This is reflected in Figure 1, which shows the development in effectively applied tariffs across different manufacturing sectors between 1990 and 2005. The subsequent increase in trade flows was substantial in both absolute and relative terms: during the 1990s, the annual growth rate of imports increased by 3.3 percentage points to 9.2 percent, and the trade-to-GDP ratio almost doubled from 17.2 to 30.6 percent [Panagariya, 2004].[9] Export restrictions were also reduced, but much

[9] This quantity increase does not fully reflect all gains from trade, as the liberalization also allowed firms and consumers alike to purchase higher-quality foreign substitutes of domestic inputs and outputs [Panagariya, 2004]. In addition to better inputs improving firm-level productivity and innovation, as in e.g. Amiti and Konings [2007], Goldberg et al. [2010a], this would also give rise to the creation of new markets and the disappearance of others. Goldberg et al. [2010b] document that a quarter of overall manufacturing growth after India's trade liberalization was driven by the creation of new products. Although we can implicitly take part of this mechanism into account in the model by assuming that the labor demand shock is a function of *quality-adjusted* prices, our level of aggregation does not allow for estimating the effect of



Note: Tariff information is retrieved from the World Integrated Trade Solution database. Each line depicts the tariff development for a two-digit manufacturing industry.

Figure 1: Change in manufacturing tariffs between 1990 and 2005

more slowly than import tariffs. Not only is the increase in exports after 1991 much less stark than that of imports, this expansion is mostly driven by growth in services exports [Panagariya, 2004].

The Indian trade liberalization offers an attractive setting in which the gains from trade can be studied, as it does not suffer from many of the endogeneity concerns that trade openings are often encumbered with. First, as mentioned before, the trade reforms were unexpected, subsiding any concerns of reverse causality. Liberalization was one of the conditions to receive the bailout from the IMF, and was therefore not planned or expected by politicians, enterprises and civilians alike [Topalova, 2010]. Second, the tariff liberalization was across the board rather than selective. According to Gang and Pandey [1996], this is because pre-reform tariffs were based on the Second Five Year Plan (1956-1961), which was drafted when the Indian economy was structured quite differently. As a result, the tariff levels before the reform do not necessarily reflect current protectionism incentives [Topalova, 2010]. Indeed, earlier work has shown that the extent of tariff reductions up until 1997 was not influenced by pre-reform sector characteristics or political economy concerns [Topalova, 2004, 2007].

Finally, there are three other important reforms that were part of the IMF bailout package . First, import licensing requirements on capital and intermediate goods were abolished, but remained in place for consumer products [Bhat, 2011]. Second, the License Raj that limited firms operating in certain sectors was slowly dismantled, and while

---

market creation and market competition separately.

there is no evidence that the delicensing of industries was influenced by expected future performance or otherwise strategic, its effect on industrial performance varied by state based on general labor regulations [Aghion et al., 2008]. Finally, there is evidence that FDI liberalization was selective, as profitable and concentrated industries were less likely to be liberalized [Chari and Gupta, 2008]. All three of these other reforms were implemented more gradually than India's import tariff reductions; nor is there any evidence that either of them were systematically related to industries' reduction in import tariffs. Thus, these other reforms are unlikely to confound the impact of India's tariff liberalization; we nevertheless take explicit account of them in our empirical analysis (see Section V).

II. India's domestic infrastructure

II.i. International gateways

The vast majority of India's international trade, both in value and volume terms, is moved by sea rather than by air or overland [NTDPC, 2014a]. Air transport was negligible, accounting for less than one percent of all freight traffic in 1990 [KPMG, 2013]. Overland trade used to be of importance during British colonial rule until Partition in 1947, when links between India and most of its neighbors were suspended and have not recovered since. According to the NTDPC [2014b], *"intra-South Asian trade, which was 18.4 percent of the total trade in 1948, dropped to 4.3 percent by 2010"*. As such, virtually all of India's international trade was going through one of its main international ports at the time of its trade liberalization.

Of these imports and exports, more than 92 percent are handled by one of India's major ports, which fall under the jurisdiction of the Ministry of Shipping [MoPSW, 2003]. In 1990, these major ports were Mumbai, Kandla, Mormugao, New Mangalore, JNTP and Cochin along the west coast, mostly engaging in trade with Europe, America, the Middle East and Africa; while the east coast's major ports – Kolkata, Haldia, Visakhapatnam, Chennai, Tuticorin and Paradip – mostly trade with Asia and the Pacific [Haralambides and Behrens, 2000].[10] The more than 150 other, minor, ports are overseen by their respective State governments. The lack of traffic for these minor ports in the 1990s is due to their generally inadequate infrastructure, both in terms of low berth and lack of loading and handling facilities as well as limited hinterland connectivity, making them especially unattractive for international trade [NTDPC, 2014c][11]. We restrict our attention to India's major ports as they are responsible for almost all of India's international trade.[12]

[10] See section IV for more detail on these ports as well as on the specific information about these ports that we use in our analysis.

[11] In fact, only about 60 of these ports were fully operational [Haralambides and Behrens, 2000].

[12] India's major ports also handled 90% of India's internal coastal trade in 1990 [MoPSW, 2003].

II.ii. Getting to the port: India's domestic transport network

India's rail and road network is the second largest in the world, but domestic transport remains a costly activity: shipping costs by rail are around 70 percent larger in India than in the US, and freight costs for trucking are 30 percent higher [Kumar, 2014].[13] An insight into the substantial domestic transport costs involved in shipping (imported) goods across India is provided in Table 1. This shows the distance, travel time and hauling costs by rail and road between selected ports and Delhi in 2008.[14] Note that these routes are some of the best developed transit corridors in India, so that these costs are most likely a substantial underestimate of the costs involved in shipping goods along other routes.

Table 1: Domestic transport costs between selected ports and Delhi

Port	Distance to/from Delhi (km)	Transit time by rail (hrs.)	Transit time by road (hrs.)	Haulage costs by rail (Rs./TEU)	Haulage costs by road (Rs./TEU)
JNPT	1,388	48	65	18,750	32,000
Mundra[15]	1,295	80	60	16,650	20,000
Visakhapatnam	1,700	67	79	22,450	66,000
Chennai	2,100	90	98	30,000	70,000

Note: Table displaying distance, travel time and haulage costs for road and rail between selected ports and Delhi in 2008. TEU is a twenty-foot equivalent unit, the standard volume of an intermodal container. For reference, 20,000 rupees in 2008 corresponds to approximately 460 US dollars. Data is retrieved from World Bank Group [2013].

The table highlights two important features about India's international market access. First, domestic transport costs are substantial. In 2008, the average cost of shipping a twenty-foot equivalent unit (TEU) container from Northern Europe to Asia was about 1,000 US dollars [UNCTAD, 2008], or 43,500 Rs. This trip is 25,000 kilometers and takes around 25 days. Yet it is only 2.3 (1.4) times as expensive as shipping the container from JNPT to Delhi by rail (road), a trip that is 18 times shorter and takes 2-3 days.

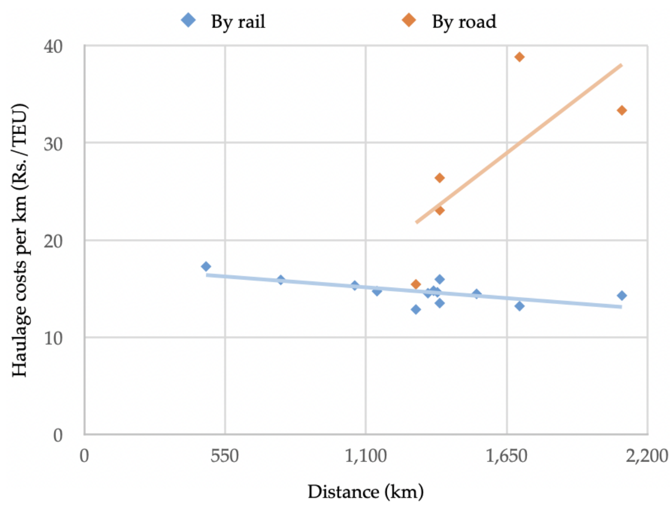
Second, long-distance transport by rail is generally faster and cheaper than by road. One important reason for this are the delays caused by the numerous checkpoints along the route for trucks. Trucks are required to stop at every state border, both to verify state tax levies and to ensure compliance with the Motor Vehicles Act (1988) [Planning Commission, 2009], causing long delays and also often involving government officials soliciting bribes [Barnwal et al., 2024]. Truckers moreover face local police check-ups and other checkpoints along the highway, such that 35 percent of travel time is typically spent on waiting at checkpoints [World Bank Group, 2005].[16]

[13] Two other potential domestic transport modes are coastal shipping and inland water transit (IWT). However, these two together account for less than 8 percent of all domestic freight, so we abstract from these modes in our analysis; more information on these transport modes is relegated to section II in the Appendix.

[14] Unfortunately, information on the exact costs of shipping in the early 1990s is very difficult to find, but, if anything, they are likely to be even higher than those shown in Table 1.

[16] World Bank Group [2005] estimated that inefficient checkpoints and their associated delays cost the

However, do note that this difference between rail and road freight costs is unlikely to carry over to shorter routes, as well as for long-distance transport between less well-served places by the Indian railways. For one, Delhi – always the final destination in Table 1 – is especially well connected to India's rail network, with freight trains departing and arriving frequently. On many other routes, trains only depart once or twice per week. Second, rail transport has less of an edge over road transport on shorter routes, both in terms of costs as well as in flexibility and spatial coverage. In fact, Figure 2 shows that haulage costs per kilometer by road are sharply increasing in the total number of kilometers traveled. For rail traffic, they fall in the total number of kilometers traveled, leveling off at about 14 Rs./TEU per km for trips beyond 1,000 kilometers. To put this into perspective: the median distance to the nearest port of a district's headquarter is 692 kilometers; the 25th and 75th percentile respectively are 447 and 1152 kilometers.



Note: This figure plots the hauling costs per kilometer as a function of total route length as described in Table 1, supplemented with data from Ministry of Railways, India [2011] on the cost of shipping a TEU by rail from Mundra or JNPT port to various ICDs in November 2009.

Figure 2: Haulage costs per kilometer for road and rail (2008/2009)

We discuss the difference (in importance) between road and rail transport in more detail below.

Rail network The Indian railway network, depicted in Figure 3, is managed by Indian Railways (IR) and was already among the largest in the world with its 62,367 kilometers in 1990. Nevertheless, the share of rail in domestic freight tonnage declined from 89 percent in 1951 to around 60 percent in 1991, reaching a new low in 2011 at 30 percent [KPMG,

Indian economy anywhere between 9 billion and 23 billion rupees a year in lost truck operating hours.

2013, NTDPC, 2014c].

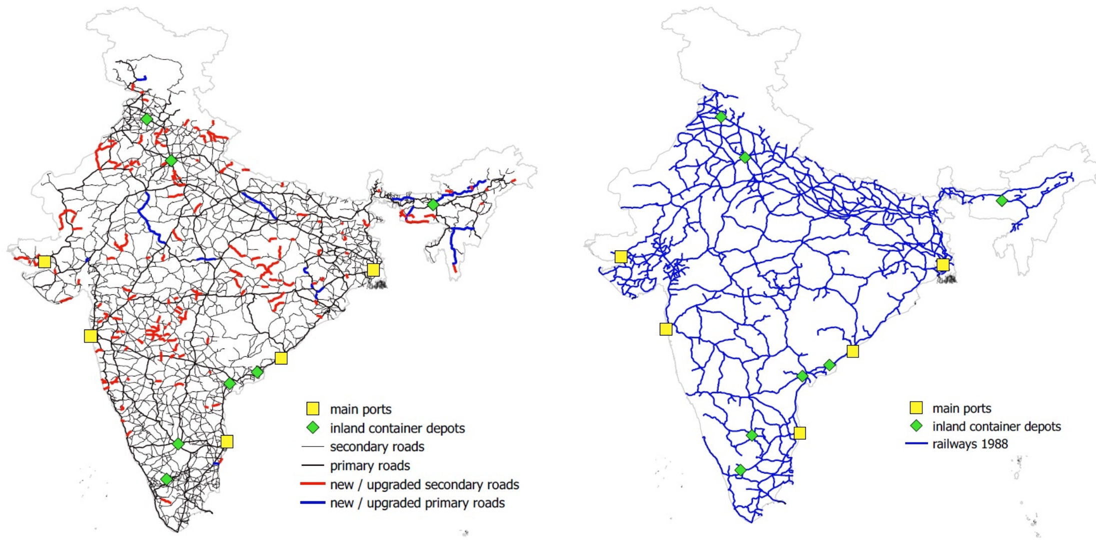
The main reason for this decline is continuous underinvestment: while rail traffic increased almost fourfold between 1951 and 1990, running track kilometers only increased by 32.5 percent [Indian Railways, 2020]. The resulting congestion reduced the average railway speed to 22.7 km/h in 1989-90, and even further on the most important routes [TERI, 2008].[17] The second factor that contributed to the decline of railways' freight market share is the increased cost. The Ministry of Railways overcharges certain commodity transports, using the additional revenue to cross-subsidize passenger transport; in 1990-91, passenger traffic used almost 60 percent of network capacity but only generated around 25 percent of IR's revenue [Indian Railways, 2020]. The aforementioned disincentives, together with its inability to provide door-to-door delivery (compare the spatial coverage of the rail network to that of the road network in Figure 3), and the often low frequency of freight train departures, mean that railway transport is mainly used for higher-value dry bulk items such as thermal coal, coking coal and iron ore rather than manufacturing goods and containers, as well as on the longer, well-traveled routes only [World Bank Group, 2013].

To make rail transport more flexible, Indian Railways invested in the creation of inland container depots (ICDs), which were meant to become well-connected hubs facilitating intermodal transport. The first ICD was established in Bangalore in 1981, followed by six other ICDs in the next seven years. They are indicated by a green diamond in Figure 3. Although the network further expanded to more than 170 functioning ICDs in 2014, the lack of supply chain management services, infrastructure access and legal risks means that intermodal transport is still the exception rather than the rule [Kumar, 2014]. Even though currently almost all ICDs are connected to the rail network, most freight traffic through ICDs is carried out by trucks [Kumar, 2014].

Road network The Indian road network, also depicted in Figure 3, comprised of 399,942 kilometers in 1951, increased to 2,327,362 kilometers in 1991, and now encompasses 6,386,297 kilometers [MoRTH, 2021]. Correspondingly, road freight transport was responsible for barely 10 percent of freight traffic movement in 1951, increasing to almost 40 percent in 1990, and reaching close to 70 percent in 2011 [KPMG, 2013]. This increase in market share is mainly due to road freight being the most flexible – it allows for door-to-door delivery and is not bound by a fixed departure/arrival schedule – and cheapest alternative on many routes. As the trucking market is very fragmented – more than 70 percent of firms possess fewer than five trucks – and therefore very competitive, firms operate with low profit margins [World Bank Group, 2005, NTDPC, 2014a].

Nevertheless, transit times in India are nearly double that of developed countries. This is partly because of the layout of the road network: while the majority of the road

[17] To illustrate, these important, or trunk, routes constituted 16 percent of the total network in 2009-10, but accommodated almost 60 percent of the freight and more than 50 percent of all passengers [NTDPC, 2014c].



(a) Roads (b) Railways

Figure 3: India's rail and road network in 1988

network constitutes unpaved rural roads, the majority of the road *usage* is limited to paved Expressways and National Highways, State Highways and Major District Roads.[18] Specifically, the National roads and State roads comprised respectively 2 and 20 percent of all roads in 1990, but carried 35 and 60 percent of all traffic [World Bank Group, 1996]. Consequently, the road network is subject to substantial congestion resulting in average speeds between 30-40 km/h on highways [KPMG, 2013].[19] Another hindrance is the aforementioned government checkpoint system, which accounts for up to 15 percent of total line-haul time [World Bank Group, 2005]. In 2013, despite intensive efforts by states and the central government to improve road infrastructure, trucks on average drove no more than 300 kilometers per day [World Bank Group, 2013].

IV. Data

I. Administrative data

First, we need to define local labor markets. Given the lack of labor mobility across different districts but a non-negligible migration from rural areas to cities within districts,

[18]National Highways are the responsibility of the Central Government, while the latter two types of roads fall under state jurisdiction.

[19]Even on uncongested roads, speeds above 40 km/h are unlikely for road freight, since trucks generally overload their capacity [World Bank Group, 2005].

we take districts as the relevant labor market [Topalova, 2010].[20] The main challenge in doing this is to accommodate the substantial number of district boundary changes over time; between 1981 and 2001, the number of districts increased by 44%. To ensure consistency over time in our definition of a local labor market, we take the 1987-1988 district boundaries as our baseline and map any new or redefined districts back to these boundaries, using population weights for each intercensal boundary change from Kumar and Somanathan [2016] if necessary.[21] Our final dataset is a balanced panel of 407 districts over the years 1990, 1998 and 2005.[22]

Our main source for local labor market outcomes is the Economic Census (EC), which is a complete count of all economic units in India except for crop production and plantation [Central Statistical Organisation, 1990]. We use the three first available rounds, which are representative of the years 1990, 1998 and 2005. The Economic Census, commissioned by the Ministry for Statistics and Programme Implementation, is conducted alongside the house listing operations of the Population Census. The advantage of this procedure is that the Economic Census also includes informal firms, even those with only one employee. Moreover, it is representative at the village-level. The natural alternative would be to use the Annual Survey of Industries (ASI), the yearly nature of which would allow us to better understand the short-run dynamics of local adjustments to the tariff shock. However, the main drawback of the ASI is that its sampling is stratified at the state-industry level, so that we cannot exploit variation in industrial composition across districts without worrying about substantial measurement error. From the EC, we obtain total employment, as well as the total number of firms, by sector and district, with sectors characterized at the three-digit National Industry Classification level.[23]

One important downside of the EC is that it does not include agricultural establishments. To that end, we supplement our dataset with information from the Primary Census Abstract (PCA), which is a subset from the decadal Population Census. We again use the three earliest available rounds (1991, 2001, 2011) and collect information about district-level population, the total labor force, the share of the labor force working in agriculture and in mining respectively, and, as an indicator of education, the share of the

[20] Note that if the assumption of workers not being able to move freely across districts was invalid, we would obtain no relationship between the district-level exposure to import tariff liberalization and subsequent labor market outcomes [Goldberg and Pavcnik, 2016].

[21] Further details and special cases are relegated to Section III of the Appendix.

[22] In 1988, India had 428 districts; we unfortunately only have data for 407 districts in the 1990 Economic Census. Fourteen districts are in Jammu and Kashmir, which was not surveyed in 1990. Furthermore, we do not have data for Hyderabad and six districts in Madhya Pradesh. According to personal correspondence with a spokesperson from the Ministry of Statistics and Programme Implementation on 6 November 2022, these datafiles were corrupted during the media conversion.

[23] Products in India are categorized according to the National Industry Classification (NIC). The Ministry of Statistics and Programme Implementation (MoSPI) first established this system in 1970, and updated it in 1987 and 1996. We use the NIC from 1996 as our baseline, since ISIC3 codes correspond directly to NIC96. We use MoSPI files to create a full concordance between all NIC rounds, leaving us with 138 distinct non-agricultural industries.

population that is literate.

II. *Tariff liberalization*

We obtain information on tariff rates at the ISIC3 level for the years 1990-2005 (with gaps) from the World Bank via the World Integrated Trade Solution database. We use the correspondence between ISIC3 and India's 1996 National Industry Classification (NIC) to match tariffs to industries in our dataset. This leaves us 72 tradable industries, of which we drop eight agricultural sectors as we have no information on employment in these sectors. Combining this with our employment information from the 1990 EC, as well as information on each industry's labor cost share from the ASI (1993-1994), we construct our main measures of *output* and *input tariff exposure*.

More specifically, we compute a district's output exposure to the 1991 trade liberalization based on Equation 3:

$$OTE_{r,1992-1990} = \sum_{i \in S} \beta_{ri} \;\; \mathrm{d}_{1992-1990} \ln\left(1+\tau_i\right) \tag{8}$$

with β_{ri} calculated as in Equation 1 using our employment data and info on labor cost shares. Given the evidence that reforms after the 1991 liberalization were more influenced by political economy considerations, we follow e.g. Topalova [2007], and only consider the initial tariff shock, or the reduction in tariffs from 1990-1992 (see Figure 1), in our analysis.[24]

Similarly, we compute our measure for input tariff exposure using Equation 6:

$$ITE_{r,1992-1990} = -\sum_{i \in S} \beta_{ri} \left[\sum_{j \in \mathcal{J}} \alpha^{j}_{i,1993} * \;\; \mathrm{d}_{1992-1990} \ln\left(1+\tau_j\right) \right], \tag{9}$$

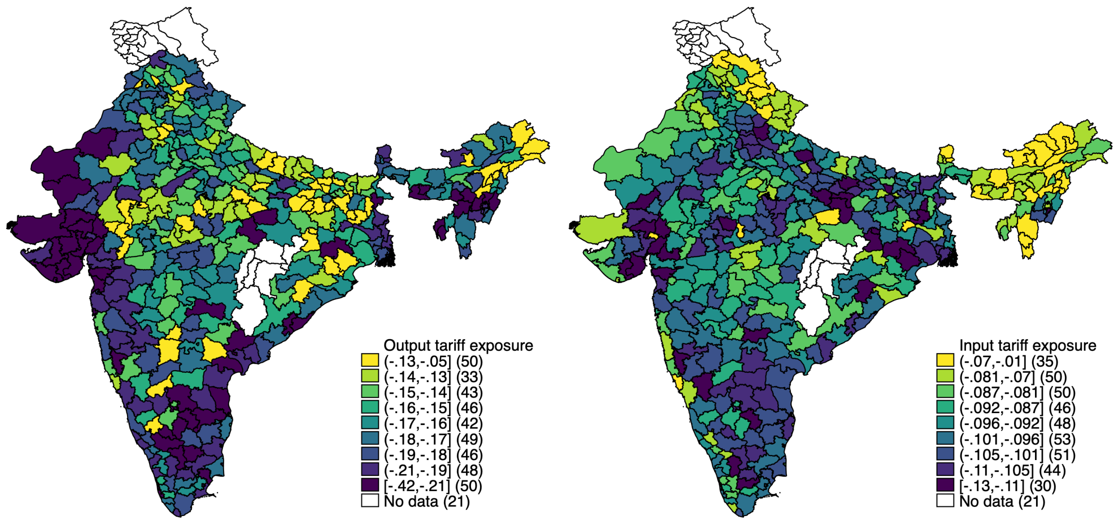
where we follow Amiti and Konings [2007] and use the earliest available Indian input-output table (1993-1994) to construct cost shares of intermediate inputs produced by industry j in the production of industry i, i.e. the $\alpha^{j}_{i,1993}$'s in Equation 9.[25]

Figure 4 shows the spatial variation in our two tariff exposure measures. No clear pattern in the spatial distribution of either exposure measure can be discerned: district with high or low Input or Output Tariff exposure can be found across India.[26]. In addition

[24]Figure 19 in Section II of the Appendix shows how this tariff shock compares to the later development of tariffs. In short: tariffs in 1992 and 1997 are highly correlated, showing that the initial liberalization was persistent; after that, further tariff harmonization weakens this correlation.

[25]For the computation of input tariff exposure, we do include e.g. agricultural tariffs and all sectors that use intermediates. We do not have information on the labor or material share for three tradable sectors in the Economic Census: extraction and agglomeration of peat, manufacture of coke oven products and manufacture of office machinery. These together account for 0.09% of employment in India; the median employment in these sectors in a district is 0.00008%, 0.004% and 0.007% respectively. We assigned those industries the average labor share of 0.081.

[26]Importantly, Figure 9 shows that more exposed districts are not concentrated near India's main ports, nor in its more inland areas.



(a) Output tariff exposure by district

(b) Input tariff exposure by district

Figure 4: Districts are differently exposed to output and input tariff reductions

to this, and importantly so, Figure 5 shows that both exposure measures are not correlated. If they would have been strongly correlated this could jeopardize our ability to separately identify the effect of India's tariff reductions increasing import competition from that of providing improved access to foreign intermediates.

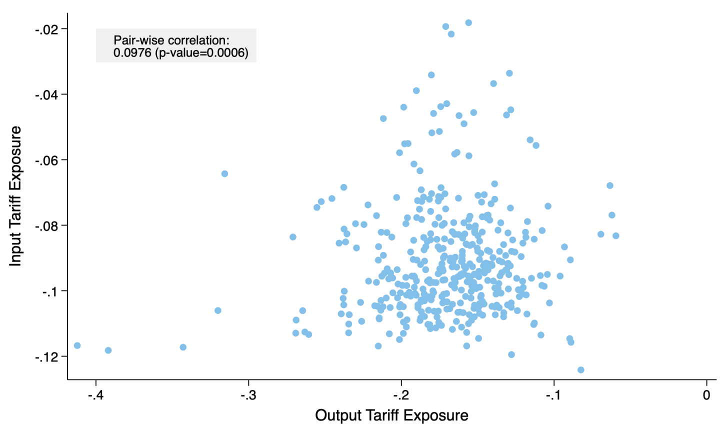


Figure 5: Relationship between Input- and Output Tariff Exposure

III. Domestic infrastructure

III.i. Road and rail data

To capture India's rail and road network before, at the time of, and after India's trade liberalization, we digitized three editions of the *Road Map of India*, created by the Indian government in 1977, 1988 and 1996 respectively, with the help of World Bank staff. First, the maps were georeferenced to create a correspondence with the administrative boundaries of India in 1988, and then the infrastructure network was digitized by hand. These maps also provide us with the location of all the 1988 district headquarters. Figure 3 already depicted the resulting (digitized) road and rail network in 1988.

This data allows us to map out the temporal variation in both road and rail infrastructure. Moreover, in case of the road network, we can also account for improvements in road *quality* over the years since the maps display three distinct types of roads: primary (Expressways and National Highways), secondary (State roads and Major District roads) and tertiary roads (rural roads).[27]

III.ii. Ports and ICDs

Next, we gathered information on India's main international gateways in 1990. Given the dominance of sea transport in India's international trade, we first generated a shapefile of India's Major Ports based on a list of ports available through the Indian Ministry of Shipping website. Only those ports that were operating in 1988 were included, resulting in 12 Major Ports. Two of these ports, Haldia and Jawaharlal Nehru Port Trust (JNPT), were initially created as an extension of the Kolkata and Mumbai port respectively. Since they are so close to these older ports, we group these newer satellite ports with their older counterparts when calculating our measures of "access to the port".

Next, we further refined this list making use of information from the India Port Report [I-Maritime, 2003]. This report provides a detailed discussion of each Major Port's hinterland connectivity, and several efficiency metrics. And, most importantly for us, it reports each port's total import volume as well as a breakdown of these imports into containers, break-bulk, oil, coal and fertilizer raw materials, from 1992 to 2002. Based on this information, we further narrowed our focus to only five of India's Major Ports: Chennai, Kandla, Kolkata, Mumbai and Vishakhapatam. Each of these ports individually handled at least 10% of India's total imports in 1992-1993. Together, they accounted for 82% of all imports and 78% of all container and break bulk imports. The other five Major ports – Paradip, Cochin, Tuticorin, Mormugao and Mangalore – are far less important, especially for the handling of break bulk and containers. Instead, they generally specialize in the import of one particular commodity (e.g., coal for Tuticorin, fertilizer raw material for Paradip, and crude oil for Cochin). The yellow squares in Figure 3 show the location

[27]See Appendix II for more details.

of the five main Indian international gateways that we consider in our main analysis.

Finally, we geolocated all seven Inland Container Depots (ICDs), India's most important intermodal shipment points, that existed in India in 1988 making use of information from the Northern Indian Steamer Agents Association. These seven ICDs are depicted by the green diamonds in Figure 3.

III.iii. Access to international markets

Given the lack of information on the transport costs involved in shipping imported goods from their port of entry to each Indian district, we use our data on India's historical road and rail network to construct proxies of each district's access to international markets. Our main measure is the shortest road or rail distance from a district's headquarter to the nearest of our five main ports. An important assumption underlying this choice is that this measure's variation across Indian districts is a good proxy for the variation across Indian districts in the actual domestic transport costs involved in shipping imported goods from the port to the district.

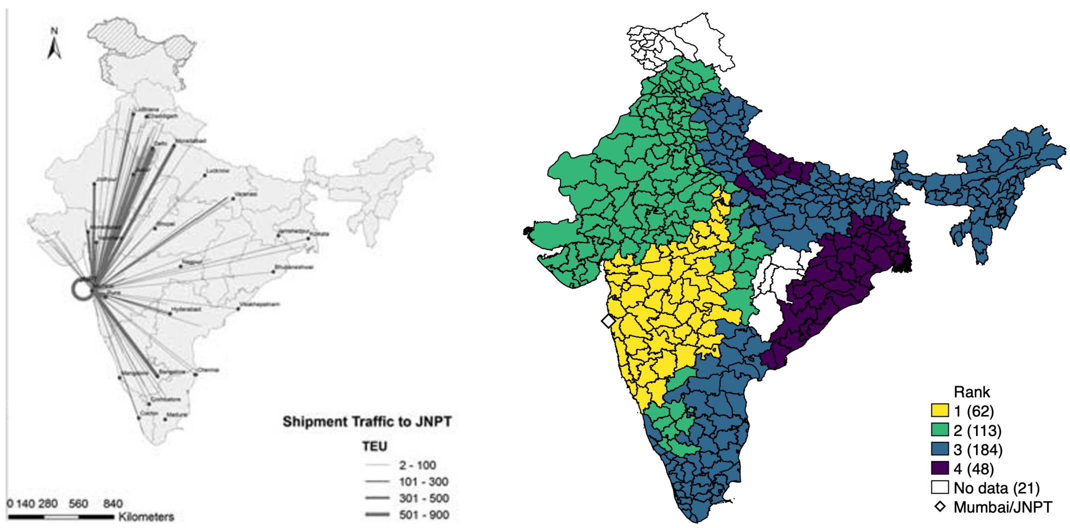
Bonadio [2022] reports that in 2015, firms on average used a port that was about 25% farther away than its nearest port, due to firms taking into account not only domestic transport costs but also differences in port efficiency. However, this interport competition on quality was practically nonexistent before the 1990s [Thill and Venkitasubramanian, 2015]. Even in the post-liberalization period, and despite the increase in international trade flows, such competition was still very limited due to strict regulations and the long distances between the different major ports [Haralambides and Behrens, 2000]. Port competition has intensified only in recent years – as reflected by increases in port quality, port expansions, construction of new (container) ports, explaining Bonadio [2022]'s findings. They do not carry over to our historical setting, as the fragmented port market would push most firms to opt for the nearest major international port.[28]

As a sanity check of our "distance to the nearest port" measure, we compare it to the hinterland analysis conducted by Thill and Venkitasubramanian [2015]. Using disaggregated exports data from India to the United States over the course of October 2006, they classify the hinterland of the twelve Major Ports that exported containers.[29]

[28]Note that focusing on the five most important major ports in our analysis also alleviates this concern. Considering the other, much smaller, more specialized major ports as well, would most likely provide a wrong, much noisier, proxy for the costs involved in shipping products to/from foreign markets for districts nearest to these smaller major ports. Most international trade of these districts would, in reality, move through one of India's five largest major ports instead. In fact, our main results come through when including the five smaller, less important Major Ports in our calculation of each district's distance to its nearest port, but they are less precisely estimated.

[29]This includes all Major Ports in our dataset except for Paradip, which does not handle any containers. It also includes the private ports of Mundra and Pipavav. These two ports commenced operations in 1998 and 2002 respectively, and only became important (container) ports for international trade after 2002 (when Mundra was linked to the Indian railway network) and 2009 respectively.

Figure 6 shows this comparison for Mumbai/JNPT – in both our and their period by far the largest port in India. Corresponding figures for Chennai (Figure 28), Kolkata/Haldia (Figure 29) and Kandla (Figure 30) are relegated to Section II in the Appendix. The left-hand side of these figures depict the “desire lines” to each port as computed by Thill and Venkitasubramanian [2015]. These describe the spread as well as the intensity of traffic from each respective port. On the right, we compare this to the ranking implied by our road distance to the nearest port measure, where the port is the nearest port for districts with rank 1, the second-nearest for districts with rank 2, et cetera. Despite the change in market structure between the 1990s, which is the focus of our study, and the mid-2000s analyzed in Thill and Venkitasubramanian [2015], their and our (simpler distance-based) pictures generate similar hinterlands of these ports.



(a) Thill and Venkitasubramanian [2015]

(b) Distance-based rank

Figure 6: Hinterland analysis for JNPT/Mumbai ports

Note also that the likelihood of using a different port than the nearest port is very likely to decrease in the distance differential between one’s nearest and second-nearest port. Thus, even if we misclassify a district’s main international gateway to be its nearest port instead of its second-nearest port, the distance to its nearest port will not be that different from that to its second-nearest port. The only possible exception here is Mumbai/JNPT port, India’s largest port handling about a third of break bulk and container imports in 1990. Its hinterland (see Figure 6) covers much of India, with the exception of its (North)-East and utmost South, and the districts close to Kolkata/Haldia and Chennai port in particular. In Section VI.iii we come back to this issue and show how sensitive our results are to excluding Mumbai or one, or more, of our other four major ports when calculating each district’s distance to its nearest port.

Finally, we need to take a stance on how distance traveled maps to domestic transport costs in India. In light of our discussion in Section III.ii, we focus on *road* distance to the nearest main port as our main proxy for a district's domestic transport costs to India's main international gateways, assuming that these costs increase (linearly) in distance to the nearest main port. With the exception of the longer rail freight corridors, road transport is the dominant mode for transporting most goods in India on most routes, given its flexibility and superior spatial coverage to rail transport. Figure 7 depicts the spatial distribution of districts' 1988 road distance to the nearest of India's five most important major ports.

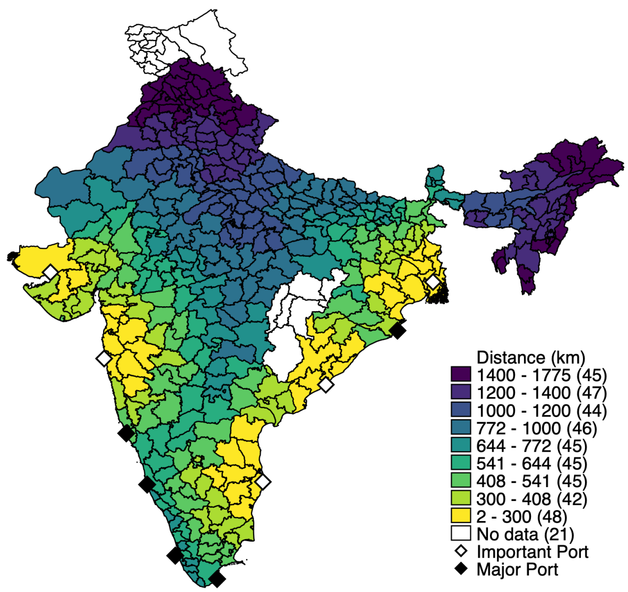


Figure 7: Distance to the nearest main international port

A concern with this specific choice is that places at a similar distance from the port can face vastly different transport costs due to unobserved route-specific characteristics, such as road quality, speed limits or congestion, but also the competitiveness of the transportation sector. The latter is especially relevant in more remote locations, which face higher marginal transport costs, higher markups, and lower service quality, hereby further reducing the passthrough of the effects of India's tariff liberalization [Allen et al., 2022]. This "triple curse of remoteness" will especially affect travel to and from the least accessible places, of which India has many. For the purpose of our analysis however, note that non-agricultural economic activity in India's districts is generally concentrated around their district headquarters or other urban centers. With only a few exceptions, these district headquarters are all connected to at least one of India's primary or secondary roads (National Highways, State roads and Major District roads), and to the Indian rail network.[30] As such, we find it unlikely that these other unobserved trade cost

[30]In our sample, only 11 (2.7%) of the headquarter *centroids* is more than 15 kilometers away from the

determinants invalidate our underlying assumption that the difference in transport costs between districts faced by (non-agricultural) firms in shipping goods to/from India's main major international ports is well-proxied by the difference in district headquarters' distance to its nearest of these ports. If villages were our unit of observation instead, this would be a much bigger concern. The substantial heterogeneity in the (often poor) quality of tertiary roads as well as in the availability and quality of transport services between villages would make a simple distance-based measure a much noisier proxy of transport costs.

We further verify the sensitivity of our results to our choice of proxy in several robustness checks. We, for example, consider alternative proxies of these costs, such as the *travel time* to the nearest important port that takes into account the quality of roads traveled[31], the distance and travel time to the nearest port *through an ICD*, or we use *rail* instead of road distance. Also, we extend our travel time measures to take the aforementioned *interstate checkpoints* into account to proxy for the costs of delays (and possible bribes) at state border crossings. Finally, since we do not know the exact mapping of distance to transport costs, we also let transport costs depend *nonlinearly* on road or rail distance, travel time or on the number of interstate checkpoints crossed on the route (see Section VI for more details).

V. Empirical strategy

To verify whether, and if so how, domestic infrastructure conditions the local labor market effects of India's trade liberalization, we adopt the following empirical framework:

$$\begin{aligned} \mathrm{d}\ln y_{r,t-1990} = {} & \alpha_r^{ER} + \beta_0 D^{port}_{r,1988} + \beta_1 OTE_{r,1992-1990} + \beta_2 (OTE_{r,1992-1990} \times D^{port}_{r,1988}) \\ & + \beta_3 ITE_{r,1992-1990} + \beta_4 (ITE_{i,1992-1990} \times D^{port}_{r,1988}) + \mathcal{X}'_{r,1990}\gamma + \epsilon_{rt}, \end{aligned} \tag{10}$$

where subscripts r and t indicate districts and years respectively. $\mathrm{d}\ln y_{r,t-1990}$ denotes the change in district r's local labor demand, which in our main specifications is taken to be the change in log non-agricultural employment taken from India's Economic Census – either between 1990 and 1998 or between 1990 and 2005. $OTE_{r,1992-1990}$ and $ITE_{r,1992-1990}$ are a district's output and input tariff exposure as defined in Equations 8 and 9. Remember that both variables are strictly below zero as they each capture the exposure to India's substantial tariff *declines*. Finally, $D^{port}_{r,1988}$ denotes a district's initial access to India's main international gateways prior to its trade reforms, where this access can be any of the measures we introduced in Section IV.III. As mentioned before, in our baseline estimations it is captured by a district's road distance to its nearest main port.

closest primary or secondary road (in 1988). The spatial coverage of India's Railways is not as extensive: 56 (14%) of the headquarters *centroids* are more than 50 kilometers away from the nearest railway line in 1988.

[31] Here we use the road classification into primary, secondary and tertiary roads, and follow World Bank Group [2005] in assuming an average speed of 40 km/h on primary roads, 30 km/h on secondary roads, and 10 km/h on tertiary roads.

We also always include *economic region fixed effects*, the α_r^{ER}, in our regressions. India has 77 economic regions, partitioning states into regions that are similar in their economic structure. Including these dummies flexibly controls for local trends in non-agricultural employment related to conditions in these economic regions that might be correlated with our variables of interest. Notably, they capture any differences between districts' state-level regulation that might be related to both a district's exposure to India's trade liberalization as well as to changes in employment, as highlighted in for example Besley and Burgess [2004]. In most specifications we also include a set of pre-liberalization control variables, $\mathcal{X}_{r,1990}$; see below for a more thorough discussion of (why to include) these variables. Finally, $\epsilon_{i,t}$ captures any remaining unobserved district-period specific idiosyncratic variables that also affect the outcome variable $\mathrm{d}\ln y_{i,t-90}$. We allow for spatial autocorrelation up to 300 kilometers away from the district headquarter, adjusting the standard errors using the method proposed by Conley [1999].[32]

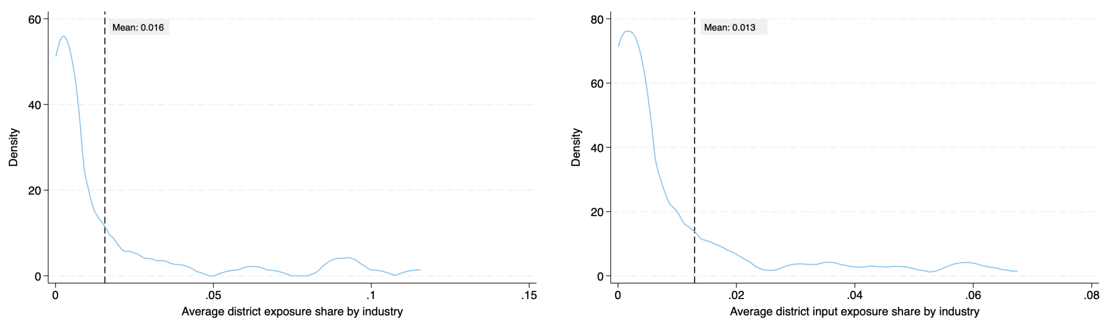
Our main coefficients of interest are β_1, β_2, β_3 and β_4: together with their estimated standard errors, they allow us test our main hypotheses regarding the mitigating effect of domestic transport costs on the effect of Input and Output Tariff Exposure on districts' local labor demand that followed from our conceptual framework.

Identification concerns Our empirical strategy effectively relies on four shift-share variables, with industry specific output- and input tariff changes as the shifts and industry employment shares, as well as districts' distance to the nearest port, as our measures of exposure. Crucial to our identification is that the shifts are as-good-as-randomly assigned and that the exclusion restriction holds [Borusyak et al., 2025]. First, this would imply that the tariff shocks are randomly assigned to different industries. Second, there should be no systematic pattern of high or low tariffs for industries that are economically important in districts that are either facing large or small (unobserved) labor market shocks, or are far from or close to a port.

A first-order threat to identification is any factor that violates the assumption that shifts are assigned as good as random. We are not concerned about the tariff liberalization itself introducing endogeneity: in Section III, we already highlighted that the incidence and size of the tariff reductions are unrelated to pre-reform industry characteristics or political economy concerns. Furthermore, given the unexpected nature India's tariff changes, reverse causality issues are very unlikely to be a threat to our identification. Finally, our identification does not rest on a few large sectors. Figure 8 shows the density plot of each industry's average employment share across districts in industries exposed to India's (a) output and (b) input tariff reductions respectively, which is proportional to the importance weight of each shift in our analysis [Borusyak et al., 2025]. The average

[32] Each district headquarter is at least located within 300 kilometers from one other district headquarter; the average headquarter is located within 300 kilometers from 15 other districts. Our results are robust to extending the distance cutoff to 500 kilometers or clustering the standard errors at the state or region level.

importance weight for the output tariff shifts is 0.016, the same as what it would be if each industry was equally represented (0.016).[33] The average district-level exposure to input tariffs by industry is 0.013, which is only slightly higher than 0.012, which would be the average importance weight if each industry were equally represented.[34]



(a) Output Tariff Exposure (b) Input Tariff Exposure

Figure 8: Density plot of shock importance weights

Finally, we need to be sure to account for other shift-level, or industry-level, shocks that occurred simultaneously with the trade liberalization. As mentioned before, three other reforms, all also part of the IMFs bailout package, immediately come to mind here: FDI liberalization, the gradual dismantling of the License Raj and the elimination of import licenses, with all three also affecting industries differently. Not taking these other reforms into account runs the risk of biasing our coefficients of interest if they are correlated with the tariff reductions and affect local labor demand. To avoid this, we obtain district-level measures of exposure to these other reforms from Topalova [2010] and include them as controls in our analysis, even though they are generally only weakly correlated with our OTE and ITE variables.[35]

A second concern is that the more inland Indian districts already faced larger domestic transport costs in shipping goods to or from India's main ports in the period before India's trade liberalization. If firms took this into account in their location decision, we might see a relationship between a district's access to India's important ports and its initial industrial composition, and thus also possibly their exposure to India's tariff reductions in the early 1990s. The extremely high import barriers in India prior to 1991 (resulting in very limited imports) make this very unlikely. If this were the case, it would make it (much) more difficult for us to identify how the effect of a district's output or input tariff exposure is conditioned by a district's domestic transport costs. For example, if, in an

[33]Moreover, the number of effective shifts, which is defined in Borusyak et al. [2025] as the inverse of the Herfindahl index of shock importance weights, is equal to 16, or a fourth of the total shifts.

[34]The number of effective shifts for input tariffs is 23, or almost 30% of the total number of shifts.

[35]The complete set of correlations – conditional on region fixed effects – is shown in Figures 24, 25 and 26 in Section II of the Appendix.

extreme case, all the variation in output or input tariff exposure were to be found near India's main ports, it would be impossible to estimate how the effect of this exposure differs in districts closer to and farther away from these ports. Figure 9 shows that there is, indeed, little concern for this, as we find similar variation in Output and Input Tariff Exposure at all distances from India's main ports.[36] Even at the industry level, the median correlation between a district's distance to its nearest major port and its employment share in each respective industry is only 0.004 (0.044) without (with) partialling out the region fixed effects. In Panel (a) of Figure 20 in Section II of the Appendix, we show the results of regressing district-level employment shares on distance to the nearest major port for each individual industry separately, controlling for region fixed effects. For only a few industries, the district-level employment share varies systematically (but not substantially) with a district's distance to its nearest main international port. Moreover, Panel (b) of the same figure shows there is little correlation between these regression coefficients and industry-level input and output tariff changes respectively.

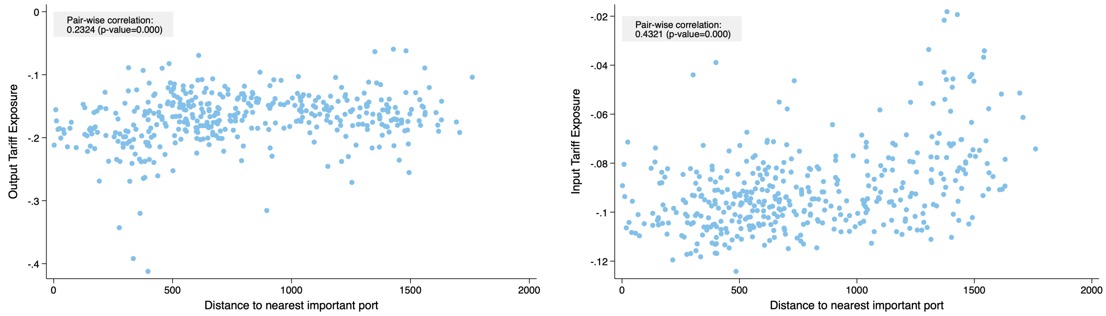
A related threat to identification would be endogenous placement of infrastructure in anticipation of trade liberalization. Specifically, if districts with relatively high or low output or input tariff exposure systematically improved their connection to the port in the pre-period, this would bias our estimates if the infrastructure improvements also affected employment in other ways. Using our data on India's road network prior to the shock, we computed the change in distance and travel time to the nearest important port for each district headquarter between 1977 and 1988, and, in Figure 10, we plot this against our measures of output- and input-tariff exposure conditional on region fixed effects.[37] Reassuringly, there is no systematic relationship between a district's change in access to India's main international trade hubs and their respective output or input tariff exposure.[38] In Section I of the Appendix, we furthermore show that, in general, there is no evidence of differential road construction or upgrading on routes to India's major ports in anticipation of its trade reforms.

Finally, note that our proxy for transport costs is district-specific. However, it is not unlikely that these transport costs also differ by industry, either because of differences between industries in transport costs unrelated to distance traveled, or in their dependence on distance traveled. Our data unfortunately does not allow us to explicitly take this into account. However, the former only jeopardizes our ability to identify our effects of interest if any such industry-specific transport cost differences are systematically related to an industry's output or input tariff reductions; there is no reason to believe that this was the case. The latter only threatens our identification if industries that faced larger

[36] Figure 22 in Section II of the Appendix shows this relationship with all economic region fixed effects partialled out; Figure 21 shows the corresponding conditional correlation between OTE and ITE.

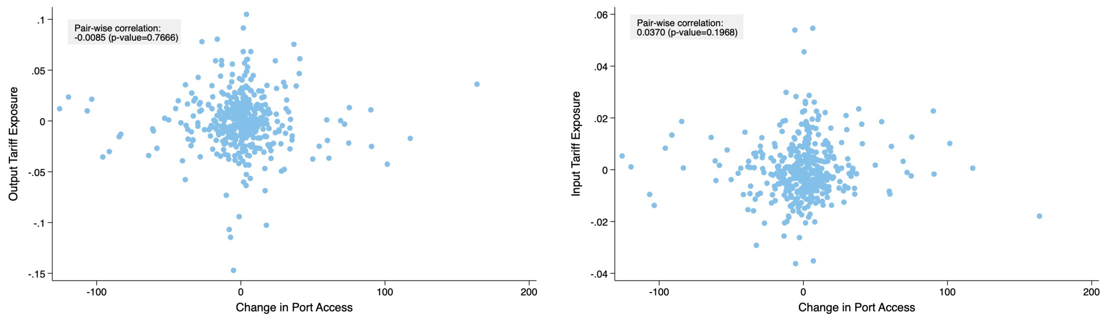
[37] This is also the case when considering the change in travel time to the nearest main port instead; see Figure 23 in Section II of the Appendix.

[38] Not surprisingly, including either the change in distance or travel time to the nearest port between 1977 and 1988 as additional controls in our regression does not affect any of our results.



(a) Output Tariff Exposure (b) Input Tariff Exposure

Figure 9: Correlation distance to the port and tariff exposure



(a) Output Tariff Exposure (b) Input Tariff Exposure

Figure 10: No evidence for anticipation tariff shock in infrastructure improvements

output or input tariff reductions also tend to be disproportionately located closer to or farther away from its main international ports. The aforementioned Panel (b) of Figure 20 in the Appendix clearly shows that this is not the case.

The main remaining identification concern is that ϵ_{rt} consists of omitted district-level variables that are correlated with changes in a district's output (*OTE*) or input tariff exposure (*ITE*), or with a district's initial access to its main international gateways. As discussed in McCaig [2011], a primary concern in this regard is the presence of underlying (three-digit) industry trends. For example, if sectors that experienced a relatively large decline in their output tariffs, and thus substantial increases in foreign import competition, were already on a differential employment trajectory before the trade liberalization, we would be unable to separately identify the effect of exposure to the trade shock from these underlying trends. Including all pre-reform three-digit industry shares would capture these trends, but remove all between-district identifying variation. Instead, we follow McCaig [2011] and control for industry trends at a more aggregate level, so that we rely on variation within these more broadly specified industry categories for identification.

Specifically, we include the district-level share of the labor force working in agriculture and mining before the reform as controls, as well as the share of the non-agricultural labor force working in non-tradable sectors; the latter also acts as our "incomplete share" control [Borusyak et al., 2025].[39] Do note that the included economic region fixed effects also substantially alleviate this concern given that economic regions group districts with a similar economic structure.

Second, our measure of access to international markets could be correlated with other variables that affect district-level employment changes. India's economic activity and population is more concentrated along the coast, as shown in Figure 27 in Section II of the Appendix. If the already larger, more developed, coastal districts were to expand relative to their inland counterparts regardless of the trade liberalization, for example due to agglomeration economies or having not only better international but also domestic market access, the estimated effect of the import competition channel (*OTE*) in coastal districts would be downward biased, whereas that of the intermediate inputs channel (*ITE*) would be biased upward. To account for these potentially different outcome trends between the typically larger coastal and smaller inland districts, we include the size of a district's pre-reform labor force[40], as well as its distance to the nearest top-15 largest city in India[41], its distance to Delhi (the country's capital and main inland population center), and its distance to Bangalore (the country's main southern growth pole), as controls in our regressions. Moreover, we include the share of the population that is literate to capture possible employment trends related to the level of education in a district as in Topalova [2010], Kis-Katos and Sparrow [2015]. Finally, note that the included economic region fixed effects also alleviate this second concern to some extent.

VI. Results

I. *Baseline findings*

Table 2 builds up to our main findings. First, we simply consider the effect of a district's Output and Input Tariff Exposure on its non-agricultural employment, without conditioning on its distance to its nearest main port. The results in columns (1) to (3), and in (7) to (9) show that both a district's Output Tariff Exposure (OTE) and its Input Tariff Exposure (ITE) significantly affect local labor demand. Moreover, as in Dix-Carneiro and Kovak [2015], comparing the 1990-1998 results to the 1990-2005 results in the Table shows that these effects are persistent: if anything, they become more pronounced over time. Also, reassuringly, results are very similar whether or not we include our full set of district-level

[39] We classify sectors without a tariff as nontradable; this includes utilities and (public) services.

[40] Including other measures of economic or population size, such as a district's total population or total non-agricultural employment, yields very similar results.

[41] As per the 1981 population census; the cities are Mumbai, Kolkata, Delhi, Chennai, Bangalore, Ahmedabad, Hyderabad, Pune, Kanpur, Nagpur, Jaipur, Lucknow, Patna, Kochi, and Surat.

controls and other policy controls.

As expected, being more exposed to import competition following India's rapid trade liberalization negatively impacts a district's employment growth. Districts facing a one standard deviation larger *decrease* in OTE, i.e. *higher* exposure to import competition, experienced a (5.1) 4.8 percentage point larger decline in employment in (1998) 2005. In contrast, having access to cheaper foreign intermediates as a result of India's trade liberalization positively impacts a district's employment growth. Districts facing a one standard deviation larger *decrease* in ITE, experienced a (7.5) 10.3 percentage point larger increase in employment in (1998) 2005. These are sizable effects considering that the average district's employment growth over the 1990-1998 and the 1990-2005 period amounted to 5.9% (SD: 30.5%) and 24.9% (SD: 32.4%) respectively.

This pattern is qualitatively similar to that observed after Brazil's trade liberalization as studied in Dix-Carneiro and Kovak [2017]: they find that a one standard deviation decrease in output tariff exposure leads to a persistent proportional decline in regional formal employment of 14 (18) percentage points between 1990 and 2000 (2005). Regarding input tariff liberalization, the closest comparison is Kis-Katos and Sparrow [2015]. Their results show no impact of either Output or Input Tariff Exposure on labor demand of medium and large (formal) manufacturing firms, either in the short (3 year) or longer run (9 years).[42] However, they measure a region's OTE and ITE as its *output-weighted* average industry tariff decline, instead of the theoretically consistent *employment-weighted* average industry tariff decline that we use, which complicates making a direct comparison to their findings.

Next, we turn to our main set of results and verify whether, and to what extent, these effects are more or less pronounced depending on a district's access to India's main ports. To do this, we estimate our baseline specification in Equation 10, which also includes our OTE and ITE measures interacted with a district's road distance to its nearest main port (in 100 kilometers). The results are shown in columns (4) to (6), and in columns (10) to (12) of Table 2.

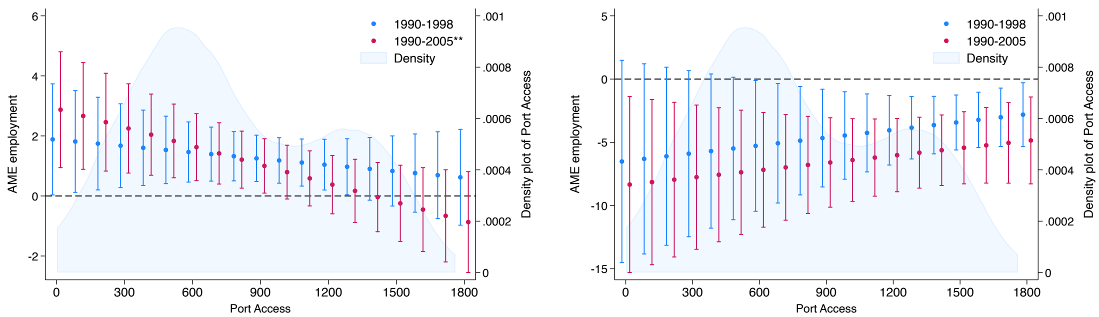
Our preferred specifications, including all controls, in columns (6) and (12) show that both the effect of import competition and of having better access to foreign inputs on local labor markets decays with distance to the port; but only the former does so significantly. Being 100 km farther away from one of India's main ports significantly reduces the effect of import competition on employment growth over the 1990-2005 period by more than 7%. The impact of access to cheaper intermediates is reduced by 2.3% with every 100 km farther away from the nearest main port, but this effect is not significant. To put these numbers into perspective, note that the median Indian district is located about 700 km from its nearest main port. This means that the effect of import competition (access to foreign intermediates) on local employment is about 51% (16.3%) smaller in the median district compared to a district right next to one of India's main ports.

[42] See their Table 10, panel B.

Table 2: Inland districts are more shielded from import competition

	Employment growth 1990-1998						Employment growth 1990-2005					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Output Tariff	1.236**	1.253**	1.243**	1.650	1.937	1.885*	1.480**	1.291**	1.165*	3.206***	3.011**	2.873**
Exposure (OTE)	(0.509)	(0.554)	(0.531)	(1.120)	(1.186)	(1.122)	(0.575)	(0.651)	(0.660)	(1.133)	(1.234)	(1.170)
Port Access				-0.044	-0.075	-0.070				-0.215*	-0.209*	-0.208**
× OTE				(0.105)	(0.107)	(0.104)				(0.110)	(0.112)	(0.105)
Input Tariff	-4.042***	-4.426*	-4.392*	-5.576*	-6.737	-6.515	-5.284***	-6.066***	-5.987***	-7.280***	-8.643**	-8.339**
Exposure (ITE)	(1.324)	(2.267)	(2.263)	(3.004)	(4.873)	(4.849)	(1.277)	(2.282)	(2.280)	(2.393)	(4.370)	(4.217)
Port Access				0.183	0.224	0.206				0.186	0.220	0.194
× ITE				(0.221)	(0.307)	(0.303)				(0.234)	(0.297)	(0.282)
District controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Policy controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407	407	407	407	407	407	407
R-squared	0.512	0.525	0.526	0.515	0.527	0.528	0.669	0.681	0.683	0.674	0.686	0.688

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 1998 or 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Port Access* is the distance in 100 kilometres to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). The regression controls include port access, the share employed in nontradable sectors (1990), log labor force (1991), share employed in agriculture (1991), share employed in mining (1991), share of population that is literate (1991), the log distance to the closest metropole, the log distance to Delhi and the log distance to Bangalore. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.



(a) Output tariff exposure

(b) Input tariff exposure

Figure 11: Average marginal effects of tariff exposure

To further facilitate interpretation of the coefficients reported in columns (6) and (12), Figure 11 shows the average marginal effects of OTE and ITE on employment growth for the full range of distances to the nearest main port.[43] These Figures show more directly at which distances from the nearest main port the effect of OTE and ITE are significantly different from zero.[44] Figure 11(a) clearly shows the declining impact of Output Tariff Exposure in districts located farther away from India's main ports. In fact, the negative effect of import competition on district-level employment strengthens over time, but only in districts close to India's main ports. Districts beyond about 900 km from these ports no longer experience significant employment declines as a result of their OTE. Figure 11(b) also shows that the positive effect of having better access to foreign intermediates is strongest in districts closer to India's main ports and persists over time; remember that this decrease in distance to the nearest port was not significant however. In the next section, we investigate how these effects differ for alternative measures of domestic transport costs.

In conclusion, India's trade liberalization persistently affected its regional employment patterns. Those districts that were more exposed to import competition saw their employment fall relative to other less-exposed districts. At the same time, districts that were more exposed to the benefits of easier access to foreign intermediates saw their employment go up relative to other less exposed districts. Both effects are strongest in districts nearest to India's main ports – signifying the importance of domestic trade costs in shaping the consequences of opening up to international markets.

[43]Note that the coefficients on OTE and ITE shown in columns (4)-(6) and (10)-(12) are showing the effect of OTE and ITE on employment changes *right next* to – i.e. zero kilometers from – one of India's main ports. The closest non-port district to one of India's main ports is Thane, located about 25 km from Mumbai's port.

[44]Do note that for the very extremes of the 'distance to the nearest port distribution' effects are often not significant due to a lack of variation.

I.i. Overall (predicted) employment effect of India's tariff liberalization

As we discussed, tariff liberalization gives rise to two opposing effects on regional non-agricultural employment patterns: one negative through increased import competition, the other positive through better access to foreign intermediates. This raises the question how the two effects compare, and together determine the overall regional employment response to India's trade reforms. To that end, we compute each district's overall predicted employment change over the period 1990-2005 given its Output and Input Tariff Exposure:

$$\begin{aligned} \mathrm{d}\ln \widehat{y_{r,2005}}_{-1990} = \hat{\beta}_1 OTE_{r,1992-1990} + \hat{\beta}_2 (OTE_{r,1992-1990} \times D^{port}_{r,1988}) \\ + \hat{\beta}_3 ITE_{r,1992-1990} + \hat{\beta}_4 (ITE_{i,1992-1990} \times D^{port}_{r,1988}), \end{aligned} \tag{11}$$

which we calculate using the estimated coefficients shown in column (12) of Table 2. Figure 12 plots these predicted employment changes as green circles against districts' road distance to the nearest main port. The solid green line shows the corresponding LOWESS curve, summarizing how the average predicted employment effect varies (nonlinearly) in a district's distance to its nearest main port.[45]

First of all, the overall predicted employment effects of India's tariff liberalization are virtually always positive: given each district's actual OTE and ITE, the negative import competition effect is always smaller than the positive effect of having better access to foreign intermediates. However, given that the negative import competition effect decreases faster in a district's distance to its nearest main port than the positive effect of having better access to foreign intermediates (see Table 2 column (12)), the overall positive employment effect of India's tariff liberalization is weakest in districts located close to India's main ports. In fact, the only two districts whose employment did not go up as a result of India's tariff liberalization, Dhanbad and Gandhinagar, are both districts with a high OTE and a low ITE that are also located relatively close to one of India's main ports (Kolkata and Kandla respectively). As a result, they face very strong import competition, while hardly benefiting from having better access to foreign intermediates, resulting in an overall decline in their non-agricultural employment. Not conditioning on a district's access to its nearest port, i.e., *assuming that differences in districts' domestic transport costs to its nearest main international port do not affect the passthrough of the effect of India's tariff reductions*, would instead predict slightly higher overall positive employment effects in districts with better access to India's main ports (as illustrated by the blue LOWESS curve in Figure 12)[46].

[45] The LOWESS curves shown here, as well as those in section IV, are generated using a bandwidth of 0.8.

[46] This is based on predicted overall employment effects calculated based on results in column (9) of Table 2.

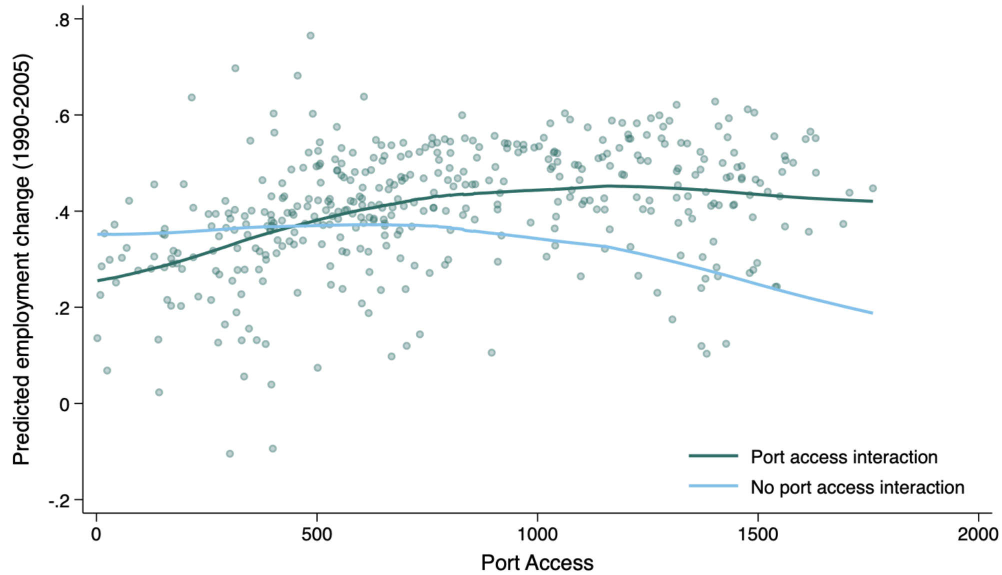


Figure 12: Predicted total employment changes 1990-2005

Given the extremely low interregional labor mobility in India (see e.g. Topalova [2010]), as well as the fact that our employment measure includes both formal and informal workers, the overall increase in districts' non-agricultural employment as a result of India's tariff reductions suggests the importance of two possible reallocation mechanisms. One, workers who lost their job as a result of increased import competition find a new job in sectors that benefit from cheaper access to foreign intermediates. Second, workers might move from agriculture to one of the expanding non-agricultural sectors that benefit from cheaper access to foreign intermediates, similar to how reduced tariff uncertainty in China induced structural transformation [Erten and Leight, 2021]. Providing conclusive evidence on the importance of these reallocation mechanisms is unfortunately very difficult, given the lack of comprehensive data tracking individual workers' (un)employment history before and after India's tariff liberalization.[47]

II. Alternative ways of incorporating domestic transport costs

In our baseline specifications, we assess whether the effect of Output and Input Tariff Exposure changes *linearly* with a district's *road* distance to its nearest main port. This follows among others Coşar and Fajgelbaum [2016], but, given that we do not know the exact mapping of transport costs to a measure of road distance to the port, our

[47] Estimating our main regression using the log change in districts' agricultural employment between 1991 and 2011 as the dependent variable does show tentative evidence that a district's agricultural employment falls in its input tariff exposure, and the more so the in districts closer to one of India's main ports.

results could change when modeling this differently. In this subsection, we do so in three different ways. First, we allow the effects of OTE and ITE to change linearly in distance- or travel time-based measures of domestic transport costs other than *road* distance. Second, we let the effects of OTE and ITE depend *nonlinearly* on different measures of a district's access to its nearest main port. And finally, we assess the sensitivity of our findings to our assumption that a district's imports arrive through its *nearest of India's five main ports*.

To not unnecessarily lengthen our discussion, we will from now on always focus on results that consider district employment changes over the period 1990-2005, including both district-level and policy controls. Results when considering the 1990-1998 period instead confirm the persistent nature of our found effects. Moreover, for the same reason, we focus on the robustness of the separate effects of OTE and ITE in our discussions. Unless explicitly stated, the results regarding their overall (combined) effect are in line with those discussed in Section VI.I.i. All results are, if necessary, available upon request.

II.i. Other domestic transport costs proxies than road distance

Table 3 shows the results of estimating our main equation with different measures of domestic transport costs. The first three columns replace our baseline proxy of the costs involved in shipping goods to a district from its main international gateway by a different distance measure. In column (1), this is a district's *rail* instead of road distance to its nearest main port, accounting for the fact that a nontrivial share of India's domestic transport takes place by rail. In column (2), we explore the possibility that some domestic trade will first go through one of the country's ICDs before reaching its final destination. To be more specific: for those districts whose *road* distance to its nearest port is larger than that to its nearest ICD, we consider those districts' distance to the port *via its nearest ICD* instead of the distance of the direct route to its nearest port. Here, based on our earlier discussion on the functioning of ICDs, we take the *rail* distance from the port to the ICD and then the road distance from the ICD to the district in question. Finally, in column (3) we further restrict India's international gateways to its three main *container* ports only – Kolkata, Chennai and Mumbai – and consider each district's *road* distance to its nearest main container port.

Allowing the effect of OTE and ITE to change linearly in these other distance-based measures of domestic transport costs, we reach very similar conclusions. Being located 100 km farther away from one of India's main ports *by rail* decreases the import competition effect by 5.2%; and when using either of the other distance-based measures this number is 7%-7.2%, very similar to what we found in our baseline. This mitigating effect of domestic transport costs on the effect of import competition is also always significant. Moreover, we find a mitigating effect of domestic transport costs on the positive effect of having better access to foreign intermediates, again mostly not significant. As in our baseline findings, being located 100 km further away from one of India's main ports reduces the positive effect of ITE by about 1.5%-2.8%.

Table 3: Attenuating effect of international market access robust to different access measures

	Distance measures			Travel time measures	
	(1) Rail	(2) Via ICD	(3) Container ports	(4) Time	(5) Incl. delays
OTE	2.592**	2.854**	3.156**	2.886**	3.150***
	(1.038)	(1.200)	(1.305)	(1.154)	(1.201)
Port Access × OTE	-0.135**	-0.199*	-0.226*	-0.07**	-0.056**
	(0.066)	(0.107)	(0.117)	(0.032)	(0.025)
ITE	-9.393**	-8.789**	-7.449*	-8.924**	-10.31**
	(3.900)	(4.129)	(4.248)	(4.120)	(4.419)
Port Access × ITE	0.246	0.248	0.111	0.085	0.094
	(0.198)	(0.271)	(0.280)	(0.098)	(0.075)
Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407
R-squared	0.690	0.688	0.689	0.688	0.688

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Rail* is the distance by rail to the nearest important port, *Via ICD* is the distance to the nearest important port via an ICD and *Container Port* is the distance to the nearest container port (Chennai, Kolkata, Mumbai/JNPT). *Travel time* is the shortest travel time in hours to an important port; *Incl. delays* adds five hours delay for every state border crossing. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The final two columns of Table 3 use a travel time-based proxy for domestic transport costs instead of the distance-based proxies we used so far. In column (4) we allow the effect of OTE and ITE to change linearly with the shortest travel time from a district to its nearest important port, taking into account road quality (see Section IV). In column (5), we also incorporate a five-hour delay when crossing a state border to (rudimentarily) account for the hours lost at state-border checkpoints. Using these travel-time-based proxies, we again find very similar results as in our baseline specification. The negative effect of being exposed to increased import competition falls significantly in a district's travel time to its nearest main port: one additional hour of travel reduces this effect by 1.8%-2.4%. The positive effect of ITE also decreases in travel time to the nearest main port – one additional hour of travel reduces its effect by about 1% – but as in our baseline results this effect is not significant. Given that the median travel time from a district to its nearest main port is about twenty hours, these findings imply that the effect of import competition (access to foreign intermediates) on local employment is about 48% (35%) smaller in the median district compared to a district right next to one of India's main ports.

II.ii. Non-linear mapping of road distance to transport costs

So far, we have only allowed the effect of Output and Input Tariff Exposure to depend *linearly* on distance or travel time to one of India's main ports. Given that we do not know the exact mapping of transport costs to distance or travel time, Table 4 shows what happens when we model this differently and simply classify some districts as "far from" and others as "close to" one of India's main ports. To be more specific, in columns (1), (2) and (3) we simply allow the effect of OTE and ITE to be different for districts located more than 500 km, 700 km, or 900 km from their nearest main port respectively.[48]

This analysis shows that the effects of output tariff exposure is always larger in those districts closest to the port; also, in the more distant districts it is always insignificant. Interestingly, when using 900 km as a distance cutoff to define districts as "close to" or "far away from" India's main ports, we even reject the hypothesis that the effect of OTE is the same for districts close to and far away from the main ports (see the reported p-values of the corresponding F-test). Being one standard deviation more exposed to import competition results in a 7.9% employment decline, but only so in districts closer than 900 kilometers to one of India's main ports.

We also find that the effect of having better access to foreign intermediates is strongest in places close to India's main ports. Interestingly, the difference in the estimated effect of ITE between districts "close to" and "far away from" the port is again largest when using a larger distance to the port threshold to define districts as "close to" or "far away from" the port. Using a threshold of 900 kilometers in column (3) shows that being one standard deviation more exposed to the benefits of having better access to foreign intermediates results in a 13.5% employment increase in districts closer than 900 km for India's main ports, almost twice as large as in districts farther away. However, note that we cannot formally reject the hypothesis that these differences are significant.

Columns (4) and (5) add to these results by taking a slightly different nonlinear perspective. Instead of basing our access measure on distance to the nearest main port, we focus on the number of state border crossings that need to be cleared before reaching a district's nearest port. As mentioned before, the substantial costs incurred at state border crossings are responsible for a large share of domestic transport costs on interstate routes [Barnwal et al., 2024].[49]

[48] These distances roughly correspond to the 30th, 50th (median) and 62.5th percentile of the distribution of distance to the nearest main port.

[49] van Leemput [2021] imputes these interstate barriers to amount up to 40 percent of the total domestic barriers to trade.

Table 4: Nonlinear measures of distance

	Distance threshold			State border threshold	
	(1) 500 km	(2) 700 km	(3) 900 km	(4) 0 or > 1	(5) 0, 1 or > 2
Low × OTE	1.531* (0.874)	1.503* (0.770)	1.926** (0.793)	2.089*** (0.560)	2.099*** (0.548)
Medium × OTE					2.023** (0.913)
High × OTE	0.884 (0.700)	1.037 (0.794)	-0.207 (0.495)	1.240 (0.783)	-0.437 (0.636)
Low × ITE	-4.994 (3.708)	-7.644** (3.254)	-7.912** (3.157)	-12.11*** (3.273)	-12.62*** (3.348)
Medium × ITE					-6.007* (3.578)
High × ITE	-5.824*** (2.060)	-5.013*** (1.856)	-5.525*** (1.854)	-5.043** (2.242)	-5.152*** (1.497)
P-values					
$\beta^{Low}_{OTE} = \beta^{High}_{OTE}$	0.521	0.626	0.038	0.323	0.003
$\beta^{Low}_{OTE} = \beta^{Medium}_{OTE}$					0.942
$\beta^{Medium}_{OTE} = \beta^{High}_{OTE}$					0.035
$\beta^{Low}_{ITE} = \beta^{High}_{ITE}$	0.761	0.301	0.385	0.01	0.005
$\beta^{Low}_{ITE} = \beta^{Medium}_{ITE}$					0.059
$\beta^{Medium}_{ITE} = \beta^{High}_{ITE}$					0.795
Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407
R-squared	0.690	0.685	0.690	0.688	0.692

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. See the notes under Table 2 for details on the controls. *p-value OTE* and *p-value ITE* report the p-value of the F-test comparing the coefficient on *Output Tariff Exposure* and *Input Tariff Exposure* for regions closer and further from the port. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

In column (4), we simply allow the effects of OTE and ITE to differ between districts in port states, which do not have to cross any state borders to access international markets, and districts that do because they are located in a state without a main port. Again, both the passthrough of import competition and cheaper access to foreign intermediates are strongest for districts with the best access to India's main ports. Both the effect of OTE and ITE are about twice as large in districts in port states, and, in case of ITE this difference is now significant. A one standard deviation increase in OTE (ITE) lowers (increases) non-agricultural employment by 8.5% (20.8%) in districts in port states. In non-port states these percentages are 5.1% and 8.7% respectively.

In column (5), we refine these results by further distinguishing between districts for which importing foreign goods requires one or (more than) two state-border crossings to reach the nearest of India's main ports respectively.[50] Interestingly, the effect of OTE is significant, positive and very similar in districts in a port state and those that only have to cross one state border to get to their nearest main port. However, it is significantly different for districts whose imports involve two or more state border crossings: in those districts, being differentially exposed to an increase in import competition does not affect employment. For ITE instead, we find that being more exposed to the benefits of better access to foreign intermediates is always associated with faster employment growth. However, this effect is significantly different and is more than twice as large in districts located in port states compared to districts that are one or more border crossings away from their nearest port.[51]

In conclusion, allowing the effect of OTE and ITE to nonlinearly depend on a district's distance to its nearest port, or on other proxies of the cost involved in interstate transport flows than road distance, gives us qualitatively very similar results as when allowing these effects to depend linearly on a district's road distance to the nearest port.[52] Both effects are weakest in districts that face the highest domestic transport costs: India's more inland districts are shielded from the negative employment effects from increased import competition, while at the same time benefiting less from having better access to foreign intermediates.

II.iii. Choice of main ports

Last but not least, we assess the sensitivity of our results to our choice of considering India's five largest main ports when calculating a district's distance to its nearest port. Among these ports, Mumbai/JNPT stands out as handling about a third of India's break bulk and container imports, followed by Chennai and Kandla, each handling roughly 13%, and Kolkata and Visakhapatnam, each handling around 10% of break bulk and container imports. These size differences are also reflected in the size of these ports' hinterlands. Mumbai/JNPT serves as an important port for much of India, with the

[50]The maximum number of state borders that one needs to cross to reach any of India's districts from its nearest main port is three.

[51]Of course, routes crossing two state borders are typically, but not necessarily, also longer than routes than those crossing one border, that are in turn typically, but not necessarily, longer than routes that do not cross any state border at all.

[52]Table 9 in the Appendix shows that this is the case when allowing the effect of OTE and ITE to depend non-linearly on the other distance- or travel time-based measures also used in Table 3.

exception of northeast and the south of India. In contrast, the other ports' hinterlands are more localized: Kandla's hinterland is confined to northwest India, Chennai's to South India, and Visakhapatnam and Kolkata mainly serve the northeast.[53] Assuming that each district imports through its nearest of these five ports might thus especially understate the importance of Mumbai/JNPT for many districts.

Table 5, accompanied by Figure 13, shows how sensitive our results are to including or excluding particular ports in our set of possible nearest main ports. Column (1) removes Mumbai from our baseline set of five main ports. The results are very similar to our baseline, but slightly less precisely estimated. Next, column (2) also excludes Kandla. Kandla is itself an important port for northwest India during our sample period[54], but since it is close to Mumbai, the correlation between the distance from a district to Mumbai and Kandla is substantial ($\rho = 0.86$). Leaving out Kandla together with Mumbai therefore arguably does a better job at assessing the importance of including Mumbai/JNPT in our set of main ports; not doing so still classifies districts in northwest India as well-connected to a main port.[55] Column (2) shows that leaving out Mumbai/JNPT and Kandla – i.e. misclassifying much of northwest India as being poorly connected to a main port – does affect our results. The interactions between distance to the port and our measures of Output and Input Tariff Exposure are *both* no longer significant and much smaller compared to our baseline findings. When we also exclude Chennai in column (3), i.e. also misclassifying much of the South as being poorly connected to a main port, the results look similar to those in column (2).

The next three columns instead consider what happens when we progressively exclude the smallest of our five main ports in our proxy for port access. When excluding Kolkata and Visakhapatnam in column (4) – and thus misclassifying India's North-East as facing much higher domestic transport costs to import foreign goods – our baseline results come through. Moreover, the same holds when we also exclude Chennai (column (5)), and even when we only consider Mumbai/JNPT (column (6)). In the latter two cases, the coefficient on the interaction between distance to the port and ITE even turns significant.

Overall, our main finding that the effects of trade liberalization are more pronounced in places that are better connected to the country's main ports is most sensitive to the inclusion of Mumbai and/or Kandla in our set of main ports. Not including these ports misclassifies India's North-West as being poorly connected to international markets; in reality, their imports reach them almost exclusively through the ports of Mumbai/JNPT, and to a lesser extent Kandla. Excluding one or more of the other three main ports from our baseline set of five main ports leaves our main findings unaffected. This is consistent with Mumbai/JNPT also being an important international gateway for the imports of

[53] See Figures 6-30, Figure 1 in Venkita Subramanian and Thill [2019], or Figure 2.3.9 in JICA [2012].

[54] The new port of Mundra, located very close to Kandla, has taken over much of Kandla's traffic since it started operating in 2001.

[55] Only leaving out Kandla gives us results that are very similar, both in terms of point estimates as well as significance, to those in our baseline.

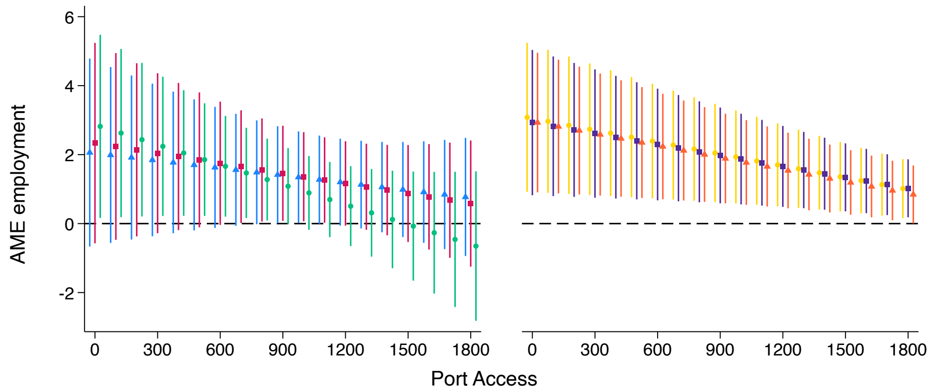
many districts located near to Chennai, Visakhapatnam, and Kolkata.

Table 5: Choice of main ports

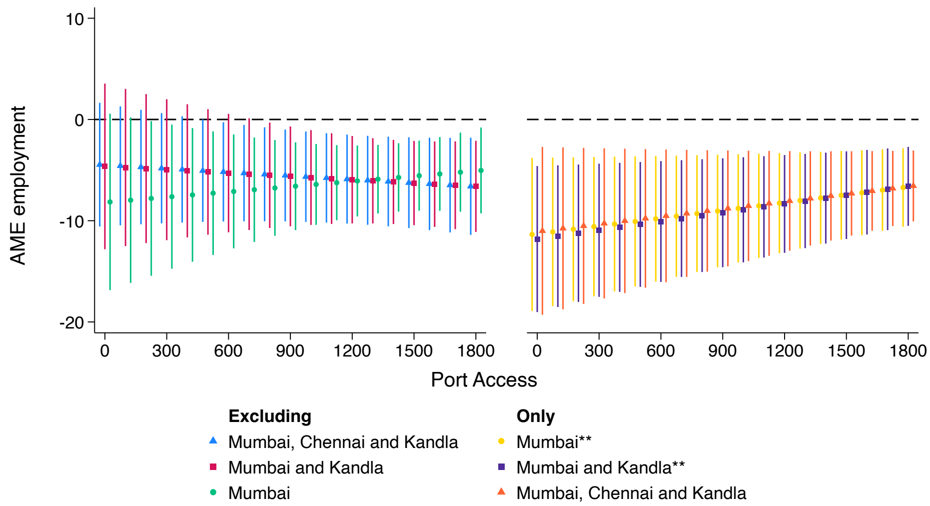
	Excluding largest ports			Only largest ports		
	(1) Mumbai	(2) Mumbai & Kandla	(3) Mumbai & Kandla & Chennai	(4) Mumbai & Kandla & Chennai	(5) Mumbai & Kandla	(6) Mumbai
OTE	2.819**	2.334	2.060	2.930***	2.929***	3.081***
	(1.350)	(1.478)	(1.385)	(1.029)	(1.071)	(1.098)
Port Access × OTE	-0.193	-0.0974	-0.0716	-0.116**	-0.106**	-0.115**
	(0.122)	(0.113)	(0.101)	(0.0477)	(0.0454)	(0.0463)
ITE	-8.145*	-4.636	-4.465	-11.00***	-11.81***	-11.36***
	(4.429)	(4.160)	(3.106)	(4.215)	(3.666)	(3.847)
Port Access × ITE	0.173	-0.110	-0.119	0.247	0.289**	0.258**
	(0.300)	(0.248)	(0.175)	(0.157)	(0.122)	(0.129)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407
R-squared	0.686	0.685	0.685	0.692	0.691	0.691

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. *Port Access* is the distance in 100 kilometers to the nearest port under different port choices. First, we exclude the three largest ports in turn: column (1) excludes Mumbai; column (2) excludes Mumbai and Kandla and column (3) excludes Mumbai, Kandla and Chennai. The last three columns only consider the three largest ports: column (4) includes Mumbai, Kandla and Chennai, (5) Mumbai and Kandla and in column (6) *Port Access* is restricted to distance to Mumbai only. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.

p $^{*} p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.



(a) Output tariff exposure



(b) Input tariff exposure

Note: Superscripts denote whether the slope is significantly different from zero at the 10% level; * only for *OTE*, ** for both *OTE* and *ITE*.

Figure 13: Average marginal effects with different ports

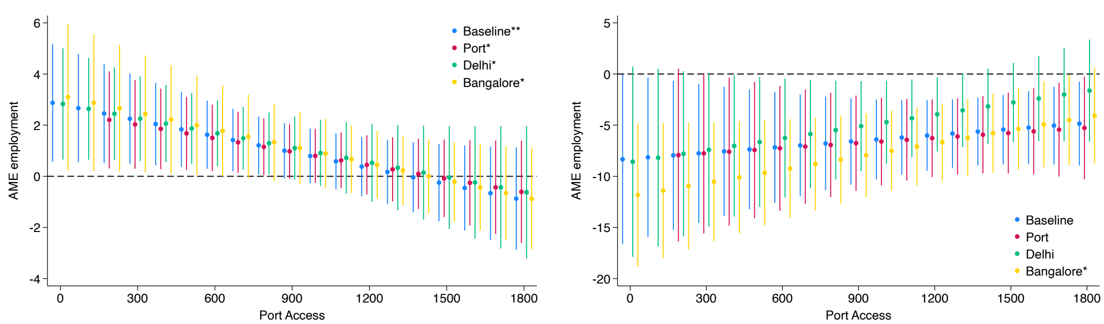
III. Robustness

Particular districts driving our findings. *First*, the attenuating effect of port access on the consequences of trade liberalization could be driven only by the districts in the immediate vicinity of India's main ports. In order to assess this, we estimate our baseline specification and exclude districts within 150 or 250 km from their nearest main port.[56] *Second*, due to the shape of the Indian subcontinent, India's southern districts are generally

[56] We choose 150 km because Kandla is located 143 kilometers from its own district's headquarter. Excluding only the five main ports' districts yields very similar results.

located closer to one of India's main ports. At the same time, Southern India experienced faster economic growth rates compared to Northern India during the period we consider. If our included control variables do not adequately capture features of these Southern districts related to both their economic growth as well as their exposure to import tariff liberalization, our findings could be biased. To that end, we exclude the 15% or 20% districts located closest to Bangalore, the South's main growth hub, from our sample; this corresponds to districts located within 600 or 800 km from Bangalore respectively. *Third*, a similar argument could be made for districts closer to Delhi, India's capital, main population center and an important industrial hub[57]. These districts are also among those furthest from India's main ports. To assess whether our results are sensitive to the inclusion of these districts, we exclude the 15% or 20% closest districts from Delhi from our sample - corresponding to those districts located within 400 or 500 km from Delhi respectively.[58]

Figure 14, plotting the dependence of the marginal effects of (a) Output and (b) Input Tariff Exposure on distance to the nearest main port, shows that either of these robustness checks gives us results that are very similar to our baseline findings. Figure 14 is based on the results shown in columns (2), (4) and (6) of Table 10 in the Appendix, obtained by estimating our baseline specification while excluding districts within 250 km from a main port, within 500 km from Delhi and within 800 km from Bangalore, respectively.[59]



(a) Output tariff exposure

(b) Input tariff exposure

Figure 14: Average marginal effects with different samples

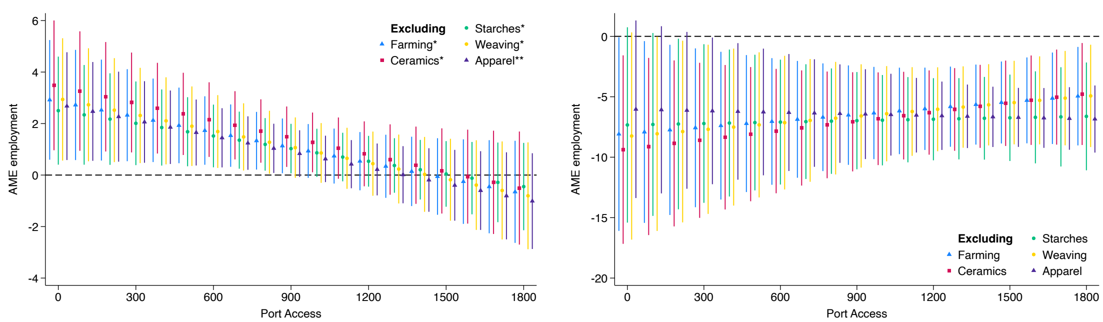
[57]Specifically, the Agra-Delhi-Kalka-Saharanpur Industrial Region.

[58]Figure 31 in the Appendix maps the districts excluded from the sample when using the various exclusion criteria outlined above.

[59]Columns (1), (3) and (5) of the same Table 10 show very similar results when excluding districts located within 150 km, 400 km and 600 km from a nearest main port, Delhi and Bangalore respectively. Also, Table 11, shows that our results using a non-linear mapping of road distance to domestic transport costs instead – as shown in Table 4 – also hold with these sample restrictions.

Particular sectors driving our findings. To confirm more explicitly that our identification does not rest on a few important industries and their respective tariff changes, we further examine how sensitive our results are to the exclusion of any of the five industries with the highest average district-level employment shares in 1990, i.e. the highest importance shares in our (shift-share) tariff exposure measures: farming of animals, manufacturing of non-metallic mineral products (this includes mainly porcelain and ceramics), grain mill products and starches, spinning and weaving of textile, and apparel respectively. Each of these industries represents about 9%-11% (5%-8%) of the average (median) district's non-agricultural employment.[60] A potential threat to identification is that the estimated coefficients on our tariff exposure measures do not only pick up the labor demand effects of India's tariff liberalization but also that of unobserved trends or shocks affecting labor demand in these specific industries, whose tariff reductions feature more prominently in our exposure measure given their higher importance shares.

To verify this, we re-estimate our baseline specification, and exclude each of these five industries from our exposure measure in turn. In these regressions, the estimated coefficients on our tariff exposure measures capture the effect of the import tariff liberalization on all tradable industries *except for the industry that is left out*. Figure 15 plots the marginal effects of these "purged" tariff exposure measures, showing how they depend on a district's distance to the port. They are based on the results shown in Table 12 in the Appendix. None of our baseline findings rest on the inclusion of any of these five industries with highest importance shares in our tariff exposure measures.



(a) Output tariff exposure

(b) Input tariff exposure

Figure 15: Average marginal effects excluding different industries

IV. Heterogeneous effects across the firm size distribution

So far, we have considered the effects of India's trade liberalization on a district's total non-agricultural labor demand. However, the extensive literature on firm heterogeneity

[60] The average employment share of the industry with the next (sixth) largest average employment share across India's districts is only 6%.

and international trade suggests that this can conceal the heterogeneous impact on firms across the productivity or size distribution – see e.g. Melitz and Redding [2014a] for an overview. First, a decrease in output tariffs would mainly affect the least productive firms, as fiercer import competition pushes them to exit the market first [Melitz, 2003]. Second, the gains from a fall in input tariffs are concentrated among the larger firms, as they are more likely to actually use these imported intermediates [Amiti and Konings, 2007, Nataraj, 2011, Topalova and Khandelwal, 2011].

To shed light on these possible heterogeneous effects across the firm size distribution, we make use of the fact that the Economic Census reports total employment at the firm level. This allows us to re-estimate Equation (10) using a district's employment growth in different bins of the firm size distribution as the dependent variable. More specifically, we distinguish four different categories: firms with fewer than 10 employees, which also corresponds to India's definition of the informal sector, small firms with 10-20 employees, medium firms with 20-100 employees and finally large firms with more than 100 employees.[61] The vast majority of firms is informal. In 1990, formal firms represented only 3% of all firms in the average district, employing 34% of all non-agricultural workers. And, of these formal workers, an average of 24%, 37%, 38% worked in one of the district's small, medium, and large firms, respectively.

Table 6 and Figure 16 show that the effects of India's trade liberalization between 1990 and 2005 are indeed very different across the firm size distribution. Moreover, Table 14 in Section III of the Appendix shows that the distance to the port already significantly conditioned the impact of OTE on district employment at small and medium-sized firms in the 1990-1998 period; an effect we did not pick up when considering a districts' total employment. For each of the four firm size bins, the first column (columns (1), (3), (5) and (7) in the Tables) displays the effect of Output and Input Tariff Exposure on the change in employment between 1990 and 2005. Next, the second column in each category (columns (2), (4), (6) and (8) in the Tables) also conditions this effect on a district's road distance to its nearest major international port.

Import competition and employment across the firm size distribution. In line with predictions of Melitz-type models, the negative labor demand effect of being more exposed to foreign import competition following India's trade liberalization is most pronounced for the smaller firms (see columns (1)-(4)). Moreover, column (5) shows that for medium-sized firms not accounting for port access would lead one to believe their employment is not affected by import competition. However, when conditioning this effect on a district's distance to the nearest port, as in column (6), both the coefficient on OTE and its interaction with this distance become significant. Employment in medium-sized firms decreases when a district is more exposed to import competition, but only significantly

[61] This classification of firms into small, medium, and large firms follows the one used in the NSSO unorganized manufacturing survey [Mehrotra and Giri, 2019].

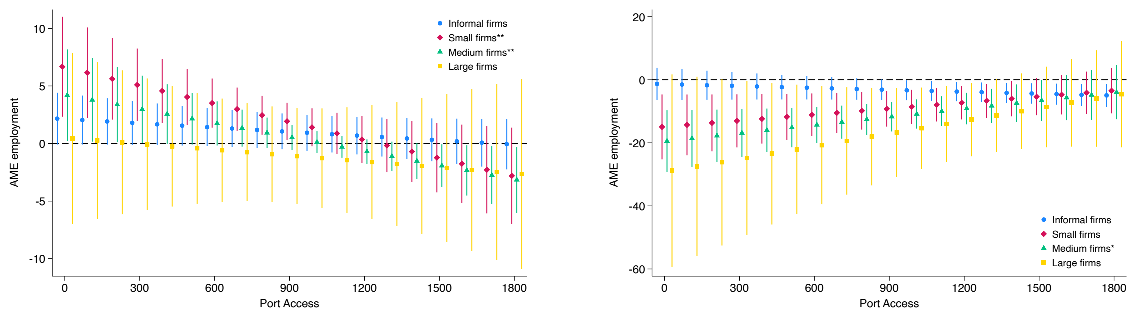
so in districts closer to one of India's main ports. This effect is even more pronounced for small firms: labor demand at smaller firms exposed to the same increase in import competition falls more, and does so significantly for farther distances from India's main ports. This is shown more clearly in Figure 16, which plots the average marginal effects based on the results in Table 6. A one standard deviation increase in Output Tariff Exposure, i.e. being more exposed to foreign import competition, reduces small firms' labor demand in 2005 by 25% (12%) in districts located 100 km (700 km) from the nearest main port, 1.6 (2.2) times as much as that by medium sized firms. In sharp contrast: labor demand of India's largest firms does not depend on Output Tariff Exposure, neither close to, nor far from, India's main ports. In Section III of the Appendix, we show the corresponding results for the number of firms, and find again that the number of small and medium firms decreases with higher exposure to the tariff shock. Consequently, average formal firm size *increases* in districts that are more exposed to import competition (see Table 16 in Section III of the Appendix).

Table 6: Heterogeneous effects of trade liberalization across different firm sizes

	Informal firms		Small firms		Medium firms		Large firms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OTE	1.304*	2.164*	2.291*	6.674***	0.629	4.197**	-1.526	0.449
	(0.766)	(1.144)	(1.208)	(2.207)	(0.937)	(2.022)	(2.099)	(3.773)
Port Access × OTE		-0.123		-0.527**		-0.408**		-0.172
		(0.0898)		(0.223)		(0.186)		(0.376)
ITE	-2.895*	-1.313	-7.391**	-14.95***	-10.11***	-19.46***	-15.78*	-28.83*
	(1.523)	(2.606)	(3.337)	(5.231)	(2.036)	(4.976)	(8.100)	(15.52)
Port Access × ITE		-0.202		0.638		0.862*		1.347
		(0.171)		(0.387)		(0.448)		(1.144)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407	395	395
R-squared	0.835	0.836	0.776	0.784	0.576	0.588	0.571	0.574

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment at informal (<10 workers), small (10-20 workers), medium (20-100 workers) or large firms (>100 workers) respectively, between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Port Access* is the distance in kilometres to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). Note that in 2005, 12 districts do not have any large firms. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Finally, informal firms' labor demand also falls most in districts that are most exposed to increased foreign import competition, but not as much as in small- or medium-sized



(a) Output tariff exposure (b) Input tariff exposure

Figure 16: Average marginal effects of tariff exposure by firm size

formal firms. This apparent contradiction to the theoretical prediction that employment in smaller, less productive firms is most negatively affected by import competition can be explained by informal employment being a form of insurance against unemployment [Dix-Carneiro and Kovak, 2019, Ponczek and Ulyssea, 2021]. In effect, many workers losing their job at (in)formal firms because of import competition will end up finding an informal job or starting a small informal business, thereby mitigating the negative effect of import competition on a district's informal employment.

Access to imported intermediates and employment across the firm size distribution. The effect of improved access to foreign intermediates also differs depending on firm size. It is smallest and hardly significant for informal firms, which are the least likely to use foreign intermediates in their production process. If at all, they only indirectly benefit from increased demand of other formal firms in their district that do use foreign intermediates [Moreno-Monroy et al., 2014, Mukim, 2014]. In contrast, formal firms do benefit from improved access to foreign intermediates, and the more so the larger the firm.

The results in columns (4), (6) and (8) of Table 6, and their corresponding AMEs depicted in panel (b) of Figure 16, clearly show that *formal* firms closer to one of India's main ports experience the largest labor demand increases, although the difference with firms located farther away is not always significant. In 2005, a one standard deviation decrease in Input Tariff Exposure, i.e. having better access to foreign intermediates, increases labor demand by India's largest firms by 47% (33%) in districts located 100 km (700 km) from one of India's main ports; 1.5 and 1.9 times as much as that by medium and small firms respectively. The number of formal firms also increases in a district's exposure to the input tariff shock. But, even though the percentage increase in small or medium-sized formal firms is lower than that in the largest formal firms, the average formal firm size *falls* in a district's exposure to improved access to foreign intermediates (see Table 16 in Section III of the Appendix). In absolute terms, the increase in smaller formal firms is largest, consistent with earlier evidence that access to cheaper inputs boosts entry of new, initially typically smaller, firms [Fiorini et al., 2021].

Overall (predicted) effect of India's tariff liberalization across the firm size distribution. The different employment effects of import competition and access to cheaper intermediates on firms of different sizes raise the question how these two compare in explaining the overall employment change at firms of different sizes following India's drastic tariff liberalization. Figure 17 shows how each separate firm size category's predicted district-level employment change between 1990 and 2005 – calculated as in Equation 11, but now using the coefficients in columns (2), (4), (6) and (8) of Table 6 respectively – varies with a district's distance from its nearest main port. Estimates for informal, small, medium and large firms are depicted by blue circles, orange triangles, green squares, and red diamonds respectively; we also plot the corresponding LOWESS curve for each firm size category.

First, the predicted impact of India's tariff liberalization on overall employment within each firm size category is generally positive, with the exception of informal employment which decreased in districts closest to one of India's main ports. Second, a district's employment increase as a result of India's tariff cuts is always predicted to be (much) more pronounced at medium-sized to large firms than at smaller or informal ones. This difference is largest in districts close to one of India's main ports. This can be explained by our earlier findings (see Table 6) that larger firms are better able to withstand the increase in foreign import competition as a result of India's tariff liberalization, while at the same time benefiting more from the better access to foreign intermediates. As the size of both these effects fall in distance to the nearest main port, we see much starker differences in the overall predicted employment changes across firms of different sizes in districts with the best access to India's main ports.

Apart from being in line with the predictions of models such as Melitz [2003] and with the notion that larger (formal) firms are more likely to use foreign intermediates Nataraj [2011], these differential effects of India's trade liberalization across the firm size distribution also provide some further nuances to our discussion on the most likely way in which labor reallocated across sectors following India's tariff liberalization at the end of Section I. Workers not only reallocated across sectors, they also moved from the smallest (least productive, least likely to use foreign intermediates) firms in a district to the larger more productive (more likely to use foreign intermediates) firms – either in the same or in another sector. And, importantly, this type of reallocation was also most pronounced in districts located closer to one of India's main international ports.

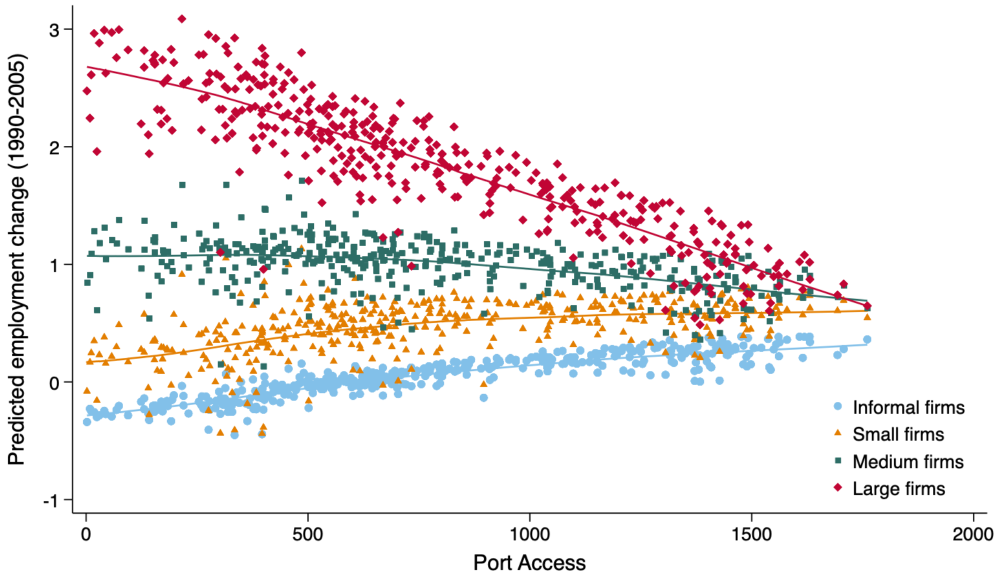


Figure 17: Predicted total employment changes by firm size, 1990-2005

VII. Conclusion

Domestic transport costs are very high in many developing countries, sometimes even exceeding the cost of shipping goods internationally. In this paper, we use detailed historical information on India's domestic and port infrastructure to document how India's high domestic transport costs have mitigated the local labor market effects of its drastic unexpected tariff liberalization in the early 1990s. These high domestic transport costs shield more inland districts from the negative employment effects of increased import competition, while at the same time reducing the positive employment effects of improved access to foreign intermediates. The overall impact of India's drastic tariff reductions on spatial employment patterns across the country then depends on the relative size of these two opposing effects at different distances from the country's main international trade hubs.

In addition, we find that employment at India's small to medium-sized firms suffered most from the increased import competition following its tariff liberalization, whereas medium-sized to large firms benefited most from improved access to foreign intermediates. These differential effects on firms of different sizes are also most pronounced in districts with better access to India's main international ports.

It would be very interesting to see how these findings carry over to other (developing) country settings. Notably, India was characterized by limited spatial labor mobility at the time of its trade liberalization. In settings where people are more mobile, labor reallocation across districts in response to the difference in opportunities arising from a country's improved access to foreign markets could lead to different, possibly even more pronounced, changes in regional employment patterns than what we find in this paper. Second, we focus on the labor market consequences of India opening its market to foreign imports. As such, our findings do not say much about how domestic transport costs condition the effect(s) of having better access to foreign export markets. There are good reasons to believe that districts with better access to a country's main international gateways are also differentially affected when a country's access to foreign export markets improves. Earlier work by Coşar and Fajgelbaum [2016] e.g. already showed that export-oriented industries are disproportionately located near China's main international ports. Levying large, plausibly exogenous changes in a country's export barriers – such as e.g. experienced by Vietnam in 2001 – in combination with detailed information about its domestic infrastructure, ports and transport sector prior to these changes, would be particularly promising to identify whether, and if so how, domestic transport costs also condition the local labor market effects of improved access to foreign export markets.

References

Japan International Cooperation Agency (JICA). Data collection survey on international logistics centered on sri lanka - final report, 2012. Retrieved from: `https://openjicareport.jica.go.jp/pdf/1000027744.pdf`.

Ministry of Ports, Shipping and Waterways and Ministry of Shipping (MoPSW). Basic port statistics 2001-2002, 2003.

P. Aghion, R. Burgess, S. J. Redding, and F. Zilibotti. The unequal effects of liberalization: Evidence from dismantling the license raj in india. *The American Economic Review*, 98(4): 1397–1412, 2008.

T. Allen, D. Atkin, S. Cantillo Cleves, and C. Eduardo Hernández. Trucks. Preliminary draft, 2022.

M. Amiti and J. Konings. Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *The American Economic Review*, 97(5):1611–1638, 2007.

D. Atkin and D. Donaldson. Who's getting globalized? the size and implications of intra-national trade costs. NBER working paper, No.21439, 2015.

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

T. Besley and R. Burgess. Can labour regulation hinder economic performance? evidence from india. *The Quarterly Journal of Economics*, 119(1):91–134, feb 2004. doi: 10.1162/003355304772839533.

T. Bhat. *Structural Changes in India's Foreign Trade: A Study Prepared as a Part of a Research Programme "Structural Changes, Industry, and Employment in the Indian Economy" : Macroeconomic Implications of Emerging Pattern*. Institute for Studies in Industrial Developmen, New Delhi, 2011.

B. Bonadio. Ports vs. roads: Infrastructure, market access and regional outcomes. Working paper, 2022.

K. Borusyak, P. Hull, and X. Jaravel. A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204, February 2025. doi: 10.1257/jep.20231370. URL `https://www.aeaweb.org/articles?id=10.1257/jep.20231370`.

M. Brülhart, C. Carrère, and F. Trionfetti. How wages and employment adjust to trade liberalization: Quasi-experimental evidence from austria. *Journal of International Economics*, 86(1):68–81, 2012.

L. Caliendo and F. Parro. Estimates of the Trade and Welfare Effects of NAFTA. *The Review of Economic Studies*, 82(1):1–44, 2015.

Central Statistical Organisation. Third economic census, 1990. Data retrieved from `https://microdata.gov.in/NADA43/index.php/catalog/142`.

A. Chari and N. Gupta. Incumbents and protectionism: The political economy of foreign entry liberalization. *Journal of Financial Economics*, 88(3):633–656, 2008.

T. Conley. Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1): 1–45, 1999. ISSN 0304-4076. doi: https://doi.org/10.1016/S0304-4076(98)00084-0. URL `https://www.sciencedirect.com/science/article/pii/S0304407698000840`.

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

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

G. Daudin, J. Héricourt, and L. Patureau. International transport costs: new findings from modeling additive costs. *Journal of Economic Geography*, 22(5):989–1044, 05 2022. ISSN 1468-2702. doi: 10.1093/jeg/lbac007. URL `https://doi.org/10.1093/jeg/lbac007`.

J. De Loecker, P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik. Prices, markups, and trade reform. *Econometrica*, 84(2):445–510, 2016.

R. Dix-Carneiro and B. K. Kovak. Trade liberalization and the skill premium: A local labor markets approach. *The American Economic Review*, 105(5):551–557, 2015.

R. Dix-Carneiro and B. K. Kovak. Trade liberalization and regional dynamics. *American Economic Review*, 107(10):2908–46, October 2017. doi: 10.1257/aer.20161214.

R. Dix-Carneiro and B. K. Kovak. Margins of labor market adjustment to trade. *Journal of International Economics*, 117:125–142, 2019.

D. Donaldson. The gains from market integration. *Annual Review of Economics*, 7(1): 619–647, 2015.

B. Erten and J. Leight. Exporting out of agriculture: The impact of wto accession on structural transformation in china. *The Review of Economics and Statistics*, 103(2):364–380, 2021.

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

J. Fan. Internal geography, labor mobility, and the distributional impacts of trade. *American Economic Journal: Macroeconomics*, 11(3):252–288, 2019.

M. Fiorini, M. Sanfilippo, and A. Sundaram. Trade liberalization, roads and firm productivity. *Journal of Development Economics*, 153:102712, 2021. ISSN 0304-3878. doi: https://doi.org/10.1016/j.jdeveco.2021.102712. URL `https://www.sciencedirect.com/science/article/pii/S0304387821000869`.

P. K. Goldberg and N. Pavcnik. The effects of trade policy. In K. Bagwell and R. W. Staiger, editors, *Handbook of Commercial Policy*, volume 1, pages 161–206. Elsevier B.V., North-Holland, 2016.

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

P. K. Goldberg, A. K. Khandelwal, N. Pavcnik, and P. Topalova. Multiproduct firms and product turnover in the developing world: Evidence from india. *Review of Economics and Statistics*, 92(4):1042–1049, 2010b.

G. H. Hanson. Regional adjustment to trade liberalization. *Regional Science and Urban Economics*, 28(4):419–444, 1998.

H. Haralambides and R. Behrens. Port restructuring in a global economy: an indian perspective. *International Journal of Transport Economics*, 27(1):19–39, 2000.

I-Maritime. India Port Report, 2003.

Indian Railways. Statistical summary, 2020. Retrieved from: `https://indianrailways.gov.in/railwayboard/uploads/directorate/stat_econ/Annual-Reports-2019-2020/Summery-sheet-Annual-Report-English_2019-20.pdf`.

K. Kis-Katos and R. Sparrow. Poverty, labor markets and trade liberalization in indonesia. *Journal of Development Economics*, 117:94–106, 2015.

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

KPMG. Logistic game changers: Transforming india's logistics industry. Technical report, CII Institute of Logistics, 2013.

P. Krugman and R. L. Elizondo. Trade policy and the third world metropolis. *Journal of development Economics*, 49(1):137–150, 1996.

A. Kumar. Freight logistics and intermodal transport: Implications for competitiveness. Technical report, Institute for Studies in Industrial Development, New Delhi, 2014. Working paper 175.

H. Kumar and R. Somanathan. State And District Boundary Changes In India (1961-2001). Working papers 248, Centre for Development Economics, Delhi School of Economics, 2016.

C. V. Martincus and J. Blyde. Shaky roads and trembling exports: Assessing the trade effects of domestic infrastructure using a natural experiment. *Journal of International Economics*, 90(1):148–161, 2013.

B. McCaig. Exporting out of poverty: Provincial poverty in vietnam and us market access. *Journal of International Economics*, 85(1):102–113, 2011.

B. McCaig and N. Pavcnik. Moving out of agriculture: structural change in vietnam. NBER working paper, No.19616, 2013.

S. Mehrotra and T. Giri. The size structure of india's enterprises: Not just the middle is missing. CSE Working Paper 25, 2019.

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

M. J. Melitz and S. J. Redding. *Heterogeneous Firms and Trade*, volume 4, pages 1–54. Elsevier, 2014a. URL `http://dx.doi.org/10.1016/B978-0-444-54314-1.00001-X`.

M. J. Melitz and S. J. Redding. Missing Gains from Trade? *The American Economic Review*, 104(5):317–321, 2014b.

Ministry of Railways, India. Report no. 34 of 2010 - compliance audit on railways, union government (railway), 2011. Retrieved from: `https://saiindia.gov.in/rly/allahabad/en/audit-report/details/3220`.

Ministry of Road Transport and Highways (MoRTH). Annual Report 2020-21, 2021.

A. I. Moreno-Monroy, J. Pieters, and A. A. Erumban. Formal sector subcontracting and informal sector employment in indian manufacturing. *IZA Journal of Labor & Development*, 3:1–17, 2014. ISSN 2193-9020. doi: 10.1186/s40175-014-0022-2. URL `https://hdl.handle.net/10419/152397`.

M. Mukim. Coagglomeration of formal and informal industry: evidence from India. *Journal of Economic Geography*, 15(2):329–351, 06 2014. ISSN 1468-2702. doi: 10.1093/jeg/lbu020. URL `https://doi.org/10.1093/jeg/lbu020`.

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

National Transport Development Policy Committee (NTDPC). India transport report: Moving india to 2032. volume 1. Technical report, Planning Commission, Routledge, New Delhi, 2014a.

National Transport Development Policy Committee (NTDPC). India transport report: Moving india to 2032. volume 2. Technical report, Planning Commission, Routledge, New Delhi, 2014b.

National Transport Development Policy Committee (NTDPC). India transport report: Moving india to 2032. volume 3. Technical report, Planning Commission, Routledge, New Delhi, 2014c.

A. Panagariya. India's trade reform: Progress, impact and future strategy. Proceedings from the India Policy Forum, 2004.

M. Pandey. Impact of trade liberalization in manufacturing industry in india in the 1980's and 1990's. ICRIER Working Paper no. 140, 2004.

Planning Commission. Report of the working group on logistics. Technical report, Government of India, New Delhi, 2009.

V. Ponczek and G. Ulyssea. Enforcement of Labour Regulation and the Labour Market Effects of Trade: Evidence from Brazil. *The Economic Journal*, 132(641):361–390, 06 2021. ISSN 0013-0133. doi: 10.1093/ej/ueab052. URL `https://doi.org/10.1093/ej/ueab052`.

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

A. Storeygard. Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *The Review of Economic Studies*, 83(3):1263–1295, 04 2016. ISSN 0034-6527. doi: 10.1093/restud/rdw020. URL `https://doi.org/10.1093/restud/rdw020`.

The Energy and Resources Institute (TERI). Competition issues in regulated industries: Case of the indian transport sector. Technical report, Competition Commission of India, New Delhi, 2008. Project Report no. 2007CP21.

J.-C. Thill and K. Venkitasubramanian. Multi-layered hinterland classification of Indian ports of containerized cargoes using GIS visualization and decision tree analysis. *Maritime Economics and Logistics*, 17:265–291, 2015. doi: 10.1057/mel.2014.24.

P. Topalova. Trade liberalization on productivity: The case of india. International Monetary Fund Working Paper 04/28, 2004.

P. Topalova. Trade liberalization, poverty and inequality: Evidence from indian districts. In *Globalization and Poverty*, pages 291–336. University of Chicago Press, 2007.

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

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

UNCTAD. Freight rates and maritime transport costs, 2008. Retrieved from: https://unctad.org/system/files/official-document/rmt2008_en.pdf.

E. van Leemput. A passage to india: Quantifying internal and external barriers to trade. *Journal of International Economics*, 131:1–21, 2021.

K. Venkita Subramanian and J.-C. Thill. Effect of privatization and inland infrastructural development on india's container port selection dynamics. *The Asian Journal of Shipping and Logistics*, 35(4):220–229, 2019. ISSN 2092-5212. doi: https://doi.org/10.1016/j.ajsl.2019.12.009. URL https://www.sciencedirect.com/science/article/pii/S209252121930077X.

C. Volpe Martincus, J. Carballo, and A. Cusolito. Roads, exports and employment: Evidence from a developing country. *Journal of Development Economics*, 125:21–39, 2017. ISSN 0304-3878. doi: https://doi.org/10.1016/j.jdeveco.2016.10.002. URL https://www.sciencedirect.com/science/article/pii/S0304387816300803.

World Bank Group. States' road infrastructure development technical assistance project: Technical annex, 1996.

World Bank Group. Road transport service efficiency study, 2005.

World Bank Group. Reforming the Indian ports sector (English), 2013. Retrieved from: http://documents.worldbank.org/curated/en/815401468049775225/Reforming-the-Indian-ports-sector.

Y. Xu and X. Yang. Access to ports and the welfare gains from domestic transportation infrastructure. *Journal of Urban Economics*, 126:1–19, 2021.

A. Online Appendix

I. *Model*

I.i. Set up

The model is a variation on the model in Caliendo and Parro [2015]. There are two countries: India and the Rest of the World, which is indexed by F. India is divided in $\mathcal{D}$ districts, indexed by d. Let $\mathbb{D} = 1, ..., \mathcal{D}, F$ be the full set of locations. Production takes place in three distinct industries: agriculture (A), manufacturing (M) and nontradables (N). Manufacturing consists of $\mathcal{S}$ sectors, indexed by superscript s. Each sector is represented by one good which has a continuum of varieties $\omega_s \in [0, 1]$. The representative consumer inelastically supplies labor to these three different industries: L_d^A, L_d^M, L_d^N, where $L_d^A + L_d^M + L_d^N = 1$.

I.ii. Production

Agriculture and nontradables follow the production function:

$$Q_d^A = T_d^A L_d^A; \quad Q_d^N = T_d^N L_d^N,$$

where T_A and T_N respectively refer to the productivity in location d at producing A or N respectively.

Production in manufacturing requires labor and an intermediate input, which is a composite of all varieties $\omega^s \in [0, 1]$:

$$q_d^s(\omega^s) = z_d^s(\omega^s)[l_d^s(\omega^s)]^{\gamma^s} \prod_{j \in S} [m_d^{s,j}(\omega^s)]^{\gamma^{s,j}},$$

where $l_d^s(\omega^s)$ is the labor used in producing ω^s, $m_d^{s,j}(\omega^s)$ represents the composite intermediate from sector j used in the production of ω^s and $\gamma^s > 0, \gamma^{s,j} \geq 0$ and $\sum_j \gamma^{s,j} = 1 - \gamma^s$. Here, γ^s is the share of labor (value added) in production of s and $\gamma^{s,j}$ reflects the share of materials from sector j used in producing good s. Solving for cost minimization yields the following marginal cost function:

$$MC = AC = H^s \frac{[w_d^s]^{\gamma^s} \prod_{j \in S} [P_d^j]^{\gamma^{s,j}}}{z_d^s(\omega^s)} = \frac{c_d^s}{z_d^s(\omega^s)}, \tag{12}$$

with $H^s \equiv \prod_j (\gamma^s)^{-\gamma^s} (\gamma^{s,j})^{-\gamma^{s,j}}$.

I.iii. Composite intermediates

The intermediate input per sector is the composite of all lowest-cost varieties available in district d:

$$Q_d^s = \left[\int_{\Omega^s} \left[r_d^s(\omega^s) \right]^{\frac{\sigma^s - 1}{\sigma^s}} \, \mathrm{d}\omega^s \right]^{\frac{\sigma^s}{\sigma^s - 1}}.$$

Here, $\sigma^s \geq 0$ is the elasticity of substitution between varieties of the good of sector s. Again, applying cost minimization gives us the demand per variety s in district d:

$$r_d^s(\omega^s) = \left[\frac{p_d^s(\omega^s)}{P_d^s}\right]^{-\sigma^s} Q_d^s,$$

where P_d^s is the price of the composite intermediate good; it is given by the Dixit-Stiglitz price index:

$$P_d^s = \left[\int_{\Omega^s} [p_d^s(\omega^s)]^{1-\sigma^s} \, \mathrm{d}\omega^s\right]^{\frac{1}{1-\sigma^s}}.$$

I.iv. Representative household

A representative consumer working in district d has the following preferences:

$$U_d = (C_d^M)^{b_d^M}(C_d^A)^{b_d^A}(C_d^N)^{b_d^N} \tag{13}$$

$$C_d^M = \prod_{j\in S}(C_d^j)^{\alpha_d^j} \tag{14}$$

$$C_d^A = \left[\sum_{i\in\mathbb{D}}(C_{di}^A)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}} \tag{15}$$

where both the industry preference shifters and the sector preference shifters sum to one, such that $b_d^M + b_d^A + b_d^N = 1$ and $\sum_j \alpha_d^j = 1$. Moreover, the elasticity of substitution on agricultural goods from different locations i is equal to $\epsilon \geq 0$.

The consumer maximizes utility subject to its budget constraint, which consists of the household's income out of the three sources outlined in subsection i:

$$E_d = w_d^M L_d^M + w_d^A L_d^A + w_d^N L_d^N$$

This means that the optimal quantity consumed in each industry A, M, N is:

$$C_d^M = b_d^M \frac{E_d}{P_d^M}; \quad C_d^A = b_d^A \frac{E_d}{P_d^A}; \quad C_d^N = b_d^N \frac{E_d}{P_d^N},$$

where P_d^t is the price index for each industry t. Next, we compute the optimal consumption in each manufacturing sector:

$$C_d^s = \alpha_d^s b_d^M \frac{E_d}{P_d^S} \tag{16}$$

and for each source of agricultural production:

$$C_{di}^A = b_d^A \frac{E_d}{P_d^A}\left(\frac{P_{di}^A}{P_d^A}\right)^{-\epsilon} \tag{17}$$

Note that this implies the following agricultural price index:

$$P_d^A = \left[\sum_{i \in \mathbb{D}} \left(\frac{w_i^A D_{di}^A}{T_i^A} \right)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \tag{18}$$

I.v. Trade costs and prices

The rest of the world is also characterized by (13) and (20). In accordance with India's limited global presence before 1991, we assume India cannot affect world prices. This implies that the districts take the unit cost of RoW as exogenously given by $c_{s,RoW}$ $\forall s$. Trade is subject to both domestic and international good-specific trade costs and follows:

$$D_{di}^s = \begin{cases} \delta_{di} \geq 1 & \text{if } d, i = 1, ..., \mathcal{D} \\ \delta_{dp} \delta_{pF} \tau_{imp}^s & \text{if } d = 1, ..., \mathcal{D}, i = F \\ \delta_{pd} \delta_{Fp} \tau_{exp}^s & \text{if } d = F, i = 1, ..., \mathcal{D} \end{cases},$$

where τ_{imp}^s and τ_{exp}^s denote import and export tariffs respectively and δ_{di} reflect domestic and international transport costs. Note that any imported or exported goods needs to go through the port p and cannot be shipped directly to or from any district in India.

Under perfect competition, only the lowest-cost producer will provide a variety. Thus, the price a consumer in district d pays for a variety ω_s equals:

$$p_d^s(\omega^s) = \min_{i \in \mathbb{D}} \{p_{di}^s(\omega^s)\} = \min_{i \in \mathbb{D}} \left\{ \frac{c_i^s D_{di}^s}{z_i^s(\omega^s)} \right\}. \tag{19}$$

To simplify the analysis, we assume that productivity at producing variety ω_s in district d follows a sector-specific Fréchet distribution:

$$z_d^s(\omega^s) \sim F_d^s(z) \equiv \exp\left(-T_d^s z^{-\theta^s}\right), \tag{20}$$

where:

1. T_d^s governs absolute advantage: a higher T_d^s implies a higher likelihood of higher productivity draws.

2. θ^s governs comparative advantage: a higher θ^s implies a lower productivity dispersion across varieties, implying that comparative advantage is less important in determining trade flows.

Then, we can compute the probability that $i \in \mathbb{D}$ provides ω_s at a price below p in district d:

$$G_{di}^s(p) = 1 - \exp\left\{ -T_i^s \left(\frac{c_i^s D_{di}^s}{p} \right)^{-\theta^s} \right\}.$$

The probability that a consumer in district d pays less than p for variety ω_s is:

$$G_d^s(p) = 1 - \exp\left\{-\Phi_d^s p^{\theta^s}\right\}, \tag{21}$$

where

$$\Phi_d^s = \sum_{i \in \mathbb{D}} T_i^s (c_i^s D_{di}^s)^{-\theta^s}. \tag{22}$$

We can rewrite the district-sector-specific price index, given that $\theta^s > \sigma^s - 1$, as follows:

$$P_d^s = g^s (\Phi_d^s)^{-\frac{1}{\theta^s}}, \tag{23}$$

where $g^s = \Gamma\left(\frac{\theta^s - (\sigma^s - 1)}{\theta^s}\right)^{\frac{1}{1-\sigma^s}}$ with $\Gamma(\cdot)$ as the gamma function.

I.vi. Expenditure shares

The probability that district i is the lowest-cost supplier of variety ω^s in district d – which because of the Law of Large Numbers equals the fraction of goods i supplies to d – is given by:

$$\pi_{di}^s = T_i^s (c_i^s D_{di}^s)^{-\theta^s} \Phi_d^s. \tag{24}$$

Total expenditure on sector s in district d is given by:

$$X_d^s = \underbrace{\alpha_d^s b_d^M E_d}_{\text{Final demand}} + \underbrace{\sum_j \gamma^{s,j} \sum_{i \in \mathbb{D}} \pi_{id}^s X_i^s}_{\text{Intermediate demand}}. \tag{25}$$

Then, bilateral sector-specific trade flows are given by:

$$X_{di}^s = \pi_{di}^s X_d^s. \tag{26}$$

Note that expenditure on agriculture and nontradables in district d equals:

$$X_d^A = b_d^A E_d; \quad X_d^N = b_d^N E_d. \tag{27}$$

The total wage income in sector s in district d is given by the cost share in production:

$$w_d^s L_d^s = \gamma^s \sum_{i \in \mathbb{D}} X_{id}^s. \tag{28}$$

Then, using (26) and our assumption of free labor mobility across sectors within a region implying $w_d^s = w_d^M$ in equilibrium, we arrive at the following condition that gives the labor market equilibrium in manufacturing:

$$w_d^M L_d^M = \sum_{s \in S} \gamma^s \sum_{i \in \mathbb{D}} \pi_{id}^s X_i^s. \tag{29}$$

Following a similar analysis, the condition for labor market equilibrium in agriculture in district d is given by:

$$w_d^A L_d^A = \sum_{i \in \mathbb{D}} \left[\frac{p_{dd}^A D_{id}^A}{P_i^A} \right]^{1-\epsilon} b_i^A E_i, \tag{30}$$

and for the nontradable sector we get:

$$w_d^N L_d^N = P_d^N Q_d^N = b_d^N E_d. \tag{31}$$

Finally, trade balance implies:

$$\sum_s \sum_d \pi_{dF}^s X_d^s + \sum_d \left[\frac{p_{FF}^A D_{dF}^A}{P_d^A} \right]^{1-\epsilon} b_d^A E_d = \sum_s \sum_d \pi_{Fd}^s X_F^s + \sum_d \left[\frac{p_{dd}^A D_{Fd}^A}{P_F^A} \right]^{1-\epsilon} b_F^A E_F$$

I.vii. Equilibrium

Equilibrium. For given $\{T_d^s\}$, $\{D_{di}^s\}$, $\{L_d^t\}$, $\{\gamma^s\}$ and $\{\gamma^{s,j}\}$, an equilibrium is a set of sector-district-specific price indices $\{P_d^s\}$, district-specific prices for agriculture and nontradables $\{P_{di}^A\}$, $\{P_d^N\}$, wages in all industries $\{w_d^j\}$ $\forall j = \{M, A, N\}$, $d, i \in \mathbb{D}$, and $\forall s \in S$ such that:

1. Sector-state specific productivities follow (20);
2. Consumers maximize utility (13);
3. Consumers buy from the lowest-cost supplier, such that (19) gives the price of the acquired varieties;
4. labor markets clear.

The equilibrium is then characterized by equations (12), (23), (18), (24), (26), (27), (28), (30) and (31). Note that labor market equilibrium implies balanced trade.

I.viii. Equilibrium in relative changes

We can now see how a reduction in the country's import tariffs affect this equilibrium. Following Dekle et al. (2008), we can express this in terms of the relative changes after a policy change from τ_{imp} to τ'_{imp}, where $\tau_{imp} > \tau'_{imp}$.

Equilibrium. Let $(\mathbf{w}, \mathbf{P})$ be the equilibrium under tariff structure τ_{imp} and $(\mathbf{w}', \mathbf{P}')$ the equilibrium under τ'_{imp}. Let $\hat{x} = x'/x$ denote the relative change in x. Then, the equilibrium conditions in relative changes satisfy:

1. Cost of the input bundles:

$$\hat{c}_d^s = (\hat{w}_d^s)^{\gamma^s} \prod_{j \in S} (\hat{P}_d^j)^{\gamma^{s,j}}. \tag{32}$$

2. The district-sector-specific price index:

$$\hat{P}_d^s = (\hat{\Phi}_d^s)^{-1/\theta^s} \equiv \left[\sum_{i \in \mathbb{D}} \pi_{di}^s (\hat{c}_i^s \hat{D}_{di}^s)^{-\theta^s} \right]^{-1/\theta^s}. \tag{33}$$

3. The agricultural price index:

$$\hat{P}_d^a = \left[\sum_{i \in \mathbb{D}} \pi_{di}^A \left(\hat{w}_i^A \hat{D}_{di}^A \right)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \tag{34}$$

4. Bilateral trade shares:

$$\hat{\pi}_{di}^s = \left[\frac{\hat{c}_i^s \hat{D}_{di}^s}{\hat{P}_d^s} \right]^{-\theta^s}. \tag{35}$$

5. Total expenditure:

$$X_d' = E_d' + \sum_j \gamma^{s,j} \sum_{i \in \mathbb{D}} \pi_{id}^{j'} X_i^{j'}. \tag{36}$$

6. Labor market clearing in each manufacturing sector:

$$w_d^{s'} L_d^{s'} = \gamma^s \sum_{i \in \mathbb{D}} X_{id}^{s'}. \tag{37}$$

7. Labor market clearing in agriculture:

$$w_d^{A'} L_d^{A'} = \sum_{i \in \mathbb{D}} \left[\frac{p_{dd}^{A'} D_{id}^{A'}}{P_i^{A'}} \right]^{1-\epsilon} b_i^{A'} E_i' \tag{38}$$

We then derive the effect of a change in international trade costs on wages to be:

$$\begin{aligned} \ln \hat{w}_d^s = & \underbrace{- \sum_s \frac{\alpha_d^s}{\theta^s \gamma^s} \ln \left[\hat{\pi}_{dF}^s \left(\frac{\hat{c}_d^s}{\hat{D}_{dF}^s} \right)^{-\theta^s} \right]}_{\text{Import competition}} \\ & \underbrace{- \sum_s \frac{\alpha_d^s}{\gamma^s} \left[\sum_{j \in S} \ln \left[[\hat{\pi}_{dF}^s]^{-\frac{1}{\theta^s}} \hat{D}_{dF}^s \right] \right]}_{\text{Imported inputs}} \\ & + \sum_s \frac{1}{\gamma^s} \ln \hat{P}_d^s. \end{aligned} \tag{39}$$

This equation shows how the local labor market effects of the country's tariff reductions is determined by two distinct channels: increased import competition on the one hand, and access to cheaper foreign intermediates on the other hand.[62] Note that if wages are sticky, this shock would affect the quantity demanded of labor. The former puts downward pressure on local wages (labor demand), the latter instead increases wages (labor demand). Both effects are stronger in districts located close to the country's international port.

Domestic infrastructure conditions the effects of import tariff liberalization To see this, consider the following thought experiment. Imagine two districts that are completely identical in every respect, with the only difference being that one district is located close to the country's international port and the other very far from it, such that the transport costs involved in importing foreign goods are much lower in the latter district: $D^s_{dF} << D^s_{iF}$. In this stylized example, it is obvious that the initial expenditure share on foreign goods, i.e. that before the country's tariff liberalization, as defined in Equation 24, is higher in the coastal district compared to the inland district: $\pi^s_{dF} > \pi^s_{iF}$.

Next, the country substantially lowers its import tariffs. This, see Equation 33, affects the district-sector-specific price index much more drastically in the coastal district, as that region was already sourcing more of its goods from abroad, so that $\hat{P}^s_d > \hat{P}^s_i$. This by Equation 35 implies that the relative change in the foreign expenditure share, following the country's tariff reduction, is higher in the coastal district, or $\hat{\pi}^s_{dF} > \hat{\pi}^s_{iF}$. With this intuition in hand, we can analyze how the international trade shock affects local labor markets.

First, it increases the degree of import competition faced in all districts, the first term in Equation 39, putting downward pressure on wages. But, since $\hat{\pi}^s_{dF}$ is larger for districts with better access to the country's international port, this means that this negative effect on wages is larger for districts close to the port. Inland districts are relatively more shielded from the import competition channel. *Second*, it lowers the costs of foreign intermediate inputs, the second term in Equation 39. Given that districts with better access to the country's international port were already sourcing a larger share of their intermediate inputs from the ROW (i.e. $\hat{\pi}^s_{dF}$ is larger), this also affects these districts more than those farther away. Specifically, firms in coastal districts profit more from input tariff declines then their more inland counterparts, which disproportionally improves their profitability and thereby the wages they pay their workers. Earlier papers have pointed out the importance of this channel in determining the gains from trade, especially through its effects on firm productivity [Amiti and Konings, 2007, Nataraj, 2011].

[62]The last term captures any general equilibrium effects.

II. *Additional details on domestic infrastructure*

Domestic infrastructure network The road and rail network both draw on the *Road Map of India* created by the Indian government. The first edition of this map was published in 1960; we use three subsequent editions printed in 1977, 1988 and 1996 respectively.[63] Scanned copies of these static maps were obtained from the World Bank. First, the maps were georeferenced to create a correspondence with the administrative boundaries of India in 1988, and then the infrastructure network was carefully digitized by hand. These maps also provided us with the location of the historical district headquarters, which we use to proxy for the district's access to international markets.

For the road network, this process resulted in three shapefiles, or spatial data layers – one for each year in which a map was available. A notable feature of these maps is the road classification system, which helped us assign appropriate travel speeds along each road. The maps display three distinct types of roads, where the higher classes provide superior facilities: primary (Expressways and National Highways), secondary (State roads and Major District roads) and tertiary roads (rural roads). Also, for secondary and tertiary roads we have information on whether the road is accessible in all weathers or merely fair weather. This distinction is of high importance since less than half of all roads are paved, and only a subset of these are estimated to be in good condition [TERI, 2008]. This hierarchy was preserved in the shapefiles. Unfortunately, due to the age of the maps and the quality of the scans, some details in the map were difficult to discern. Therefore, some assumptions on the continuity of roads had to be made based on how the roads appeared. The primary roads were cross-referenced with the location of modern highways on Indian maps; for the secondary and tertiary roads, such information was unavailable. Another consideration that was taken into account were nodes, or points where two or more shapefile segments intersect and connect to each other. We assumed that any roads that intersected on the map also intersected in real life (as opposed to overpasses), and that all intersections would be mapped as connection nodes.

Mapping out the temporal variation in the railway network yielded two shapefiles, depicting the rail coverage in different years: 1988 and 1996. All railway segments are given the same hierarchy. Much like the road data, breaks in the railway lines on the maps posed issues to data creation; this rendered the network in 1977 unusable. To overcome this difficulty, we verify that the routes on the rail network correspond to the road network – if there is a rail, there is a way. This means we can use a pared-down version of the road network to verify rail infrastructure investments in the period before the reform.

Figure 18 shows the level of detail on the map for the area around Kolkata.

Ports and waterways Besides international trade, ports could also facilitate in interstate freight traffic. The two modes that would be particularly suitable given India's geography

[63]See Appendix II for more details.

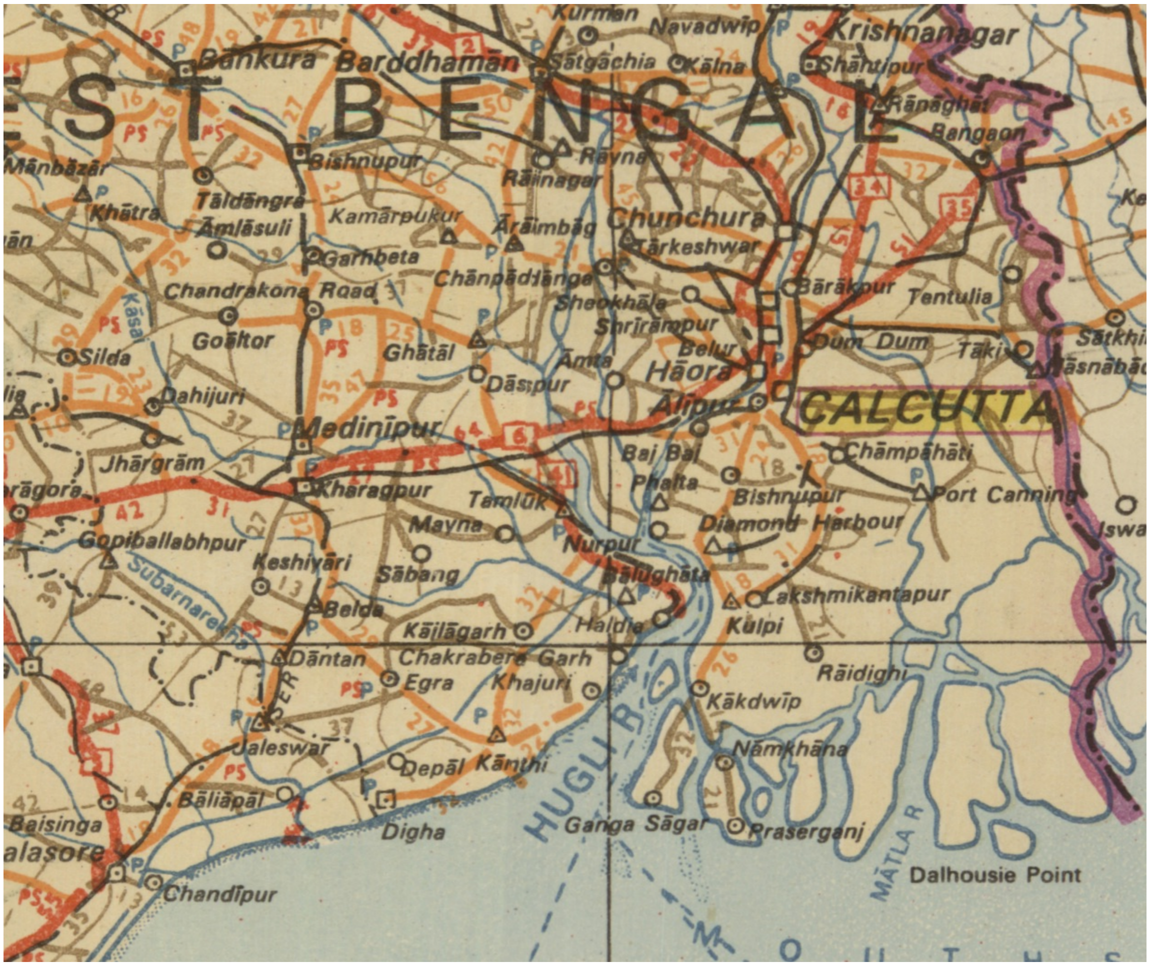


Figure 18: Infrastructure in the area around Kolkata on the *Road Map of India* (1977)

are coastal shipping and inland water transit (IWT). Coastal shipping is up to 25 percent more affordable than conventional modes of transport, alleviates congestion on roads and rail and is more environmentally friendly. India's coastline of more than 7,500 km has provided many trade opportunities in the past, with coastal shipping tonnage as a fraction of total tonnage equalling 57 percent in 1950 [NTDPC, 2014b]. However, due to limited investment in port technology and ships, this has dropped to 8.7 percent in 1990. As mentioned before, Minor ports suffer from inadequate equipment and infrastructure, meaning that most of coastal shipping relies on Major ports. The complication is that Major ports operate under constraints: in 1990-91, capacity utilization at Major ports equalled 95 percent; the average turnaround time in the same period was 8.1 days [NTDPC, 2014b, Kumar, 2014]. India's inland water transit is also relatively underdeveloped, because of both geographical constraints and the lack of public investment to overcome these difficulties [NTDPC, 2014c]. The majority of Indian Waterways has insufficient depth for larger vessels, which limits the economies of scale that can be achieved through IWT. According to the NTDPC [2014c], there are also multiple bridges that prevent vessels from coming through and therefore not enough vessels in the first place.

III. Mapping districts across census years

This subsection describes how exactly we ensured India's districts were comparable over time despite the large increase in the number of districts and substantial transfers over time.[64] We chose 1987-88 as our baseline, which corresponds to the NSS round 43. To construct the district-level panel dataset, we thus need to map all future districts into these 1987 boundaries.

There are two types of boundary changes we need to take into account: the formation of a new district and a transfer to another district. Ideally, we would reconstruct the baseline districts based on economic activity; this is however not possible from the available data. Since population is at least partially correlated with economic activity, we opt for population weights instead. Kumar and Somanathan [2016] have compiled such population weights for every intercensal boundary change between 1971 and 2001 in terms of both the new and old district. Note that these weights are calculated with respect to the population in the closest subsequent census year; this is the most precise statistic we can achieve with the available data. We complement this with exact information on the timing of these boundary changes to ensure our district-level panel is as precise as possible. For all newly created districts, this information was readily available from the 2011 District Census Handbooks provided by the Ministry of Home Affairs. However, since only 'notable' transfers were mentioned in these census documents, just 17 out of 34 transfers could be dated with certainty. Five more transfers could be dated using other sources.[65] For the remaining transfers, we have to venture into educated guesses. Given that boundary changes within a state are often clustered, we could allocate a date to 12 more transfers based on another confirmed transfer in the same state. The remaining transfer is assigned the year in which the most boundary changes in the whole of India occurred. A complete overview of these transfers is provided in Table 7.

III.i. General strategy

In most cases, the mapping is straightforward. In what follows, we refer to any district that was carved out of another district(s) a 'child' district; the original district is called the 'parent'. We classify districts into three categories: districts with unchanged boundaries, districts that split from an existing district and districts whose current boundaries contained multiple districts in the previous census-year. Note that the latter category contains

[64]We consider a transfer substantial if the transferred area corresponds to more than one percent of the sending or receiving district's population.

[65]The main source for this is the website Statoids, which is a supplement to *Administrative Subdivisions of Countries* (1999) by Gwillim Law. The other source is the official website of the Idukki district.

both new districts with multiple 'parents' and districts that were on the receiving end of a transfer. The first category poses no difficulty: districts whose boundaries have remained unchanged between 1983 and 2005 are assigned a weight equal to one. Districts in the second category are addressed based on whether it was subdivided before or after 1987. The five districts in our dataset that were partitioned between 1983 and our baseline year 1987 are assigned to the 1987 baseline districts based on population weights. For example, Anna split from Madurai in 1985. Since the population share that remained in Madurai is equal to 65.51%, we assign 34.49% of Madurai to Anna between 1983 and 1985. The new districts that originated from a simple partition after 1987 are fully assigned to their 'parent' districts to recover the population of the 1987 boundaries corresponding to the parent district. Thus, if in 1988 Siddarth Nagar splits from Basti, both districts are fully allocated to Basti in any dataset from 1988 onwards.

For the third category of districts, those whose current boundaries encompass multiple districts in the previous census year, the reconstruction of the district is more complex. Most transfers, be it to a new or existing district, consist of a part of a tehsil or taluk. From the General Population Tables we can observe how many villages were in the transferred area and how many people lived there, but not which exact villages were part of the transfer. Thus, we cannot track the economic activity in the transferred district across multiple census years. This is most important when a district (1) transfers to more than one district or (2) both sends and receives transfers. To circumvent this issue we make two assumptions. First, we assume that the characteristics of interest are (sufficiently) distributed in proportion to population across the parent and the child district; this mitigates the complication that the transferred district is not clearly defined. Similarly, we assume that the population growth in connected child and parent districts is the same over time, which ensures that we can use the same population weights across multiple census years.

Given these restrictive but necessary assumptions, mapping these districts to the 1987 baseline proceeds as follows. First, consider a child district that has multiple parents. We use the share of each parent's original population in the child district's population in the subsequent census year to divide the child district over these parent districts. Given our assumptions, this ensures that the 1987 boundaries are recovered. For example, in 1989 Firozabad was formed from a part of Agra (corresponding to 20.89% of Agra's population) and a part of Mainpuri (38.48% of its population). According to the 1991 census, this translates to 47.29% of Firozabad originally belonging to Agra, and 52.71% to Mainpuri. Therefore, a share of 0.5271 of Firozabad is assigned to Mainpuri and the remaining 0.4729 is assigned to Agra from 1989 onwards. In case of transfers, we adopt a similar methodology. In 1989, a part of West Siang, corresponding to 7.88% of its population, was transferred to East Siang. This transfer is equivalent to 7.66% of East Siang's population in 1991, the closest census year. To restore the 1987 boundaries, this 7.66% of East Siang is reassigned to West Siang in any dataset after 1989. Again, we assume the population growth rates to be equivalent in interconnected districts, such that these weights remain the same over time. Table 7 shows the substantial transfers, which encompass at least one percent of the population in either the sending or receiving district, that we account for in our sample creation.

III.ii. Special cases

This subsection discusses certain special cases where the aforementioned general strategy falls short:

(i) **Haryana**: This state is subject to the largest number of boundary changes between 1981 and 2001, whereof two-thirds consists of transfers. Six districts exchange territories between themselves so often that we had to simplify the transfer patterns substantially.[66] Specifically, we only correct for *net* transfers when two districts exchange parts of land, and only register transfers that fall outside the 1987 boundaries. For example, Panipat was formed in 1989 from almost half of Karnal and a small part of Jind. Then, in 1996, 14.1% of Panipat was transferred to Karnal, while 1.91% of Karnal was transferred to Panipat. Normally, we would calculate the size of the net transfer and assign this such that the 1987 boundaries are restored. However, it is very likely that these transfers only concern the grounds of the 1987 Karnal district. Therefore, we do not correct for this transfer in the dataset.

(ii) **Gujarat**: As mentioned before, round 38 and 50 of the NSS only provide data on the regional level. Since seven districts in Gujarat cover two economic regions, they are assigned two separate identifiers in all the rounds to guarantee internal consistency.[67]. However, this set-up is not compatible with the ASI and EC. To that end, we have merged the two NSS codes for these districts and mapped it into the ASI and EC datasets. Then, we have assigned the economic regions to each merged NSS identifier based on the relative populations in each region in 1987. This allows us to still use these region-level datasets.

(iii) **EC05**: For new districts since 2001, there is no existing data on the population weights of the parent and the child districts. For three new districts, which had multiple parents, we had to create these ourselves. Using the District Census Handbooks we could match the individual tehsils that made up the new district to the parents and calculate the population weights based on the tehsil-level demographics provided in the 2011 census.

[66]These districts are Jind, Kaithal, Karnal, Kurukshetra, Rohtak and Sonipat.

[67]These districts are Bharuch, Mehesana, Panch Mahals, Sabarkantha, Surat, Vadodara and Valsad

Table 7: Transfers between 1981-2001

Year	State	Sending district	Share of sending district	Receiving district	Share of receiving district
1986	Nagaland	Kohima	4.54	Zunheboto	15.66
1986	Nagaland	Tuensang	9.99	Mon	16.17
1986	Rajasthan	Kota	1.61	Bundi	4.09
1987$^{\diamond}$	Tripura	North Tripura	1.49	South Tripura	1.48
1988$^{\diamond}$	Uttar Pradesh	Garwhal	1.45	Chamoli	2.48
1989	Arunachal Pradesh	West Siang	7.88	East Siang	7.66
1989†	Haryana	Faridabad	1.54	Gurgaon	1.78
1989†	Haryana	Karnal	1.67	Kurukshetra	4.05
1989	Haryana	Sonipat	30.96	Rohtak	17.10
1989$^{\diamond}$	Maharasthra	Solapur	0.73	Osmanabad	1.84
1992$^{\diamond}$	Megalaya	East Khasi Hills	0.35	West Khasi Hills	1.06
1993	Assam	Dubri	0.61	Kokrajhar	1.00
1996	Gujarat	Ahmadabad	4.46	Gandhinagar	19.89
1996†	Gujarat	Amreli	13.24	Junagadh	7.93
1996	Gujarat	Bhavnagar	9.69	Amreli	16.97
1996	Gujarat	Kheda	0.33	Gandhinagar	1.07
1996	Gujarat	Mahasana	15.07	Gandhinagar	41.08
1996†	Haryana	Jind	1.05	Kaithal	1.29
1996†	Haryana	Jind	0.87	Karnal	0.81
1996†	Haryana	Kaithal	3.82	Jind	3.20
1996†	Haryana	Kaithal	0.96	Karnal	0.76
1996†	Haryana	Kaithal	1.66	Kurukshetra	2.03
1996†	Haryana	Karnal	1.91	Panipat	2.43
1996†	Haryana	Karnal	0.24	Kurukshetra	0.31
1996	Haryana	Panipat	18.28	Karnal	14.71
1996	Haryana	Rohtak	16.05	Sonipat	27.77
1996†	Haryana	Rohtak	1.43	Bhiwani	2.22
1996†	Haryana	Yamuna Nagar	1.9	Kurukshetra	2.33
1996	Madhya Pradesh	Gwalior	8.43	Datia	23.10
1996‡	Punjab	Sangrur	1.44	Ludhiana	1.02
1996$^{\diamond}$	Tamil Nadu	Ramanathapuram	2.18	Siviganga	2.26
1996	Uttar Pradesh	Barabanki	12.81	Faizabad	18.42
1996	Uttar Pradesh	Kanpur Dehat	39.05	Kanpur Nagar	25.67
1998$^{\diamond}$	Kerala	Idukki	2.14	Ernakulam	0.81

This table lists all transfers between Indian districts over the period 1981-2001, provided that they were sufficiently large: the transfer size as a share of the population of either the sending or receiving district exceeds one percent. This share is based on the population size of the districts concerned in the previous census year. For the unmarked transfers, the exact date was retrieved from the District Census Handbooks. Transfers marked with $\diamond$ are dated based on other sources. Transfers marked with † are dated based on other transfers in the same state in the same census period and the transfer marked with ‡ is assigned the year in the census period with the most boundary changes.

B. Appendix

I. Additional tables

Even though we do not necessarily rely on the exogeneity of our transport cost proxy to estimate the coefficients of interest, we can use our data on India's road network prior to the shock to test for endogenous placement of infrastructure in anticipation of its liberalization.[68] To be more specific, we use our dataset of all routes between district headquarters, ICDs and ports in 1977 and 1988, and calculate the change in each route's distance, and travel time, as a result of either newly constructed roads or the upgrading of roads from e.g. a secondary to a primary road. Using this information, we then estimate the following equation:

$$\begin{aligned} \mathrm{d}\ln D_{ij}^{1988-1977} = {} & \beta_1 IP_{ij} + \beta_2 ICD_{ij} + \beta_3 AB_{ij} + \beta_4 AA_{ij} + \beta_5 Met_{ij}^1 + \beta_6 Met_{ij}^2 \\ & + \beta_7 \ln D_{ij}^{1977} + \alpha_1 S_{ij}^1 + \alpha_2 S_{ij}^2 + \epsilon_{ij}, \end{aligned} \tag{40}$$

where i and j are the end nodes on a route and $\mathrm{d}\ln D_{ij}^{1988-1977}$ is the 1977-1988 change in the route's log shortest distance, or travel time. IP_{ij} and ICD_{ij} are dummy variables equal to one if one of the route's end notes is an important port or ICD respectively. AB_{ij} and AA_{ij} are dummy variables that equal one if one or two of the route's end nodes respectively have above median population; Met_{ij}^1 and Met_{ij}^2 are dummy variables equal to one if one or two of the route's end nodes respectively are among the fifteen largest cities in India in 1981. We furthermore control for the initial route-specific log shortest distance or travel time, $\ln D_{ij}^{1977}$, and add route-specific state fixed effects, S_{ij}^1 and S_{ij}^2.[69] ϵ_{ij} captures any remaining unobserved route-specific idiosyncratic variables affecting changes in $\mathrm{d}\ln D_{ij}^{1988-1977}$.

Results in columns (1) and (2) in Table 8 show that there is no evidence of differential road construction or upgrading on routes to the port in anticipation of India's trade reforms. Instead, in the years leading up to India's trade liberalization, novel road construction primarily focused on (better) connecting India's more populated districts, whereas road upgrading occurred mostly in the lesser populated districts of the country. This is not surprising as roads connecting the larger districts were often already of better quality. More specifically: routes with an important port at one of the end nodes saw a relative decrease in accessibility over the 1977-1998 period. All other routes saw their distance decrease by 0.8 percent more as a result of novel road construction, and their travel time by 0.1 percent more than routes connected to ports, albeit not significantly so.

[68] India's rail network did not change substantially over the years prior to its trade liberalization.

[69] More specifically we add route-specific dummies denoting whether none, one (S_{ij}^1) or both of a route's end nodes (S_{ij}^2) are in a certain state.

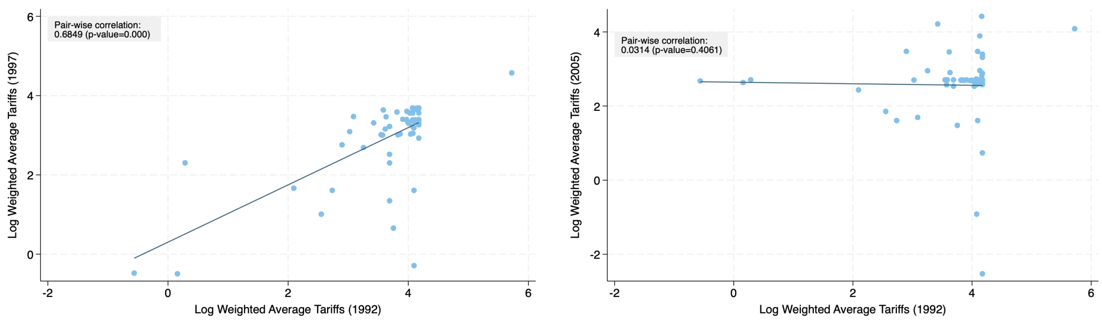
Table 8: Ports are relatively inaccessible

	Log difference 1977-1988	
	(1) Distance	(2) Time
Important Port	0.008***	0.001
	(0.001)	(0.001)
ICD	0.001	-0.001
	(0.0009)	(0.001)
Above-below median population	-0.004***	0.001*
	(0.001)	(0.001)
Above-above median population	-0.006***	0.003***
	(0.001)	(0.001)
One metropole	0.00002	0.002**
	(0.001)	(0.001)
Both metropoles	-0.0005	0.009*
	(0.004)	(0.005)
Log distance 1977	-0.017***	
	(0.001)	
Log time 1977		-0.026***
		(0.001)
Route-specific State FE	Yes	Yes
Observations	91378	91378
R-squared	0.108	0.205

The dependent variable is either the log difference in travel distance or time between 1977 and 1988 on all routes between district headquarters, ports and ICDs. Each route is uniquely characterized by its origin and destination combination, which means we only consider the upper triangular (excluding the diagonal) of the full route matrix. The *Important Ports* (Kolkata, Visakhapatnam, Kandla, Mumbai and Chennai) each unloaded more than ten percent of imports in 1992. *Above-below median pop.* equals one if one end node of the route had an above median population, and *Above-above median pop.* is one if both nodes on the route had an above median population in 1991, while *One* or *Both metropoles* denotes whether none, one or both nodes were among the ten largest cities in India in 1981. Finally, *Log distance 1977* and *Log time 1977* control for the route accessibility in 1977; the same holds for 1988. State fixed effects are determined at the node level, as we control for whether one or both nodes are in a specific state. Robust standard errors are in parentheses.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

This pattern is corroborated in Figure 3, which depicts the spatial placement of infrastructure construction or upgrading over the 1977-1998 period. It shows the location, in red and blue, of new or upgraded secondary and primary roads in 1988 respectively. As can be seen, a sizable share of infrastructure investments in the pre-liberalization period was undertaken in the middle of India and in its northern and northeastern regions, and not on routes connecting its main ports.

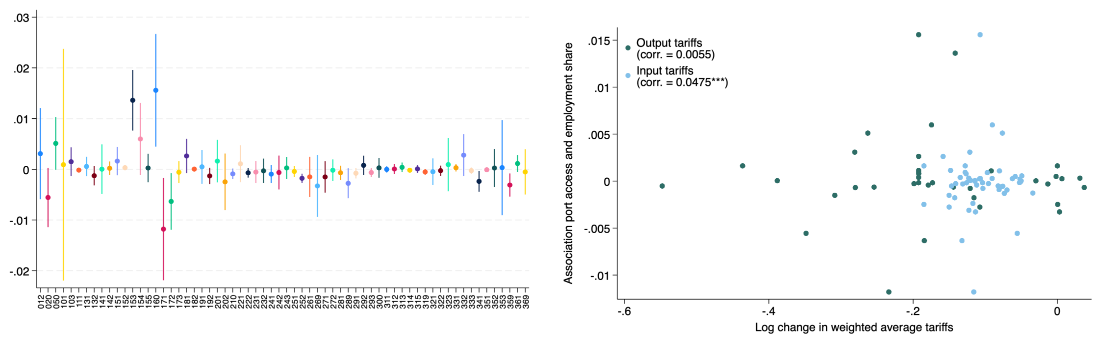
II. *Additional figures*



(a) 1992-1997

(b) 1992-2005

Figure 19: Correlation between tariffs one year post-reform and EC rounds



(a) Association between port access and employment shares

(b) Correlations with tariff changes

Figure 20: No evidence for correlation location decisions and tariff shock

II.i. Correlations between shifts and shares

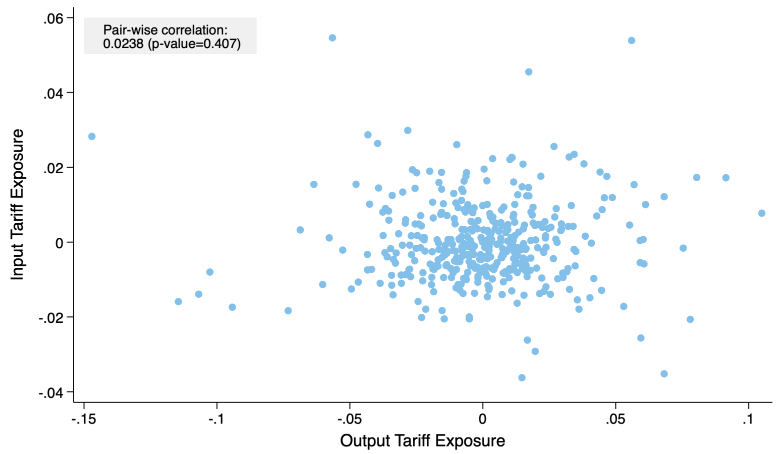
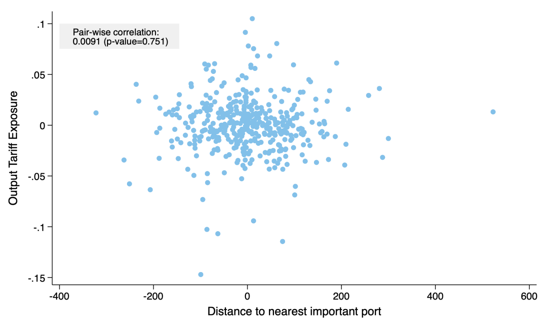
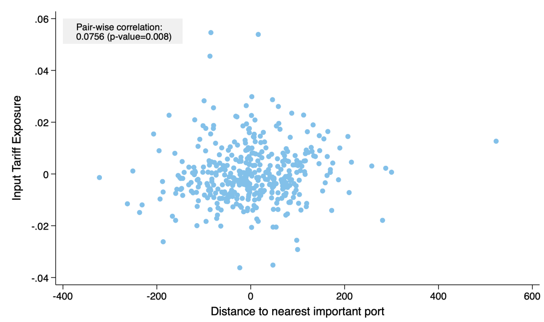


Figure 21: Correlation between OTE and ITE conditional on region FE

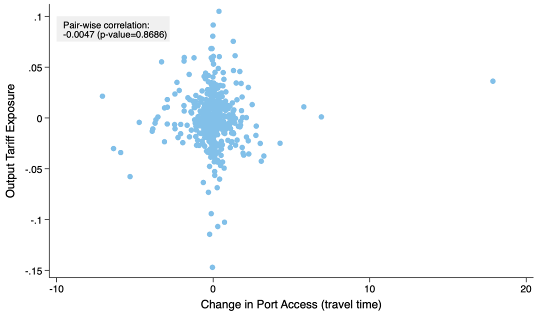


(a) Output Tariff Exposure

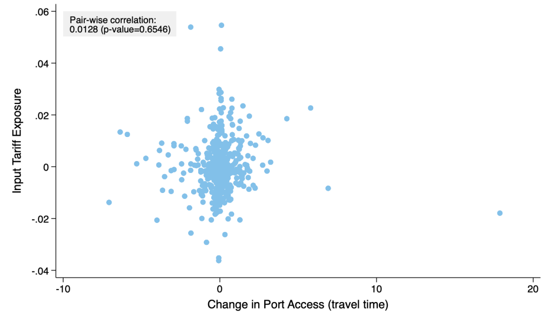


(b) Input Tariff Exposure

Figure 22: Scatterplot of distance to the port and tariff measures conditional on region FE



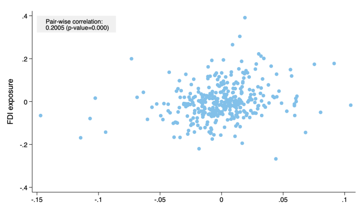
(a) Output Tariff Exposure



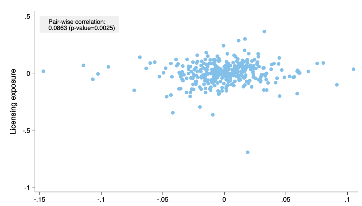
(b) Input Tariff Exposure

Figure 23: Scatterplot of change in travel time and tariff measures conditional on region FE

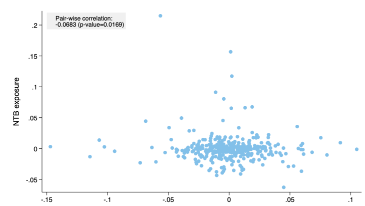
II.ii. Correlations with other policies conditional on region fixed effects



(a) FDI liberalization

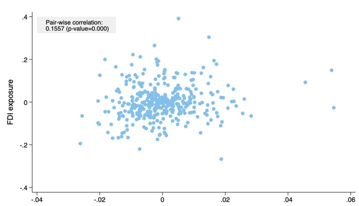


(b) License deregulation

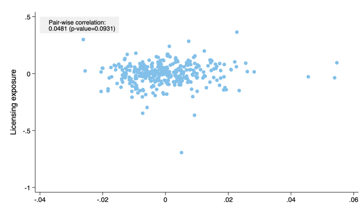


(c) Abolishing import licenses

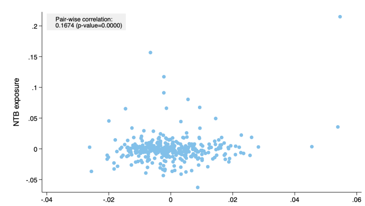
Figure 24: Conditional correlations with Output Tariff Exposure



(a) FDI liberalization

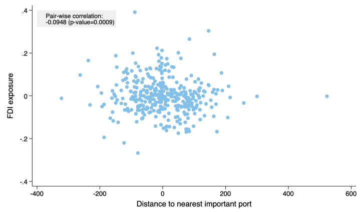


(b) License deregulation

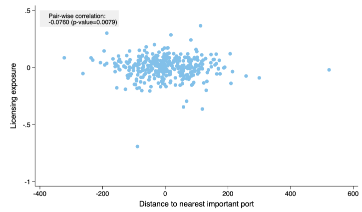


(c) Abolishing import licenses

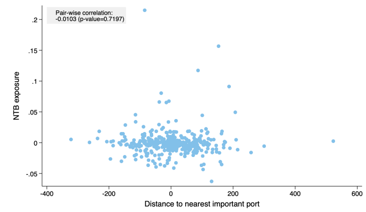
Figure 25: Conditional correlations with Input Tariff Exposure



(a) FDI liberalization



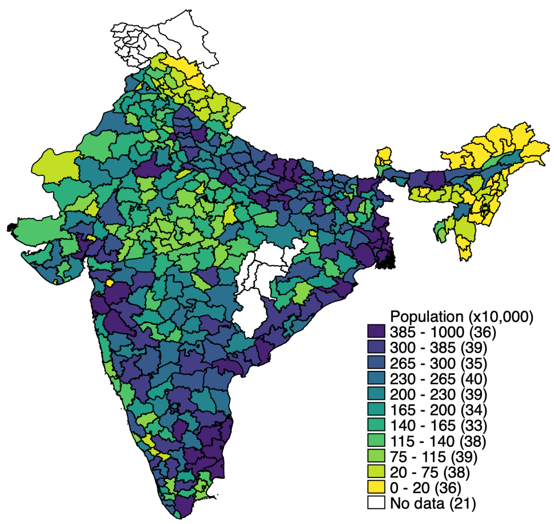
(b) License deregulation



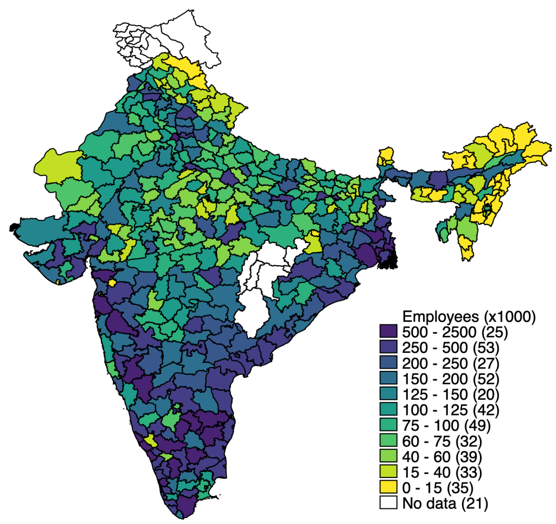
(c) Abolishing import licenses

Figure 26: Conditional correlations with distance to the nearest port

II.iii. Additional maps

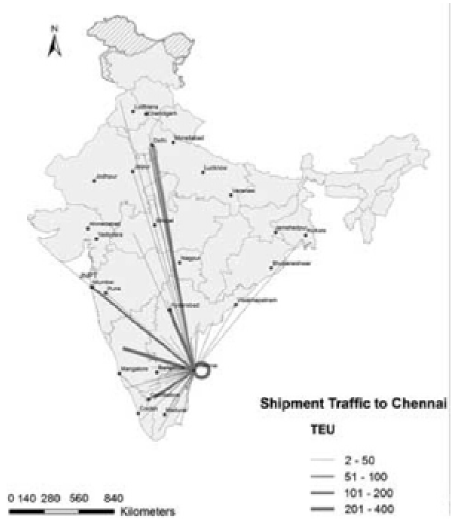


(a) Total population

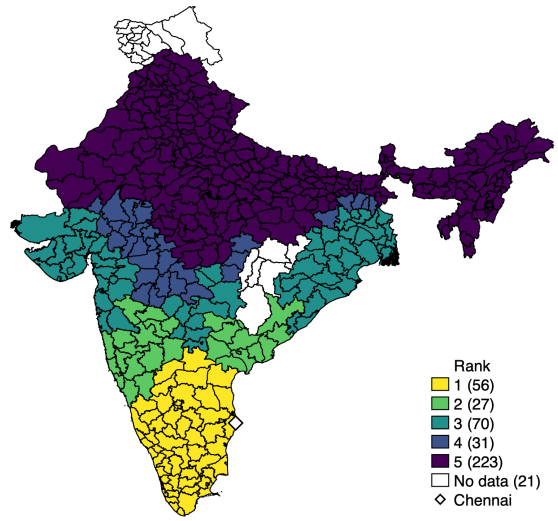


(b) Total non-agricultural employment

Figure 27: Coastal districts

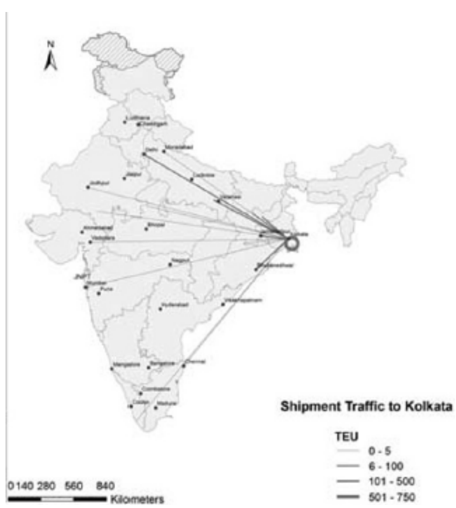


(a) Thill and Venkitasubramanian [2015]

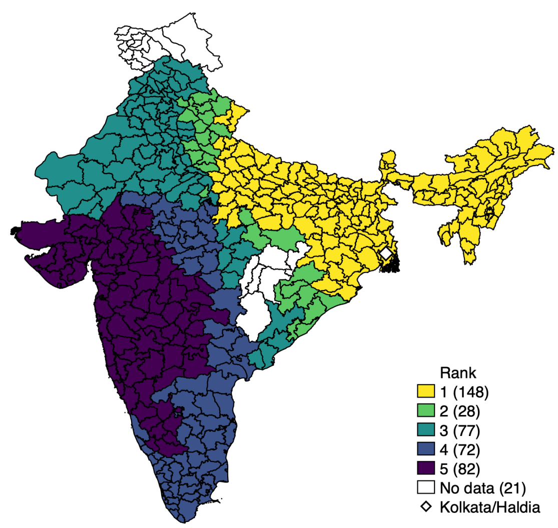


(b) Distance-based rank

Figure 28: Hinterland analysis for Chennai port

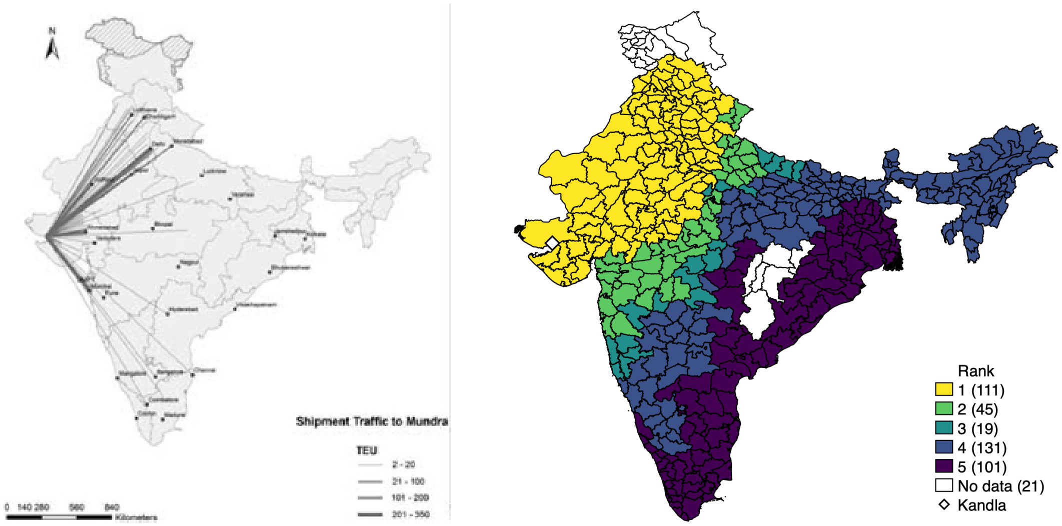


(a) Thill and Venkitasubramanian [2015]



(b) Distance-based rank

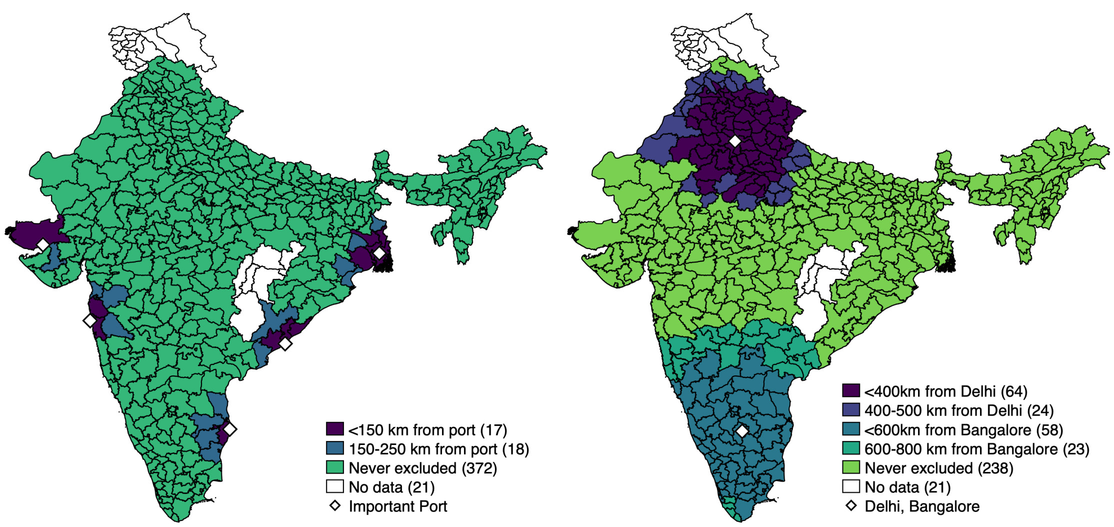
Figure 29: Hinterland analysis for Kolkata port



(a) Thill and Venkitasubramanian [2015] (b) Distance-based rank

Note: The Thill and Venkitasubramanian [2015] figure for Kandla is that for Mundra. This port did not yet operate in 1990 but is located very close to Kandla.

Figure 30: Hinterland analysis for Kandla port



(a) Based on distance from the port (b) Based on distance from Delhi or Bangalore

Figure 31: Excluded districts in robustness check

III. *Additional results*

Table 9: Nonlinear measures of distance - different distance and travel time measures

	Distance measures			Travel time measures	
	(1) Rail	(2) Via ICD	(3) Container ports	(4) Time	(5) Incl. delays
Low × OTE	1.587**	1.909**	1.927**	1.949**	1.488*
	(0.716)	(0.796)	(0.824)	(0.770)	(0.765)
High × OTE	0.749	-0.0767	0.112	-0.162	0.961
	(0.728)	(0.480)	(0.657)	(0.508)	(0.893)
Low × ITE	-8.405***	-7.963**	-7.420**	-8.124**	-6.825**
	(3.137)	(3.146)	(2.880)	(3.152)	(3.212)
High × ITE	-4.283**	-5.257***	-5.695***	-5.243***	-5.743***
	(1.856)	(1.870)	(2.011)	(1.812)	(2.073)
P-values					
$\beta_{OTE}^{Low} = \beta_{OTE}^{High}$	0.328	0.0384	0.0966	0.0402	0.662
$\beta_{ITE}^{Low} = \beta_{ITE}^{High}$	0.138	0.311	0.489	0.294	0.713
Controls	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407
R-squared	0.686	0.689	0.688	0.690	0.686

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Log distance* is the log distance to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). Unless specified otherwise, *Low* and *High* classify each observation as below and above median respectively. *Rail* is the distance by rail to the nearest important port, *Via ICD* is the distance via an ICD to the nearest important port and *Container Port* is the distance to the nearest container port (Chennai, Kolkata, Mumbai/JNPT). *Travel time* is the shortest travel time in hours to an important port; *Incl. delays* adds five hours delay for every state border crossing. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 10: Attenuating effect port access on impacts of trade liberalization not driven by very accessible locations

	Distance to port		Distance from Delhi		Distance from Bangalore	
	(1)	(2)	(3)	(4)	(5)	(6)
	> 150 km	> 250 km	> 400 km	> 500 km	> 600 km	> 800 km
OTE	2.793**	2.557**	2.839**	2.829**	2.725**	3.099**
	(1.203)	(1.144)	(1.096)	(1.105)	(1.227)	(1.459)
Port Access	-0.194*	-0.176*	-0.203**	-0.192*	-0.187*	-0.221*
$\times$ OTE	(0.112)	(0.104)	(0.102)	(0.116)	(0.0995)	(0.114)
ITE	-8.305*	-8.257*	-8.243*	-8.575*	-11.53***	-11.83***
	(4.668)	(4.918)	(4.410)	(4.721)	(3.349)	(3.551)
Port Access	0.171	0.166	0.252	0.386	0.397*	0.431*
$\times$ ITE	(0.331)	(0.346)	(0.303)	(0.339)	(0.228)	(0.245)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	390	372	343	319	349	326
R-squared	0.686	0.677	0.662	0.655	0.744	0.735

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. In the first two columns, the sample excludes districts within 150 or 250 kilometres of a port. In the third and fourth column, districts within 400 and 500 kilometers from Delhi are excluded; the last two columns respectively do not include districts within 600 and 800 kilometers of Bangalore. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Port Access* is the distance in 100 kilometres to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

Table 11: Nonlinear measures of distance – excluding very accessible districts

	Distance to port		Distance from Delhi		Distance from Bangalore	
	(1)	(2)	(3)	(4)	(5)	(6)
	$>$ 150 km	$>$ 250 km	$>$ 400 km	$>$ 500 km	$>$ 600 km	$>$ 800 km
Low $\times$ OTE	1.949**	1.782**	1.975**	1.964***	1.904**	2.010**
	(0.779)	(0.786)	(0.780)	(0.745)	(0.823)	(0.958)
High $\times$ OTE	-0.162	-0.219	-0.254	-0.307	-0.310	-0.288
	(0.491)	(0.498)	(0.498)	(0.653)	(0.519)	(0.519)
Low $\times$ ITE	-8.145**	-7.798**	-7.705**	-7.652**	-9.866***	-9.744***
	(3.242)	(3.402)	(3.202)	(3.367)	(2.806)	(2.975)
High $\times$ ITE	-5.773***	-6.054***	-4.438**	-2.578	-6.349***	-6.200***
	(1.933)	(2.039)	(2.175)	(2.030)	(1.905)	(1.973)
P-values						
$\beta^{Low}_{OTE} = \beta^{High}_{OTE}$	0.0376	0.0527	0.0226	0.0372	0.0319	0.0412
$\beta^{Low}_{ITE} = \beta^{High}_{ITE}$	0.409	0.548	0.262	0.110	0.150	0.160
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	390	372	343	319	349	326
R-squared	0.689	0.681	0.665	0.660	0.751	0.741

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Low* and *High* here mean a distance to the nearest important port below or above 900 kilometers respectively. See the notes under Table 2 for details on the controls, and Table 10 for the sample classifications. *p-value OTE* and *p-value ITE* report the p-value of the F-test comparing the coefficient on *Output Tariff Exposure* and *Input Tariff Exposure* for regions below and above median distance from the port. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.

$^{*} p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.

Table 12: Attenuating effect port access on impacts of trade liberalization not driven by large industries

	Excluding industry					
	(1) None	(2) Farming	(3) Ceramics	(4) Starches	(5) Weaving	(6) Apparel
OTE	2.873**	2.920**	3.484***	2.502**	2.938**	2.677**
	(1.170)	(1.181)	(1.281)	(1.066)	(1.207)	(1.064)
Port Access	-0.208**	-0.198*	-0.222*	-0.164*	-0.208*	-0.205**
× OTE	(0.105)	(0.106)	(0.116)	(0.0930)	(0.107)	(0.0994)
ITE	-8.339**	-8.089**	-9.370**	-7.330*	-8.245*	-6.033
	(4.217)	(4.070)	(3.964)	(4.107)	(4.358)	(3.734)
Port Access	0.194	0.174	0.255	0.0393	0.184	-0.0449
× ITE	(0.282)	(0.273)	(0.268)	(0.290)	(0.296)	(0.233)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407
R-squared	0.688	0.689	0.693	0.694	0.688	0.700

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 13: Nonlinear measures of distance – excluding large industries

	Excluding industry					
	(1) None	(2) Farming	(3) Ceramics	(4) Starches	(5) Weaving	(6) Apparel
Low × OTE	1.926**	2.031**	2.523***	1.832**	2.028**	1.719**
	(0.793)	(0.792)	(0.867)	(0.765)	(0.832)	(0.677)
High × OTE	-0.207	-0.0691	0.143	0.0333	-0.110	-0.149
	(0.495)	(0.493)	(0.609)	(0.415)	(0.532)	(0.518)
Low × ITE	-7.912**	-7.739**	-8.696***	-7.954***	-7.864**	-6.869**
	(3.157)	(3.077)	(2.965)	(3.053)	(3.236)	(2.894)
High × ITE	-5.525***	-5.605***	-5.713***	-6.236***	-5.563***	-6.294***
	(1.854)	(1.849)	(1.932)	(1.798)	(1.892)	(1.538)
p-value $\beta_{OTE}^{Low} = \beta_{OTE}^{High}$	0.0379	0.0422	0.0319	0.0632	0.0409	0.0433
p-value $\beta_{ITE}^{Low} = \beta_{ITE}^{High}$	0.385	0.429	0.259	0.527	0.418	0.818
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407
R-squared	0.690	0.692	0.696	0.694	0.691	0.701

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Low* and *High* here mean a distance to the nearest important port below or above 900 kilometers respectively. See the notes under Table 2 for details on the controls. *p-value OTE* and *p-value ITE* report the p-value of the F-test comparing the coefficient on *Output Tariff Exposure* and *Input Tariff Exposure* for regions below and above said distance from the port. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 14: Heterogeneous effects of trade liberalization across different firm sizes (1998)

	Informal firms		Small firms		Medium firms		Large firms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OTE	1.289*	1.709	1.539**	3.461**	1.001	3.974**	0.0729	-0.622
	(0.669)	(1.074)	(0.688)	(1.602)	(0.725)	(1.985)	(1.842)	(3.636)
Port Access		-0.0537		-0.233*		-0.359*		0.134
× OTE		(0.0893)		(0.134)		(0.197)		(0.298)
ITE	-1.848	-2.086	-7.231***	-10.17***	-10.25***	-15.11***	-9.240	-15.98
	(1.298)	(2.229)	(2.594)	(3.661)	(2.714)	(5.511)	(6.138)	(15.44)
Port Access		0.00917		0.240		0.404		0.754
× ITE		(0.0906)		(0.218)		(0.373)		(1.184)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407	393	393
R-squared	0.562	0.562	0.360	0.367	0.294	0.306	0.510	0.511

The unit of observation is an Indian district; the dependent variable is the district-level change in log nonagricultural employment at informal (<10 workers), small (10-20 workers), medium (20-100 workers) or large firms (>100 workers) respectively, between 1990 and 1998 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Port Access* is the distance in kilometres to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). Note that in 1998, 14 districts do not have any large firms. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

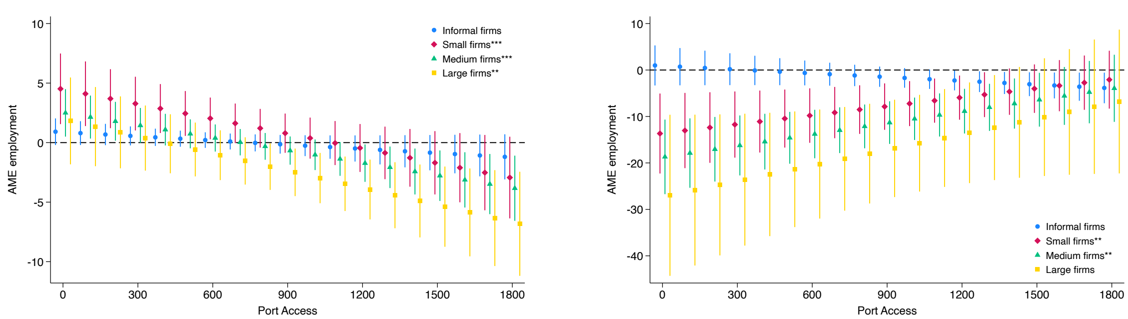
Trade liberalization and the number of firms across the firm size distribution. Table 15 and Figure 32 show that the small-to-medium sized firms not only reduce their labor demand most due to increased import competition: they are also the ones most likely to go out of business altogether, and significantly more so in districts nearer to one of India's main ports. In fact, we only(!) uncover this effect of OTE when allowing for a differential passthrough of the trade shock by conditioning on district's distance to the nearest main port – see columns (4) and (6). A one standard deviation decrease in Output Tariff Exposure decreases the number of small firms in a district located 100 km (700 km) from a nearest major port by 17% (6.7%) and the number of medium-sized firms by 8.8% (0.1%).

Moreover, better access to foreign intermediates only significantly increases the number of formal firms, and the number of larger formal firms in particular. This effect is also (significantly) more pronounced for firms located in districts closer to one of India's five main ports - see columns (4) and (6). A one standard deviation decrease in Input Tariff Exposure increases the number of small firms in districts located 100 km (700 km) from a nearest port by 22% (16%), the number of medium-sized firms by 31% (22%), and the number of large firms by 44% (32%). By contrast, and as expected, the number of informal firms is unresponsive to changes in access to foreign intermediates.

Table 15: Heterogeneous effects of trade liberalization on number of firms by size category

	Informal firms		Small firms		Medium firms		Large firms	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OTE	0.117	0.917	0.974	4.520***	-0.643	2.497**	-2.588**	1.825
	(0.301)	(0.568)	(1.067)	(1.507)	(0.649)	(1.013)	(1.233)	(1.853)
Port Access × OTE		-0.117		-0.414***		-0.352***		-0.480**
		(0.0758)		(0.155)		(0.117)		(0.197)
ITE	-1.335	0.987	-6.363**	-13.67***	-9.888***	-18.70***	-15.75**	-26.97***
	(1.020)	(2.193)	(2.981)	(4.375)	(1.941)	(4.079)	(6.881)	(8.813)
Port Access × ITE		-0.269		0.644**		0.820**		1.122
		(0.178)		(0.317)		(0.345)		(0.715)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	407	407	407	407	407	407	395	395
R-squared	0.896	0.897	0.820	0.825	0.665	0.674	0.590	0.603

The unit of observation is an Indian district; the dependent variable is the district-level change in the log number of nonagricultural firms with 0-10 workers (informal), 10-20 workers (small), 20-100 workers (medium) or >100 workers (large) respectively, between 1990 and 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Port Access* is the distance in kilometres to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). Note that in 2005, 12 districts do not have any large firms. See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$.



(a) Output tariff exposure

(b) Input tariff exposure

Figure 32: Average marginal effects of tariff exposure on number of firms by firm size

Table 16: Heterogeneous effects of trade liberalization on average formal firm size

	1998		2005	
	(1)	(2)	(3)	(4)
OTE	0.468	-0.724	-0.251	-1.554***
	(0.354)	(0.579)	(0.271)	(0.542)
Port Access		0.139***		0.157***
$\times$ OTE		(0.0497)		(0.0522)
ITE	3.611*	6.486**	2.998**	5.332***
	(2.060)	(2.783)	(1.230)	(1.516)
Port Access		-0.258		-0.197
$\times$ ITE		(0.224)		(0.120)
Controls	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	407	407	407	407
R-squared	0.841	0.843	0.848	0.850

The unit of observation is an Indian district; the dependent variable is the district-level change in the log average formal firm size between 1990 and 1998 or 2005 as recorded in the Economic Census. *Output Tariff Exposure* and *Input Tariff Exposure* quantify the district-level exposure to the initial 1991 tariff liberalization. *Port Access* is the distance in kilometres to the nearest important port (Chennai, Kandla, Kolkata, Mumbai/JNPT or Visakhapatnam). See the notes under Table 2 for details on the controls. Conley standard errors, robust to spatial autocorrelation up to 300 kilometres from the district's headquarters, are in parentheses.
$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.