

I've Been Waiting on the Railroad: The Effects of Congestion on Firm Production

John Firth*

MIT

Job Market Paper

14 November 2017

[CLICK FOR LATEST VERSION](#)

Abstract

Transportation networks worldwide suffer from heavy congestion. This paper provides the first estimates of congestion's effect on the production side of the economy, combining firm survey data with traffic data from Indian Railways. Geographic variation in congestion comes from a recent wave of passenger trains which were planned according to certain rigid rules, making it possible to identify the costs the additional traffic imposes on firms using the railways to ship goods. In estimating this "congestion externality", the empirical strategy accounts for both direct and spillover effects of congestion. It also draws on a traffic model from operations research to disentangle a mean effect (congestion makes the average shipment slower) from a variance effect (congestion makes shipping times less predictable). In response especially to the unpredictability, firms simplify operations in several ways, leading to lower productivity and substantial revenue loss. While affected firms suffer, however, I draw on a general equilibrium model of competition to identify gains to their competitors. Policy implications of these results concern both the management of traffic on existing infrastructure, and the construction of new infrastructure.

*jfirth@mit.edu. I thank Daron Acemoglu, Abhijit Banerjee, and Ben Olken for invaluable guidance and support. For helpful comments, I thank Nikhil Agarwal, Josh Angrist, Costas Arkolakis, Sam Asher, David Atkin, Arnaud Costinot, Dave Donaldson, Esther Duflo, Ben Faber, Siddharth George, Nick Hagerty, Peter Hull, Donghee Jo, Gabriel Kreindler, Ernest Liu, Matt Lowe, Ben Marx, Rachael Meager, Yuhei Miyauchi, Scott Nelson, Brendan Price, Narayan Rangaraj, Otis Reid, Natalia Rigol, Shripad Salsingikar, Frank Schilbach, Arvind Subramanian, Marco Tabellini, Rob Townsend, Nick Tsivanidis, and participants at the MIT Development Lunch, and MIT Applied Micro Lunch. For financial support, I thank the NSF Graduate Research Fellowship and the Shultz Fund. Any errors are my own.

1 Introduction

Transportation networks worldwide suffer from heavy congestion. In economics, most existing research on congestion treats it as an urban problem, affecting personal commutes. Yet congestion also affects long distance goods shipments, with firms and policymakers alike claiming that this poses a major barrier to firms' productive efficiency and growth.

To visualize why congestion might affect firm production, consider a manufacturer waiting on its inputs to ship from Mumbai to New Delhi, along one of India's busiest rail corridors. The distance is 870 miles. With a clear railway line, a freight train running at normal speed could make the trip in less than a day. In practice, it can take two weeks. Walking from Mumbai to Delhi would be faster – 11 days by Google Maps estimates. Financing and depreciation costs might accumulate while the goods are in transit, and the slow shipping might limit the manufacturer's ability to adapt to changing conditions. But slow shipping is only part of the potential problem, as congestion also makes shipping unpredictable: the goods might find a clear path on the rails and arrive in a couple of days, before the manufacturer has any use for them, or bad traffic might push the wait to weeks, disrupting the manufacturer's supply chain and slowing productivity. Factors like these provide the basis for firms' complaints about congestion, and the justification for major infrastructure investments such as India's Dedicated Freight Corridor, a proposed network of freight-only railroad tracks aiming to boost firm productivity by freeing freight shipments from congestion.

At the same time, two economic factors suggest congestion might not prove very costly after all. First, firms can try to insulate themselves from congestion, for instance by holding inventories to guard against stockouts. The cost of congestion for an individual firm depends on the availability of these insulating measures, and on the extent to which the measures bring costs of their own. Second, even if congestion hurts some firms, its net effect depends on the ability of these firms' competitors to steal their business and replace the lost revenue. In light of these possibilities, the magnitude of congestion's cost is empirically ambiguous, and so too are the benefits of policies and infrastructure projects aimed at congestion relief.

To settle these empirical questions, I compile a unique dataset linking firm surveys with detailed measures of congestion and shipping times on the Indian Railways. The data reveal staggering amounts of congestion, with more than half of the major lines running beyond the capacity prescribed by international engineering norms. Consistent with operations research models of railway traffic, travel on these congested lines is slow and arrival times vary widely, both across routes and over time for a given route. The main reason for the congestion is that the Indian Railways, a political Ministry answering to voters' demands, often introduces new passenger trains, and does so without regard for the effects on overall traffic flows. In countries such as the United States, where freight operators own the tracks, passenger trains often need to stop and wait for freight trains. But India does the opposite: passenger trains almost always get first priority on the rails. As a result, heavy passenger traffic slows freight shipments, and looking at the addition of passenger trains is an ideal way to study the effects of congestion on firms using railway freight.

I focus, in particular, on a major recent passenger train program, the Duronto trains, exhibiting two features crucial for identification. First, Durontos adhere to a rigid rule of taking the shortest possible path between

endpoints, ruling out endogenous selection of the path. Second, Durontos are supposed to make no stops between their endpoints, ruling out any effects of the trains on the intermediate rail lines other than through congesting these lines and disrupting freight shipments in the area. To avoid confounds from selection or effects unrelated to congestion, I only focus on these intermediate districts, excluding from my analysis the endpoint cities targeted by the Duronto program. Several pieces of evidence indicate that, conditional on being located between two cities considered for the Duronto program, actually having a Duronto pass through a given district is as good as random.

Even with as-good-as-random shocks to congestion, an important identification challenge remains, having to do with spillover effects. Specifically, when one railway line becomes more congested, some of its traffic moves to neighboring lines, increasing congestion there. Thus, the neighboring lines are not a suitable control group. In the language of Rubin (1980), using the neighbors as controls would violate the stable unit treatment value assumption (SUTVA). Finding “pure controls” which satisfy SUTVA is a challenging problem in spatial economics, especially as relates to infrastructure projects, and the literature offers few convincing solutions (Donaldson, 2015; Redding and Turner, 2015). The fundamental dilemma is that control units need to simultaneously (a) satisfy SUTVA, which is more likely if they are far away or different in kind from the treatment units, and (b) serve as a counterfactual for the treatment units, which is more likely when they are nearby or similar.

I overcome this dilemma by using data on rail traffic patterns to identify exactly which districts will receive spillover traffic from the Durontos. For each Duronto route and each pair of stations along the route, I identify all the paths taken by at least one regularly scheduled train traveling between these stations. I refer to this set of paths as the “spillover routes” for the Duronto in question. To show that the districts on these spillover routes are exactly the ones exposed to spillover traffic, I conduct a “zero-th stage” analysis of traffic patterns. It shows that when a Duronto is introduced on a given rail line, this increases traffic both on that line and on the spillover routes I have identified. The spillovers do not extend any farther, however, as there is no traffic increase on the “second order” spillover-routes-of-the-spillover-routes. I therefore know the set of districts exposed to spillovers, and can account for this in studying the effects on these districts’ firms.

Reduced form results show that Duronto traffic leads to increased costs, less efficient production, and ultimately a substantial revenue loss for firms in rail using industries. For each new line of Duronto service passing through a district, local factory revenue falls by 1.9 percent. The preferred specification also includes a control for each district’s exposure to spillover traffic, which serves two purposes. First, it shows that spillover traffic also causes revenue loss, with each spillover route passing through a district leading to a 1.1 percent loss in factory revenue for rail using firms. Second, it serves to remove bias in the estimates of the Duronto main effect, by controlling for an omitted variable. Spillover traffic through a district is negatively correlated with Duronto traffic through that district, because the spillover routes tend to run in parallel to the main Duronto routes. Since spillovers themselves have a negative effect, failing to control from them would bias estimates of the Duronto main effect toward zero. In the end, the local revenue loss associated with Duronto traffic is substantial, and we realize its full magnitude only by accounting for spillovers.

Why do affected firms lose revenue from Duronto-associated congestion? In settings other than Indian Railways, congestion could affect firms through a variety of mechanisms, including increased shipping prices,

reduced availability of freight shipment, or disruptions to passenger travel. In this setting, however, the evidence points toward a single, clear mechanism: the Durgam traffic disrupts freight shipments for rail-using firms, making these shipments slower and less predictable, which raises the effective costs of producing output and delivering it to consumers.

Several institutional details specific to Indian Railways support the claim that this is the mechanism at work. The Indian Railways fixes freight rates as a function of distance, independent of congestion, ruling out costs associated with firms facing higher shipping expenses. It also does not ration freight trains or change the schedules of existing passenger trains as a result of new trains like the Durgams, ruling out effects associated with freight availability or movements of passengers for labor or consumption. Moreover, each of these institutional details finds validation in the empirical evidence, as Durgam traffic has no effect on firms' reported nominal shipping expenses, the number of freight trains run, or the number of passenger trips. So while congestion, in general, works through a bundle of possible channels, this paper's institutional setting provides an opportunity to isolate its effects working through freight shipping times.

Apart from these institutional factors, several aspects of the firm's observed empirical response to Durgam congestion also point toward effects associated with shipping times. First, Durgam traffic leads affected firms to hold larger inventories. The theory of inventory management, in the tradition of [Arrow, Harris and Marschak \(1951\)](#), holds that for a firm trying to guard against stockouts, a key determinant of the optimal inventory level is the time it takes inputs to arrive. Second, Durgam congestion leads firms to alter their product mix, making fewer products per factory, and switching to products which are less time sensitive and which have more predictable demand. These responses are consistent with firms trying to remove uncertainty in the production process, in order to offset the increased uncertainty about shipment times on a congested transport network.

Given that shipping times are the reason Durgam traffic creates problems, I distinguish two specific aspects of this problem: a mean effect (congestion slows average shipping times) and a variance effect (congestion makes shipping times less predictable). To understand the link between congestion and shipping times and, indeed, why congestion is so central to transportation economics, consider a single hypothetical railway line. In principle, arbitrarily many trains can run on the line at arbitrarily high speeds, and with no variance in arrival time, if they are dispatched one after another, running at the same speed and in the same direction. With train speed differentials and different directions of travel, however, trains meet and delay each other. It is because of these potential meetings between heterogeneous trains that congestion becomes a problem: each new train is a possible source of delay for the trains already on the line. Mean travel times increase with congestion due to the higher number of expected train meetings, and variance increases because with more congestion there are more possible meetings, each of which might or might not happen. These effects are, moreover, worse when there is more heterogeneity in train speeds. Averaging 70 to 80 kilometers per hour, the Durgams are among the fastest trains on Indian Railways, while freight trains, typically running at 25 kilometers per hour, are the slowest. It therefore comes as no surprise that the Durgams cause substantial increases in both the mean and the variance of freight shipping times.

In terms of the economic implications for firms, the relative importance of shipping time mean versus variance is, *a priori*, ambiguous. Slow shipping, in the sense of high mean shipping times, might prove unproblematic

if firms simply need to place orders farther in advance. Or it might cause major problems if production and demand are uncertain. For instance, a car manufacturing firm might forecast high demand for red cars and place an order for red paint, only to find that by the time its paint arrives all of its recent orders are for blue cars and it is stuck with the wrong color. Variance of shipping time becomes a problem when, for instance, a firm's input orders arrive later than anticipated, forcing it to stop production because it lacks a key input. On the other hand, variance might matter less if firms can costlessly guard against stockouts with measures such as inventories, or if they can forecast the arrival time of a particular shipment and plan accordingly.

The empirical challenge is to obtain independent variation in the mean and variance of shipping time, which I accomplish by drawing on an operations research model of railway travel times. [Chen and Harker \(1990\)](#) and [Harker and Hong \(1990\)](#) model travel times on a railway line where trains are dispatched according to a given distribution of departure times and train characteristics. I extend their model to show how travel times change with the introduction of additional trains. Both mean and variance increase with additional traffic, but the model's key result is that at higher congestion levels the variance diverges from the mean. Intuitively, this divergence comes from "knock-on effects": on a congested line trains might adhere to schedule on days when none of them are delayed; but when one train gets delayed, this delays other trains which need to wait for it, which is in turn correlated with delays between other pairs of trains, yielding an especially high variance of travel time. I thus instrument for mean and variance using the Duronto shock along with flexible interactions of this shock with pre-Duronto congestion levels.

Two stage least squares results show differing effects of the slowness and the unpredictability, with unpredictability proving the more costly. Consistent with models of inventory management, increases in both mean and variance of shipping time prompt firms to increase their inventory holdings, as a guard against stockout risk. While these measures may provide firms with some insulation, they do not fully buffer against the costs of unpredictable shipping. For each 10 percent increase in the variance of shipping times, average costs increase by 0.3 percent, and in turn, revenue falls by 1.1 percent.

The magnitude of these revenue losses is substantial, raising a puzzle: how can just a few new passenger trains cause such large losses for affected firms? I distinguish two basic explanations. One possibility is that Duronto congestion has a large "cost effect": it substantially raises an affected firm's production costs by making freight shipments slower and less reliable, and thereby disrupting its supply chain. Large cost effects imply that if every firm in the economy suffered an increase in congestion, large losses in output and welfare would follow. Several features of the setting make a large cost effect seem like a plausible explanation. The median district in the sample has only one relatively small railway line, so adding even a single Duronto route through it consumes a large amount of its line capacity. The associated increases in both the mean and variance of freight shipment times forces firms into responses such as higher inventory holdings and more conservative production processes. With these movements away from efficient production, firms ultimately exhibit higher average costs and lower revenue productivity.

The alternate possibility, however, is that Durontos' cost effect is actually small, and firm revenue losses owe more to stiff competition: in competitive markets, firms with even a small cost increase can fall behind their competitors and see revenue plummet. It is possible, moreover, that when these competitors steal the business of congestion-affected firms, this is a relocation of output, but not a large net loss. Distinguishing

between these cost and competition based explanations is essential because if the cost effect is in fact small, then increasing congestion for every firm in the economy could lead to negligible aggregate losses.

In distinguishing these explanations, I first turn to evidence isolating the cost effect, based on the finding that Duronto traffic leads to increases in factories' average cost per unit of production. A simple prediction of firm optimization, for a very general class of production functions, and even in the presence of competition effects, is that observed increases in average cost provide a lower bound on the magnitude of the shift in the firm's cost function. Intuitively, any competition effects reduce a firm's output, pushing it down its cost function, which, with decreasing returns, serves to reduce average cost.¹ So observed effects of congestion on average cost reflect the actual outward shift in the firm's cost function, offset by this downward force from competition. The methodological appeal of this result is that it leverages firm data to isolate basic features of the firm problem which are invariant to any competition effect, making it possible to identify the cost effect without strong dependence on general equilibrium model assumptions.

Given the cost increase for affected firms, the ultimate implications for firm revenue depend on the nature of competition, which is characterized by two main empirical results. First, exposure of a firm's competitors to Duronto traffic leads to increased sales for that firm, indicating ready substitution between the products of the affected and unaffected firms. Second, the negative effect of Durontos on firm revenue is concentrated in industries with high elasticities of substitution, indicating that the degree of substitutability magnifies the consequences of the cost effect. To interpret these results and quantify the aggregate effects of the congestion shock, net of business stealing, I draw on a model of general equilibrium interactions between firms competing in a given industry. Based on [Rotemberg \(2017\)](#), the model predicts how a cost shock will translate into revenue loss for the affected firm, as a function of elasticities of substitution and the exposure of the firm's competitors to similar cost shocks. The associated parameter estimates imply that competitors' gains from business stealing are almost as large as the losses suffered by Duronto-affected firms, adding up to only a minimal aggregate loss from running the Duronto trains.

In applying these results to policy, I first consider the implications for traffic management on existing infrastructure. Currently, Indian Railways maintains a uniform priority for passenger trains over freight traffic, and does little to increase the speed of lagging freight trains. Since variable shipping times are the main source of cost for freight using firms, however, the Railways could greatly help these firms by granting higher priority to a freight train which has already met some delay. One such priority scheme would be a backpressure routing algorithm ([Neely, 2010](#)), which routes traffic on a network to minimize a sum of squared delays, and so reduced the probability of extreme delays and thereby the variance. More concretely, the Indian Railways is experimenting with running freight trains on fixed time tables, in contrast to the current policy of scheduling them on an ad hoc basis, in between the running of passenger trains, and with no promised arrival times. Implementing fixed time tables and more predictable freight shipping is a particular priority for the Dedicated Freight Corridor, and the costs associated with variable shipping times once again indicate that these policies could yield substantial gains for affected firms.

Apart from these implications for traffic management on existing infrastructure, congestion also bears on

¹In the presence of fixed costs, this movement might not reduce average cost, but would reduce average variable cost. The effects I find on average variable cost are similar to those on average cost.

the construction of new infrastructure. In this vein, a second policy application considers the choice between two hypothetical rail construction projects. The first project adds a new rail line between Mumbai and Delhi, a corridor with several lines serving it already, but suffering from heavy congestion. The second hypothetical project is a new line between Amaravati, the newly planned capital of the state of Andhra Pradesh, and Raipur, the capital of neighboring Chhattisgarh. Currently the route connecting these cities is a circuitous 873 kilometers, even though the straight-line distance is only 406 kilometers. Assuming the Amaravati-Raipur route is congestion-free, how much of a reduction in its length would deliver the same benefits as building a new line to decongest the Mumbai-Delhi corridor?

In the logic of [Fogel \(1964\)](#) and of the least-cost path approach common in the contemporary trade literature ([Donaldson, 2017](#); [Donaldson and Hornbeck, 2016](#)), the benefit of the new Mumbai-Delhi line is minimal, because it acts as a substitute for the existing lines. Accounting for congestion, however, two factors point to larger benefits for the Mumbai-Delhi line. First, given the convexity of congestion costs, large benefits result from decongesting an already congested line. Second, decongesting Mumbai-Delhi also relieves congestion on neighboring lines, leading to gains from spillover effects. To model the implications of these factors for welfare, I use a version of the [Allen and Arkolakis \(2016\)](#) framework for characterizing how welfare is affected by infrastructure improvements. Combining this model framework with my empirical estimates, I find that building the new Mumbai-Delhi line would yield the same benefits as shortening the Amaravati-Raipur rail route to the physically impossible distance of 384 kilometers. Decongestion indeed has possible advantages over simply shortening travel distances.

This paper’s analysis of congestion offers an empirical supplement to a recent literature modeling optimal infrastructure investment in the presence of congestion ([Fajgelbaum and Schaal, 2017](#); [Allen and Arkolakis, 2016](#)). These papers model trade costs as a function of quantities shipped along a trade link, departing from the conventional assumption of iceberg trade costs. [Fajgelbaum and Schaal \(2017\)](#) show that incorporating congestion in this manner shuts down complementarities in infrastructure investment, convexifying the optimization problem of a social planner choosing infrastructure investments, goods movements, and economic quantities. This convexification ensures a unique solution and simplifies the procedure for finding it. The assumption behind this modeling device is that additional traffic on the transportation network increases costs for other users, and my results provide empirical support for this assumption.

More broadly, this paper adds to a burgeoning literature on the micro-foundations of trade costs. As surveyed in [Anderson and Van Wincoop \(2004\)](#), both domestic and international trade costs depend on a variety of frictions, from nominal freight prices, to policy barriers, among many others. More recent work uses a combination of theory and micro data to show exactly how these frictions depend on the economics of, for instance, imperfect information ([Allen, 2014](#); [Startz, 2016](#)), the organization of production networks ([Hillberry and Hummels, 2008](#)), and contractual relationships ([Macchiavello and Morjaria, 2015](#)). Closest to my paper is the [Hummels and Schaur \(2013\)](#) study using exporters’ revealed preference for shipments by air versus ocean, in order to estimate the value of time in trade. I go beyond their findings by identifying congestion as an important source of the variation in shipping times, then demonstrating the causal effect on firms, the mechanism of the firm response, and how these effects differ, separately, for changes in the mean and variance of shipping time. My ability to take these extra steps owes to advantageous features

of my setting, in which freight rates and distances are fixed, enabling me to isolate the effects of shipping times. Characterizing trade costs as a function of shipping times offers a useful way to predict the effects of infrastructure projects, since the planning of most projects involves ready engineering estimates of how the project will affect travel times.

Indeed, my emphasis on congestion bears special relevance to modern infrastructure projects. Existing infrastructure papers tend to focus on historical projects (Fogel, 1964; Donaldson, 2017; Banerjee, Duflo and Qian, 2012), or in any case on the establishment of large-scale transport systems (Baum-Snow, 2007; Duranton and Turner, 2012; Faber, 2014), aiming to speak to the old debate about the importance of railroads and national highway systems in countries' development. Today, when a developing country like India spends 3 percent of annual GDP on infrastructure, it typically is not constructing *new* transport systems, but more often widening existing highways, adding links to an already-dense railway network, or otherwise addressing the problem that the *existing* transport systems are inefficient, unreliable, and indeed, congested. These issues require firms to make complex logistical adjustments, making it essential to understand why and for which firms the adjustments are most costly.

At the broadest level, I contribute to the literature on the determinants of low firm productivity in developing countries. Policy debates cite poor infrastructure as a major impediment to productivity (World Bank, 1994; Bajaj, 2010; Ziobro, 2017), and point specifically to inventory holdings as a symptom (Guasch and Kogan, 2003; Datta, 2012; Li and Li, 2013). I provide causal evidence on the mechanisms of this firm response to poor infrastructure, its origins in the unpredictability as well as the slowness of shipping, and its implications for productivity. While some papers consider the effects of uncertainty on productivity (Allcott, Collard-Wexler and O'Connell, 2016) and misallocation (Asker, Collard-Wexler and De Loecker, 2014; David, Hopenhayn and Venkateswaran, 2016), disentangling uncertainty from the adverse events that often accompany it is difficult in practice (with one notable effort being Bloom, 2009), but I provide some of the first causal microeconomic evidence drawing this distinction.

2 Context and data

This section describes the Indian Railways context, and the data used to study the effects of congestion in this setting. The Indian Railways are an important carrier of both passengers and goods, but Railways traffic data shows overwhelming congestion on most of its lines, leading to slow and unreliable freight shipments. A major source of congestion is the indiscriminate adding of new passenger trains, so to study how this hurts freight using firms, I highlight one particular wave of new passenger trains, the Durontos, which were introduced according to certain rigid rules proving useful for identification. A basic contribution of this paper is linking data on these railway traffic patterns, with detailed data on firm outcomes, which I draw from India's Annual Survey of Industries (ASI).

2.1 Indian Railways

The official slogan of the Indian Railways, “Lifeline to the Nation”, speaks to the perceived economic importance of the Railways. India has the world’s third largest railway network by track length, and trails only Japan in passenger volume, handling over 8 billion trips per year. India especially excels at making passenger travel affordable. The average Indian passenger fare amounts to 0.6 US cents per kilometer, compared with 2.4 US cents per kilometer in China, and far higher rates in developed countries: 12.6 cents per kilometer in Germany, for instance, and 19.0 cents per kilometer in Japan.²

The convenience and affordability of passenger travel comes, however, at a cost. Passenger fares are insufficient to cover operating expenses, so the Railways’s financing of passenger travel relies on a cross-subsidy from freight shipments. As a result, Indian freight rates are, in nominal terms, 49 percent *higher* than those in China, and on par with those in developed countries. Adjusted for PPP, Indian freight rates are approximately twice as high as those in both China and the United States ([Ministry of Railways, 2015a](#)). Apart from passengers’ financial burden on freight, passengers also consume the scarce track space shared by the two forms of traffic. Unlike many countries, India does not have separate tracks for passenger and freight trains.

In allocating track space, moreover, India accords highest priority to passenger travel. Passenger trains run on fixed schedules and new trains are frequently introduced by politicians to gain their constituents’ favor. Freight trains, on the other hand, have no fixed schedules, running on an as-needed basis. When a customer wants to make a freight shipment, the customer files a request with the Railways, and a railway manager tries to find a time to dispatch the freight train, in between the scheduled running of the passenger trains ([Ministry of Railways, 2008](#)). As a result, freight shipments often need to wait before beginning their journey, then even once they are en route, stop and wait again for passenger traffic to clear. So freight shipments are slow and unreliable, and the key determinant of freight shipping performance in a given area is the amount of passenger traffic there.

The role of passenger trains in affecting freight shipments motivates this paper’s focus on the introduction of new passenger trains as its source of variation in congestion. I focus on one particular set of passenger trains, the Durgas, introduced by Rail Minister Mamata Banerjee in 2009 ([Banerjee, 2009](#); [Ministry of Railways, 2009](#)). The Durgas trains aim to provide nonstop service on the shortest possible routes between 12 of the largest cities in India. The decisions about where to introduce Durgas trains were based on passenger demand for travel between these major cities. The intermediate districts on the Durgas routes receive congestion as a by-product, and this is the identifying variation I use.

Over time, the heavy passenger congestion on the Railways has pushed freight traffic off of the rails, and on to other modes of transportation such as roads. In 1950, Railways carried 89 percent of India’s freight traffic, measured by weight, but by 2016, this share fell to 31 percent. Of the freight traffic remaining on the Railways, an overwhelming majority, 87 percent, comes from just a few “rail goods”: coal, iron, steel,

²Adjusted for PPP, these countries’ passenger fares, relative to those in India, are 2.7 times higher in China, 6.2 times higher in Germany, and 9.4 times higher in Japan.

fertilizers, cement, mineral oils, and food grains (see Figure 4a). Conversely, these goods rely heavily on the rails, as indicated by the modal shares reported in Figure 4b. In particular, the railways carry 80 percent of India’s coal shipments and more than 50 percent of its iron, steel, and cement (Ministry of Railways, 2011). Producers of these goods have little choice but to ship by rail, since the goods’ bulk makes them difficult, or far more costly, or in some cases unsafe, to transport by road. Given this clear specialization of the rails, I focus my analysis on firms in “rail-using” industries, which I define to be those producing a rail good, or with rail goods comprising at least 5 percent of their input cost share, based on the industry input-output structure in the years prior to the introduction of Durontos.

2.2 Data sources

2.2.1 Railways data

To study geographic patterns in these train movements and the associated congestion effects, I collect data from the Indian Railways, consisting of three parts.

First, Line Capacity data describes the structure of the railway network and the traffic on each part of the network. It comes from annual Line Capacity Reports, prepared by each of the 17 zonal authorities on Indian Railways. These reports provide traffic data based on a division of the railway network into 1218 track sections, where the median section length is 35 kilometers. For each section and each year, the reports indicate (i) how many passenger, freight, and other trains run there in an average 24 hour period, (ii) the types of signaling, electrification, and other physical capabilities present on the section, and (iii) the theoretical capacity of the section. The theoretical capacity is an engineering estimate of the number of trains which can safely run on a section in a 24 hour period. It is based on Scott’s Formula, which accounts for the physical features of the track, the type of equipment present, and a range of other factors. Railway operators regard line capacity numbers as a rough guideline and often run trains in excess of these numbers, perhaps causing some loss in safety or travel time efficiency. Indeed, Figure 1 shows a histogram of the traffic on each section as a fraction of the section’s line capacity, and its most striking feature is that half of the track sections on Indian Railways operate beyond their prescribed capacity.

Second, shipping times data indicates the mean and variance of railway freight shipments. This data is available starting in 2011, the year in which the Railways adopted its current computerized train database. For this project, annual summary data was extracted on freight shipments for all possible origin-destination pairs from a sample of 179 major stations. These 179 stations consist of the 109 most important freight shipment points, and a random sample of 70 additional stations. For each origin-destination pair, the data reports the annual number of freight trains run, and the mean and variance of the running time for these trains. Figure 2 illustrates three key lessons from these data. First, freight shipments are slow in general, with even relatively short shipments often taking five to ten days, and most shipments taking longer than the Railways’s benchmark shipping speed. Second, some routes experience extremely slow shipping, with average times stretching well beyond ten days. Finally, even conditional on track distance, there is considerable cross-route variation in shipping time; the R^2 from regressing shipping time on distance in the cross section is

only 0.19.³ So shipping times and the associated costs depend on factors other than distance, including, as I will show, congestion.

Finally, geographic data on all railway stations and the routes of all passenger trains is scraped from the [IndiaRailInfo](#) website, a sample of which appears in Figure 3. The resulting data details a variety of train characteristics, and lists the stations which each train passes, including stations where the train does not stop. I use this data for two main purposes. First, it identifies the districts and track sections crossed by the Duronto trains and exposed to the associated congestion. Second, since it includes the universe of passenger trains, it provides the set of reasonable routes that a train might follow between two stations. As described in Section 3, this information on traffic patterns helps me identify the areas subject to spillover congestion when the Durontos are introduced.

2.2.2 Firm data

The main source of outcome data is India’s Annual Survey of Industries (ASI), which has been widely used in the economics literature. ASI data comes from an annual government survey of manufacturers, and includes factory-level measures of output, input use, and a variety of other firm characteristics, such as inventory holdings, which can help explain how firms adapt to congested infrastructure. Output data appears separately for each product, making it possible to observe changes in the factory’s product mix. Input data includes detailed measures of capital and labor, as well as materials use, disaggregated by the commodity category of each material. This disaggregation is useful, both for observing how firms alter their product and input mixes in response to congested railways, and for identifying the firms which use heavy inputs typically shipped by rail and which are therefore most likely to be affected by railway congestion.

The ASI data includes all manufacturing establishments above a certain employment threshold which varies by year, and a random sample of smaller establishments. It is provided to researchers in two forms. The first, ASI Panel data, includes factory identifiers, making it possible to link a factory’s data across years and form a panel. While ASI Panel does not include district identifiers, a separate version of the data, ASI Geo, contains all of the same firms and indicators, with the addition of district identifiers, but with the factory identifiers excluded. To construct panel data with geographic identifiers, I use observable characteristics to match the entries in the ASI Panel with those in ASI Geo for 2009-10, the year in which ASI Geo was discontinued. I then use the factory identifiers to link across years, ascertaining the district of factories in the ASI Panel data for 2010-11 through 2012-13. Constructing this geographically identified panel dataset enables me to use district-level geographic variation in congestion while still running regressions with factory fixed effects, and enables me to link the geographically identified firm data with Railways shipping time data which is available only starting in 2011. Table 1 provides additional descriptive statistics from both the ASI data, and the data on Railway traffic patterns.

³There is, likewise, considerable variation across shipments for a given route, as indicated by the annual route-wise data on variance of shipping time.

3 Reduced form effects of congestion on firm revenue

This section identifies reduced form effects of Durgam passenger traffic on rail-using firms. I argue that for certain intermediate districts, having a Durgam run through the district is as good as random. This argument rests both on institutional facts about the Durgam trains, which by rule follow the shortest nonstop path between their endpoints, and on empirical checks of balance and parallel trends. To address SUTVA concerns, the empirical strategy accounts for spillover effects arising because Durgams divert traffic onto neighboring lines. After 3.1 outlines the basic strategy, subsection 3.2 describes the approach to spillovers, and 3.3 presents results showing that Durgam traffic disrupts firm operations, raising production costs and leading to revenue loss.

3.1 Basic empirical strategy

Figure 5a illustrates the empirical strategy, using a comparison between representative “treatment” district Rourkela and representative “control” district Bokaro. Both Rourkela and Bokaro are important steel-producing districts with populations around 500,000. Neither is a major urban center, though, so neither was under consideration to receive Durgam passenger service. Rourkela happens to lie on the shortest rail path between Mumbai and Kolkata, so the Mumbai-Kolkata Durgam passes through. Bokaro lies on a similarly important rail line which is part of the shortest path between Ahmedabad and Kolkata. Other Durgam trains serve Ahmedabad and others serve Kolkata, but the specific Ahmedabad-Kolkata route does not receive Durgam service, so no Durgam passes through Bokaro. The fact that a Durgam passes through Rourkela but not Bokaro is an incidental consequence of the Railways’ intention to provide nonstop service between Mumbai and Kolkata, unrelated to any other differences between Rourkela and Bokaro. This observation supports the empirical strategy’s core identifying assumption, which is that firms in the two districts are comparable via differences-in-differences: in the absence of any Durgams, changes in firm outcomes in Rourkela would have been the same as those in Bokaro.

Some institutional details provide further support for this assumption. First, Durgams make no intermediate stops between their endpoints. This eliminates any possibility that the Durgam routes were chosen to serve or not serve the passengers of places such as Rourkela and Bokaro. A remaining concern is that planners might have chosen Durgams’ paths to avoid congesting Bokaro’s rail lines, for instance because these lines were already too congested or because this congestion would interfere with positive economic trends in Bokaro. A second institutional detail helps allay this concern: the Durgams by rule follow the shortest path between their endpoints. While other trains’ routes might be planned to avoid congesting favored or fast-growing areas, the Durgams’ shortest-path rule ensures that Durgam routes are not chosen based on these characteristics of the intermediate areas. A final remaining concern is endogenous choice of the entire Durgam route, for example because planners favor the firms between Ahmedabad and Kolkata, as a group, over the firms between Mumbai and Kolkata. This possibility is difficult to falsify, but is inconsistent with the motives of the Railways planners, whose explicit goal was to facilitate passenger travel between the target cities, and is also allayed, as I will show, by parallel pre-trends in firm outcomes in the districts with and without Durgam traffic.

The comparison between these districts motivates the basic specification

$$y_{it} = \beta_D D_{dt} + \beta_S S_{dt} + \gamma_i + \gamma_{t \times s} + \gamma_{t \times k} + \epsilon_{it}, \tag{1}$$

where y_{it} is an outcome of interest in year t for factory i , which operates in industry k and is located in district d of state s , D_{dt} is the number of Duronto trains passing through d as of year t .⁴ The sample is limited to intermediate districts which lie between the 12 major urban centers served by the Duronto program. This sample definition excludes two types of districts. First, it excludes the 12 urban centers targeted by the Duronto program, so all of the sample’s variation in Duronto traffic depends on which cities a district happens to lie between, not on any explicit intention to target or avoid that city. The main results are robust to also excluding a “donut” of districts bordering the urban centers. Second, the sample excludes remote districts not lying between any of the 12 major cities. Results are, however, robust to including all districts in India as controls.

While the initial Duronto plan involved nonstop service on the shortest paths between endpoints, later adjustments involved certain Durontos making stops or deviating from the shortest path. These changes pose little threat to identification, since most of them happened after the end of my sample period in 2012, and in any case deviations were minimal. As of 2016, the average Duronto makes 2.4 stops, and travels on a route 2.9 percent longer than the shortest possible route. Still, to avoid concerns that these deviations might have been endogenous, I construct all Duronto treatment variables based on the shortest path between the Duronto’s endpoints. Figure 6 shows district-wise treatment status, measured as the total number of Duronto routes passing through each district as of 2012.

The final ingredient in (1) is S_{dt} , a measure of exposure to spillover traffic. It serves two purposes. First, controlling for spillover effects is essential for identifying the causal effect of Duronto traffic, relative to a counterfactual of no Durontos. Second, estimating the spillover effects is of inherent interest, since measuring the full cost of the Durontos requires accounting for these effects.

3.2 Spillovers from diversion of traffic onto alternate routes

Spillovers arise because when a Duronto train passes through one district, the congestion it creates there diverts traffic to neighboring districts. Figure 5b illustrates this possibility, with Durontos running through Rourkela leading to diversion of traffic and possible congestion effects in Bokaro. Since this spillover traffic flows onto lines other than the main Duronto lines, a reasonable expectation is that spillover traffic through a district is negatively correlated with main Duronto traffic through that district, and that failing to account for the spillovers will therefore lead to downward bias in estimates of the Duronto main effect. In principle, however, the opposite bias is also possible, if the lines with Duronto traffic are geographically concentrated,

⁴Looking at the number of routes actually passing through a district may seem like a simplistic measure of exposure to the added congestion, relative to measures accounting for (a) traffic on the lines a firm uses to ship its goods, or (b) for changes in market access (Redding and Sturm, 2008; Donaldson and Hornbeck, 2016). In Appendix D, however, I consider versions of these alternate measures. Empirically, they do not provide explanatory power beyond the actual passage of Duronto and spillover routes through a given district, so I focus the analysis on the simpler and easily interpretable measures of passage through a district.

so that one Duronto’s spillover traffic flows onto the other Duronto lines, leading to positive correlation between Durontos and spillovers. Which type of bias prevails is therefore an empirical matter.

To account for these spillovers, I use information on the Railways’ typical traffic patterns, drawn from the data on the universe of passenger train routes. For each Duronto route and each pair of stations along the route, I identify all of the paths taken by at least one passenger train traveling between these stations. I refer to this set of routes as the “spillover routes” for the Duronto in question. Figure 5b illustrates this construction with stations A and C lying on the Mumbai-Kolkata line, and certain non-Duronto trains traveling from A to C via Bokaro. So although Bokaro is not directly affected by Duronto traffic, it is on an alternate route for trains traveling between points on the Duronto route, and is therefore subject to a possible spillover effect.

To validate this definition of the spillover routes, I conduct a “zero-th stage” analysis of how Durontos affect traffic patterns. Unlike the main district-level regressions in this paper, the zero-th stage analysis is at the level of track section n . Specifically, it studies how traffic on the section, Traffic_{nt} , measured as the average daily number of trains of a given type, responds to Duronto and alternate-route spillover traffic running on the section:

$$\text{Traffic}_{nt} = \alpha_D D_{nt} + \alpha_S S_{nt} + \gamma_n + \gamma_{t \times \text{Sample}} + \epsilon_{nt}. \quad (2)$$

Here, D_{nt} is the number of Duronto trains running on section s as of year t and S_{nt} is the number of Duronto trains for which s is on the spillover route. If the outcome is total number of trains running on the section, and all other traffic is held fixed on a line when a Duronto is introduced, we would find $\hat{\alpha}_D = 1$ and $\hat{\alpha}_S = 0$. If, on the other hand, one Duronto train leads to displacement of exactly one train onto each section of its alternate route, we would find $\hat{\alpha}_D = 1$ and $\hat{\alpha}_S = 0$.

Table 2 reports results of this regression. Column (1) shows that for each additional Duronto scheduled to run along a line, the total number of passenger trains running on that line increases by 0.61. This increase is less than one, because some traffic is diverted onto the spillover routes. Each spillover route receives 0.22 additional passenger trains, and as Column (2) indicates, 0.23 additional freight trains. Column (3) shows that $\hat{\alpha}_D$ and $\hat{\alpha}_S$ add to one, meaning the total amount of traffic is unchanged.⁵ Columns (4) through (6) show that while the spillover routes as defined above receive traffic as a result of the Durontos, there is no change in traffic on the “second-order” spillover routes of the spillover routes. Thus, I conclude that the possible traffic-diversion spillover effects extend to the routes I have identified, but no farther.⁶

3.3 Reduced form results

This subsection describes the reduced form results, beginning with empirical checks of balance and pre-trends, then proceeding to the main reduced form effects of Duronto traffic, and a body of supplemental evidence which helps clarify the mechanism behind these effects.

⁵For each kilometer of Duronto route, there is 1.02km of alternate route, so the amount of train-kilometers diverted onto alternate routes is approximately the same as the amount of train-kilometers from the introduced Durontos.

⁶The empirical results are robust to using alternate definitions of the spillover routes, for example restricting to alternate routes for trains traveling between the same endpoints as the Durontos, and restricting to spillover routes within a 200 kilometer radius of the Duronto main route.

Throughout the analysis, I focus on four main outcomes: revenue, productivity (TFPR), average cost, and total inventory holding. I consider the natural logarithm of each of these variables, so estimated effects can be interpreted as percent changes. Effects on revenue represent an overall effect of the Duronto congestion, inclusive of any associated increases in production costs or losses in sales to competitors due to poor shipping performance. Effects on TFPR show how Durontos affect productivity: this effect could result from the Duronto congestion disrupting the production processes, though it also includes effects due to changes in the price of the firm’s product.

Studying average cost removes effects of these changes in output price. For single-product factories in the data, measures of average cost come from dividing total costs by the data’s reported quantities. For a multi-product factory i making products $\{1, \dots, K\}$, the average cost measure is

$$AC_i = \frac{\text{Total Cost}}{\sum_{k=1}^K \bar{p}_k q_{ik}}, \quad (3)$$

where q_{ik} is i ’s quantity of k produced, and \bar{p} is the median all-India price of k . Using a fixed product price \bar{p} acts simply to weight across the factory’s product-level output quantities.⁷

Finally, inventories represent the response firms take to insulate themselves from the costs of congestion. Of course, firms take many insulating measures apart from inventories, and Tables A2 and A3 detail some of these responses. But inventories appear repeatedly in the literature as a key response to poor infrastructure (Guasch and Kogan, 2003; Datta, 2012; Li and Li, 2013) and to uncertainty more generally (Fafchamps, Gunning and Oostendorp, 2000). Models of optimal inventory management trace their roots to Edgeworth (1888) and the Newsvendor Problem of Arrow, Harris and Marschak (1951). Appendix E presents a modern version of these models, in which a firm holds inventory to guard against stockout risk arising from uncertain lead times and demand fluctuations. As the model shows, the firm should hold larger inventories in response to increases in either the mean or variance of lead time. The model also predicts larger inventory responses for goods with higher value added, higher penalty of stockout, and higher demand uncertainty. These predictions provide an interpretation for the inventory effects of Duronto congestion, and its associated effects on the mean and variance of shipping time.

3.3.1 Balance and pre-trend checks

In addition to the institutional features supporting the empirical strategy, the data also show evidence of balance and parallel trends. Table 1 shows that intermediate districts receiving more Duronto traffic are similar to those receiving less. Districts set to receive more total Duronto and spillover traffic by 2012

⁷Changes in average cost as measured in (3) could be correlated with changes in product quality or by changes in the relative prices of the firm’s products. This correlation could lead to bias in regressions of average cost on Duronto running, if Durontos affect quality or these relative prices. However, the main results on average cost are robust to alternate measures of \bar{p} , such as using fixed 2008 median prices to remove the effect of changes in relative product price, or 2008 firm-specific prices to account for fixed firm-specific differences in relative product quality. A disadvantage to using 2008 prices, and the reason I avoid it as the preferred definition of \bar{p} , is that the ASI’s product classifications change in 2010 from ASICC to NPCMS, and the ASICC to NPCMS concordance is more exact in some industries than other, leading to differential changes in measured average costs between 2009-10 and 2010-11.

do exhibit slightly lower revenue productivity and reliance on rail goods as inputs, though this difference is economically small and significant at only the 10 percent level. More important from the perspective of the difference-in-difference strategy is that Figure 8 shows parallel trends across Duronto-affected and unaffected districts, both in terms of congestion levels, and in terms of each of the four main outcomes of interest: revenue, productivity, average cost, and inventory holding. This evidence lends empirical support to the identification strategy’s most basic assumption that these districts would have continued to follow parallel trends in the absence of the Duronto program.

3.3.2 Main results

Table 3 presents the main reduced form effects of running Duronto trains. First consider Panel A, showing results from the preferred specification which accounts for the effects of both Durontos and the traffic spillovers. As Column (1) shows, one two-way Duronto route running through a district leads to a 1.9 percent decrease in revenue for the rail-using factories in that district.⁸ This revenue effect is large. For perspective, one Duronto route amounts to approximately 7 percent of the chartered line capacity in the median district. Thus, scaling the revenue effect implies that if a district went from a completely clear railway line, with no passenger traffic, to having Duronto trains use its full line capacity, factories would suffer a revenue loss of $1.9/0.07 = 27$ percent. Of course, this calculation perhaps represents an upper bound on the effect of having a line become completely full, since such a large increase in congestion might prompt firms into larger reorganizations to offset the congestion effect.

Still, the large revenue effect might appear surprising on its surface: why should some passenger trains speeding through a district lead to such losses for firms? Part of the answer appears in Columns (2) through (4), which show effects on rail using factories’ productivity, production costs, and inventory holdings. Each Duronto route reduces TFPR by 1.1 percent, a smaller magnitude than the revenue loss, indicating that input use falls, but by less than the decrease in revenue. Whereas the revenue and TFPR effects both depend on the firm’s output price, the effects on average cost reflect cost per unit of output, independent of this price. As Column (3) shows, each Duronto route increases average cost by 0.8 percent. These cost increases could come from a variety of sources, including financing costs as goods shipments become slower, or risk of input stockout with uncertain arrival times. Inventory stocks, though they bring holding costs of their own, help insulate firms against these costs, and as Column (4) shows, each Duronto route increases firm inventory holding by 1.0 percent. This absolute increase in inventory holdings comes despite a scaling down in firm revenue, entailing an even larger increase in inventory holdings as a fraction of firm revenue.

Panel A also shows evidence of spillover effects. In particular, each alternate route through a district decreases rail-using firms’ revenue by 1.1 percent and increases average cost by 0.7 percent. These effects are smaller in magnitude than the Duronto main effect, though measured with less precision, making them statistically indistinguishable from the main effects.

⁸As noted above, the rail-using firms are those in industries which either produce one of the goods typically shipped by rail (coal, iron, steel, fertilizers, food grains, cement, and mineral oils), or have these goods amount to at least 5 percent of their input cost share, as per the 2007-08 input-output table.

Apart from the economic importance of their effects, the spillovers also play a role in identifying the main effect of Duronto traffic. As Panel B shows, the estimated magnitudes of the Duronto main effects are far smaller than in Panel A, as a result of omitting the spillover controls. The revenue loss, for instance, is only 1.3 percent. While this estimate is not quite statistically distinguishable at the 10 percent level from the Panel A estimate of 1.9 percent, the difference between these point estimates is economically meaningful. The reason for the difference is that spillover traffic through a district is negatively correlated with Duronto main line traffic, and the spillover effects themselves work in the same direction as the main effects. Thus, omitting the spillover control leads to downward bias.

As a placebo test, Table 4 shows no effect on firms in non rail using industries. Theoretically, these firms might have experienced Duronto effects, either due to congestion spillovers as traffic moves from the congested rail lines to roads, or due to general equilibrium effects, for instance if they compete or do transactions with rail-using firms. While these possibilities make the placebo test imperfect, another way to interpret Table 4 is as a falsification of the hypothesis that Duronto-affected districts were, even without the Duronto congestion, set to embark on different economic trends from the unaffected districts. This could happen, as discussed above, if the Duronto running patterns were correlated with planners' broader policy favoritism of certain districts. In such a case, we would expect to find effects even on the non rail using firms in Duronto-affected districts. Yet we see no such effects.

3.3.3 Additional evidence on mechanism

To detail the mechanism behind these reduced form effects, I provide evidence on several additional outcomes (reported in detail in Appendix A) and on heterogeneity (reported in Appendix B). This evidence tells a clear story: the congestion arising from new Duronto passenger trains disrupts freight shipments for rail-using firms, making these shipments slower and less predictable, which raises the effective costs of producing output and delivering it to consumers.

In support of this story, we should find, first of all, that Duronto congestion has its largest effects on firms relying most intensively on railway shipment of freight. The contrast between the effects for rail-using firms and lack of effect for non rail-using firms shows already that the extensive margin of rail use matters for finding congestion effects. The intensive margin also matters, as Table B1 shows larger effects for industries with a greater fraction of their input cost share coming from goods shipped by rail and larger effects for industries actually producing these rail goods.⁹ So the costs associated with railway congestion scale with the firm's reliance on railway shipments, and this congestion seems to create problems both for the receipt of inputs and for the delivery of output.

While firms' use of railway freight at least partly explains the reduced form effect, it remains possible that another part of the effect comes from changes in passenger movement as a result of the Durontos. In particular, one possibility is that when the Durontos ease passenger movements between the major cities,

⁹Among the rail using industries in the sample, producing a rail good is positively correlated with the industry's rail good input cost share.

this has some effect on the intermediate sample districts, either because citizens there can take advantage of the Durontos by traveling via the major cities, or because the advent of Durontos brings economic benefits to the major cities which then spill over to nearby intermediate districts. Ruling out this possibility, Table C4 shows that the main results do not change when we exclude from the sample a 100 kilometer “donut” of districts surrounding the major cities, and Table C1 shows that there is no effect of simply being located close to a major city receiving Duronto service.

Another possibility related to passenger movements is that the Durontos affect local passenger train movements in the intermediate sample districts, either by creating too much congestion for the local trains, or by causing substitution away from other long distance trains which, unlike the Durontos, make stops in the intermediate districts. In practice, the Indian Railways rarely cancels or significantly modifies the schedules of existing passenger trains, making this possibility unlikely. Empirical results in Table A1 show that, indeed, Duronto traffic through an intermediate district affects neither the number of passenger departures from that district’s stations, nor the number of arrivals.¹⁰ So the evidence points, once again, toward the conclusion that the Durontos’ reduced form effects arise not from effects on passenger movements, but from disruptions of freight shipment.

It remains to establish *why* Duronto traffic disrupts freight shipment: is it actually by affecting freight shipping times, as my hypothesis suggests, or via some other channel such as reduced availability of freight trains? Four sets of facts point toward the shipping time story. First, if Duronto congestion led to reduced availability of freight trains, we would expect to see reductions in railway freight shipment volumes. Empirically, however, Columns (3) and (4) of Table A1 shows that the number of freight train arrivals and departures from a district is unaffected by Duronto traffic.¹¹ This finding is consistent with the institutional detail that congestion does not necessarily lead to rationing of freight train slots (Ministry of Railways, 2008): firms can, and empirically do, still present their goods for shipment when the rail lines become congested, but they simply need to wait longer for these shipments.

Second, we might suspect that, apart from any effects on shipping time, congestion raises firms’ freight shipment prices. While congestion could affect freight prices in many settings, for instance because it is associated with shifts in the supply or demand for freight shipments, none of these effects are likely to occur on Indian Railways, where the government fixes freight rates as a function of the good shipped and distance covered. Empirically, Table A2 also shows no effects of Duronto traffic on the amounts rail-using firms report paying in distribution costs.

Third, we might suspect that congestion leads firms to ship their goods by roads instead of rail. Such behavior could lead firms to incur additional costs associated with highway shipment, or provide certain advantages, perhaps without affecting shipment times. Table B2 shows, however, that the Duronto effects are no different in states with greater availability of road shipping options, as measured by the density of

¹⁰These results on the total number of arrivals and departures include trips both within and outside the district, but results similarly show no effect when disaggregated by arrivals and departures within and outside the district.

¹¹The data shows the number of freight train departures, and not the weight or value of goods aboard the trains. So it is possible that congestion-affected firms are in fact sending and receiving less material. But the continued running of the same number of freight trains shows, at least, that congestion does not make these trains unavailable if the firms want to put goods aboard them.

national highways. Indeed, given that most of the goods shipped by rail are heavy materials for which road shipment is unsafe or impractical, it makes sense that firms shipping these goods cannot use road substitution as an insulator against rail congestion, and are instead at the mercy of the shipping times available on the railways.

Finally, the ways in which firms cope with congestion are consistent with an effort to make their production process simpler and more predictable, in the face of slower and less reliable goods shipments. Apart from increasing inventory holdings as discussed already, firms also hedge against uncertainty with adjustments in their product mix. Table A3 reports these adjustments. Most basically, the firms make fewer products per factory, as reported in Column (1).¹² Column (2) shows that they also substitute toward less time sensitive products, where the measure of time sensitivity comes from the revealed preference of internationally trading firms to pay for fast air shipping, as studied in [Hummels and Schaur \(2013\)](#). Column (3) shows substitution toward products with more predictable demand, based on a standard volatility measure from the literature.¹³ Column (4) shows evidence, though significant only at the 10 percent level, that firms substitute toward products with less complex production processes, where the measure of complexity, as in [Levchenko \(2007\)](#), is the (inverse) Herfindahl index of the product’s input cost shares according to US input-output tables. Each of these adjustments potentially streamlines the firm’s production process, amounting to exactly the firm response we should expect if the main problem caused by congestion is to make freight shipments slower and less reliable.

4 The effect of shipping times: mean versus variance

Building on the argument that Duronto effects result from a disruption of freight shipments, this section asks whether the problem is that shipments become slower (mean effect) or that they become less predictable (variance effect). To explain how the Durontos affect shipping times, I draw on an operations research model of railway traffic, and leverage the model’s predictions to instrument separately for the mean and the variance of freight shipping times. Section 4.1 details this empirical strategy. As the results in Section 4.2 show, the Duronto effects owe primarily to the variance of shipping time, which adds uncertainty to the already uncertain world that firms face.

¹²This result indicates that factories exit from certain product markets, though as reported in Table 1, the Duronto and spillover traffic do not have large enough effects to prompt exit at the plant level.

¹³Demand uncertainty, estimated at the product level using pre-2005 ASI data, is measured as standard deviation of $\nu_{i,t}$, the unpredictable part of log changes in (product-level) sales:

$$\Delta \ln PY_{i,t} = \rho_{0,i} + \rho_{1,i} \Delta \ln PY_{i,t-i} + \nu_{i,t}.$$

This method of estimating demand uncertainty follows papers such as [McConnell and Perez-Quiros \(2000\)](#) and [Blanchard and Simon \(2001\)](#).

4.1 Model and empirical strategy

We are interested in estimating an equation of the form

$$y_{it} = \beta_M \ln M_{dt} + \beta_V \ln V_{dt} + \gamma_i + \gamma_{t \times s} + \gamma_{t \times k} + \eta_{it}, \quad (4)$$

where y_{it} is an outcome of interest for factory i , M_{dt} is the mean shipping time for shipments to and from district d , and V_{dt} is the variance.¹⁴ The empirical challenge is to obtain independent empirical variation in this mean and variance which is not correlated with the error term η_{it} .

To separately identify these mean and variance effects, I draw on [Chen and Harker \(1990\)](#) and [Harker and Hong \(1990\)](#), who model two-way traffic on a single rail line, with trains dispatched according to a given distribution. Trains i and j meet with probability q_{ij} , in which case i experiences delay d_{ij} , which is random. The mean and variance of travel time are

$$E(t_i) = FR_i + \sum_j q_{ij} E(d_{ij}) \quad (5)$$

$$Var(t_i) = \sum_j [q_{ij} Var(d_{ij}) + q_{ij}(1 - q_{ij})E^2(d_{ij})] + \sum_h \sum_k Cov(q_{ih}d_{ih}, q_{ik}d_{ik}), \quad (6)$$

where FR is free-running time. Solving for the expectation and variance of each t_i requires numerical methods, and [Appendix F](#) elaborates on this solution, but equations (5) and (6) reveal a key prediction. To first order, the effect of adding more trains on $E(t_i)$ is simply that each new train j imposes some expected delay, $q_{ij}E(d_{ij})$, on train i . For $Var(t_i)$, however, there is both a direct effect of this additional train j , reflected in the first sum in (6), and an additional effect arising from the covariance of the meeting times for all possible pairs of trains on the line. The extra dimension of these pairwise interactions makes the covariance term, and ultimately $Var(t_i)$, scale more rapidly when there are many trains on the line. The implication is that the effect on $Var(t_i)$ of adding an additional train to the line, relative to the effect on $E(t_i)$, is greater for lines which already have high congestion, than for low congestion lines. The intuition for this prediction is based on “knock-on effects”. Even on a congested line, all trains might run on schedule and reach their destinations quickly. But once one train is delayed, it meets other trains and makes them delayed, starting a chain reaction and possibly very slow travel for all trains involved. The variance blows up at high congestion levels because of this difference between everything running on schedule and everything falling in to the chain reaction.

¹⁴To calculate this mean and variance for the empirical application, I restrict focus to those point-to-point shipping routes with at least two shipments in each of the sample years for which data is available (2011 and 2012). This ensures that changes in mean and variance are due to shipping becoming slower and less predictable for a fixed set of routes, rather than due to changes in the composition of shipping origins and destinations. Having at least two shipments is necessary, because otherwise the variance is undefined. For district-level measures of mean and variance, I average across the routes serving each district, weighting by the number of trains run on that route over all the years for which data is available (2011 to 2015). I also normalize each value by the route’s track distance, to avoid giving more weight to longer routes, though doing the analysis without this normalization does not substantially change the results.

This prediction serves as a basis for the first-stage equations

$$\begin{aligned}\ln M_{dt} &= \pi_1^M D_{dt} + \pi_2^M (D_{dt} \times T_{d,t=t_0}) + \pi_3^M S_{dt} + \gamma_d + \gamma_y + \epsilon_{dt}^M \\ \ln V_{dt} &= \pi_1^V D_{dt} + \pi_2^V (D_{dt} \times T_{d,t=t_0}) + \pi_3^V S_{dt} + \gamma_d + \gamma_y + \epsilon_{dt}^V.\end{aligned}\tag{7}$$

Here, D_{dy} is the number of Durontos affecting district d , S_{dy} is the spillover control, and $T_{d,y=y_0}$ is the amount of traffic on the local railway lines in 2008, the year prior to the introduction of Durontos. So the idea of the identification strategy is that when a Duronto hits a given railway line, it has some effect on mean shipping times which is relatively independent of the amount of pre-existing congestion on that line. For variability of shipping times, however, pre-existing congestion and the associated interaction term matters: a Duronto hitting a low-congestion line has some small effect on shipping variance, while a Duronto hitting a high congestion line sets off knock-on effects that entail a much greater increase in the variance. Figure 9 provides an empirical illustration of this mechanism, binning all track sections by their pre-existing congestion levels, and plotting the bin-specific effects of Durontos on mean and variance of shipping time. The divergence it shows between these two curves represents the source of identifying variation.

Even with random variation in the introduction of Durontos and with the controls for spillovers, this identification strategy requires an exclusion restriction: the Duronto and interaction instruments affect firm outcomes only through their effect on mean and variance of freight shipping times. One possible violation of the restriction would occur if Durontos work through channels other than shipping times. As discussed in Section 3.3.3, both institutional and empirical details help to rule out these possibilities. Because Durontos run non-stop and local passenger train schedules are unaffected, Durontos have no effects on local labor movement. Also, because Indian Railways fixes freight rates based on the type of good and distance traveled, congestion arising from the Durontos has no effect on shipping prices or on the quantities that can be shipped; the shippers simply need to wait longer.

Even if Durontos work only through shipping times, another possible violation of the exclusion restriction would occur if Durontos work through some feature of the shipping time distribution other than the mean and variance. For example, congestion might fatten the tails of this distribution, increasing the probability of extremely long and disruptive delays. Because my shipping time data includes only mean and variance statistics, I cannot directly test for this. However, I construct an over-identification test (Hansen, 1982) to help address the concern. Specifically, assume D and $D \times T$ are valid instruments. Testing whether higher-order interactions such as $D^2 \times T$ are correlated with the $\hat{\eta}$ is a way of testing whether the track conditions and congestion affect the outcomes through a channel other than mean and variance of shipping time. These tests do not reject the hypothesis that the extended set of instruments are correlated with the errors, lending support to the claim that they actually are not producing effects through higher-order moments of the shipping time distribution.¹⁵

¹⁵I also employ similar tests using D interacted with the district's pre-Duronto congestion bin, for instance $D \times \mathbf{1}[50\% < C \leq 60\%]$. These tests similarly do not reject the exclusion restriction. In an earlier version of the paper, I estimated linear effects of M and V , instead of the log functional form used here. This linear formulation enables a Wald-style specification test (Godfrey, 1988) whose formulation is similar to the over-identification test but which follows a different logic. Suppose, regardless of whether D and $D \times T$ are valid instruments for M and V , that D , $D \times \mathbf{1}[50 < C \leq 60]$, $D \times \mathbf{1}[60 < C \leq 70]$, and so forth are valid instrument for an extended set of endogenous variables including not only M and V , but polynomial terms M^2 and MV . Estimating this extended model and constructing the Wald test shows that it is actually the mean and variance,

Another threat to identification comes from the use of pre-existing congestion in the interaction term. Areas with higher pre-existing congestion are different from less congested areas, and might be on different time trends. Figure 10 addresses this concern, showing that high-congestion lines receiving Duronto trains are not on a differential trend.

4.2 Results of the shipping times IV

Table 5 presents first-stage effects of the Duronto trains on the mean and variance of freight shipment times. Column (1) and shows that Durontos increase mean shipping times, but as per the small point estimates on the interaction term, this effect is no greater for Durontos hitting high-congestion lines. Column (2) shows that the effect of Durontos on the variance of shipping times increases with the pre-existing congestion in that district. For each additional 10 percent of pre-existing line capacity utilization, a Duronto route through a district leads to 2.1 percent greater variance of shipping time.

Table 6 presents two-stage least squares estimates of the effects of mean and variance of shipping times. Column (1) shows that increasing the variance of shipping times by 10 percent reduces firm revenue by 1.1 percent. Mean shipping times, on the other hand, do not have a statistically significant effect on revenue, and we can easily reject the hypothesis that the mean and variance are equal, in favor of the alternative that the variance effect is greater. Estimates for revenue productivity, reported in Column (2), show negative point estimates which are comparable in magnitude, though only the effect of variance is significant at the 10 percent level. Column (3) shows how shipping times affect production costs, with a 10 percent increase in variance leading to a 0.3 percent increase in average cost, compared with almost no effect coming from the mean. Finally, Column (4) indicates that both mean and variance contribute to increases in inventory holdings, consistent with the predictions of canonical inventory models as described in Appendix E.

5 Explaining the revenue loss: costs versus competition

To determine the economic reason for Durontos' effect on firm revenue, this section models how firm production and competition are affected by the running of Duronto trains. As the reduced form results in Section 3 show, a rail-using firm suffers substantial losses when one of these trains passes through its district. But this revenue loss could occur for two basic reasons. One possibility, which I call a "cost effect", is that the Duronto traffic greatly disrupts firm operations and increases production costs. Large cost effects entail that if every firm in the economy suffered an increase in congestion, large losses in aggregate output would follow. A second possibility, however, is that Durontos' revenue effects owe more to simple market competition: the disruption caused by Durontos is, perhaps, only very small, but because the disrupted firms compete with other firms less exposed to traffic, even a small cost increase can force them out of business. Distinguishing

and not these polynomial terms, producing the effects on firms. Effects of M^2 being the true source of the estimated effects of variance is a particular concern given that the variance scales with the mean or with mean-squared for many random variable distributions. However, that does not seem to be what is happening here.

between these possibilities is essential both because it bears on the net effect of the Duronto program, and because if the cost effect is in fact small, then increasing congestion for every firm in the economy could lead to negligible aggregate losses, while a large cost effect implies large losses from nationwide congestion.

5.1 Model and empirical strategy

The following model serves to isolate the pure cost effect in the presence of competitive forces, and to provide empirical estimating equations. As in Rotemberg (2017), the economy has K sectors, and a consumer with income I has utility

$$U = \sum_{k=1}^K Q_k^\phi + c, \quad (8)$$

where Q_k is sectoral output and c is consumption of an outside good, whose price is normalized to one.

Consumer optimization implies that sectoral revenue is

$$P_k Q_k = \left(\frac{P_k}{\phi} \right)^{\frac{\phi}{\phi-1}}. \quad (9)$$

Sectoral production is a CES aggregate of the output quantities q_{jk} of each firm j in the sector:

$$Q_k = \left(\sum_{j=1}^N a_{jk} S_{jk} q_{jk}^{\frac{\sigma_k-1}{\sigma_k}} \right)^{\frac{\sigma_k}{\sigma_k-1}}, \quad (10)$$

where a_{jk} is quality, and S_{jk} is the share of output going to consumption. Sectoral prices come from profit maximization of the sector's final good producer:

$$P_k = \left(\sum_{j=1}^N p_{jk}^{1-\sigma_k} \right)^{\frac{1}{1-\sigma_k}}. \quad (11)$$

The elasticity of substitution across varieties, σ_k , is an important parameter for determining the effects of exposing some firms in the sector to congestion. Low substitutability σ_k will mean that when some firms are exposed, these firms can raise their prices to offset the associated costs, without losing much business to their competitors. Their ability to retain sales could result from their making specialized products, or from the geography of production, for instance because their customers are local and distant competitors have difficulty reaching these customers. High σ_k , on the other hand, means that if congestion forces affected firms into even a small price increase, customers will switch to the competitors.

Production of each firm's variety is Cobb-Douglas:

$$q_{jk} = A_{jk} K_{jk}^{\alpha_K} L_{jk}^{\alpha_L} R_{jk}^{\alpha_R} N_{jk}^{\alpha_N}, \quad (12)$$

where A is firm-specific TFP, and production uses capital K , labor L , “rail good” materials R , and “non rail good” materials N . In particular, R is a composite of the rail goods specified above (coal, iron, steel, cement, fertilizers, foodgrains, and mineral oils), while N is a composite of all other materials. For ease of notation, index inputs by \mathcal{I} . Returns to scale are reflected by $\gamma \equiv \sum_{\mathcal{I} \in \{K,L,R,N\}} \alpha_{\mathcal{I}}$. For now, assume constant returns to scale ($\gamma = 1$).

As in [Hsieh and Klenow \(2009\)](#), production is subject to firm-specific distortions affecting the marginal product of each input: $\tau_{K,j}$, $\tau_{L,j}$, $\tau_{R,j}$, and $\tau_{N,j}$. The literature proposes many possible sources of these distortions, from credit constraints to political connections; [Hopenhayn \(2014\)](#) provides a useful survey. Taking the pre-existing distortions as given, transport congestion could increase the distortions through a variety of channels. For instance, slow shipping on a congested rail network could force the firm to incur some financing or depreciation costs for each unit of rail input used, or uncertainty in input arrival times could distort another input, such as labor, if workers tasks become less productive or more difficult to coordinate as a result of the uncertainty. So I model congestion as potentially affecting each of the distortions, and will show how firm behavior responds to these changes in distortions.¹⁶

The firm takes the overall price index as given and maximizes profits

$$\pi_{jk} = p_{jk}q_{jk} - \sum_{\mathcal{I}} (1 + \tau_{\mathcal{I},j}) p_{\mathcal{I}} \mathcal{I}_j, \quad (13)$$

implying that it sets price at a constant markup over marginal cost:

$$p_{jk} = \frac{\sigma_k}{\sigma_k - 1} \cdot \prod_{\mathcal{I}} \left(\frac{p_{\mathcal{I}}}{\alpha_{\mathcal{I}}} \right)^{\alpha_{\mathcal{I}}} \cdot \frac{\prod_{\mathcal{I}} (1 + \tau_{\mathcal{I},j})^{\alpha_{\mathcal{I}}}}{A_{jk}}. \quad (14)$$

Firm revenue is

$$y_{jk} = p_{jk}q_{jk} = (p_{jk}^{1-\sigma_k})(P_{jk}^{\sigma_k-1}) \left(\frac{P_k}{\phi} \right)^{\frac{\phi}{\phi-1}}. \quad (15)$$

Allow firm productivity to grow according to

$$\widehat{A}_{jk} = -\epsilon_{jk}, \quad (16)$$

where ϵ_{jk} is mean-zero and normally distributed. Then, combining (15) with (14), (11), and (16), changes in firm revenue are

$$\widehat{y}_{jk} = \underbrace{(1 - \sigma_k) \left(\sum_{\mathcal{I}} \alpha_{\mathcal{I}} (1 + \widehat{\tau}_{\mathcal{I},j}) + \epsilon_{jk} \right)}_{\text{direct effect}} + \underbrace{\left(\sigma_k - \frac{1}{1 - \phi} \right) \sum_{j'=1}^{N_k} \left[\left(\sum_{\mathcal{I}} \alpha_{\mathcal{I}} (1 + \widehat{\tau}_{\mathcal{I},j'}) + \epsilon_{j'k} \right) \frac{y_{j'k}}{Y_k} \right]}_{\text{stealing effect}}. \quad (17)$$

¹⁶Apart from the channels mentioned here, congestion could also yield effects similar to an “output distortion”, for instance if slow shipping makes consumers buy less of the firm’s product at a given price. As [Hsieh and Klenow \(2009\)](#) note, however, the effects of changing this output distortion are equivalent to the effects of changing all of the input distortions equally. I thus omit an explicit output distortion, though I note for purposes of interpretation that the input distortions I study could also reflect these channels related to output distortion.

As I elaborate below, the first term in (17) is the direct effect of exposure to congestion on firm revenue loss, while the second term captures the firms gains from stealing the business of competitors exposed to congestion. Let $\psi_{\mathcal{I}}$ be the effect of one Duronto route on a firm's input \mathcal{I} distortion:

$$(1 + \widehat{\tau_{\mathcal{I},j}}) = \psi_{\mathcal{I}} D_j. \quad (18)$$

We can now write a simplified version of equation (17):

$$\widehat{y}_{jk} = \beta \Psi D_j + \chi \Psi \mu_k + \tilde{\epsilon}_{jk}, \quad (19)$$

where

$$\begin{aligned} \beta &\equiv 1 - \sigma_k \\ \Psi &\equiv \sum_{\mathcal{I}} \alpha_{\mathcal{I}} \psi_{\mathcal{I}} \\ \chi &\equiv \sigma_k - \frac{1}{1 - \phi} \\ \mu_k &\equiv \frac{\sum_{j'} D_j y_{jk}}{Y_k} \\ \tilde{\epsilon}_{jk} &\equiv (1 - \sigma_k) \epsilon_{jk} + \left(\sigma_k - \frac{1}{1 - \phi} \right) \sum_{j'=1}^{N_k} \epsilon_{j'k} \frac{y_{j'k}}{Y_k}. \end{aligned}$$

Here, β reflects the direct effect of increasing distortions. This effect is largest if the elasticity of substitution σ_k is high, since this means that firm varieties in the sector are close substitutes, so even a small distortion to one firm's costs will cause it to lose a large amount of business to its competitors. While β captures the effect of the distortions themselves, Ψ captures how these distortions respond to Duronto routes D_j through the firm's district. In particular, Ψ is a sum of the Durontos' effect, $\psi_{\mathcal{I}}$, on each input distortion, weighted by each input's cost share $\alpha_{\mathcal{I}}$.

Just as congestion can lead to revenue losses for a given firm, it also presents an opportunity for the firm to steal the business of its competitors who experience congestion of their own. The magnitude of this stealing depends on crowd-out parameter χ . It is largest when the firm's product is a ready substitute for its competitors' products (high σ_k), and when the sector as a whole is less replaceable by other sectors (low $\frac{1}{1-\phi}$). The measure of sectoral exposure, μ_k is an output-weighted average of exposure to Duronto congestion for all the firms in the sector. Finally, the disturbance ϵ'_{jk} is normally distributed with mean zero.

An additional prediction comes from the observation that the revenue effect increases with the elasticity of substitution. In particular, re-write (19) as

$$\widehat{y}_{jk} = \Psi_1 D_j - \Psi_2 (\sigma_k \times D_j) + \chi \Psi \mu_k + \epsilon'_{jk}. \quad (20)$$

Here, Ψ_1 reflects the cost effect for firms in low σ industries, while Ψ_2 reflects that revenue losses become greater for firms in more competitive industries. Below, I use industry level estimates of σ to estimate (20).

Note that if the model is correct and σ is measured perfectly, we should find $\Psi_1 = \Psi_2 = \Psi$.

The aggregate effect on sectoral output comes from summing across all firms in (17), yielding

$$\hat{Y}_k = (\beta + \chi)\Psi\mu_k + \epsilon_k, \quad (21)$$

where the disturbance $\epsilon_k \equiv \sum_{j=1}^{N_k} \tilde{\epsilon}_{jk} \frac{y_{jk}}{Y_k}$ is a weighted average of the firm-level disturbances. Equation (21) nicely breaks the effect of the Duronto congestion shock into three parts. First, firms in the sector face some exposure to the congestion, as measured by μ_k . Second, this disrupts firm operations, leading to some total distortion Ψ , which reflects pure ‘‘cost effect’’ of the Durontos, independent of any output market competition. Finally, $\beta + \chi$ reflects how the previous two components, working through market competition, lead to an ultimate effect on sectoral revenue. The aggregate effect depends on whether the direct losses to firms, reflected by β , are large relative to the ability of other firms in the sector to replace the lost output, as reflected by χ .¹⁷

To isolate the pure cost effect Ψ , note that Cobb-Douglas production (12) with constant returns to scale entails that the firm’s average cost equals marginal cost:

$$AC_j = MC_j = \prod_{\mathcal{I}} \left(\frac{p_{\mathcal{I}}}{\alpha_{\mathcal{I}}} \right)^{\alpha_{\mathcal{I}}} \cdot \frac{\prod_{\mathcal{I}} (1 + \tau_{\mathcal{I},j})^{\alpha_{\mathcal{I}}}}{A_{jk}}. \quad (22)$$

It follows that

$$\widehat{AC}_j = \Psi D_j + \widehat{\epsilon}_{AC,j}, \quad (23)$$

with

$$\epsilon_{AC} \equiv \prod_{\mathcal{I}} \left(\frac{p_{\mathcal{I}}}{\alpha_{\mathcal{I}}} \right)^{\alpha_{\mathcal{I}}} \cdot \frac{1}{A_{jk}}. \quad (24)$$

Under the assumption that Duronto running is uncorrelated with changes in factor prices $p_{\mathcal{I}}$ and the physical TFP A_{jk} in the firm’s production function, regressions of the form (23) identify the effect of Durontos on costs. In other words, observed changes in average cost reflect an actual cost effect, independent of any competition effect.

A more general version of this statement holds for generic production functions with non-increasing returns. Let $C(q)$ be the cost function, which is unknown, but assumed to satisfy $C'(q) > 0$ and $C''(q) \geq 0$. Also assume there are no fixed costs ($\lim_{q \rightarrow 0^+} C(q) = C(0) = 0$).¹⁸ Suppose costs increase by shifting outward, so the new cost function is $\tilde{C}(q) = (1 + \tilde{\tau})C(q)$. How can we identify $(1 + \tilde{\tau})$?

¹⁷It is straightforward to extend the above discussion to account for the effects of spillover traffic. Letting S_j be the amount of spillover traffic in the district of firm j , Ψ^S the total cost effect of spillover routes, and μ^S the exposure of other firms in the sector, we obtain analogues of equations (19) and (21):

$$\begin{aligned} \hat{y}_{jk} &= \beta(\Psi D_j + \Psi^S S_j) + \chi(\Psi \mu_k + \Psi^S \mu_k^S) + \tilde{\epsilon}_{jk} \\ \hat{Y}_k &= (\beta + \chi)\Psi \mu_k + (\beta + \chi)\Psi^S \mu_k^S + \epsilon_k. \end{aligned}$$

¹⁸Even for a production with fixed costs, a version of (27) holds, with average variable cost, rather than average cost, as the object of interest.

Equating marginal revenue with marginal cost, firm optimization entails

$$\sigma + 1 = (1 + \tilde{\tau})C'(q). \quad (25)$$

Differentiating with respect to $\tilde{\tau}$, we see that

$$\frac{\partial q}{\partial \tilde{\tau}} = -\frac{C'(q)}{C''(q)} \frac{1}{1 + \tilde{\tau}} < 0. \quad (26)$$

Finally, noting that $AC_j = (1 + \tilde{\tau})C(q)$ and considering the effect of changing $\tilde{\tau}$, it follows that

$$\widehat{AC}_j = \widehat{(1 + \tilde{\tau})} + \frac{\partial[C(q)/q]}{\partial q} \frac{\partial q}{\partial \tilde{\tau}} \widehat{(1 + \tilde{\tau})}. \quad (27)$$

So the effect of the cost shift $\tilde{\tau}$ on average costs is, first, a direct increase in costs, $\widehat{(1 + \tilde{\tau})}$. But as the second term reflects, the cost shift also pushes the firm down its cost function ($\frac{\partial q}{\partial \tilde{\tau}} < 0$), which with non-increasing returns has the effect of reducing average costs ($\frac{\partial[C(q)/q]}{\partial q} > 0$). Thus, since the second term in (27) is negative, observed changes in average cost \widehat{AC}_j are a lower bound on the cost shift $\widehat{(1 + \tilde{\tau})}$.

5.2 Empirical application of the model

To identify the effect of competitors' exposure to Duronto congestion, I estimate an empirical counterpart of (19):

$$y_{it} = a_1 D_{dt} + a_2 S_{dt} + a_3 \mu_{sk} + a_4 \mu_{sk}^S + \gamma_i + \gamma_{t \times s} + \gamma_{t \times k} + \epsilon_{it}, \quad (28)$$

where μ_{sk} and μ_{sk}^S are the exposure to Duronto and spillover traffic, respectively, of factories in the same state s and four-digit NIC industry k as factory i . All exposure measures are calculated based on the Duronto routes in service as of year t , but the 2008 district locations of each industry's output. As above, all regressions include fixed effects for each firm, and year-specific effects for each state and industry.

Table 7 presents results of this regression. Column (1) shows, first, that the main effect of a Duronto route is a 3.1 percent loss in revenue for rail-using factories. This is greater than the revenue loss estimated in the basic reduced form regression of Table 3, because Duronto traffic is positively correlated with the exposure of competitors to Duronto traffic, and this exposure μ_{sk}^D itself has a positive effect on a firm's own revenue. In particular, if each of a firm's competitors is exposed to an additional Duronto route, that firm gains 2.5 percent in revenue.

In the context of the model, the sum of these revenue coefficients, $\hat{a}_1 + \hat{a}_3$, provides an estimate of $(\beta + \chi)\Psi$, which indicates the aggregate effect of Duronto exposure on firm revenue. I cannot statistically reject the hypothesis that the sum of these coefficients is greater than or equal to zero, against the alternative that it is negative; the p-value on this test is 0.29. So it is not possible to rule out that the competitors replace all, or at least a large portion, of the output lost by congestion-affected firms. Estimates of the spillover and state-industry spillover exposure effects offer less precision, but yield a similar qualitative conclusion.

Columns (2), (3), and (4) of Table 7 show that competitors’ exposure to congestion does not affect a firm’s revenue productivity, average cost, or inventory holding. These results are unsurprising: while competitors’ exposure enables a firm to steal the business of these competitors, it does not affect the firm’s own logistical operations or production costs. In principle, competitors’ exposure to congestion might have affected revenue productivity through price effects, though revenue productivity depends not only on prices but on physical productivity, which is likely to remain unaffected. The main effects on these three variables remain the same as in the reduced form, however, with each Duronto route still leading to a 0.8 percent increase in average costs. As per equation (23), this effect on average costs is interpretable as an estimate of the pure cost effect Ψ under Cobb-Douglas production, and more generally as a lower bound on the shift in the cost function as illustrated in (27). So Duronto congestion does lead to some disruption of firm production and pure cost effect which, though magnified by competition, is nontrivial on its own.

Table 8 shows support for the additional prediction of equation (20) that revenue effects scale with the elasticity of substitution. In an industry with inelastic demand, Duronto congestion causes little revenue loss: the 10th percentile elasticity is $\sigma = 2.9$, implying the Duronto effect on revenue is a 2.0 percent loss. Intuitively, the low elasticity means that when congestion increases costs for these firms, consumers still buy their products. For high elasticity industries, on the other hand, the congestion effect leads customers to substitute to other sellers, and affected firms suffer a larger revenue loss: the 90th percentile industry has $\sigma = 6.2$, implying a 4.2 percent revenue loss. The estimated coefficient on the Duronto main effect $\hat{\Psi}_1 = -0.0014$ and that on the elasticity interaction $\hat{\Psi}_2 = -0.0065$ do not explicitly validate the theoretical prediction that $\hat{\Psi}_1 = -\hat{\Psi}_2$, though the confidence interval on $\hat{\Psi}_1$ is wide enough that we also cannot reject this prediction. One likely reason for the difference between $\hat{\Psi}_1$ and $-\hat{\Psi}_2$ is measurement error in the elasticities σ . The economically relevant elasticity concerns substitution between a firm’s variety and the varieties of other firms in the state-industry, but the elasticities in the data reflect substitution patterns between Harmonized Standard 6-digit products.¹⁹ Still, this type of measurement error would attenuate estimates of $\hat{\Psi}_2$. So we should expect the true magnitude of Ψ_2 to be larger than estimated, and the basic conclusion still holds: Duronto congestion effects are worst for firms facing stiff competition.

While the estimates so far use firm-level data to estimate the parameters that matter for aggregate revenue effects, a direct test for aggregate effects is also possible, using an empirical counterpart of (21):

$$Y_{skt} = b_1\mu_{skt} + b_2\mu_{skt}^S + \gamma_{t \times s} + \gamma_{t \times k} + \epsilon_{skt}, \quad (29)$$

where Y_{skt} is aggregate output for industry k firms in state s in year t and the exposure measures are calculated as above. The results in Table 9 show negative but statistically insignificant effects of exposure to Duronto and spillover traffic. The magnitudes of these estimates nevertheless fall within the same range as the implied aggregate effects from the firm level regression. In particular, the implied value of $(\beta + \chi)\Psi$ from Table 7 is -0.006 , while the Duronto exposure effects in Table 9, which estimate the same parameter,

¹⁹The product categories with the lowest elasticities are those with specialized products: uranium and thorium ores, manufacture of cement plaster, manufacture of electricity distribution, and control apparatus, and manufacture of electric motors. Those with higher elasticities include more substitutable products: iron ores, soft drinks, alcohol, and animal feeds. So although the elasticities from Broda, Greenfield and Weinstein (2006) do not measure the relevant cross-firm-variety elasticity, they reflect the interchangeability of products within categories, and in this sense proxy well for the relevant elasticities.

range between -0.002 and -0.014 .²⁰

Taken together, the empirical results in this section show that the reduced form revenue loss owes, in large part, to firms losing their edge against competitors, who in turn take advantage of the opportunity and mitigate aggregate revenue loss. Still, congestion affected firms do experience a genuine disruption to their operations and an increase in production cost, which would imply some losses in aggregate productivity if all firms in an economy experienced a congestion increase.

6 Policy

Congestion bears on infrastructure policy for two distinct reasons. First, it has implications for traffic management on existing infrastructure. Only with notions of capacity and congestion can we conceptualize the economic benefits from congestion pricing and prioritization of different types of traffic. Second, decisions about how and where to construct new infrastructure need to account for congestion. Doing so overturns some commonly held intuitions about the form of optimal investment.

6.1 Traffic management on existing infrastructure

6.1.1 Congestion externality from running additional traffic

The first traffic management issue is how to account for congestion in setting prices or restricting quantities. Currently, Indian Railways does not increase prices with congestion. In calculating how to set congestion pricing, an essential input is a measure of the cost externality the running of one train imposes on other users of the rail network. My estimated Duronto effects provide a measure of this externality. Of course, running one Duronto train may impose externalities on the passengers in other trains, in addition to the effects on freight-using firms. My estimates capture the effects on freight alone, and in this sense are a lower bound on the total externality.

A naive way to measure firm losses from introducing one Duronto route is to look at the revenue loss for Duronto-affected firms, relative to firms in districts unaffected by Duronto traffic. To calculate this loss, I sum the 2008 revenue of all rail-using firms in the path of each Duronto train, and multiply by the estimated revenue loss coefficient from Table 7. As reported in Column (1) of Table 11, the introduction of the average Duronto route leads to a firm revenue loss of INR 461 million in the districts it passes through, plus an additional INR 155 million in districts subject to spillover traffic, for a total loss of INR 616 million (USD 12.7 million at 2008 exchange rates). For comparison, this loss amounts to 60 percent of the estimated INR 1,024 million annual passenger revenue from running one Duronto route.²¹ Railway passenger services

²⁰Even if Duronto traffic did not produce aggregate effects on revenue, it might affect gross value added by relocating business to less productive factories which would not have produced as much output in the absence of the congestion increase. Table 10 tests for these effects on gross value added, again finding coefficients which are negative but statistically insignificant.

²¹I do not have detailed data on fare revenues, but derive estimates by using the limited number of per-journey revenue amounts reported in [Ministry of Railways \(2015b\)](#), and multiplying by the annual number of journeys for each route.

already operate at a loss, with operating costs twice as high as the fare revenue collected ([Ministry of Railways, 2015a](#)), and this externality adds an additional cost on top.

At the same time, consistent with a central theme of this paper, the negative externality for certain firms leads to a positive externality for the firms which steal their business. As Column (2) of Table 11 reports, competitors in the same state and industry as Duronto and spillover affected firms gain a total of INR 567 million for each Duronto route introduced. Thus, the net firm revenue loss as reported in Column (3) is INR 49 million, or only about 5 percent of the route’s passenger fare revenue.

While the thought experiment so far considers the effects of running Duronto trains through some districts but not others, it leaves open an important economic question: what would be the effects of a nationwide increase in congestion? Apart from the economic interest in answering this question, it is also relevant to real policies the Railways might consider, such as uniform limits on the amount of passenger traffic congesting a given line, uniform increases in the track priority of freight relative to passenger traffic, or the construction of the proposed nationwide network of Dedicated Freight Corridors, aiming to improve freight performance for all firms.

The effects of such a nationwide change in congestion depend on the extent to which congestion disrupts firm production, as reflected in the “cost effect” discussed in Section 5. The model there shows that if we assume Duronto congestion increases production costs by some proportion $1 + \tilde{\tau}$, then estimated effects on average cost provide a lower bound on $\tilde{\tau}$. Under perfect competition, multiplying each firm’s cost function by $1 + \tilde{\tau}$, equivalent to multiplying aggregate supply by $1 + \tilde{\tau}$, will lead to a $100 \cdot \tilde{\tau}$ percent reduction in output, and a $100 \cdot \tilde{\tau}$ percent reduction in total surplus. As Column (4) of Table 11 reports, exposing all rail-using firms to this cost shock would lead to an output loss amounting to INR 94,962 million (USD 2.0 billion). Whereas this represents the effect of exposing every rail using firm to Duronto traffic, Column (5) reports the effect of exposing every manufacturing firm, rail-using or not, to a similar cost shock, resulting perhaps from a Duronto-sized congestion increase on its preferred mode of transportation, whether that be rails, roads, or otherwise. This effect amounts to INR 258,551 (USD 5.3 billion).

Of course, this extrapolation to non rail using firms assumes that these firms are as sensitive to congestion as the rail using firms studied in my empirical analysis. While it is possible that these non rail using industries are less sensitive to congestion, two factors suggest that, in fact, they could be more sensitive. First, in terms of selection, the industries choosing to remain on the rails despite the high congestion are likely industries for which this congestion is less of a problem. Second, the goods that rail-using firms ship on the railways are typically homogeneous commodities like coal, iron, and cement. Whereas these firms might succeed in buffering themselves against congestion by holding large inventories of the homogeneous commodities, we might expect worse effects of congestion for other firms shipping more specialized inputs that need to arrive quickly and predictably. So in both of these regards, my estimates of the rail-specific congestion effects are perhaps lower bounds on the effects of congestion for the productive economy as a whole.

6.1.2 Priority of traffic

A second traffic management issue is how to prioritize different types of traffic. Daily operations on Indian Railways are handled by managers who decide which trains are allowed to run first on an open track, and how to accelerate or decelerate trains so they arrive at certain times. Currently, these managers' protocol is to give the highest priority to passenger trains, making them adhere as well as possible to their schedule. An alternative would be to increase the priority for freight trains, either running the freight trains on fixed schedules, or granting higher priority to a freight train once it has met a certain amount of delay. The latter notion is the idea behind back-pressure routing (Neely, 2010), which is an approach to maximizing throughput based on minimizing a sum of squares of units' backlogs. By using backpressure routing or another prioritization objective function which helps lagging traffic catch up, railways managers could reduce the variance of travel times. Whether this strategy yields economic benefits depends on whether the variance of travel times leads to economic costs, and my estimates indicate that it does.

6.2 New infrastructure

Congestion also factors into planners' decisions about how and where to build new infrastructure. India, with \$12 billion in financing from the World Bank, is now in the process of constructing Dedicated Freight Corridors, which will be a set of higher speed railway tracks exclusively for freight shipment. Policymakers see congestion relief as a chief goal of these projects, and argue that this relief will provide great help to manufacturing growth (Ministry of Finance, 2015). One corridor is under construction along the west coast, between Mumbai and Delhi, with another in progress running from Punjab to West Bengal. Several other branches in other parts of the country are under consideration. But which of these lines to actually build remains an open question.

To see the implications of congestion in answering this question, consider a choice between two hypothetical rail construction projects. The first project, like the actual Dedicated Freight Corridor under construction, adds a new rail line between Mumbai and Delhi, a corridor with several lines serving it already, but suffering from heavy congestion. Figure 11a depicts a stylized version of this project. In a least-cost path approach to specifying trade costs, as is typical in the empirical literature on infrastructure (Donaldson, 2017; Donaldson and Hornbeck, 2016), the cost of moving from M to D, τ_{MD} , is a function of the length of the shortest path between M and D.²² If the new line and the existing shortest path between M and D are of similar length and quality, and we have no notion of capacity or congestion, then adding the new line will not reduce the trade cost τ_{MD} . The new line is simply a close substitute for the existing line.

The intuition that comparable links in a transportation network serve as substitutes for one another has a long and influential intellectual tradition, going back to Fogel (1964). Fogel's main insight was that, although the American railways carried large volumes of freight shipments, the railroads in fact made a

²²The approach does, often, show sophistication in allowing for different prices per distance traveled on different parts of the network or for different modes, or might incorporate other aspects of trade costs. But the core idea of a least-cost path approach is to find the minimum travel distance between points. Applying such a least cost path approach to the Indian Railways would almost certainly specify costs as a function of distance, since, on the Railways, the distance determines the price.

small contribution to economic growth, because even in the absence of railroads, shippers would have been able to use a close substitute: the waterways. The intuition of substitutability between two different lines or modes of transport is, perhaps, correct in a context like the American railroads, if there is little congestion relative to the level of capacity.

In a congested network, however, this intuition breaks down. First, the new line between M and D shares the traffic load with the existing line, reducing congestion and the associated trade cost between M and D. Second, due to traffic spillovers, the new line will reduce trade costs for trips to and from the neighboring city X. In particular, if there are some traders who previously traveled from M to D via X in order to avoid congestion on the short path between M and D, these traders can now move to the new, less congested short path between M and D, reducing congestion along the line passing through X. For these reasons, the new line is not a perfect substitute for the existing lines, but acts as a sort of complement, in that it helps carry the burden of traffic.

To quantify the possible advantages of adding this link in a congested area, consider the comparison with a more classical infrastructure project connecting previously unconnected cities, as depicted in Figure 11b. This project, which in principle could be constructed as another Dedicated Freight Corridor, builds a new line between Amaravati, the newly planned capital of Andhra Pradesh, and Raipur, the capital of neighboring Chhattisgarh. Currently the route connecting these cities is a circuitous 873 kilometers, even though the straight-line distance is only 406 kilometers, leaving considerable scope to build a shorter line. Assume that, in this area with less passenger traffic, there is no congestion, so that, as in the classical approach, trade costs between two points are proportional to the minimum distance between this points on the transport network. How short would we need to make the new distance-reducing Amaravati-Raipur line, in order to achieve the same gains as the new congestion-reducing Mumbai-Delhi line?

To answer this question, I draw on a basic result from the general equilibrium trade model in Allen and Arkolakis (2016), which is that the welfare effects of reducing travel costs along a link, (i, j) in a transport network can be expressed as

$$\frac{d \ln W}{d \ln t_{ij}} = \sum_{k=1}^N \sum_{l=1}^N \frac{d \ln W}{d \ln \tau_{kl}} \times \frac{d \ln \tau_{kl}}{d \ln t_{ij}}. \quad (30)$$

Here, W is aggregate welfare, τ_{kl} is the average cost of trading between k and l , which is determined by the paths that various traders take. Along these paths between k and l traders incur costs t_{ij} , of moving between each directly connected pair of cities i and j . Accounting for congestion, these costs can depend on the total amount of trader traffic between i and j . As (30) shows, building new infrastructure between i and j lowers trade costs τ_{kl} for each pair of trading cities whose routes pass through i and j . In turn, these reductions in τ_{kl} affect welfare W according to standard trade model predictions. Specifically, in the economic geography version of Allen and Arkolakis (2016) with mobile labor, a straightforward application of the envelope theorem shows that the reduction in trade cost between two cities is proportional to the bilateral trade flow between the cities:

$$\frac{d \ln W}{d \ln \tau_{kl}} = -\frac{X_{kl}}{Y^W}, \quad (31)$$

where X_{kl} is the bilateral trade flow, and Y^W is world income.

To apply this model to the hypothetical comparison between the two projects, I rely on a stylized version of this comparison, in order to abstract from the many real world differences between Mumbai-Delhi and Amaravati-Raipur, including in particular geographic and political barriers to building between Amaravati-Raipur, and differences in the sizes and composition of the economies in these areas. I instead focus on the conceptual factors relevant to congestion.

This stylized comparison requires several assumptions. First, let $\Delta \ln \tau_{kl}$ be the effect on (k, l) trade costs of building the new project in question, and assume all trade costs are symmetric. Second, assume that the project in the congested area has one “main” effect $\tilde{\tau}_m$ on trade costs between the cities directly connected ($-\tilde{\tau}_m = \Delta \ln \tau_{MD} = \Delta \ln \tau_{DM}$), and a uniform effect $\tilde{\tau}_s$ on trade costs involving the “spillover” city X ($-\tilde{\tau}_s = \Delta \ln \tau_{MX} = \Delta \ln \tau_{XM} = \Delta \ln \tau_{DX} = \Delta \ln \tau_{XD}$). Third, assume that goods traded to and from the neighboring cities X and Y always travel directly on the line between these cities and the endpoints (M, D, A, or R), while goods traveling between the endpoint cities sometimes take the longer route through X or Y.²³ It follows that the A-R line does not affect trade costs for Y ($\Delta \ln \tau_{kl} = 0$ if $i = Y$ or $j = Y$). Fourth, normalize world income to one, and assume that the total bilateral flow between the endpoints, f_m is equal in each of the scenarios ($f_m \equiv X_{MD} + X_{DM} = X_{AR} + X_{RA}$), as is the total flow f_s from each of the neighboring spillover cities to each of the endpoints ($f_s \equiv X_{MX} + X_{XM} = X_{DX} + X_{XD} = X_{AY} + X_{YA} = X_{RY} + X_{YR}$). Fifth, let $\Delta d \equiv 1 - \frac{d_{AR}}{d_{AY} + d_{YR}}$ be the proportional reduction in A-R travel distance from building the new line between A and R; here, d_{ij} is the physical distance between i and j . Finally, let the area of each of the projects be a closed economy, so building the project in this area does not lead to business stealing from other areas, and we can abstract from the competitive mechanism studied earlier in this paper, allowing for a more straightforward application of the [Allen and Arkolakis \(2016\)](#) model.²⁴

Based on (30) and (31), the welfare effects of the two lines are

$$\Delta \ln W_{(\text{M-D line})} = f_m \tilde{\tau}_m + 2f_s \tilde{\tau}_s \quad (32)$$

$$\Delta \ln W_{(\text{A-R line})} = f_m \Delta d. \quad (33)$$

The effects of building in the congested area depends on both the direct effects on M and D, and the spillover effect onto neighboring X. The effect of building in the uncongested area, on the other hand, depends only on the reduction in travel distance between A and R, with no indirect effect. From (32) and (33), it follows that benefits of building in the congested M-D area are greater just in case

$$\Delta d < \tilde{\tau}_m + 2 \frac{f_s}{f_m} \tilde{\tau}_s. \quad (34)$$

This expression is intuitive. Building a new line in the congested M-D area is more beneficial when this has a large effect on trade costs between M and D (high $\tilde{\tau}_m$), when there are large spillover effects on trade costs

²³In a [Wardrop \(1952\)](#) equilibrium on a congested network, travelers between given endpoints will equalize the cost of travel across routes between these endpoints. The Wardrop equilibrium concept, concerned with decentralized travelers, is perhaps not entirely applicable in the Indian Railways setting of centrally planned traffic, though might be applicable to a model of total travel and congestion between the endpoints, inclusive of rail users and road users who make decentralized travel decisions. In any event, even in optimal centrally planned traffic flows, having some traders take each route typically requires there being more congestion on the shorter route so that its cost of travel is approximately equal to the cost of travel on the longer route.

²⁴The equilibrium in this model and the derived welfare effects still, of course, account for competition between the firms in each of the cities within the area being treated as a closed economy.

in the neighboring city ($\tilde{\tau}_s$), and when there is a relatively high volume of economic activity exposed to these spillover gains (high f_s/f_m).

My empirical estimates put magnitudes on the relevant variables in (34). First, the effect of Durontos on average costs, as argued above, provides a lower bound estimate of “cost effect” $\tilde{\tau}$. Such an increase in per-unit production cost will affect a firm in the same way as an increase in trade cost $\tilde{\tau}_m$. Recall that running one Duronto route leads to a 0.8 percent increase in average cost, and that this one route is 7 percent of the line capacity in the median district. If the new M-D line reduces line utilization between M and D from 100 percent to 50 percent, and the effects of this decongestion are proportional to the effects of adding Durontos, then the effect of the new line on the main line trade cost is $\tilde{\tau}_m = 0.8(\frac{100-50}{7}) = 5.7$ percent. Next, the estimates show that one Duronto route increases costs in the neighboring spillover areas by 0.7 percent. If these lines experience similar decongestion effects, then $\tilde{\tau}_s = 0.7(\frac{100-50}{7}) = 5.0$ percent.²⁵ The comparison in (34) depends, at this point, on the relative amount of economic activity exposed to the spillovers, f_s/f_m . This quantity could be small if, as is the case with Mumbai-Delhi, the link between the endpoints is an important trade route. It could also be large if, as is also the case with Mumbai-Delhi, there is a great amount of economic activity in the endpoints’ neighboring areas which are exposed to spillover traffic. Assuming for instance that $f_s/f_m = 1$, it follows that building the M-D line achieves the same gains as shortening the A-R line by 15.7 percent.

While this figure compares, in a stylized vacuum, the effects of de-congesting versus shortening, the numbers are more stark accounting for the actual levels of economic activity in the real-life comparison cities. In particular, letting f_{m_1} be bilateral trade between Mumbai and Delhi and f_{m_2} be bilateral trade between Amaravati and Raipur, the comparison in (34) becomes

$$\Delta d \leq \frac{f_{m_1}}{f_{m_2}} \tilde{\tau}_m + 2 \frac{f_s}{f_{m_2}} \tilde{\tau}_s. \quad (35)$$

In particular, the economic advantage to building the Mumbai-Delhi line becomes greater if there is more trade between these cities than between Amaravati and Raipur (high f_{m_1}/f_{m_2}), or if there is more activity in the spillover areas relative to that between Amaravati and Raipur (high f_s/f_{m_2}).²⁶ Based on rail shipment volumes along these routes and along the set of spillover routes, I obtain estimates of $f_{m_1}/f_{m_2} = 2.1$ and $f_s/f_{m_2} = 4.4$, implying that building the Mumbai-Delhi line achieves the same gains as shortening the Amaravati-Raipur distance by 56 percent.²⁷ This shortening would require building a 384 kilometer line between Amaravati-Raipur, which is, of course, physically impossible, given that the straight line distance is 406 kilometers. So this comparison does reveal possible benefits to building new lines in already served but congested areas.

²⁵This could also lead to some offsetting of the gains from decongesting the main line, as traffic previously on the long route moves back to this main line. In the extreme, the “fundamental law” of road congestion (Duranton and Turner, 2011) holds that building a new route can have no effect on travel times, as travelers fill the new route and increase its travel times. Such an extreme possibility is unlikely in the case of centrally managed rail traffic, though offsetting mechanism will likely occur, to some extent, depending on how the Railways re-routes traffic.

²⁶An additional complication in the real-life comparison is that building the Amaravati-Raipur line gives rail connection to previously unconnected districts between these cities. At the same time, the districts between Mumbai and Delhi gain some congestion relief from accessing the Dedicated Freight Corridor.

²⁷In obtaining these figures, I take the Amaravati area to include the adjoining Krishna district, which contains the city of Vijayawada.

7 Conclusion

The example of the Duronto trains shows that while running additional traffic on a transport network benefits those involved with that traffic, it also imposes externalities on certain other users of the transport system. These externalities work in large part by increasing the variance of shipment times, adding uncertainty to an already uncertain world faced by developing country firms. These uncertainty effects are difficult to disentangle from negative shocks in most other settings, but here I show that they create a significant drag on the productivity of the affected firms. At the same time, one firm's loss is a competitor's gain, which helps to offset the affected firms' losses in terms of congestion's net effect.

This analysis also points to some interesting further questions. First, a full welfare analysis of the Duronto trains would depend not only on how they affect firms in intermediate districts, but also on how they benefit the passengers riding them between the endpoint districts, and on how they create congestion for other passenger trains. While these effects are beyond the scope of this paper, related work studies the benefits on the passenger side, using Railways data to study patterns of seasonal migration from rural areas to labor markets (Firth, Forster and Imbert, 2017). Second, over the long run, firms can make locational adjustments in response to conditions on the transportation network. For example, Gulyani (2001) reports that Indian automakers respond to transportation problems by clustering geographically, and thus limiting their reliance on transport infrastructure. In this light, another related paper studies how certain distortions in railway freight pricing contributed, over the long run, to agglomeration of closely related industries in certain regions of India (Firth and Liu, 2017).

References

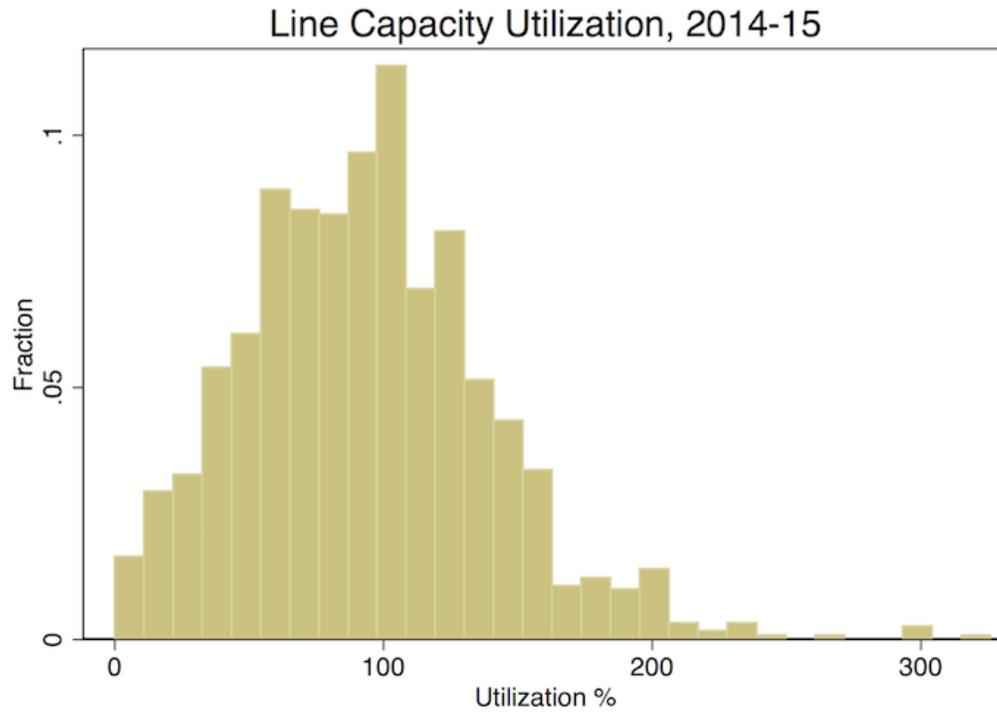
- Allcott, Hunt, Allan Collard-Wexler, and Stephen D O’Connell.** 2016. “How do electricity shortages affect industry? Evidence from India.” *The American Economic Review*, 106(3): 587–624.
- Allen, Treb.** 2014. “Information frictions in trade.” *Econometrica*, 82(6): 2041–2083.
- Allen, Treb, and Costas Arkolakis.** 2016. “The Welfare Effects of Transportation Infrastructure.”
- Anderson, James E, and Eric Van Wincoop.** 2004. “Trade costs.” *Journal of Economic Literature*, 42(3): 691–751.
- Arrow, Kenneth J, Theodore Harris, and Jacob Marschak.** 1951. “Optimal inventory policy.” *Econometrica*, 250–272.
- Asker, John, Allan Collard-Wexler, and Jan De Loecker.** 2014. “Dynamic inputs and resource (mis)allocation.” *Journal of Political Economy*, 122(5): 1013–1063.
- Bajaj, Vikas.** 2010. “Clogged Rail Lines Slow India’s Development.” *The New York Times*.
- Banerjee, Abhijit, Esther Duflo, and Nancy Qian.** 2012. “On the Road: Access To Transportation Infrastructure and Economic Growth in China.”
- Banerjee, Mamata.** 2009. “Speech Introducing the Railway Budget 2009-2010.”
- Baum-Snow, Nathaniel.** 2007. “Did highways cause suburbanization?” *The Quarterly Journal of Economics*, 122(2): 775–805.
- Blanchard, Olivier, and John Simon.** 2001. “The long and large decline in US output volatility.” *Brookings Papers on Economic Activity*, 2001(1): 135–164.
- Bloom, Nicholas.** 2009. “The impact of uncertainty shocks.” *Econometrica*, 77(3): 623–685.
- Broda, Christian, Joshua Greenfield, and David Weinstein.** 2006. “From groundnuts to globalization: A structural estimate of trade and growth.”
- Broyden, C. G.** 1965. “A Class of Methods for Solving Nonlinear Simultaneous Equations.” *Mathematics of Computation*, 19(92): 577–593.
- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller.** 2011. “Robust inference with multiway clustering.” *Journal of Business & Economic Statistics*, 29(2): 238–249.
- Chen, Bintong, and Patrick T Harker.** 1990. “Two moments estimation of the delay on single-track rail lines with scheduled traffic.” *Transportation Science*, 24(4): 261–275.
- Datta, Saugato.** 2012. “The impact of improved highways on Indian firms.” *Journal of Development Economics*, 99(1): 46–57.
- David, Joel M, Hugo A Hopenhayn, and Venky Venkateswaran.** 2016. “Information, misallocation, and aggregate productivity.” *The Quarterly Journal of Economics*, 131(2): 943–1005.

- Donaldson, Dave.** 2015. “The gains from market integration.” *Annual Review of Economics*, 7(1): 619–647.
- Donaldson, Dave.** 2017. “Railroads of the Raj: Estimating the impact of transportation infrastructure.” *American Economic Review*.
- Donaldson, Dave, and Richard Hornbeck.** 2016. “Railroads and American economic growth: A ‘market access’ approach.” *Quarterly Journal of Economics*, 131(2): 799–858.
- Duranton, Gilles, and Matthew A. Turner.** 2011. “The Fundamental Law of Road Congestion: Evidence from US Cities.” *American Economic Review*, 101(6): 2616–52.
- Duranton, Gilles, and Matthew A Turner.** 2012. “Urban growth and transportation.” *Review of Economic Studies*, 79(4): 1407–1440.
- Eaton, Jonathan, and Samuel Kortum.** 2002. “Technology, geography, and trade.” *Econometrica*, 70(5): 1741–1779.
- Edgeworth, Francis Y.** 1888. “The mathematical theory of banking.” *Journal of the Royal Statistical Society*, 51(1): 113–127.
- Faber, Benjamin.** 2014. “Trade integration, market size, and industrialization: evidence from China’s National Trunk Highway System.” *Review of Economic Studies*, 81(3): 1046–1070.
- Fafchamps, Marcel, Jan Willem Gunning, and Remco Oostendorp.** 2000. “Inventories and risk in African manufacturing.” *The Economic Journal*, 110(466): 861–893.
- Fajgelbaum, Pablo D, and Edouard Schaal.** 2017. “Optimal Transport Networks in Spatial Equilibrium.”
- Firth, John, and Ernest Liu.** 2017. “Manufacturing underdevelopment: India’s Freight Equalization Scheme, and the long-term effects of distortions on the geography of production.”
- Firth, John, Felix Forster, and Clement Imbert.** 2017. “Internal migration in India: New evidence from rail passenger travel data.”
- Fogel, Robert William.** 1964. *Railroads and American Economic Growth*. Cambridge University Press.
- Godfrey, L. G.** 1988. *Misspecification Tests in Econometrics*. Cambridge University Press.
- Guasch, J Luis, and Joseph Kogan.** 2003. “Just-in-Case Inventories: A Cross-Country Analysis.”
- Gulyani, Sumila.** 2001. “Effects of Poor Transportation on Lean Production and Industrial Clustering: Evidence from the Indian Auto Industry.” *World Development*, 29(7): 1157–1177.
- Hadley, G., and Thomson M. Whitin.** 1963. *Analysis of Inventory Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1963.
- Hansen, Lars Peter.** 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica*, 50(4): 1029–1054.

- Harker, Patrick T, and Sungwook Hong.** 1990. “Two moments estimation of the delay on a partially double-track rail line with scheduled traffic.” *Journal of the Transportation Research Forum*, 31(1).
- Hillberry, Russell, and David Hummels.** 2008. “Trade responses to geographic frictions: A decomposition using micro-data.” *European Economic Review*, 52(3): 527–550.
- Hopenhayn, Hugo A.** 2014. “Firms, misallocation, and aggregate productivity: A review.” *Annual Review of Economics*, 6(1): 735–770.
- Hsieh, Chang-Tai, and Peter Klenow.** 2009. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics*, 124(4): 1403–1448.
- Hummels, David, and Georg Schaur.** 2013. “Time as a Trade Barrier.” *American Economic Review*, 103(7): 2935–2959.
- Levchenko, Andrei A.** 2007. “Institutional quality and international trade.” *The Review of Economic Studies*, 74(3): 791–819.
- Li, Han, and Zhigang Li.** 2013. “Road investments and inventory reduction: Firm level evidence from China.” *Journal of Urban Economics*, 76: 43–52.
- Macchiavello, Rocco, and Ameet Morjaria.** 2015. “The value of relationships: evidence from a supply shock to Kenyan rose exports.” *The American Economic Review*, 105(9): 2911–2945.
- McConnell, Margaret M, and Gabriel Perez-Quiros.** 2000. “Output Fluctuations in the United States: What Has Changed since the Early 1980’s?” *American Economic Review*, 90(5): 1464–1476.
- Ministry of Finance.** 2015. “Economic Survey of India for 2014-15.”
- Ministry of Railways.** 2008. *Operating Manual for Indian Railways*.
- Ministry of Railways.** 2009. *Railway Budget 2009-2010*.
- Ministry of Railways.** 2011. “Indian Railways: Annual Report and Accounts 2010-11.”
- Ministry of Railways.** 2015a. “Indian Railways: A White Paper.”
- Ministry of Railways.** 2015b. “Report of the Committee for Mobilization of Resources for Major Railway Projects and Restructuring of Railway Ministry and Railway Board.”
- Nahmias, Steven.** 2001. *Production and Operations Analysis*. McGraw-Hill.
- Neely, Michael J.** 2010. “Stochastic network optimization with application to communication and queueing systems.” *Synthesis Lectures on Communication Networks*, 3(1): 1–211.
- Petersen, ER.** 1974. “Over-the-road transit time for a single track railway.” *Transportation Science*, 8(1): 65–74.
- Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez.** 2016. “Trade, domestic frictions, and scale effects.” *The American Economic Review*, 106(10): 3159–3184.

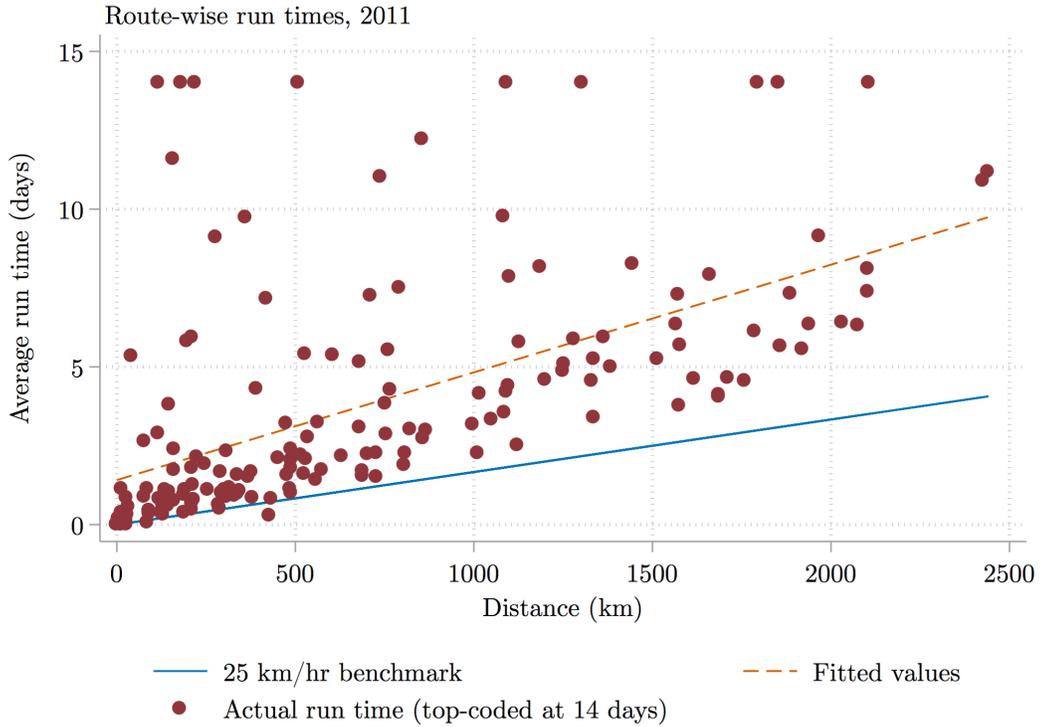
- Redding, Stephen J., and Daniel M Sturm.** 2008. "The costs of remoteness: Evidence from German division and reunification." *The American Economic Review*, 98(5): 1766–1797.
- Redding, Stephen J., and Matthew A. Turner.** 2015. "Transportation Costs and the Spatial Organization of Economic Activity." *Handbook of Regional and Urban Economics*, 5(20): 1339–1398.
- Rotemberg, Martin.** 2017. "Equilibrium Effects of Firm Subsidies."
- Rubin, Donald B.** 1980. "Randomization analysis of experimental data: The Fisher randomization test comment." *Journal of the American Statistical Association*, 75(371): 591–593.
- Startz, Meredith.** 2016. "The value of face-to-face: Search and contracting problems in Nigerian trade."
- Wardrop, John Glen.** 1952. "Some Theoretical Aspects of Road Traffic Research." *Proceedings of the Institution of Civil Engineers*, 1(3): 325–362.
- World Bank.** 1994. "World Development Report 1994: Infrastructure for Development."
- Ziobro, Paul.** 2017. "Trains in Vain: Epic CSX Traffic Jam Snarls Deliveries, From Coal to Fries." *The Wall Street Journal*.

Figure 1: Histogram of line capacity utilization



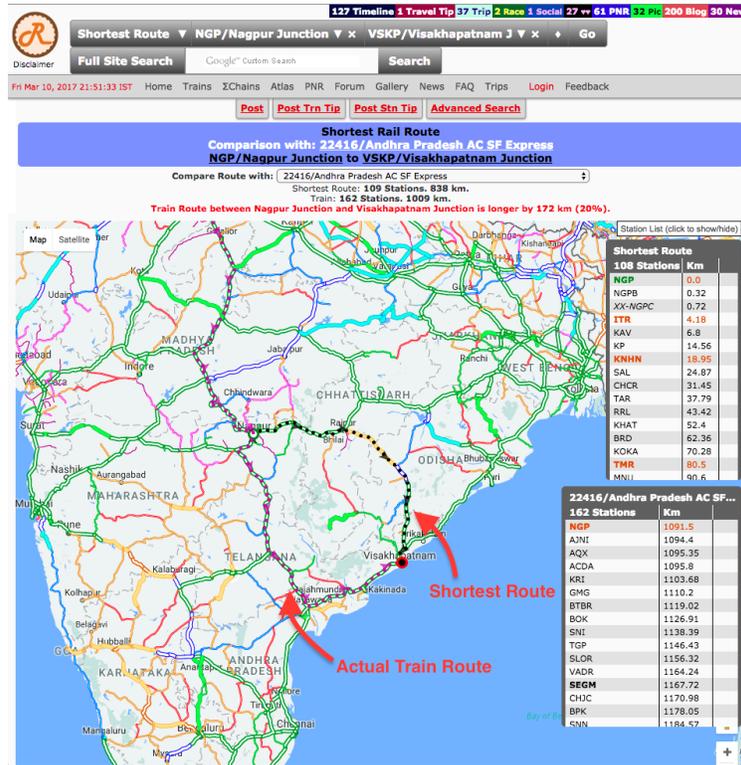
Notes: This figure depicts the capacity utilization of the track sections on Indian Railways. The utilization percentage is measured as the average daily number of trains passing on the section, divided by the prescribed amount of traffic for that section, which is based on an engineering rule known as Scott's Formula. Source: Indian Railways Line Capacity Charts.

Figure 2: Route-wise average freight shipment times



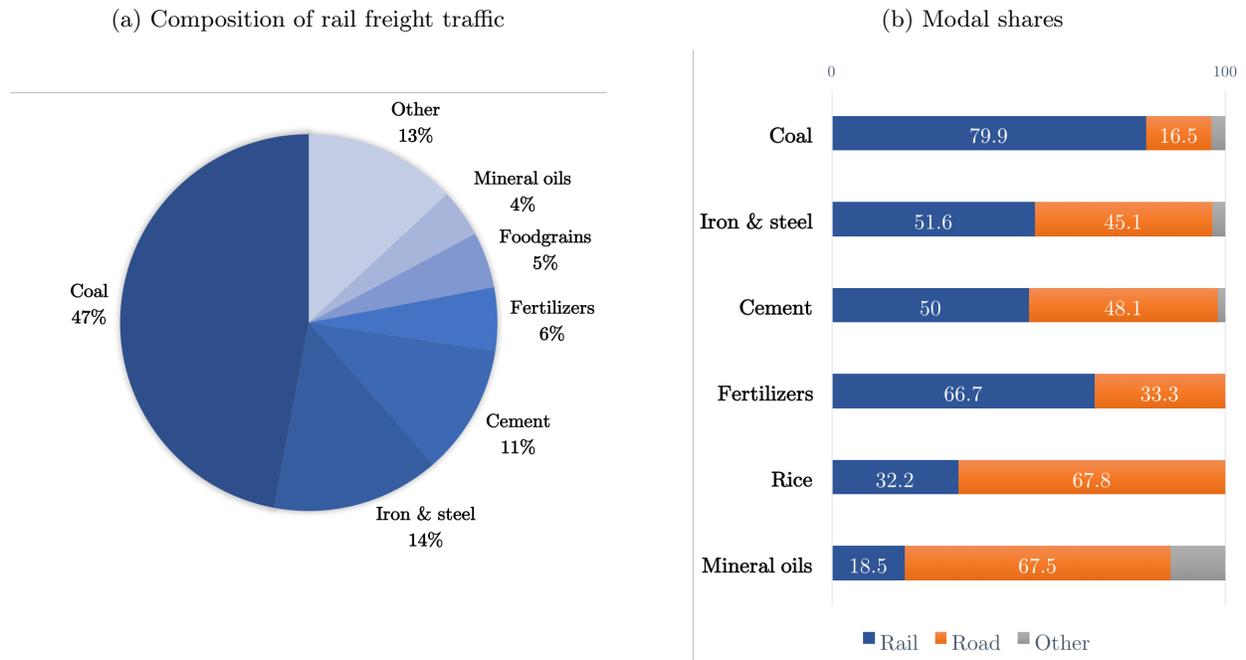
Notes: This figure plots cross-sectional route-wise average run times for freight shipments, against the track distance of the route. Each point reflects the annual average run time, in days, for an origin-destination pair between which freight shipment takes place. These points are compared, first, against the benchmark time it would take if shipments maintained the standard freight shipment speed of 25 kilometers per hour. The other line is a best-fit of run time as a function of distance. The R^2 from regressing run time on distance is 0.19. Source: Indian Railways freight shipment database.

Figure 3: Sample from scraped website with data on train routes



Notes: This figure shows a sample of the information about each train which is scraped from the website [IndiaRailInfo](http://IndiaRailInfo.com). For each train, data is collected on the actual route the train follows, and the shortest possible path between its endpoints.

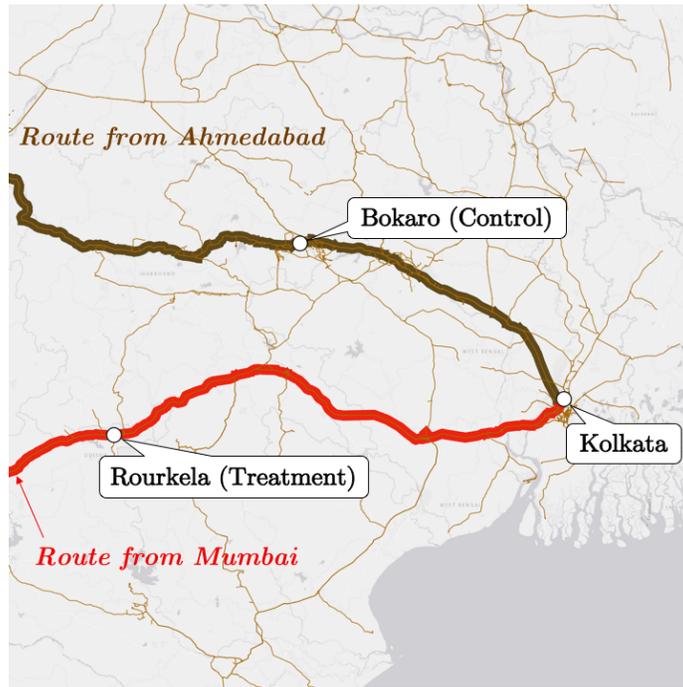
Figure 4: Goods shipped by rail in India



Notes: This figure shows that a certain set of “rail goods” account for the bulk of railway freight traffic in India, and conversely that these goods rely heavily on the rails rather than other modes of transportation. Panel (a) shows the commodity-wise composition by weight of goods shipped by rail; the composition by value is similar. Panel (b) shows, for each of these commodity categories with available data, the fractions of freight shipped by Rail, by Road, and by Other modes of transport. Source: [Ministry of Railways \(2011\)](#).

Figure 5: Reduced form empirical strategy, accounting for spillover effects

(a) Basic reduced form



(b) Spillovers from diversion of traffic

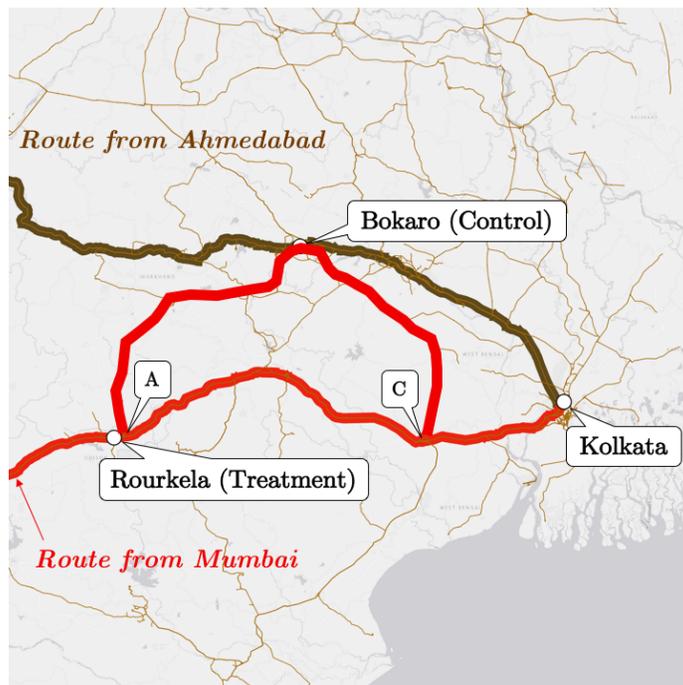
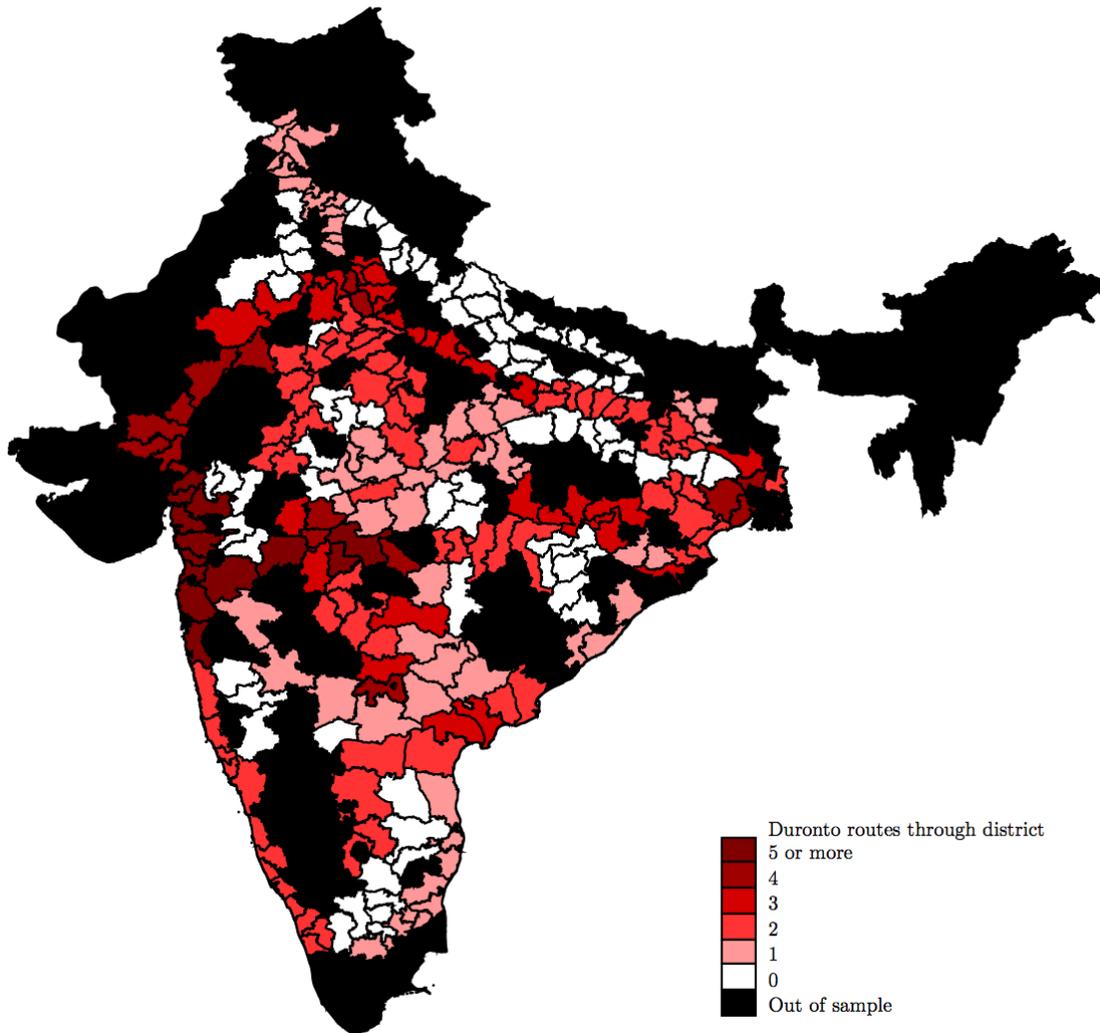
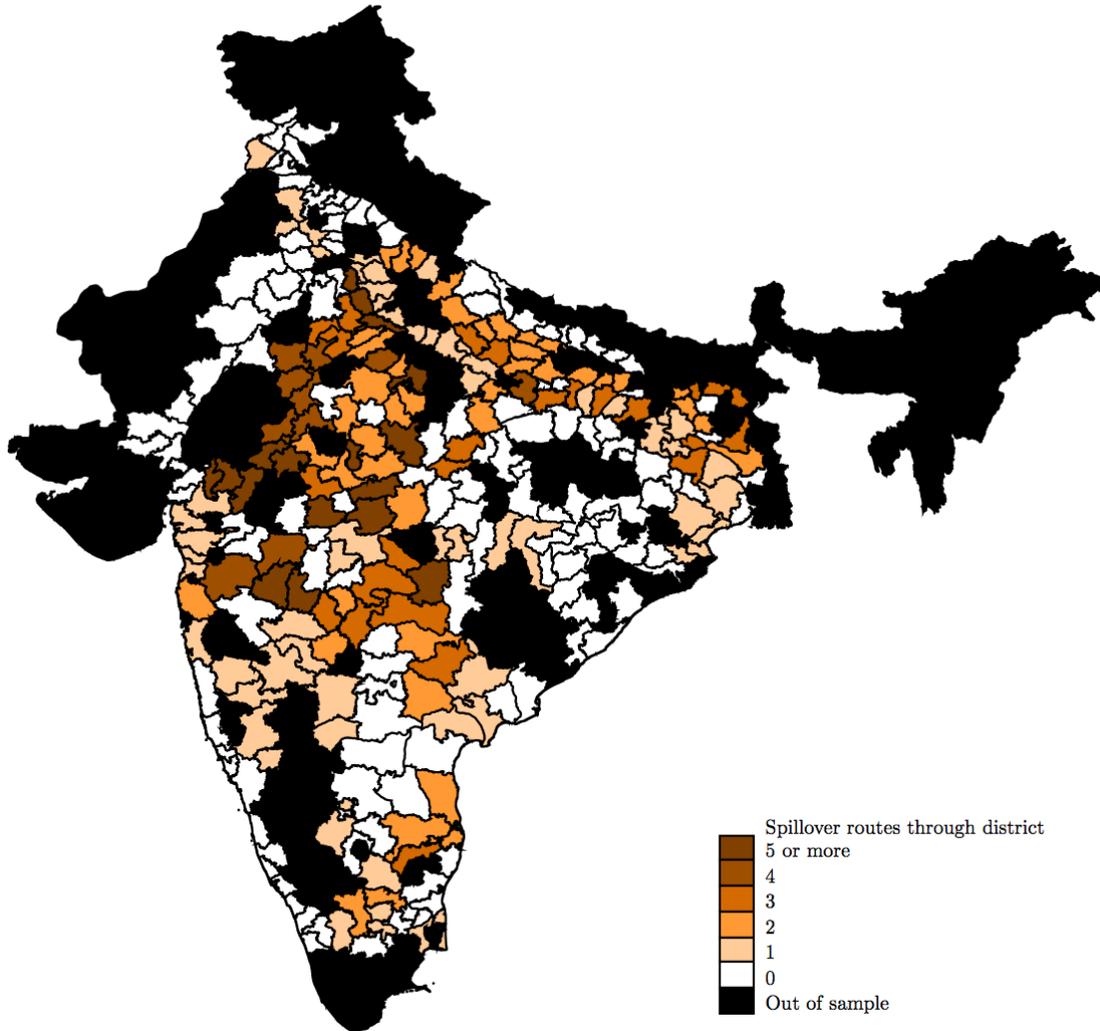


Figure 6: District-wise exposure to Duronto routes



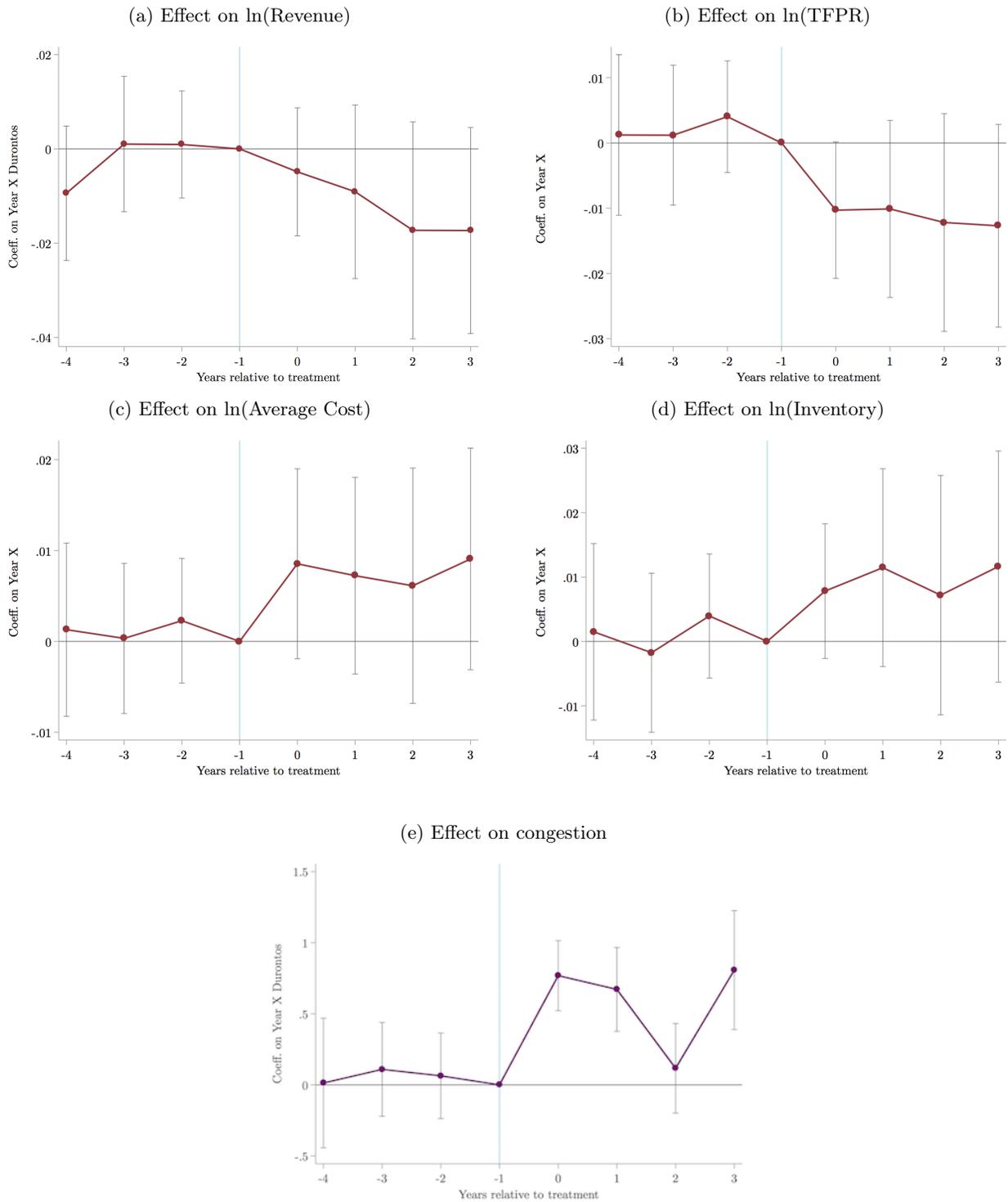
Notes: This figure depicts each district's exposure to the Duronto treatment, defined, as in the text, as the number of two-way Duronto shortest-path routes passing through the district. It shows the cumulative treatment as of 2012, including all trains added between 2009 and 2012. The sample is restricted to districts on the shortest path between the major cities connected by the Duronto program, and which were therefore places that the Durontos conceivably could have run. The out of sample areas, including the endpoint districts actually served by the Durontos, are shaded in black.

Figure 7: District-wise exposure to spillover routes



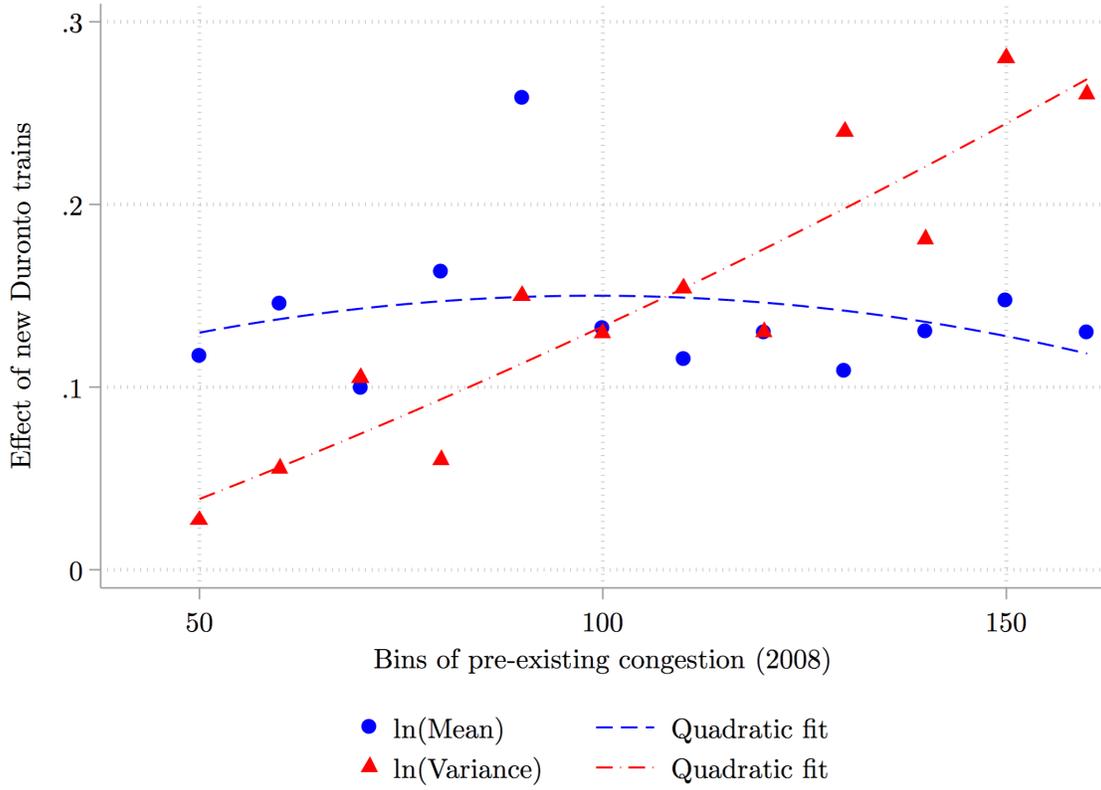
Notes: This figure depicts each district's exposure to the spillover traffic from the Duronto treatment, defined, as in the text, as the number of Duronto routes for which the district lies on a "diversion" route. It shows the cumulative treatment as of 2012, including all trains added between 2009 and 2012. The sample is restricted to districts on the shortest path between the major cities connected by the Duronto program, and which were therefore places that the Durontos conceivably could have run. The out of sample areas, including the endpoint districts actually served by the Durontos, are shaded in black.

Figure 8: Event study for effect of Durontos



Notes: In panels (a) to (d), this figure shows event studies for the effect of introducing a Duronto route on each of the four main firm outcomes of interest. Panel (e) presents a “zero-th stage event study”, showing at the track section level, the effect of a new Duronto route on the amount of traffic running on that section.

Figure 9: Mean and variance response to Durontos, as a function of pre-existing congestion

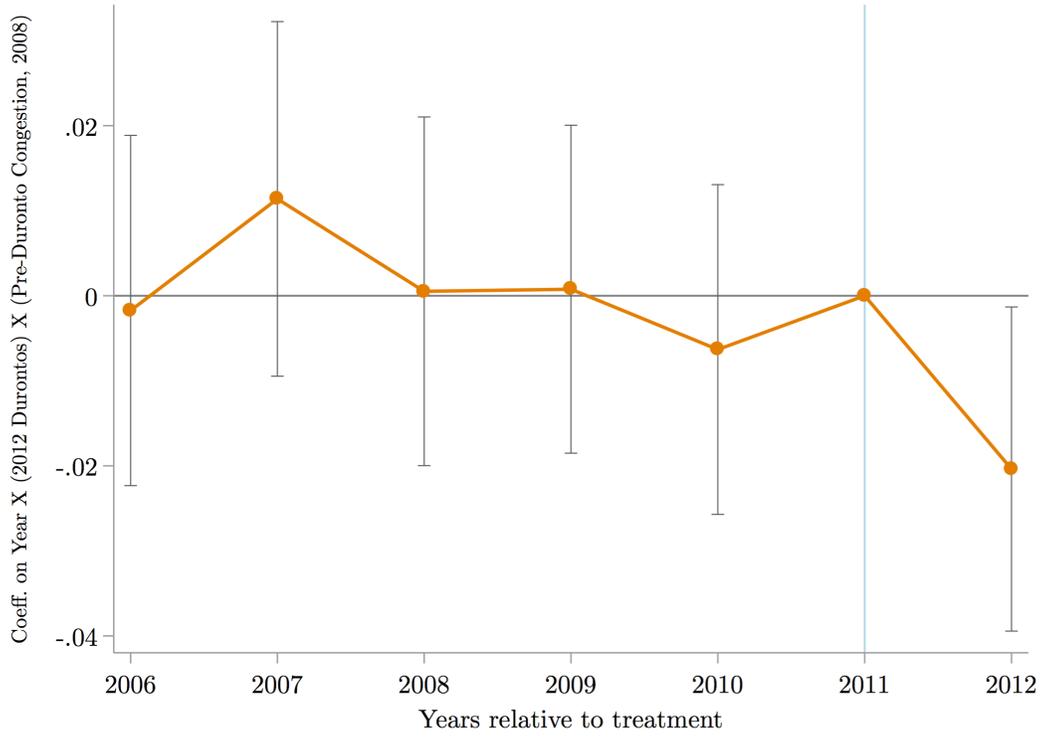


Notes: This figure shows how shipping times respond to increased traffic, consistent with the railway model from operations research. Specifically, it plots the β_c coefficients from the regressions

$$\ln M_{dy} = \sum_{c=50}^{160} \beta_c^M (D_{dy} \times \mathbf{1}[c \leq C_{d,2008} < c + 10]) + \gamma_d + \gamma_y + \epsilon_{dy}$$

$$\ln V_{dy} = \sum_{c=50}^{160} \beta_c^V (D_{dy} \times \mathbf{1}[c \leq C_{d,2008} < c + 10]) + \gamma_d + \gamma_y + \epsilon_{dy}.$$

Figure 10: Event study for the effect of $D_{dy} \times T_{d,y=y_0}$ on revenue



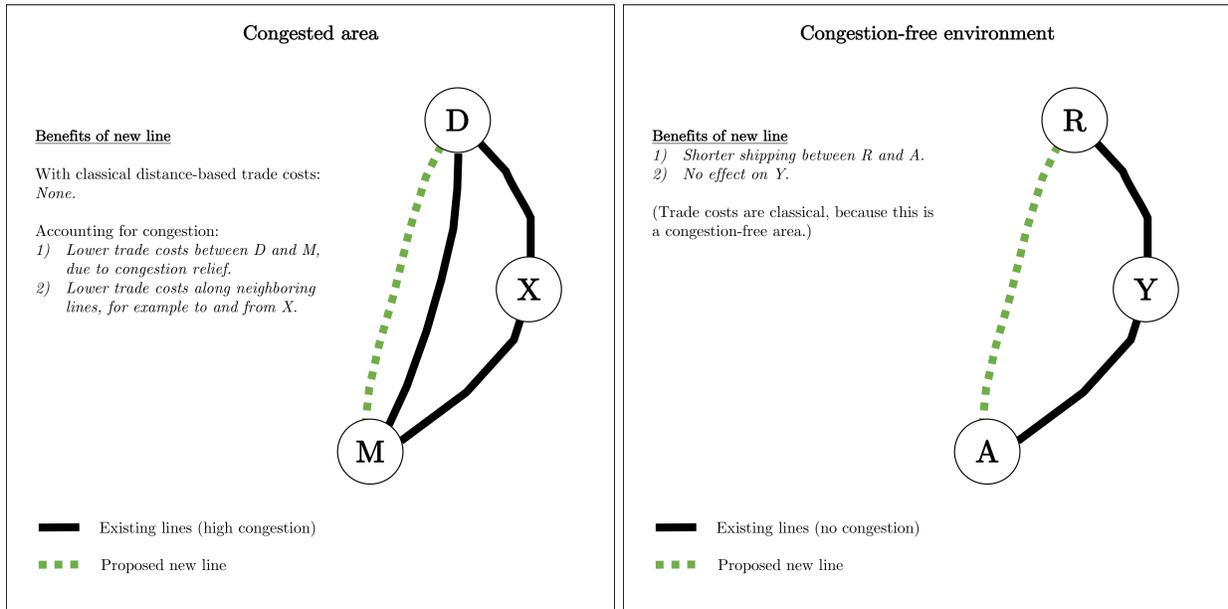
Notes: This figure shows an event study for the effect on revenue of the interaction of Duronto traffic with district pre-existing congestion. Specifically, it plots the β_y coefficients in

$$\ln(\text{Revenue})_{it} = \sum_{y=2006}^{2012} \beta_y (D_{d,2012} \times T_{d,2008} \times \mathbf{1}[t = y]) + D_{dt} + S_{dt} + \gamma_i + \gamma_{t \times s} + \gamma_{t \times k} + \epsilon_{it}$$

Figure 11: Effects of two hypothetical construction projects

(a) Stylized Mumbai-Delhi corridor

(b) Stylized Amaravati-Raipur corridor



(c) Map of cities involved



Table 1: Descriptive statistics for factories in rail using industries

	Mean (1)	St. Dev. (2)	Δ by eventual treatment	
			Duronto (3)	Spillover (4)
<i>Firm variables, at factory level</i>				
Revenue (million INR)	1251.7	3096.1	29.304 (20.634)	15.31 (26.368)
ln(TFPR)	2.389	0.864	-0.022* (0.013)	0.001 (0.008)
Average cost	1.013	1.136	-0.006 (0.008)	0.013 (0.015)
Total inventory (million INR)	185.5	480.8	3.254 (3.399)	-0.713 (4.268)
Inputs	108.9	278.2	1.713 (2.054)	-1.142 (2.462)
Finished goods	76.3	190.0	1.437 (1.308)	0.776 (1.759)
Input share of rail goods	0.265	0.189	-0.005* (0.003)	-0.003 (0.004)
Makes rail good (dummy)	0.631	0.483	0.003 (0.004)	0.011* (0.006)
Survival until 2012	0.534	0.499	-0.011 (0.007)	0.002 (0.012)
<i>Rail traffic variables</i>				
Line capacity (trains per day)	32.4	28.2	-1.055 (2.193)	0.997 (2.312)
Line capacity utilization %	95.2	11.8	0.406 (0.452)	0.744 (0.539)
% passenger traffic	66.3	16.1	0.263 (0.588)	-0.560 (0.582)
% freight traffic	28.1	15.3	-0.157 (0.573)	0.403 (0.683)
% other traffic	5.5	4.8	-0.110 (0.076)	0.137 (0.242)
Mean freight ship time, days (normalized to 1000km)	5.11	4.10	0.624 (0.810)	-0.255 (0.898)
Variance of freight time, days (normalized to 1000km)	6.39	11.19	-0.037 (0.722)	-0.402 (0.734)
Factories in sample	8281			
Districts in sample	248			

Notes: This table presents descriptive statistics for ASI factories in rail-using industries, defined as those which either (a) produce a good commonly shipped by rail (coal, iron, steel, cement, fertilizers, food grains, mineral oils), or (b) whose input cost share for the median firm in pre-2009 data is at least 5 percent. The rail traffic variables are district-level measures for the rail lines in the districts containing at least one of these rail using factories. The rail shipping time variables are calculated as a weighted average over all of the shipping routes going to and from the district, weighted by the number of freight trains run on each route. Sources: ASI, Indian Railways Line Capacity data, Indian Railways, Freight Shipment data.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Effects of Durontos on railway line traffic patterns

	Main specification			With second-order spillovers		
	Passenger trains (1)	Freight trains (2)	Total trains (3)	Passenger trains (4)	Freight trains (5)	Total trains (6)
Duronto routes	0.611*** (0.131)	0.0208 (0.118)	0.666*** (0.177)	0.602*** (0.130)	0.0286 (0.120)	0.666*** (0.178)
Spillover exposure (alternate routes)	0.221** (0.0917)	0.227** (0.114)	0.408*** (0.152)	0.210** (0.0907)	0.239** (0.115)	0.411*** (0.154)
Second-order spillovers				0.0485 (0.0389)	-0.0499* (0.0256)	-0.00979 (0.0473)
Mean of dep. var.	27.43	13.6	43.57	25.83	12.71	40.97
R^2 (adjusted, within)	0.042	0.006	0.031	0.040	0.007	0.030
Observations	2198	2198	2198	2494	2494	2494
Section FE	✓	✓	✓	✓	✓	✓
Yr \times Sample FE for {Dur,Alt}	✓	✓	✓	✓	✓	✓
Yr \times Sample FE for S-O				✓	✓	✓

Notes: This table presents estimates of equation (2), showing the “zero-th stage” effect of Duronto trains on railway congestion. It is estimated at the level of the track section, where the dependent variable is the annual daily average number of trains of each type running on the section. The first independent variable is the number of Duronto trains (based on the shortest path between endpoints) scheduled to run on the section as of that year. The next independent variable, spillover exposure is the number of introduced Duronto trains for which the section lies on a spillover alternate route, as defined in the text. The second order spillovers variable, considered only in Columns (4) through (6), indicates the exposure of the district to the alternate routes of these alternate routes, showing that traffic spillovers do not extend quite this far. Standard errors in parentheses clustered by track section. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Reduced form effects of Duronto trains on rail using firms

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
<i>Panel A. Preferred specification</i>				
Duronto routes through district	-0.0194*** (0.0050)	-0.0111*** (0.0041)	0.0081** (0.0032)	0.0097* (0.0053)
Spillover routes through district	-0.0110* (0.0063)	-0.0064 (0.0042)	0.0072* (0.0042)	0.0006 (0.0064)
<i>Panel B. Without spillover control</i>				
Duronto routes through district	-0.0125** (0.0049)	-0.0075* (0.0041)	0.0036 (0.0030)	0.0093* (0.0053)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), for factories in rail-using industries. The dependent variables are the four main outcomes of interest as defined in the text. The regressors are the number of two-way Duronto routes (based on shortest path) passing through the district as of the current year, and the number of introduced Duronto trains for which the district lies on a spillover alternate route, as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year (Cameron, Gelbach and Miller, 2011). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Placebo effects on non rail using firms

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
<i>Panel A. Preferred specification</i>				
Duronto routes through district	-0.0006 (0.0044)	-0.0031 (0.0019)	0.0011 (0.0032)	0.0021 (0.0049)
Spillover routes through district	-0.0023 (0.0059)	-0.0024 (0.0028)	0.0012 (0.0041)	-0.0015 (0.0060)
<i>Panel B. Without spillover control</i>				
Duronto routes through district	0.0003 (0.0044)	-0.0019 (0.0019)	0.0004 (0.0033)	0.0031 (0.0049)
Observations	50483	48101	37688	45012
Clusters 1 (factories)	10844	10420	8664	9690
Clusters 2 (district \times year)	2329	2293	2248	2305

Notes: This table presents estimates of equation (1), for factories in non rail using industries. All other details are as in Table 3 above. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: First stage effects of Duronto traffic on freight shipment times

	ln(Mean) (1)	ln(Variance) (2)
Duronto routes through district	0.113*** (0.028)	0.039 (0.026)
(Duronto routes) × (2008 congestion)	-0.021 (0.031)	0.211*** (0.042)
Observations	6896	6896
Clusters (districts)	174	174
<i>F</i> statistic	19.22	27.43
Control for spillovers	✓	✓

Notes: This table presents estimates of equations (7), indicating the first stage effect of Duronto traffic through a district on the (log) mean and (log) variance of annual shipping times to and from the district. The district level shipping time measures are calculated using the set of freight routes which remain in operation, with at least one train running in each year, throughout the sample period. The measure of 2008 congestion is the total amount of traffic on all of the railway lines in the district, divided by the prescribed line capacity. Both regressions include fixed effects for district and year. Robust standard errors in parentheses, with clustering by district. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: 2SLS estimates of mean and variance effects

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
<i>Panel A. 2SLS</i>				
ln(Mean)	-0.029 (0.026)	-0.025 (0.016)	0.001 (0.018)	0.032* (0.019)
ln(Variance)	-0.107*** (0.031)	-0.033* (0.019)	0.034* (0.019)	0.042* (0.023)
<i>Panel B. Reduced form</i>				
Duronto routes through district	-0.007 (0.007)	-0.004 (0.003)	0.001 (0.003)	0.005* (0.003)
(Duronto routes) × (2008 congestion)	-0.022*** (0.006)	-0.006 (0.004)	0.007* (0.004)	0.008** (0.004)
Observations	6896	6682	6390	6676
Clusters 1 (factories)	3448	3341	3195	3338
Clusters 2 (district × year)	348	348	344	348
Control for spillovers, exposure	✓	✓	✓	✓

Notes: Panel A of this table presents second stage estimates of equation (4), showing the effects of mean and variance of shipping time on the four main firm outcomes of interest. Panel B presents reduced form estimates of these outcomes on the instruments specified in (7). All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Model estimates of cost and competition effects

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0308*** (0.0050)	-0.0115*** (0.0037)	0.0079** (0.0035)	0.0094** (0.0046)
Spillover routes through district	-0.0104* (0.0061)	-0.0079* (0.0046)	0.0069* (0.0037)	0.0004 (0.0064)
Exposure of (State \times Industry) to Duronto routes	0.0249** (0.0106)	0.0035 (0.0079)	0.0012 (0.0110)	0.0011 (0.0131)
Exposure of (State \times Industry) to spillover routes	0.0186 (0.0125)	-0.0108 (0.0082)	-0.0036 (0.0127)	-0.0014 (0.0147)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (28), for factories in rail-using industries. The dependent variables are the four main outcomes of interest as defined in the text. The regressors are the number of Duronto and spillover routes passing through the district, along with the exposure of other district competitors in the same state and 4-digit NIC industry to Duronto and spillover traffic, weighted by the 2008 industry revenue in the competing district. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Model estimates of cost and competition effects, with elasticity interactions

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0014 (0.0052)	-0.0081** (0.0039)	0.0080** (0.0037)	0.0089** (0.0041)
(Duronto routes) \times σ	-0.0065*** (0.0004)	-0.0006** (0.0003)	-0.0001 (0.0004)	0.0002 (0.0006)
Spillover routes through district	-0.0046 (0.0064)	-0.0056 (0.0052)	0.0067* (0.0039)	0.0004 (0.0073)
(Spillover routes) \times σ	-0.0012* (0.0007)	-0.0003 (0.0005)	0.0002 (0.0005)	0.0003 (0.0007)
Exposure of (State \times Industry) to Duronto routes	0.0231* (0.0119)	0.0039 (0.0084)	0.0022 (0.0117)	0.0019 (0.0140)
Exposure of (State \times Industry) to spillover routes	0.0201 (0.0129)	-0.0094 (0.0088)	-0.0034 (0.0140)	-0.0008 (0.0151)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (28) for factories in rail-using industries, adding regressors capturing the interaction between Duronto and spillover traffic and the industry elasticity of substitution coming from Broda et al. (2006). All regressions include fixed effects for factory, year by state, and year by NIC industry. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Aggregate effects of Duronto congestion on revenue, at state-industry level

	Dependent variable: $\ln PY$			
	(1)	(2)	(3)	(4)
Exposure of (State \times Industry) to Duronto routes	-0.0132 (0.0159)	-0.0020 (0.0227)	-0.0139 (0.0155)	-0.0028 (0.0224)
Exposure of (State \times Industry) to spillover routes			-0.0048 (0.0097)	-0.0051 (0.0206)
Observations	7901	7883	7901	7883
Clusters (state \times industry)	1932	1914	1932	1914
State \times Industry FE	✓	✓	✓	✓
Year FE	✓		✓	
State \times Yr FE, Ind \times Yr FE		✓		✓

Notes: This table presents estimates of equation (29), showing the effect of Duronto and spillover traffic exposure on aggregate sales for each state by 4-digit NIC industry. All regressions include fixed effects for year and state by industry, with Columns (2) and (4) adding effects for year by state and year by industry. Robust standard errors in parentheses, with clustering by state times industry. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Aggregate effects of Duronto congestion on gross value added, at state-industry level

	Dependent variable: $\ln GVA$			
	(1)	(2)	(3)	(4)
Exposure of (State \times Industry) to Duronto routes	-0.0149 (0.0164)	-0.0058 (0.0231)	-0.0098 (0.0160)	-0.0073 (0.0230)
Exposure of (State \times Industry) to spillover routes			-0.0056 (0.0100)	-0.0043 (0.0208)
Observations	7850	7831	7850	7831
Clusters (state \times industry)	1909	1901	1909	1901
State \times Industry FE	✓	✓	✓	✓
Year FE	✓		✓	
State \times Yr FE, Ind \times Yr FE		✓		✓

Notes: This table presents estimates of equation (29), showing the effect of Duronto and spillover traffic exposure on aggregate (log) gross value added for each state by 4-digit NIC industry. All regressions include fixed effects for year and state by industry, with Columns (2) and (4) adding effects for year by state and year by industry. Robust standard errors in parentheses, with clustering by state times industry. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: The cost of running one Duronto train

<i>Panel A: Cost of running one Duronto route, imperfect competition</i>			
	Loss for affected firms (1)	Gain for competitors (2)	Net effect (3)
Duronto direct effects	-461.2	372	-89.2
Spillover effects	-154.9	195.4	40.5
Total (million INR)	-616.1	567.4	-48.7

<i>Panel B: All firms experience congestion increase equivalent to one Duronto, perfect competition</i>		
	Rail-using firms (4)	All manufacturing (5)
Total output loss (million INR)	94,962	258,551

Notes: All figures in this table are in millions of 2008 Indian rupees (nominal exchange rate is 48.5 INR per USD; exchange rate at PPP is 12.9 INR per USD). Calculations are as described in the text, with Panel A reporting the estimated revenue loss for rail-using firms of running one Duronto train, inclusive of direct losses to affected firms and gains to their competitors. A point of comparison for these figures is the annual passenger fare revenue from one of these routes, which I estimate at INR 1,024 million. Panel B reports the aggregate effects of exposing all firms to a cost shock equivalent to that estimated for the Duronto-affected firms.

A Additional firm outcomes

Table A1: Effects on district railway traffic

	Passenger trips		Freight trains	
	Originating (1)	Terminating (2)	Originating (3)	Terminating (4)
Durontos	0.0062 (0.0113)	0.0029 (0.0111)	-0.0081 (0.0121)	0.0012 (0.0169)
Spillovers	0.0014 (0.0122)	0.0017 (0.0139)	-0.0058 (0.0151)	-0.0119 (0.0149)
Observations	304	312	224	224
Clusters (districts)	152	156	112	112

Notes: This table presents a district level regression of passenger and freight trips, on the amount of Duronto and spillover traffic introduced through the district. The dependent variables, all in natural logarithms, measure the number of trips of each type either originating or terminating in the district. The regression sample includes only 2011 to 2012, the first years for which passenger and freight traffic data are available. All regressions include fixed effects for district, and year by state. Robust standard errors in parentheses, with clustering by district. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Effects on firm logistics

	Input inventory (1)	Output inventory (2)	Transport expenses (3)	Transport equipment (4)
Durontos	0.0173*** (0.0052)	0.0023 (0.0063)	-0.0081 (0.0124)	0.0012 (0.0164)
Spillovers	0.0015 (0.0066)	0.0004 (0.0074)	-0.0058 (0.0161)	-0.0119 (0.0199)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), for factories in rail-using industries. The dependent variables, all in natural logarithms, provide measures of the factory’s logistical response to railway congestion. Column (1) shows effects on holdings of input inventory, while Column (2) shows effects on holdings of finished goods inventory. Column (3) shows effects on the firm’s “other distributional expenses”, the expense category into which shipping expense falls. Column (4) shows effects on the amount of transport equipment owned. The regressors are the number of two-way Duronto routes (based on shortest path) passing through the district as of the current year, and the number of introduced Duronto trains for which the district lies on a spillover alternate route, as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Effects on firm product mix

	Number of products made (1)	Time sensitivity (2)	Demand uncertainty (3)	Product complexity (4)
Durontos	-0.0114*** (0.0030)	-0.0022* (0.0012)	-0.0119*** (0.0042)	-0.0057* (0.0029)
Spillovers	-0.0049 (0.0042)	-0.0041** (0.0018)	-0.0055 (0.0046)	-0.0041 (0.0038)
Observations	22704	22615	23138	22614
Clusters 1 (factories)	5491	5241	5313	5241
Clusters 2 (district \times year)	1825	1876	1885	1876

Notes: This table presents estimates of equation (1), for factories in rail-using industries. The dependent variables, all in natural logarithms, provide measures of the factory's product mix. Column (1) shows effects on the number of distinct products produced. Column (2) shows effects on the average time sensitivity of the products made, weighed across product by output value. As described in the text, product level measures of time sensitivity come from [Hummels and Schaur \(2013\)](#). Column (3) shows effects on the average demand uncertainty of the products made, again weighted by product output value. The measure of demand uncertainty is as described in the text and is as used in [Blanchard and Simon \(2001\)](#). Column (4) shows effects on product complexity, measured as in [Levchenko \(2007\)](#) as the (inverse) Herfindahl index of the inputs used to make the product according to US input-output tables. The regressors are the number of two-way Duronto routes (based on shortest path) passing through the district as of the current year, and the number of introduced Duronto trains for which the district lies on a spillover alternate route, as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Other heterogeneity

Table B1: Heterogeneity by use of rail goods as inputs, and production of rail goods as output

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Durontos	-0.0040 (0.0067)	-0.0021 (0.0044)	-0.0029 (0.0078)	0.0017 (0.0066)
(Durontos) × (Rail input share)	-0.0351* (0.0204)	-0.0287* (0.0165)	0.0258 (0.0191)	0.0291** (0.0145)
(Durontos) × (Makes rail good)	-0.0173** (0.0077)	-0.0124** (0.0056)	0.0102 (0.0092)	-0.0011 (0.0084)
Spillovers	-0.0099 (0.0089)	-0.0050 (0.0055)	0.0046 (0.0106)	-0.0008 (0.0097)
(Spillovers) × (Rail input share)	-0.0038 (0.0243)	-0.0026 (0.0221)	0.0061 (0.0210)	0.0025 (0.0289)
(Spillovers) × (Makes rail good)	-0.0052 (0.0102)	-0.0030 (0.0079)	0.0184* (0.0107)	0.0018 (0.0112)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district × year)	1932	1914	1866	1928

Notes: This table shows heterogeneity in the Duronto and spillover effects, based on whether the factory is likely to rely on the rails for sourcing inputs or for delivering output. The rail input share is the industry's total input cost share of the goods typically shipped by rail in India (coal, iron, steel, cement, fertilizers, foodgrains, and mineral oils). The mean of this input share is 0.27 in the regression sample. Makes rail good is an indicator of whether the factory produces one of the rail goods. Its mean in the regression sample is 0.53. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Heterogeneity by road density

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Durontos	-0.0188*** (0.0060)	-0.0113** (0.0047)	-0.0082** (0.0041)	0.0094 (0.0061)
(Durontos) \times (Road density)	0.0031 (0.0075)	0.0048 (0.0060)	-0.0046 (0.0088)	0.0025 (0.0090)
Spillovers	-0.0117* (0.0061)	-0.0079* (0.0043)	0.0076* (0.0043)	0.0009 (0.0062)
(Spillovers) \times (Road density)	-0.0023 (0.0068)	-0.0073 (0.0047)	0.0059 (0.0079)	-0.0011 (0.0083)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table shows heterogeneity in the Duronto and spillover effects, as a function of state road density, measured as kilometers of national highway per square kilometer of area. This road density variable is standardized so it has mean 0 and standard deviation 1. Its raw mean is 0.036 and raw standard deviation is 0.015. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Robustness

Table C1: Reduced form estimates, controlling for distance to cities served by Durosantos

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0181*** (0.0057)	-0.0102** (0.0043)	0.0078** (0.0036)	0.0098* (0.0054)
Spillover routes through district	-0.0109* (0.0065)	-0.0051 (0.0044)	0.0074* (0.0043)	0.0017 (0.0067)
Distance to Duronto endpoint (hundred km.)	-0.0012 (0.0021)	-0.0025 (0.0018)	0.0036 (0.0029)	-0.0002 (0.0028)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), adding a control for the distance to the nearest city with a new Duronto train serving it in that year, measured in hundreds of kilometers. The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C2: Reduced form estimates, controlling for Duronto and spillover traffic on shipping lines

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0186*** (0.0061)	-0.0133*** (0.0046)	0.0071* (0.0039)	0.0014 (0.0057)
Spillover routes through district	-0.0124* (0.0071)	-0.0055 (0.0044)	0.0079* (0.0042)	0.0025 (0.0069)
Duronto traffic on shipping lines	-0.2422 (0.5750)	-0.3206 (0.4421)	0.0094 (0.3892)	0.8749* (0.4855)
Spillover traffic on shipping lines	-0.4519 (0.7041)	-0.2110 (0.5698)	-0.1002 (0.5224)	-0.0439 (0.6200)
Observations	21582	21034	19770	20993
Clusters 1 (factories)	3007	2986	2789	2994
Clusters 2 (district \times year)	972	960	908	964

Notes: This table presents estimates of equation (1), adding a control for the amount of Duronto and spillover traffic introduced along the tracks used for railway shipments to and from each district. This traffic is measured as a fraction of the tracks' line capacity, and averaged over all the routes used for the district's shipments, weighted by the total number of shipments, between 2011 and 2015 (which is all years with available data). The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C3: Reduced form estimates, controlling for changes in market access

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0178*** (0.0054)	-0.0123*** (0.0039)	0.0074** (0.0033)	0.0025 (0.0053)
Spillover routes through district	-0.0119* (0.0064)	-0.0056 (0.0041)	0.0076* (0.0039)	0.0033 (0.0065)
ln(Market access)	-0.0834 (0.1588)	-0.1036 (0.1624)	0.0036 (0.1622)	-0.1589 (0.1667)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), adding a control for the amount of Duronto and spillover traffic introduced along the tracks used for railway shipments to and from each district. This traffic is measured as a fraction of the tracks' line capacity, and averaged over all the routes used for the district's shipments, weighted by the total number of shipments, between 2011 and 2015 (which is all years with available data). The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C4: Reduced form estimates, with sample including all districts in mainland India

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0193*** (0.0045)	-0.0103*** (0.0037)	0.0074** (0.0029)	0.0099** (0.0048)
Spillover routes through district	-0.0103* (0.0057)	-0.0058 (0.0038)	0.0074* (0.0038)	0.0006 (0.0058)
Observations	37651	36708	30412	20993
Clusters 1 (factories)	8311	8150	7179	8012
Clusters 2 (district \times year)	2855	2841	2772	2843

Notes: This table presents estimates of equation (1), where the sample includes all districts in mainland India as possible controls, not only the districts located between major cities as in the preferred specification. The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C5: Reduced form estimates, with sample excluding “donut” around Duronto endpoints

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0189*** (0.0053)	-0.0106** (0.0044)	0.0081** (0.0035)	0.0096* (0.0055)
Spillover routes through district	-0.0110* (0.0066)	-0.0062 (0.0045)	0.0072 (0.0045)	0.0009 (0.0069)
Observations	25874	25488	19760	25791
Clusters 1 (factories)	5740	5723	5286	5731
Clusters 2 (district \times year)	1791	1780	1683	1775

Notes: This table presents estimates of equation (1), where the sample excludes all districts within 100km of the districts receiving Duronto passenger train service. The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C6: Reduced form estimates, with narrower definition of spillover route

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0194*** (0.0052)	-0.0109** (0.0047)	0.0080*** (0.0029)	0.0097 (0.0063)
Spillover routes through district	-0.0109* (0.0064)	-0.0065 (0.0041)	0.0073* (0.0044)	0.0034 (0.0059)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), with spillover routes defined to include only those paths crossed by trains traveling between the Duronto endpoints, instead of between any pair of points on the Duronto route. This traffic is measured as a fraction of the tracks' line capacity, and averaged over all the routes used for the district's shipments, weighted by the total number of shipments, between 2011 and 2015 (which is all years with available data). The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C7: Reduced form estimates, with wider definition of spillovers, including second-order

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0179*** (0.0046)	-0.0116*** (0.0042)	0.0074** (0.0030)	0.0091* (0.0055)
Spillover routes through district	-0.0103 (0.0069)	-0.0063 (0.0043)	0.0068 (0.0043)	0.0006 (0.0067)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), with spillover routes defined to include not only the standard spillover routes as described in the text, but also the second-order spillover routes. This traffic is measured as a fraction of the tracks' line capacity, and averaged over all the routes used for the district's shipments, weighted by the total number of shipments, between 2011 and 2015 (which is all years with available data). The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C8: Reduced form estimates, with spillovers restricted to 200km

	ln(Revenue) (1)	ln(TFPR) (2)	ln(Avg cost) (3)	ln(Inventory) (4)
Duronto routes through district	-0.0191*** (0.0055)	-0.0101** (0.0046)	0.0076** (0.0033)	0.0094 (0.0058)
Spillover routes through district	-0.0099 (0.0062)	-0.0059 (0.0046)	0.0068 (0.0045)	0.0005 (0.0058)
Observations	27558	26896	21624	26618
Clusters 1 (factories)	6191	6074	5238	5964
Clusters 2 (district \times year)	1932	1914	1866	1928

Notes: This table presents estimates of equation (1), with spillover routes restricted to include only lines within 200km of the main Duronto route. This traffic is measured as a fraction of the tracks' line capacity, and averaged over all the routes used for the district's shipments, weighted by the total number of shipments, between 2011 and 2015 (which is all years with available data). The dependent variables are the four main outcomes of interest as defined in the text. All regressions include fixed effects for factory, year by state, and year by NIC industry. Robust standard errors in parentheses, with clustering by factory and district-year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Alternate measures of exposure to Duronto congestion

D.1 Model-based measure of changes in market access

Another possible channel for the effects of Duronto traffic is through changes in a district’s “market access” (Redding and Sturm, 2008; Donaldson and Hornbeck, 2016), or the cost, net of congestion, of sending and receiving shipments to and from other districts with large product or input markets. Returning to the illustration of Figure 5, even if Bokaro and all of its shipping lanes were untouched by the Durontos or by the traffic pushed onto alternate routes, it might still experience gains or losses as a result of the Durontos disrupting production in other areas and prices adjusting accordingly. For example, suppose Rourkela and Bokaro compete to supply goods to some third district, Ranchi. If the Duronto train affects production in Rourkela or disrupts Rourkela-Ranchi shipments, Bokaro firms gain an advantage supplying to Ranchi and might thus enjoy increased sales. On the other hand, these effects might harm Bokaro rather than benefit it. For example, if certain Rourkela firms supply inputs to Bokaro, or to Bokaro firms’ suppliers in towns like Ranchi, then the Durontos could raise the price of these inputs.

To account for these effects, I use a measure of market access in the spirit of Donaldson and Hornbeck (2016). Based on Eaton and Kortum (2002), the Donaldson and Hornbeck (2016) model shows that all of the positive and negative general equilibrium welfare effects of an infrastructure project are captured in statistic termed market access. Under symmetric trade costs, the market access of a district d is

$$MA_d = k \sum_{d'} \tau_{dd'}^{-\theta} MA_{d'}^{-1} Y_{d'},$$

where k is a constant, d' indexes the other districts, $\tau_{dd'}$ is a bilateral trade cost, θ is a trade elasticity, and $Y_{d'}$ is the real output of d' . Intuitively, a district o has greater market access when it has lower costs of trading with other districts which have large local economies but relatively little access to other districts which might compete with o for business. Donaldson and Hornbeck (2016) show that this market access term is highly correlated with a first-order approximation which is better suited to empirical applications. Through a similar derivation, I arrive at the approximation

$$MA_d \approx \sum_{d'} \tau_{dd'}^{-\theta} Y_{d'}. \quad (36)$$

I parametrize $\tau_{dd'}$ by

$$\tau_{dd'} \equiv (Traf)_{dd'}^{\psi_T} Dist_{dd'}^{\psi_D}, \quad (37)$$

where $(Traf)_{dd'} \equiv (\text{Traffic}_{dd',2008} + \text{Durontos}_{dd'}) / \text{Traffic}_{dd',2008}$ is the amount of Duronto traffic on the railway lines connecting d to d' , expressed as a fraction of 2008 traffic, and $(Dist)_{dd'}$ is the distance from d to d' . The elasticity of trade costs with respect to distance is ψ_D . Following Ramondo, Rodríguez-Clare and Saborío-Rodríguez (2016), I set $\psi_D = 0.27$ and $\theta = 4$. Finally, I set $\psi_T = 0.02$, so the effect of congestion on trade costs is assumed, conservatively, to be less than one-tenth that of distance. These parameter choices are necessarily somewhat ad hoc, representing a limitation to using the market access approach in my setting.

With richer data on the actual shipments firms make or on spatial price gaps, it would be possible to find trade cost equivalents of given amounts of congestion to make these calculations more precise.

Controlling for $\ln MA_{dt}$ captures the effect of Duronto-associated changes in market access, which are greatest for districts which have Durontos running on the lines between them and other districts which are large or growing.²⁸ Table C3 shows results of factory regressions with this market access term, in addition to the main Duronto and spillover effects. This in fact does not change the estimates of the Duronto and spillover effects, suggesting that these variables suffice to capture most of the economic effect of the added traffic, without having to turn to models involving notions like market access.

D.2 Traffic in shipping lanes

Another form of exposure occurs if Duronto or spillover traffic, though not traveling through a district, nevertheless travels on the railway lines used to carry goods to or from that district. To account for these effects, I use data on freight shipment patterns. For each district, I identify the set of routes used for its shipments, and calculate the amount of Duronto traffic on these routes. First, for each section of track n , let

$$D_{nt} \equiv \frac{\#\text{Durontos}_{nt}}{\text{Capacity}_{n,t=2008}}$$

be the number of Durontos running on that section as of year t , expressed as a fraction of the section's 2008 capacity. This section might be part of one or more origin-destination shipping routes, r . A measure of Duronto congestion on r is

$$\text{Route Duronto Traffic}_{rt} \equiv \sum_n \rho_{nr} D_{nt}$$

with weights ρ_{nr} equal to the length of section n divided by the total length of route r . So $(\text{Route Duronto Traffic})_{rt}$ is a weighted average of congestion on all of the sections n which make up route r . Finally, for a district-level measure of Durontos' effect on shipments, I take a weighted average of route congestion over all of the routes serving the district:

$$\text{District's Duronto Shipping Exposure}_{dt} \equiv \sum_r \omega_{rd} (\text{Route Duronto Exposure})_{rt},$$

with weights ω_{rd} equal to the total number of route r shipments serving district d , divided by the total number of shipments serving d . A limitation of these measures is that data on freight shipping patterns is available for only a limited number of districts. Nevertheless, Table C2 presents reduced form results incorporating these controls for the available district, and showing again that this does not produce an additional effect conditional on the amount of Duronto and spillover traffic passing directly through the district.

²⁸An identification concern arises because MA_{dt} and $Y_{d't}$ are endogenously determined. However, the results do not change substantially from following the alternate approach of [Donaldson and Hornbeck \(2016\)](#) which holds district output fixed at 2008 levels.

E Inventory model

Consider a firm solving a classic economic-order quantity (EOQ) problem: the firm's inventories deplete as it meets customer demand, then when the inventory reaches some reorder level R , the firm places a restocking order of quantity Q . Demand d at any instant is random, with cdf F and mean μ_d , variance σ_d^2 . Time t is measured in years (so μ_d is the average annual demand). Expected inventory level varies between s and $s + Q$, where safety stock s is defined as the expected inventory level just before a restocking order arrives.

$$s = R - \mu_d \mu_\tau,$$

where μ_τ is expected shipping time.

The economically relevant portion of the problem is time between the placement of the restocking order and the arrival of the order, since this is where the stockout risk arises. To focus on this portion of the problem and to simplify, we fix Q . A version of the problem without this simplification appears in [Nahmias \(2001\)](#). Note that Q will be approximately $R - s$, since the firm will not want its inventory level when orders arrive to be systematically greater than R , nor systematically less.

Stockout occurs if inventory on hand at the time of order (R) is less than total demand while the order is in transit, denoted D . Total demand D depends on both the demand realizations at each instant, and shipping time τ , which is random with cdf G , mean μ_τ , and variance σ_τ^2 . Denote the cdf of D by H .

Lemma. *Assume F, G are independent. Then,*

$$\begin{aligned}\mu_D &\equiv E[D] = \mu_d \mu_\tau \\ \sigma_D^2 &\equiv V[D] = \mu_\tau \sigma_d^2 + \mu_d^2 \sigma_\tau^2.\end{aligned}\tag{38}$$

Proof. See [Hadley and Whitin \(1963\)](#). □

When stockout occurs, the firm pays penalty p for each unit of unmet demand. The expected stockout penalty in one reorder cycle is therefore

$$n(R) = p \int_R^\infty (x - R)h(x)dx.$$

The number of annual cycles is μ_d/Q , so the expected annual penalty is

$$\frac{\mu_d}{Q}n(R).\tag{39}$$

Inventory holding costs depend on the interest rate i . Normalizing by the price of output and letting v be value added, the price of an input unit is $1 - v$, so the cost of holding it in inventory is $i(1 - v)$. Average inventory holdings are

$$\bar{I} = s + \frac{1}{2}Q = R - \mu_d \mu_\tau + \frac{1}{2}Q.$$

So expected annual holding costs are

$$i(1-v)(R - \mu_d\mu_\tau + \frac{1}{2}Q). \quad (40)$$

Combining expected penalty (39) with holding costs (40), total costs are

$$C(R) = \frac{\mu_d}{Q}n(R) + i(1-v)(R - \mu_d\mu_\tau + \frac{1}{2}Q). \quad (41)$$

The firm choose R to minimize (41), yielding first order condition

$$0 = -\frac{\mu_d}{Q}p(1 - H(R)) + i(1-v).$$

and optimum

$$R^* = \rho\sigma_D + \mu_d\mu_\tau, \quad (42)$$

where $\rho \equiv \Phi^{-1}(1 - \frac{i(1-v)}{p\mu_d}Q)$ is constant with respect to μ_τ and σ_τ .

As (42) makes clear, average inventory level is increasing in both the mean and variance of shipping time. Also, these effects of shipping times are larger for goods with higher value added v , higher penalty of stockout p , and higher demand uncertainty σ_d :

$$\begin{aligned} \frac{\partial^2 R^*}{\partial \mu_\tau \partial v} &> 0 & \frac{\partial^2 R^*}{\partial \sigma_\tau \partial v} &> 0 \\ \frac{\partial^2 R^*}{\partial \mu_\tau \partial p} &> 0 & \frac{\partial^2 R^*}{\partial \sigma_\tau \partial p} &> 0 \\ \frac{\partial^2 R^*}{\partial \mu_\tau \partial \sigma_d} &> 0 & \frac{\partial^2 R^*}{\partial \sigma_\tau \partial \sigma_d} &> 0. \end{aligned}$$

These results serve as the basis for the predictions of heterogeneous effects, as discussed in section 3.3.3.

F Shipping times model

Consider a single section of railway track, with a given set of K trains running eastbound and \bar{K} trains running westbound. Each train i has scheduled departure and arrival times D_i and A_i and a free-running time FR_i , which is the time the train would take to traverse the section without any interference from other trains.

Interference delays arise when trains meet, because one of the trains needs to stop to let the other pass. Let q_{ij} be the probability that trains i and j meet; naturally, $q_{ij} = q_{ji}$. This probability is not exogenously specified, but will need to be solved for, given the departure times and stoppages on the line. Conditional on i and j meeting, let P_{ij} be the probability that train i is delayed, and assume $P_{ij} = 1 - P_{ji}$. Conditional on experiencing such a delay, train i needs to stop for d_{ij} minutes. Thus, the total delay, t_{ij} equals d_{ij} if i and j meet, and zero otherwise. It follows that the total travel time for train i is

$$t_i = FR_i + \sum_{j \in K \cup \bar{K}} t_{ij}. \quad (43)$$

There are two possible sources of randomness in this expression for travel time. First, allow the actual departure time d_i to be a random variable centered around D_i . As in the real world, a given train might leave later than expected, causing disruption to the schedules of other trains. This randomness in departure time is not essential, however, as there is also another source of uncertainty, coming from the random length of delay conditional on two trains meeting. In particular, d_{ij} is a random variable, which can depend, as in the formulation of [Petersen \(1974\)](#) on the length of the track section, the availability of sidings for trains to pull aside, the amount of time taken for trains to make this switch, and the speed differential between meeting trains. There are different ways to specify the distribution of d_{ij} as a function of these factors ([Petersen, 1974](#); [Chen and Harker, 1990](#); [Harker and Hong, 1990](#)), though the exact specification is not important for the derivation that follows.

Taking the expectation and variance on both sides of (43), it follows that

$$E(t_i) = FR_i + \sum_{j \in K \cup \bar{K}} q_{ij} E(d_{ij}) \quad (44)$$

$$Var(t_i) = \sum_{j \in K \cup \bar{K}} [q_{ij} Var(d_{ij}) + q_{ij}(1 - q_{ij}) E^2(d_{ij})] + \sum_{h, k \in K \cup \bar{K}, h \neq k} Cov(t_{ih}, t_{ik}). \quad (45)$$

The derivation of these expressions, related to the derivation in [Chen and Harker \(1990\)](#), requires solving for the mean, variance, and covariance of t_{ij} . Since $t_{ij} = d_{ij}$ when i and j meet and zero otherwise, it follows that $E(t_{ij}) = q_{ij} E(d_{ij})$ and $E(t_{ij}^2) = q_{ij} E(d_{ij}^2)$. The expression for the mean in (44) follows immediately. For the variance,

$$Var(t_{ij}) = q_{ij} Var(d_{ij}) + q_{ij}(1 - q_{ij}) E^2(d_{ij}), \quad (46)$$

which yields first term in (45). The calculation of $Var(t_i)$ also requires solving for the covariance of the t_{ij} . These covariances of train meetings are the reason that travel time variance increases by so much with

the addition of trains to an already-congested track. In particular, on a congested track, train i meeting with train h makes it more likely that train i will also meet train k , since train i will have been stopped and thrown off schedule by the first meeting. To help see this, we solve for these covariances, which can be written as

$$Cov(t_{ih}, t_{ik}) = E(t_{ih}t_{ik}) - E(t_{ih})E(t_{ik}). \quad (47)$$

Expanding this expression requires computing the probability of train i interfering with both of trains h and k . Doing so requires, in general, considering three possible cases: (a) i is running in one direction, with h and k in the opposite direction, (b) i and k are running in one direction, with h in the opposite direction, and (c) all three of i , h , and k are running in the same direction.

With uncertainty over departure times, the covariance term for each of these cases depends on the distribution of departure times. Let $g_i(\cdot)$, $g_h(\cdot)$, and $g_k(\cdot)$ be the departure time density functions and letting $g_{h-i,k-i}(\cdot)$ be the joint distribution of differences in departure time. It follows, in case (a), that the probability of i interfering with both h and k is the probability that it meets both of them head-on, which is

$$\int_{D_k-\tau}^{D_k+\tau} \int_{D_h-\tau}^{D_h+\tau} \int_{D_i-\tau}^{D_i+\tau} g_i(x)g_h(y)g_k(z)f_1(x,y,z) dx dy dz. \quad (48)$$

Here τ is the cycle window within which all trains depart—perhaps one day if we are considering a daily schedule. The probability of i meeting both other trains given the departure times is $f_1(x,y,z)$.

In case (b), the probability of i meeting both other trains is the probability that i overtakes h and meets k head-on, plus the probability that it is overtaken by h and meets k head-on. This yields

$$\begin{aligned} & \int_0^{D_h-D_i+\tau} \int_{D_k-D_i-\tau}^{D_k-D_i-\tau} g_{h-i,k-i}(y,z)[f_2(y,z) + f_3(y,z)] dz dy \\ & + \int_{D_h-D_i-\tau}^0 \int_{D_k-D_i-\tau}^{D_k-D_i-\tau} g_{h-i,k-i}(y,z)[f_4(y,z) + f_5(y,z)] dz dy. \end{aligned} \quad (49)$$

Here, f_2 , f_3 , f_4 , and f_5 give the relevant probabilities of meetings and overtakings given the departure times.

In case (c) the probability of i meeting both other trains is given by the probability that it overtakes both of them, plus the probability both of them overtake it, plus the probability that it overtakes one and is overtaken by the other. In particular, this probability can be written

$$\begin{aligned} & \int_0^{D_h-D_i+\tau} \int_0^{D_k-D_i+\tau} g_{h-i,k-i}(y,z)f_6(y,z) dz dy \\ & + \int_0^{D_h-D_i+\tau} \int_{D_k-D_i-\tau}^0 g_{h-i,k-i}(y,z)f_7(y,z) dz dy \\ & + \int_{D_h-D_i-\tau}^0 \int_0^{D_k-D_i+\tau} g_{h-i,k-i}(y,z)f_8(y,z) dz dy \\ & + \int_{D_h-D_i-\tau}^0 \int_{D_k-D_i-\tau}^0 g_{h-i,k-i}(y,z)f_9(y,z) dz dy. \end{aligned} \quad (50)$$

Here again, $f_6, f_7, f_8,$ and f_9 give the probabilities of the meetings and overtakings as a function of realized departure times. Explicit calculations of f_1 to f_9 appear in [Chen and Harker \(1990\)](#) and [Harker and Hong \(1990\)](#), yielding the expression for the covariance term.

A simpler calculation follows from assuming no uncertainty in departure times, and focusing on the first two cases for train meetings.²⁹ To simplify notation, write the difference in train departure times as

$$D_i(j) = \begin{cases} D_j - D_i & \text{if } |D_j - D_i| \leq \tau/2 \\ D_j - D_i - \tau & \text{if } D_j - D_i > \tau/2 \\ D_j - D_i + \tau & \text{if } D_j - D_i < -\tau/2. \end{cases} \quad (51)$$

Note that a meeting between trains occurs if $t_i > |D_i(j)|$. Thus, $q_{ij} = P(t_i \geq D_i(j))$ if $D_i(j) \geq 0$ and $q_{ij} = P(t_i \geq -D_i(j))$ if $D_i(j) < 0$. Taking (47), and plugging in the probabilities from cases (a) and (b) along with these expressions for q_{ij} , we obtain an expression for the expected travel time of train i

$$E(t_i) = FR_i + \sum_{j:D_i(j)<0} P(t_i \geq -D_i(j))E(d_{ij}) + \sum_{j:D_i(j)\geq 0} P(t_i \geq D_i(j))E(d_{ij}). \quad (52)$$

Similarly, we obtain an expression for the variance:

$$\begin{aligned} Var(t_i) = & \sum_{j:D_i(j)<0} P(t_i \geq -D_i(j))\{Var(d_{ij}) + [1 - P(t_i \geq -D_i(j))]E^2(d_{ij})\} \\ & + \sum_{j:D_i(j)\geq 0} \left\{ P(t_i \geq D_i(j))Var(d_{ij}) + \{P(t_i \geq D_i(j)) \cdot A + [1 - P(t_i \geq D_i(j))] \cdot B\}E(d_{ij}) \right\}, \end{aligned} \quad (53)$$

where

$$\begin{aligned} A &\equiv \sum_{k:D_i(j)\geq D_i(k)\geq 0} [1 - P(t_i \geq D_i(k))]E(d_{ik}) \\ B &\equiv \sum_{k:D_i(k)\geq D_i(j)} P(t_i \geq D_i(k))E(d_{ik}). \end{aligned} \quad (54)$$

The equations (52) and (53) hold for a generic distribution of travel times. To simplify the calculation, assume as in [Chen and Harker \(1990\)](#) and [Harker and Hong \(1990\)](#) that t_i is normally distributed with mean T_i and variance V_i . A normal distribution of travel times should be a realistic approximation on a congested line where the number of train interferences is large, though this is not a precise application of the central limit theorem, since it is not the case that all variables here are independent. It is now straightforward to obtain the expression

$$P(t_i \geq D_i(j)) = \frac{1}{\sqrt{2\pi}} \int_{D_i(j)}^{\infty} \frac{1}{\sqrt{2\pi V_i}} e^{-(t-T_i)^2/2V_i} dt. \quad (55)$$

Substituting based on (55) in (52) and (53) and given expressions for $E(d_{ij})$ and $Var(d_{ij})$, we obtain a system of $2K$ nonlinear equations in $2K$ unknowns $\{T_1, \dots, T_K, V_1, \dots, V_K\}$.

²⁹This restriction does not substantively change the results, and [Chen and Harker \(1990\)](#) provide an adjustment which accounts for this additional case after finding the solution based on only the first two cases.

Algorithms for solving these equations, such as the Newton-Raphson method, require calculating a Jacobian matrix and applying a series of iterative approximations until convergence. Letting $f_i^T(\mathbf{T}, \mathbf{V})$ and $f_i^V(\mathbf{T}, \mathbf{V})$ be the equations for the mean and variance of travel time i , it is straightforward to the terms of the Jacobian, $\frac{\partial f_i^T(\mathbf{T}, \mathbf{V})}{\partial T_j}$, $\frac{\partial f_i^V(\mathbf{T}, \mathbf{V})}{\partial T_j}$, $\frac{\partial f_i^T(\mathbf{T}, \mathbf{V})}{\partial V_j}$, and $\frac{\partial f_i^V(\mathbf{T}, \mathbf{V})}{\partial V_j}$, for each i and j , by plugging in the derivatives

$$\begin{aligned} \frac{\partial P(t_i \geq D_i(j))}{\partial T_i} &= \frac{1}{\sqrt{2\pi V_i}} e^{-(D_i(j)-T_i)^2/2V_i} \\ \frac{\partial P(t_i \geq D_i(j))}{\partial V_i} &= \frac{D_i(j) - T_i}{2\sqrt{2\pi V_i^3}} e^{-(D_i(j)-T_i)^2/2V_i}. \end{aligned} \quad (56)$$

In solving these equations, I use a quasi-Newton method involving full calculation of the Jacobian only at the first step (Broyden, 1965), which achieves the same results and faster convergence than the Newton-Raphson and successive approximations methods used in Chen and Harker (1990) and Harker and Hong (1990).

Train i	Chen & Harker		Replication	
	$E(t_i)$	$SD(t_i)$	$E(t_i)$	$SD(t_i)$
1	2:33	0:05	2:33	0:05
2	2:35	0:05	2:35	0:05
3	1:58	0:03	1:58	0:03
4	2:18	0:00	2:18	0:00
5	2:22	0:00	2:22	0:00
6	2:31	0:05	2:28	0:05
7	2:52	0:09	2:50	0:09
8	2:39	0:05	2:38	0:04
9	2:47	0:08	2:47	0:08
10	3:18	0:14	3:14	0:13
11	2:55	0:06	2:56	0:06
12	3:00	0:08	2:59	0:08
13	2:20	0:04	2:21	0:04
14	2:27	0:07	2:26	0:07
15	2:12	0:03	2:18	0:03
16	2:19	0:04	2:19	0:04
17	3:33	0:13	3:30	0:12
18	2:31	0:07	2:31	0:07
19	2:01	0:01	2:01	0:02
20	3:11	0:09	3:10	0:09
21	3:11	0:09	3:04	0:09
22	2:19	0:05	2:15	0:05

Table F1: Replication of Chen and Harker (1990)

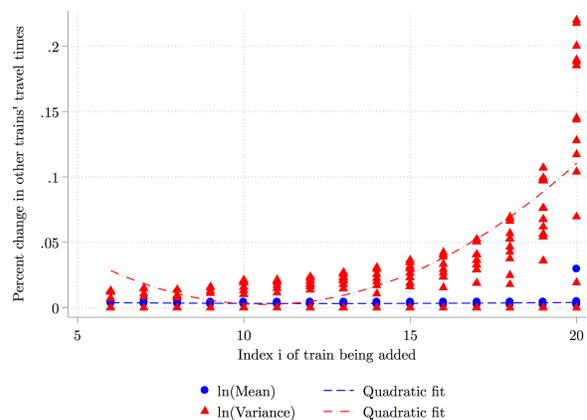


Figure F1: Effects of congestion on travel times

Following this procedure, I am able to nearly replicate the results of these papers. Specifically, Figure F1 shows the estimated mean and variance of travel time, in hours and minutes, for a line containing 22 trains studied by Chen and Harker (1990). The estimates in my replication show some minor differences from their original results, likely due to minor differences in the procedure for simplifying the meeting and overtaking probabilities. While I could not discern this exact procedure from the details provided in their paper, the

results do show a very close match, which suffices for practical purposes.

Figure F1 shows, then, the effect of adding trains to the line. Starting with the first five trains on the list, I add the succeeding trains one by one, and solve the system of equations after adding each new train, to see how this affects the travel times of the other trains. Consistent with the theoretical predictions and with the empirical strategy as described in Section 4, there is a divergence for mean and variance of travel time as the congestion level increases. Specifically, the effect of a new train on mean travel time is the same small amount regardless of whether the new train is the fifth or the twenty-second on the line. For the variance, on the other hand, there is little effect from adding the first few trains, but a far greater effect when traffic is heavy.³⁰ These results provide the basis for the empirical strategy using the congestion levels to instrument separately for the mean and the variance of shipping times.

³⁰While this escalation of the variance happens for almost all of the trains on the line, there a couple of trains (corresponding to the lowest sequence of red triangles in the graph), for which the additional traffic has little effect on the variance. These are the trains with highest priority on the rails, which do not need to stop and wait when they meet other trains. In reality, of course, the freight trains on Indian Railways, which receive the lowest priority will not exhibit this pattern. Rather, the additional traffic will increase the variance of their travel times in the manner discussed.