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

John Firth*
MIT
Job Market Paper
31 October 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. I cannot reject that these gains are as large as the affected firms’ losses, meaning congestion leads to a relocation of output, but not necessarily a net loss. 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 Dufo, 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, 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 costly after all. First, firms can try to insulate themselves from congestion, for instance by holding inventories to guard against stockout risk. Whether congestion proves costly 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 actual costs of congestion are 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 railway 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 of the paths taken by at least one passenger 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 substantial revenue losses 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 firms lose so much revenue from Duronto-associated congestion? I distinguish two basic explanations. One possibility is that Duronto congestion has a large “cost effect”: it 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. I show, consistent with the predictions of railway traffic models, that adding Duronto traffic through these districts increases both the mean and variance of freight shipment times. Firms dependent on the railways for shipments of heavy goods like coal, iron, and steel respond by holding larger inventories to guard against stockout risk, and simplifying their operations, for instance by making fewer products. With these movements away from efficient production, firms ultimately exhibit higher average costs and lower revenue productivity.

Alternatively, a second explanation 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.

To distinguish between the roles of costs and competition, I use a model of general equilibrium interactions between firms competing in a given industry. Drawing 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. Two empirical results, interpreted through the model, point to the importance of the competition mechanism. First, exposure of a firm’s competitors to Duronto traffic leads to increased sales for that firm, indicating substitution. Second, the negative effect of Durontos on firm revenue is concentrated in industries with high elasticities of substitution. In other words, any cost effects are magnified by stiff competition. Based on the parameter estimates associated with these results, I cannot reject the hypothesis that, due to competitors’ business stealing, there is no aggregate loss in firm production from running the Duronto trains.

With some firms gaining and other firms losing, we still have not identified the magnitude of the cost effect: does Duronto-associated congestion cause only a tiny cost increase for affected firms, who then lose ground to their competitors, or does it greatly disrupt their operations and increase costs? The evidence points to the latter – a genuine cost effect – which I am able to identify by looking at the response of the firm’s cost structure to Duronto congestion. Specifically, 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 costs 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 costs. So observed effects of congestion on average costs reflect the actual

---

2 In the presence of fixed costs, this movement might not reduce average costs, but would reduce average variable cost. The effects I find on average variable cost are similar to those on average cost.
outward shift in the firm’s cost function, offset by this downward force from competition. Empirically, I find that average cost increases when a firm is exposed to Duronto traffic, showing that the cost effect plays some role. The methodological appeal of this approach is that it leverages firm data to isolate basic features of the firm problem which are invariant to any general equilibrium competition effect, making it possible to identify the cost effect without strong dependence on general equilibrium model assumptions.

To unpack the sources of the cost effect, I distinguish two channels through which congestion might disrupt firm production: 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 Durontos 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 Durontos cause a great disruption to 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

---

3Consistent with theoretical predictions, I find no such effect on a firm’s own average cost coming from competitors’ exposure to Duronto traffic.
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.

This empirical strategy relies on an exclusion restriction holding that Duronto traffic affects firm outcomes only through its effects on mean and variance of shipping time. In settings other than Indian Railways, transport congestion could affect firms through a variety of channels, including increased shipping prices, reduced availability of freight shipment, or disruptions to passenger travel. However, I present a series of institutional and empirical facts which show that in the Indian Railways setting, none of these factors contribute to the congestion effect. Rather, firms’ responses to congestion and the associated revenue losses arise solely because the Duronto passenger traffic makes freight shipments slower and less reliable.

Two stage least squares results show differing effects of the slowness and the unpredictability, with unpredictability proving the more costly. Consistent with a Newsvendor model of inventory management, in the tradition of Arrow, Harris and Marschak (1951), 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.

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 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), 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 center per kilometer in China, and far higher rates in developed countries, for instance, 12.6 cents per kilometer in Germany, and 19.0 cents per kilometer in Japan.4

4 Adjusted 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.
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 passengers and freight.

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 Durontos, introduced by Rail Minister Mamata Banerjee in 2009 (Banerjee, 2009; Ministry of Railways, 2009). The Duronto 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 Duronto trains were based on passenger demand for travel between these major cities. The intermediate districts on the Duronto 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, 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 how many passenger, freight, and other trains run there in an average 24 hour period, what types of signaling, electrification, and other physical capabilities are present on that section, and what is 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. 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 India Rail Info (IRI) website, a sample of which appears in Figure 3. This data lists the stations passed by each train, even for 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.

---

5There is, likewise, considerable variation across shipments for a given route, as indicated by the annual route-wise data on variance of shipping time.
2.2.2 ASI

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, 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 presents additional descriptive statistics from both the ASI data, and the data on Railway traffic patterns.

3 Reduced form effects of congestion on firm revenue

This section identifies reduced form effects of Duronto passenger traffic on rail-using firms. I argue that for certain intermediate districts, having a Duronto run through the district is as good as random. This argument rests both on institutional facts about the Duronto 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 Durontos 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 Duronto 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 Duronto passenger service. Rourkela happens to lie on the shortest rail path between Mumbai and Kolkata, so the Mumbai-Kolkata Duronto passes through. Bokaro lies on a similarly important rail line which is part of the shortest path between Ahmedabad and Kolkata. Other Duronto trains serve Ahmedabad and others serve Kolkata, but the specific Ahmedabad-Kolkata route does not receive Duronto service, so no Duronto passes through Bokaro. The fact that a Duronto 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 Durontos, changes in firm outcomes in Rourkela would have been the same as those in Bokaro.

Two institutional details provide further support for this assumption. First, Durontos make no intermediate stops between their endpoints. This eliminates any possibility that the Duronto routes were chosen to serve or not serve the passengers of places such as Rourkela and Bokaro. It also entails that the only effect of Duronto traffic on Rourkela’s local economy is a clogging of the railway lines, interfering with shipment of freight. A remaining concern is that planners might have chosen Durontos’ 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 Durontos 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 Durontos’ shortest-path rule ensures that Duronto routes are not chosen based on these characteristics of the intermediate areas. A final remaining concern is endogenous choice of the entire Duronto 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 Duronto traffic.

The comparison between these districts motivates the basic specification

$$y_{it} = \beta_D D_{dt} + \beta S S_{dt} + \gamma_{t} + \gamma_{t \times s} + \gamma_{t \times k} + \epsilon_{it},$$

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 $t$. 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, 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 $B$ lying on the Mumbai-Kolkata line, and certain non-Duronto trains traveling from $A$ to $B$ 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 \( s \). Specifically, it studies how traffic on the section, Traffic\(_{st}\), measured as the average daily number of trains of a given type, responds to Duronto and alternate-route spillover traffic running on the section:

\[
\text{Traffic}_{st} = \alpha_D D_{st} + \alpha_S S_{st} + \gamma_s + \gamma_{t, \text{Sample}} + \epsilon_{st}. \tag{2}
\]

Here, \( D_{st} \) is the number of Duronto trains running on section \( s \) as of year \( t \) and \( S_{st} \) 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.\(^6\) 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.\(^7\)

### 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 with spillover controls.

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 \( \{1, \ldots, K\} \), the average cost measure is

\[
AC_i = \frac{\text{Total Cost}}{\sum_{k=1}^{K} p_k q_{ik}}, \tag{3}
\]

\( ^6 \)For each kilometer of Duronto route, there is 1.02 km 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.

\( ^7 \)The 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.
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.\(^8\)

Finally, inventory serves a measure of the response firms take to insulate themselves from the costs of congestion. Of course, firms take many insulating measures apart from inventories, and Table A2 details 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 D 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. Figure 8 shows parallel trends across affected and unaffected districts, both in terms of congestion levels, and in terms of firm revenue. This evidence lends empirical support to the core identifying assumption.

### 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.\(^9\) This revenue effect is large. For perspective, one Duronto route amounts to approximately 7 percent of the charted line capacity in the median district. Thus, scaling the revenue effect implies that if a district went from a completely clear railway line,

---

\(^8\)Changes 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.

\(^9\)As 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.
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 its cost per unit of output, independent of this price. As Column (3) shows, each Duronto route increases average cost by 0.8 percent. As Section 4.1.2 will elaborate, 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.

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.
4 Explaining the revenue loss: costs versus competition

To determine the reason for the revenue loss, 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 their 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 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.

4.1 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.

4.1.1 Isolating cost effects in the presence of competition

As in Rotemberg (2017), the economy has $K$ sectors, and a consumer with income $I$ has utility

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

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

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

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

$$Q_k = \left( \sum_{j=1}^{N} a_{jk} S_{jk} q_{jk}^{\gamma_k^x} \right)^{1/\gamma_k^x},$$

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:

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

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

\[ q_{jk} = A_{jk} K_{jk}^{\alpha_k} L_{jk}^{\alpha_L} R_{jk}^{\alpha_R} N_{jk}^{\alpha_N}, \]  

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 \( I \). Returns to scale are reflected by \( \gamma \equiv \sum_{I \in \{K,L,R,N\}} \alpha_I \). 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.\(^{10}\)

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

\[ \pi_{jk} = p_{jk} q_{jk} - \sum_{I} (1 + \tau_{I,j}) p_I I_j, \]  

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

\[ p_{jk} = \frac{\sigma_k}{\sigma_k - 1} \cdot \frac{\prod_{I} (p_I)^{\alpha_I} \cdot \prod_{I} (1 + \tau_{I,j})^{\alpha_I}}{A_{jk}}. \]

Firm revenue is

\[ y_{jk} = p_{jk} q_{jk} = (p_{jk}^{1-\sigma_k}) (P_k^{\sigma_k - 1}) \left( \frac{P_k}{\phi} \right)^{\frac{\sigma_k}{\sigma_k - 1}}. \]

\(^{10}\)Apart 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.
Allow firm productivity to grow according to

$$\tilde{A}_{jk} = -\epsilon_{jk},$$

(12)

where $\epsilon_{jk}$ is mean-zero and normally distributed. Then, combining (11) with (10), (7), and (12), changes in firm revenue are

$$\tilde{y}_{jk} = (1 - \sigma_k) \left( \sum_I \alpha_I (1 + \tau_{I,j}) + \epsilon_{jk} \right) + \left( \sigma_k - \frac{1}{1 - \phi} \right) \sum_{j' = 1}^{N_k} \left[ \left( \sum_I \alpha_I (1 + \tau_{I,j'}) + \epsilon_{j'k} \right) \frac{y_{j'k}}{Y_k} \right].$$

(13)

As I elaborate below, the first term in (13) 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 $\psi_I$ be the effect of one Duronto route on a firm’s input $I$ distortion:

$$(1 + \tau_{I,j}) = \psi_ID_j.$$  

(14)

We can now write a simplified version of equation (13):

$$\tilde{y}_{jk} = \beta \Psi D_j + \chi \Psi \mu_k + \tilde{\epsilon}_{jk},$$

(15)

where

$$\beta \equiv 1 - \sigma_k,$$

$$\Psi \equiv \sum_I \alpha_I \psi_I,$$

$$\chi \equiv \sigma_k - \frac{1}{1 - \phi},$$

$$\mu_k \equiv \frac{\sum_{j'} D_j y_{jk}}{Y_k},$$

$$\tilde{\epsilon}_{jk} \equiv (1 - \sigma_k)\epsilon_{jk} + (\sigma_k - \frac{1}{1 - \phi}) \sum_{j' = 1}^{N_k} \epsilon_{j'k} \frac{y_{j'k}}{Y_k}.$$

Here, $\beta$ reflects the direct effect of increasing distortions. This effect is largest if the elasticity of substitution $\sigma_k$ 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 $\beta$ captures the effect of the distortions themselves, $\Psi$ captures how these distortions respond to Duronto routes $D_j$ through the firm’s district. In particular, $\Psi$ is a sum of the Durontos’ effect, $\psi_I$, on each input distortion, weighted by each input’s cost share $\alpha_I$.

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 $\chi$. It is largest when the firm’s product is a ready substitute for
its competitors’ products (high $\sigma_k$), and when the sector as a whole is less replaceable by other sectors (low
\begin{equation}
\frac{1}{1-\phi}.
\end{equation}
The measure of sectoral exposure, $\mu_k$ is an output-weighted average of exposure to Duronto congestion for all the firms in the sector. Finally, the disturbance $\epsilon'_{jk}$ 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 (15) as
\begin{equation}
\bar{y}_{jk} = \Psi_1 D_j - \Psi_2 (\sigma_k \times D_j) + \chi \Psi \mu_k + \epsilon'_{jk}.
\end{equation}
Here, $\Psi_1$ reflects the cost effect for firms in low $\sigma$ industries, while $\Psi_2$ reflects that revenue losses become greater for firms in more competitive industries. Below, I use industry level estimates of $\sigma$ to estimate (16). Note that if the model is correct and $\sigma$ is measured perfectly, we should find $\Psi_1 = \Psi_2 = \Psi$.

The aggregate effect on sectoral output comes from summing across all firms in (13), yielding
\begin{equation}
\hat{Y}_k = (\beta \chi) \Psi \mu_k + \epsilon_k,
\end{equation}
where the disturbance $\epsilon_k \equiv \sum_{j=1}^{N_k} \bar{y}_{jk} \frac{\Psi (\Psi S)_{jk}}{k_{jk}}$ is a weighted average of the firm-level disturbances. Equation (17) 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 $\mu_k$. Second, this disrupts firm operations, leading to some total distortion $\Psi$, 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 $\beta$, are large relative to the ability of other firms in the sector to replace the lost output, as reflected by $\chi$.

4.1.2 Direct analysis of cost effects

The above analysis works to distinguish the effects of Durontos on firm production costs from associated competition effects as the firm loses business to its competitors. We can also infer these cost effects using the richness of available firm data. The idea is to examine features of the firm’s problem which are invariant to the competition effect. Below I present two complementary ways to identify these effects: the first, based on firm first-order conditions identifies distortions specific to each input, while the second, looking only at average costs, identifies a total distortion, allowing for a very general form to the firm’s production function.

\footnote{\textsuperscript{11}It 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 $j$, $\Psi^S$ the total cost effect of spillover routes, and $\mu^S$ the exposure of other firms in the sector, we obtain analogues of equations (15) and (17):
\begin{align*}
\bar{y}_{jk} &= \beta (\Psi D_j + \Psi^S S_j) + \chi (\Psi \mu_k + \Psi^S \mu^S_k) + \epsilon_{jk} \\
\hat{Y}_k &= (\beta + \chi) \Psi \mu_k + (\beta + \chi) \Psi^S \mu^S_k + \epsilon_k.
\end{align*}
Using first-order conditions. Examining the firm’s first-order conditions offers one approach to studying the cost effects, and specifically the wedges characterized above. Specifically, when the firm maximizes profits as in (9), the first-order conditions, for each $I \in \{K, L, R, N\}$, entail

$$1 + \tau_{I,j} = \frac{\sigma - 1}{\sigma} \frac{\alpha_I q_j}{p_I L_j}. \tag{18}$$

The firm variables on the right-hand side of (18) are all observed in data, so we can obtain empirical estimates $\hat{\psi}_I$ of the parameters in (14) with regressions of the form

$$\ln \frac{p_I q_j}{p_I L_j} = \psi_I D_j + \epsilon_j. \tag{19}$$

Using average costs. The Cobb-Douglas production function (8) entails that the firm’s average cost equals marginal cost:

$$AC_j = MC_j = \prod_I \left( \frac{p_I}{\alpha_I} \right)^{\alpha_I} \cdot \frac{\prod_I (1 + \tau_{I,j})^{\alpha_I}}{A_{jk}}. \tag{20}$$

It follows that

$$\hat{AC}_j = \Psi D_j + \hat{\epsilon}_{AC,j}, \tag{21}$$

with

$$\hat{\epsilon}_{AC} = \prod_I \left( \frac{p_I}{\alpha_I} \right)^{\alpha_I} \cdot \frac{1}{A_{jk}}. \tag{22}$$

Under the assumption that Duronto running is uncorrelated with changes in factor prices $p_I$ and the physical TFP $A_{jk}$ in the firm’s production function, regressions of the form (21) 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 ($\lim_{q \to 0^+} = C(0) = 0$).\footnote{Even for a production with fixed costs, a version of (25) holds, with average variable cost, rather than average cost, as the object of interest.} Suppose costs increase by shifting outward, so the new cost function is $\check{C}(q) = (1 + \check{\tau})C(q)$. How can we identify $(1 + \check{\tau})$?

Equating marginal revenue with marginal cost, firm optimization entails

$$\sigma + 1 = (1 + \check{\tau})C'(q). \tag{23}$$

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

$$\frac{\partial g}{\partial \check{\tau}} = -\frac{C'(q)}{C''(q)} \frac{1}{1 + \check{\tau}} < 0. \tag{24}$$
Finally, noting that $AC_j = (1 + \tau)C(q)$ and considering the effect of changing $\tau$, it follows that

$$AC_j = (1 + \tau) + \frac{\partial C(q)}{\partial q} \frac{\partial q}{\partial \tau} (1 + \tau).$$

So the effect of the cost shift $\tau$ on average costs is, first, a direct increase in costs, $(1 + \tau)$. But as the second term reflects, the cost shift also pushes the firm down its cost function ($\frac{\partial C(q)}{\partial q} < 0$), which with non-increasing returns has the effect of reducing average costs ($\frac{\partial C(q)}{\partial q} > 0$). Thus, since the second term in (25) is negative, observed changes in average cost $\Delta C_j$ are a lower bound on the cost shift $(1 + \tau)$.

### 4.2 Empirical application of the model

To identify the effect of competitors’ exposure to Duronto congestion, I estimate an empirical counterpart of (15):

$$y_{it} = a_1 D_{dt} + a_2 S_{dt} + a_3 \mu_{sk} + a_4 \nu_{sk} + \gamma_i + \gamma_t \times s + \gamma_t \times k + \epsilon_{it},$$

where $\mu_{sk}$ and $\nu_{sk}$ 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 5 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 $\mu_{D_{sk}}$ itself has a positive effect on a firm’s own revenue. In particular, if each of a firm’s competitors is exposed to on additional Duronto route, that firm gains 2.5 percent in revenue.

In the context of the model, the sum of these revenue coefficients, $\bar{a}_1 + \bar{a}_3$, provides an estimate of $(\beta + \chi)\Psi$, which indicates the aggregate effect of Duronto exposure on firm revenue. I cannot 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 5 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 (\ref{eq:cost}), this effect on average costs is interpretable as an estimate of the pure cost effect $\Psi$ under Cobb-Douglas production, and more generally as a lower bound on the shift in the cost function as illustrated in (\ref{eq:shift}). 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 6 shows support for the additional prediction of equation (\ref{eq:revenue}) 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 $\sigma$. 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. Still, this type of measurement error would attenuate estimates of $\hat{\Psi}_2$. So we should expect the true magnitude of $\Psi_2$ 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 (\ref{eq:aggregate}):

$$Y_{skt} = b_1 \mu_{skt} + b_2 \mu_{skt}^S + \gamma_{txs} + \gamma_{t\times k} + \epsilon_{skt},$$

(\ref{eq:aggregate2})

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 7 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 5 is $-0.006$, while the Duronto exposure effects in Table 7, which estimate the same parameter, range between $-0.002$ and $-0.014$.

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. So far, however, the analysis leaves open the source of this cost effect: what aspects of congestion disrupt firm production and increase costs?
5  Shipping times as the source of increased costs

This section examines why congestion increases production costs. I argue, in particular, that in this setting congestion effects work only by making freight shipment times slower and less predictable. I then draw on an operations research model of railway traffic to distinguish how the source of the cost effect owes to this slowness (mean shipping time) as opposed to the unpredictability (variance).

5.1 Model and empirical strategy

To link the production cost increase with the discussion of shipping times and other possible disruptions associated with congestion, define normalized mean $M_j = m_j/M$ and variance $V_j = v_j/V$, where $m_j$ and $v_j$ are the raw mean and variance of shipping time, in days, shipments in firm $j$’s district, and $M > 0$ and $V > 0$ are minimum feasible times such that $m_j \geq M$ and $v_j \geq V$.\(^\text{13}\) We parametrize

$$1 + \tilde{\tau} = M_j^{\delta_M} V_j^{\delta_V} \eta_j.$$  

(28)

Here, $\eta_j$ reflects all factors unrelated to shipping times. Note that if $\delta_M = \delta_V = 0$, then $1 + \tilde{\tau} = \eta_j$, which is to say that shipping time effects do not affect costs. Similarly, as $M_j$ and $V_j$ approach $M$ and $V$, respectively, $1 + \tilde{\tau}$ approaches $\eta_j$. I will use runnings of Duronto trains as instruments for mean and variance of shipping time, so the exclusion restriction will require that the runnings of these trains are uncorrelated with changes in $\eta_j$.

Taking logs in the expression for the cost shift, (28), we obtain

$$\ln (1 + \tilde{\tau}) = \delta_M \ln M_j + \delta_V \ln V_j + \ln \eta_j.$$  

(29)

To estimate a version of this equation, we will need suitable empirical variation in mean $\ln M_j$ and variance $\ln V_j$, which is the focus of the operations research portion of the model.

To separately identify the effects of mean and variance, 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

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

(30)

$$Var(t_i) = \sum_j [q_{ij} Var(d_{ij}) + q_{ij}(1 - q_{ij})E^2(d_{ij})] + \sum_{h,k} Cov(q_{ih}d_{ih}, q_{ik}d_{ik}),$$  

(31)

where $FR$ is free-running time. Solving for the expectation and variance of each $t_i$ requires numerical methods, and Appendix E elaborates on this solution, but equations (30) and (31) reveal a key prediction.

\(^{13}\)This normalization simply ensures the tax measures calculated below are bounded, and will prove inconsequential, given that we study logged changes in shipping times.
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 (31), 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.

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

$$
\ln M_{dt} = \pi_1^M D_{dt} + \pi_2^M (D_{dt} \times T_{d,t=t_0}) + \pi_3^M S_{dt} + \gamma_d + \gamma_y + \epsilon_d^M
$$

$$
\ln V_{dt} = \pi_1^V D_{dt} + \pi_2^V (D_{dt} \times T_{d,t=t_0}) + \pi_3^V S_{dt} + \gamma_d + \gamma_y + \epsilon_d^V.
$$

(32)

Here, $D_{dy}$ is the number of Durontos affecting district $d$, $S_{dy}$ is the spillover control, and $T_{d,y=0}$ 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. Some institutional details help rule this out. 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. Moreover, empirical evidence also supports the story told by these institutional details: as Table A2 shows, Duronto traffic has no effect on local passenger volumes, factories’ expenses on shipping and distribution, or the local number of freight trains run.

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 $\hat{\epsilon}$ 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.

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.

5.2 Results of the shipping times IV

Table 8 presents first-stage effects of the Duronto trains on the mean and variance of freight shipment times. Column (1) 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 9 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. Based on the discussion in Section 4.1.2, effects on average cost provide a lower bound on the pure cost effect of Duronto congestion. Column (3) reports these estimates, showing that a 10 percent increase in shipping time variance leads to a 0.3 percent increase in average cost, compared with almost no effect coming from mean shipping times. Finally, Column (4) indicates that both mean and variance contribute to increases in firm inventory holdings, consistent with the predictions of canonical inventory models of the type presented in Appendix D.
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 5. As reported in Column (1) of Table 10, 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.\footnote{I do not have precise data on fare revenues, but derive estimates by using the limited number of per-journey revenue amounts reported in Ministry of Railways (2015a), and multiplying by the annual number of journeys for each route.} Railway passenger services 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 10 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 4. 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 10 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 firms using the roads or other modes of transportation to ship 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, $\tau_{MD}$, is a function of the length of the shortest path between M and D.\(^{15}\) 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 $\tau_{MD}$. 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 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

---

\(^{15}\)The 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.
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

\[
\frac{d \ln W}{d \ln t_{ij}} = \sum_{k=1}^{N} \sum_{l=1}^{N} \frac{d \ln W}{d \ln \tau_{kl}} \times \frac{d \ln \tau_{kl}}{d \ln t_{ij}}. 
\]  

(33)

Here, \(W\) is aggregate welfare, \(\tau_{kl}\) 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 (33) shows, building new infrastructure between \(i\) and \(j\) lowers trade costs \(\tau_{kl}\) for each pair of trading cities whose routes pass through \(i\) and \(j\). In turn, these reductions in \(\tau_{kl}\) 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:

\[
\frac{d \ln W}{d \ln \tau_{kl}} = -\frac{X_{kl}}{Y W}, 
\]  

(34)

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.\textsuperscript{16} It follows that the A-R line does not affect trade costs for Y ($\Delta \ln \tau_{Y} = 0$ 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 ($f_{m} = X_{MD} + X_{DM} = X_{AR} + X_{RA}$), as is the total flow $f_{s}$ from each of the neighboring spillover cities to each of the endpoints ($f_{s} = X_{MX} + X_{XM} = X_{DX} + X_{XD} = X_{AY} + X_{YA} = X_{RY} + X_{YR}$). Fifth, let $\Delta d = 1 - \frac{d_{AR}}{d_{AY} + d_{YR}}$ 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.\textsuperscript{17}

Based on (33) and (34), the welfare effects of the two lines are

$$\Delta \ln W_{(M-D \ line)} = f_{m} \tilde{\tau}_{m} + 2f_{s}\tilde{\tau}_{s}$$

(35)

$$\Delta \ln W_{(A-R \ line)} = f_{m}\Delta d.$$  

(36)

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 (35) and (36), it follows that benefits of building in the congested M-D area are greater just in case

$$\Delta d < \tilde{\tau}_{m} + 2f_{s}\frac{\tilde{\tau}_{s}}{f_{m}}.$$  

(37)

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 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 (37). 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 $\tilde{\tau}_{m} = 0.8(\frac{100-50}{5}) = 5.7$ percent. Next, the estimates show that one Duronto route increases costs in the neighboring spillover areas

\textsuperscript{16}In 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.

\textsuperscript{17}The 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.
by 0.7 percent. If these lines experience similar decongestion effects, then $\tau_s = 0.7 \left( \frac{100 - 50}{7} \right) = 5.0$ percent.\(^{18}\)

The comparison in (37) 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 $f_{m1}$ be bilateral trade between Mumbai and Delhi and $f_{m2}$ be bilateral trade between Amaravati and Raipur, the comparison in (37) becomes

$$\Delta d \approx \frac{f_{m1}}{f_{m2}} \tau_m + 2 \frac{f_s}{f_{m2}} \tau_s.$$  \hspace{1cm} (38)

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 $f_{m1}/f_{m2}$), or if there is more activity in the spillover areas relative to that between Amaravati and Raipur (high $f_s/f_{m2}$).\(^{19}\) Based on rail shipment volumes along these routes and along the set of spillover routes, I obtain rough estimates of $f_{m1}/f_{m2} = 2.1$ and $f_{m1}/f_{m2} = 4.4$, implying that building the Mumbai-Delhi line achieves the same gains as shortening the Amaravati-Raipur distance by 56 percent.\(^{20}\) 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.

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.

\(^{18}\)This 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.

\(^{19}\)An 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.

\(^{20}\)In obtaining these figures, I take the Amaravati area to include the adjoining Krishna district, which contains the city of Vijayawada.
This analysis also leaves open some important 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


Figure 1: Histogram of line capacity utilization
Figure 2: Route-wise average freight shipment times

Route-wise run times, 2011

- 25 km/hr benchmark
- Actual run time (top-coded at 14 days)
- Fitted values
Figure 3: Sample from scraped website with data on train routes
Figure 4: Goods shipped by rail

(a) Composition of rail freight traffic

(b) Modal shares

- Coal: 47%
- Iron & steel: 14%
- Fertilizers: 6%
- Cement: 11%
- Mineral oils: 4%
- Other: 13%

<table>
<thead>
<tr>
<th></th>
<th>Rail</th>
<th>Road</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal</td>
<td>79.9</td>
<td>16.5</td>
<td></td>
</tr>
<tr>
<td>Iron &amp; steel</td>
<td>51.6</td>
<td>45.1</td>
<td></td>
</tr>
<tr>
<td>Cement</td>
<td>50</td>
<td>48.1</td>
<td></td>
</tr>
<tr>
<td>Fertilizers</td>
<td>66.7</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>Mineral oils</td>
<td>18.5</td>
<td>67.5</td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>32.2</td>
<td>67.8</td>
<td></td>
</tr>
</tbody>
</table>
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

(a) Effect on congestion

(b) Effect on ln(Revenue)
Notes: This figure shows how shipping times respond to increased traffic, consistent with the railway model from operations research. Specifically, it plots the $\beta_c$ coefficients from the regressions

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

$$\ln V_{dy} = \sum_{c=50}^{160} \beta_c^V (D_{dy} \times 1[c \leq C_{d,2008} < c + 10]) + \gamma_d + \gamma_y + \epsilon_{dy}.$$
Figure 10: Event study for the coefficient on $D_{dy} \times T_{d,y=30}$
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

<table>
<thead>
<tr>
<th>Firm variables, at factory level</th>
<th>Mean (1)</th>
<th>St. Dev. (2)</th>
<th>Δ by eventual treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue (million INR)</td>
<td>1251.7</td>
<td>3096.1</td>
<td>29.304</td>
</tr>
<tr>
<td></td>
<td>(20.634)</td>
<td>(26.368)</td>
<td>15.31</td>
</tr>
<tr>
<td>ln(TFPR)</td>
<td>2.389</td>
<td>0.864</td>
<td>-0.022*</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.008)</td>
<td>0.001</td>
</tr>
<tr>
<td>Average cost</td>
<td>1.013</td>
<td>1.136</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.015)</td>
<td>0.013</td>
</tr>
<tr>
<td>Total inventory (million INR)</td>
<td>185.5</td>
<td>480.8</td>
<td>3.254</td>
</tr>
<tr>
<td></td>
<td>(3.399)</td>
<td>(4.268)</td>
<td>-0.713</td>
</tr>
<tr>
<td>Inputs</td>
<td>108.9</td>
<td>278.2</td>
<td>1.713</td>
</tr>
<tr>
<td></td>
<td>(2.054)</td>
<td>(2.462)</td>
<td>-1.142</td>
</tr>
<tr>
<td>Finished goods</td>
<td>76.3</td>
<td>190.0</td>
<td>1.437</td>
</tr>
<tr>
<td></td>
<td>(1.308)</td>
<td>(1.759)</td>
<td>0.776</td>
</tr>
<tr>
<td>Input share of rail goods</td>
<td>0.265</td>
<td>0.189</td>
<td>-0.005*</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>-0.003</td>
</tr>
<tr>
<td>Makes rail good (dummy)</td>
<td>0.631</td>
<td>0.483</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>0.011*</td>
</tr>
<tr>
<td>Survival until 2012</td>
<td>0.534</td>
<td>0.499</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.012)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rail traffic variables</th>
<th>Mean (1)</th>
<th>St. Dev. (2)</th>
<th>Δ by eventual treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line capacity (trains per day)</td>
<td>32.4</td>
<td>28.2</td>
<td>-1.055</td>
</tr>
<tr>
<td></td>
<td>(2.193)</td>
<td>(2.312)</td>
<td>0.997</td>
</tr>
<tr>
<td>Line capacity utilization %</td>
<td>95.2</td>
<td>11.8</td>
<td>0.406</td>
</tr>
<tr>
<td></td>
<td>(0.452)</td>
<td>(0.539)</td>
<td>0.744</td>
</tr>
<tr>
<td>% passenger traffic</td>
<td>66.3</td>
<td>16.1</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>(0.588)</td>
<td>(0.582)</td>
<td>-0.560</td>
</tr>
<tr>
<td>% freight traffic</td>
<td>28.1</td>
<td>15.3</td>
<td>-0.157</td>
</tr>
<tr>
<td></td>
<td>(0.573)</td>
<td>(0.683)</td>
<td>0.403</td>
</tr>
<tr>
<td>% other traffic</td>
<td>5.5</td>
<td>4.8</td>
<td>-0.110</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.242)</td>
<td>0.137</td>
</tr>
<tr>
<td>Mean freight ship time, days</td>
<td>5.11</td>
<td>4.10</td>
<td>0.624</td>
</tr>
<tr>
<td>(normalized to 1000km)</td>
<td>(0.810)</td>
<td>(0.898)</td>
<td>-0.255</td>
</tr>
<tr>
<td>Variance of freight time, days</td>
<td>6.39</td>
<td>11.19</td>
<td>-0.037</td>
</tr>
<tr>
<td>(normalized to 1000km)</td>
<td>(0.722)</td>
<td>(0.734)</td>
<td>-0.402</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Main specification</th>
<th>With second-order spillovers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Passenger trains</td>
<td>Freight trains</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Duronto routes</td>
<td>0.611***</td>
<td>0.0208</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Spillover exposure</td>
<td>0.221**</td>
<td>0.227**</td>
</tr>
<tr>
<td>(alternate routes)</td>
<td>(0.0917)</td>
<td>(0.114)</td>
</tr>
<tr>
<td>Second-order</td>
<td></td>
<td></td>
</tr>
<tr>
<td>spillovers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 × Sample FE for {Dur,Alt} | ✓                  | ✓                | ✓            | ✓                  | ✓                | ✓            |
Yr × 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. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. 

48
Table 3: Reduced form effects of Duronto trains on rail using firms

<table>
<thead>
<tr>
<th></th>
<th>ln(Revenue)</th>
<th>ln(TFPR)</th>
<th>ln(Avg cost)</th>
<th>ln(Inventory)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Panel A. Preferred specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duronto routes through district</td>
<td>$-0.0194^{***}$</td>
<td>$-0.0111^{***}$</td>
<td>$0.0081^{**}$</td>
<td>$0.0097^{*}$</td>
</tr>
<tr>
<td></td>
<td>$(0.0050)$</td>
<td>$(0.0041)$</td>
<td>$(0.0032)$</td>
<td>$(0.0053)$</td>
</tr>
<tr>
<td>Spillover routes through district</td>
<td>$-0.0110^{*}$</td>
<td>$-0.0064$</td>
<td>$0.0072^{*}$</td>
<td>$0.0006$</td>
</tr>
<tr>
<td></td>
<td>$(0.0063)$</td>
<td>$(0.0042)$</td>
<td>$(0.0042)$</td>
<td>$(0.0064)$</td>
</tr>
<tr>
<td><strong>Panel B. Without spillover control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duronto routes through district</td>
<td>$-0.0125^{**}$</td>
<td>$-0.0075^{*}$</td>
<td>$0.0036$</td>
<td>$0.0093^{*}$</td>
</tr>
<tr>
<td></td>
<td>$(0.0049)$</td>
<td>$(0.0041)$</td>
<td>$(0.0030)$</td>
<td>$(0.0053)$</td>
</tr>
</tbody>
</table>

Observations: 27558, 26896, 21624, 26618
Clusters 1 (factories): 6191, 6074, 5238, 5964
Clusters 2 (district × 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

<table>
<thead>
<tr>
<th></th>
<th>ln(Revenue)</th>
<th>ln(TFPR)</th>
<th>ln(Avg cost)</th>
<th>ln(Inventory)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Panel A. Preferred specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duronto routes through district</td>
<td>$-0.0006$</td>
<td>$-0.0031$</td>
<td>$0.0011$</td>
<td>$0.0021$</td>
</tr>
<tr>
<td></td>
<td>$(0.0044)$</td>
<td>$(0.0019)$</td>
<td>$(0.0032)$</td>
<td>$(0.0049)$</td>
</tr>
<tr>
<td>Spillover routes through district</td>
<td>$-0.0023$</td>
<td>$-0.0024$</td>
<td>$0.0012$</td>
<td>$-0.0015$</td>
</tr>
<tr>
<td></td>
<td>$(0.0059)$</td>
<td>$(0.0028)$</td>
<td>$(0.0041)$</td>
<td>$(0.0060)$</td>
</tr>
<tr>
<td><strong>Panel B. Without spillover control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duronto routes through district</td>
<td>$0.0003$</td>
<td>$-0.0019$</td>
<td>$0.0004$</td>
<td>$0.0031$</td>
</tr>
<tr>
<td></td>
<td>$(0.0044)$</td>
<td>$(0.0019)$</td>
<td>$(0.0033)$</td>
<td>$(0.0049)$</td>
</tr>
</tbody>
</table>

Observations: 50483, 48101, 37688, 45012
Clusters 1 (factories): 10844, 10420, 8664, 9690
Clusters 2 (district × 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 

49
Table 5: Model estimates of cost and competition effects

<table>
<thead>
<tr>
<th></th>
<th>ln(Revenue)</th>
<th>ln(TFPR)</th>
<th>ln(Avg cost)</th>
<th>ln(Inventory)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Duronto routes through district</td>
<td>−0.0308***</td>
<td>−0.0115***</td>
<td>0.0079**</td>
<td>0.0094**</td>
</tr>
<tr>
<td></td>
<td>(0.0050)</td>
<td>(0.0037)</td>
<td>(0.0035)</td>
<td>(0.0046)</td>
</tr>
<tr>
<td>Spillover routes through district</td>
<td>−0.0104*</td>
<td>−0.0079*</td>
<td>0.0069*</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0061)</td>
<td>(0.0046)</td>
<td>(0.0037)</td>
<td>(0.0064)</td>
</tr>
<tr>
<td>Exposure of (State × Industry) to Duronto routes</td>
<td>0.0249**</td>
<td>0.0035</td>
<td>0.0012</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>(0.0106)</td>
<td>(0.0079)</td>
<td>(0.0110)</td>
<td>(0.0131)</td>
</tr>
<tr>
<td>Exposure of (State × Industry) to spillover routes</td>
<td>0.0186</td>
<td>−0.0108</td>
<td>−0.0036</td>
<td>−0.0014</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0082)</td>
<td>(0.0127)</td>
<td>(0.0147)</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of equation (26), 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 6: Model estimates of cost and competition effects

<table>
<thead>
<tr>
<th></th>
<th>ln(Revenue)</th>
<th>ln(TFPR)</th>
<th>ln(Avg cost)</th>
<th>ln(Inventory)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Duronto routes through district</td>
<td>−0.0014</td>
<td>−0.0081**</td>
<td>0.0080**</td>
<td>0.0089**</td>
</tr>
<tr>
<td></td>
<td>(0.0052)</td>
<td>(0.0039)</td>
<td>(0.0037)</td>
<td>(0.0041)</td>
</tr>
<tr>
<td>(Duronto routes) × σ</td>
<td>−0.0065***</td>
<td>−0.0006**</td>
<td>−0.0001</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0003)</td>
<td>(0.0004)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>Spillover routes through district</td>
<td>−0.0046</td>
<td>−0.0056</td>
<td>0.0067*</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0064)</td>
<td>(0.0052)</td>
<td>(0.0039)</td>
<td>(0.0073)</td>
</tr>
<tr>
<td>(Alternate routes) × σ</td>
<td>−0.0012*</td>
<td>−0.0003</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0005)</td>
<td>(0.0005)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>Exposure of (State × Industry) to Duronto routes</td>
<td>0.0231*</td>
<td>0.0039</td>
<td>0.0022</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>(0.0119)</td>
<td>(0.0084)</td>
<td>(0.0117)</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>Exposure of (State × Industry) to spillover routes</td>
<td>0.0201</td>
<td>−0.0094</td>
<td>−0.0034</td>
<td>−0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td>(0.0088)</td>
<td>(0.0140)</td>
<td>(0.0151)</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of equation (26) 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 7: Aggregate effects of Duronto congestion on state level industry revenue

<table>
<thead>
<tr>
<th>Exposure of (State × Industry) to Duronto routes</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0132</td>
<td>-0.0020</td>
<td>-0.0139</td>
<td>-0.0028</td>
</tr>
<tr>
<td></td>
<td>(0.0159)</td>
<td>(0.0227)</td>
<td>(0.0155)</td>
<td>(0.0224)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exposure of (State × Industry) to spillover routes</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0048</td>
<td>-0.0051</td>
</tr>
<tr>
<td></td>
<td>(0.0097)</td>
<td>(0.0206)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observations</th>
<th>7901</th>
<th>7882</th>
<th>7901</th>
<th>7883</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clusters (state × industry)</td>
<td>1932</td>
<td>1914</td>
<td>1866</td>
<td>1928</td>
</tr>
<tr>
<td>State × Industry FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>State × Yr FE, Ind × Yr FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of equation (27), 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 8: First stage effects of Duronto traffic on freight shipment times

<table>
<thead>
<tr>
<th></th>
<th>ln(Mean) (1)</th>
<th>ln(Variance) (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duronto routes through district</td>
<td>0.113*** (0.028)</td>
<td>0.039 (0.026)</td>
</tr>
<tr>
<td>(Duronto routes) × (2008 congestion)</td>
<td>-0.021 (0.031)</td>
<td>0.211*** (0.042)</td>
</tr>
<tr>
<td>Observations</td>
<td>6896</td>
<td>6896</td>
</tr>
<tr>
<td>Clusters (districts)</td>
<td>174</td>
<td>174</td>
</tr>
<tr>
<td>F statistic</td>
<td>19.22</td>
<td>27.43</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of equations (32), 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 9: 2SLS estimates of mean and variance effects

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

Notes: Panel A of this table presents second stage estimates of equation (29), 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 (32). 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 10: The cost of running one Duronto train

Panel A: Cost of running one Duronto route, imperfect competition

<table>
<thead>
<tr>
<th></th>
<th>Loss for affected firms (1)</th>
<th>Gain for competitors (2)</th>
<th>Net effect (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duronto direct effects</td>
<td>-461.2</td>
<td>372</td>
<td>-89.2</td>
</tr>
<tr>
<td>Spillover effects</td>
<td>-154.9</td>
<td>195.4</td>
<td>40.5</td>
</tr>
<tr>
<td>Total (million INR)</td>
<td>-616.1</td>
<td>567.4</td>
<td>-48.7</td>
</tr>
</tbody>
</table>

Panel B: All firms experience congestion increase equivalent to one Duronto, perfect competition

<table>
<thead>
<tr>
<th></th>
<th>Rail-using firms (4)</th>
<th>All manufacturing (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total output loss (million INR)</td>
<td>94,962</td>
<td>258,551</td>
</tr>
</tbody>
</table>

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.
Online Appendix

For the online appendix to this paper, see http://economics.mit.edu/grad/jfirth/research.