Railroads and the Raj: The Economic Impact of Transportation Infrastructure

David Donaldson*
London School of Economics

PRELIMINARY AND INCOMPLETE

July 18, 2008

Abstract

I estimate the economic impact of the construction of colonial India's railroad network from 1861-1930. Using newly collected district-level data on annual output, prices and internal trade flows I find that the railroad network had the following effects: (1) Railroads caused transport costs along optimal routes to fall by 73 percent for an average shipment. (2) The lower transport costs caused by railroads significantly increased India's interregional and international trade. (3) The responsiveness of a region's agricultural prices to its own rainfall shocks fell sharply after it was connected to the railroad network, but its responsiveness to shocks in other regions on the network rose. (4) Railroads raised real agricultural income by 18 percent in the districts where they were built. I find similar results using instrumental variables based on British military and famine-prevention motives for building railroad lines; and I find no effect, for any of the above four outcomes, in a 'placebo' specification that uses railroad lines that were approved and surveyed, but were never actually constructed. Finally, I estimate the structural parameters of a Ricardian trade model using data on salt prices and trade flows only. The calibrated model explains 88 percent of the real income impact of railroads, suggesting that railroads raised welfare primarily because they enabled regions to exploit their comparative advantages by trading with one another.

*Department of Economics, STICERD and India Observatory, London School of Economics. E-mail: `d.j.donaldson@lse.ac.uk`. Web: `http://personal.lse.ac.uk/donald1s`. I am extremely grateful to Timothy Besley, Robin Burgess and Stephen Redding for their encouragement and support throughout this project. Costas Arkiolakis, David Atkin, Oriana Bandiera, Abhijit Banerjee, Erlend Berg, Marianne Bertrand, Esther Duflo, Maitreesh Ghatak, James Harrigan, Elhanan Helpman, Sam Kortum, Tim Leunig, Rocco Macchiavello, Guy Michaels, Ted Miguel, Sharun Mukand, Inger Munk, Ralph Ossa, Gerard Padro-i-Miguel, Rohini Pande, Torsten Persson, Nancy Qian, Imran Rasul, James Robinson, Tirthankar Roy, Philipp Schmidt-Dengler, Daniel Sturm and Tony Venables provided valuable advice and helpful comments. I am grateful to Erasmus Ermgassen, Rashmi Harimohan, Sritha Reddy and Structured Concepts for their help with the data, and to the British Academy, the Nuffield Foundation, the Royal Economic Society, and STICERD for financial support for data collection.

1 Introduction

In 2007, almost 20 percent of World Bank lending was allocated to transportation infrastructure projects, a larger share than that of education, health and social services combined (World Bank 2007). According to most theories of international and interregional trade, the reduced transport costs brought about by these projects can be expected to facilitate trade and thereby increase real income levels. A related body of theoretical work emphasizes a second effect that may also be important in predominantly agricultural, low-income economies: lower transport costs may reduce the dependence of real incomes on local climatic variation and hence reduce the variance of real incomes. Yet despite the importance of reducing transport costs in contemporary aid priorities and economic theory, we lack estimates of the magnitudes of the mean-increasing and variance-reducing effects of transportation infrastructure, and an empirical understanding of the role that trade plays in driving these effects.

In this paper I evaluate the effect of a large improvement in transportation infrastructure on the level and variance of real income, and attempt to quantify the role of increased opportunities to trade in explaining these gains. The setting I consider is the construction of a vast railroad network from 1861-1930 in colonial India (the region comprising modern-day Bangladesh, India and Pakistan, which I henceforth refer to as India). Prior to the railroad age, trading throughout much of India took place on dirt roads, primarily by bullocks moving at no more than 3 km/h (Derbyshire 1985, Deloche 1994). Bullock-powered road transport was costly, seasonal, unpredictable and often damaged goods in transit. By contrast, steam-powered railroad shipments moved about ten times faster, on a regular year-round schedule, and were sheltered from the elements. The arrival of the railroad network in an inland Indian region brought it out of near-autarky, connecting it with the rest of India and the world. I use this shock to evaluate the economic impact of a dramatic improvements in a region's transportation infrastructure.

To structure my empirical analysis I develop a general equilibrium model with many regions and goods, where trade occurs at a cost. Because of geographical heterogeneity, regions have differing productivity levels across goods, which creates incentives to trade to exploit comparative advantage. The large number of interacting product and factor markets in the model creates a complex general equilibrium problem, but the analysis is made feasible by the use of the (Eaton and Kortum 2002) functional form on the distribution of productivities. A new railroad link between two regions lowers their bilateral transport costs, allowing consumers to buy goods from the cheapest location, and allowing producers to produce more of what they are best at.

The model makes a series of four predictions that I use to drive my empirical analysis:

1. If two regions actually trade a commodity, then that commodity's price differential between the two regions is equal to their bilateral trade cost for that good. This price differential will therefore fall as these regions are connected by a railroad line, the extent that railroads lowered transport costs.

2. Bilateral trade flows between regions (both within India and internationally) take the form of the well-known gravity equation.

3. A region's retail prices will become less responsive to local supply shocks, and more responsive to neighboring regions' supply shocks, when it is connected to these neighboring regions by railroad lines.

4. A region's average real income (averaging over time) should increase, and the variance of this income should decrease, as the region is connected to the railroad network. However, the opposite will be true if one of the region's neighbors is connected to a third region by a railroad line.

To evaluate the impact of the railroad network and test the predictions of my theoretical framework I have collected new data on prices, output and interregional and international trade in India. The dataset consists of annual district-level observations (with 29 districts in my sample), for each of 17 commodities (85 commodities on trade flows), from 1861-1930. Data on bilateral region-to-region interregional trade is rarely available to researchers, especially during a time of dramatic improvements to the transportation technology. Yet observing the trades that happen within a country, and how they respond to an improvement in the transportation technology, is an important component of understanding the nature of this change. The final component of my analysis is a newly constructed GIS database on the Indian transportation network in every year from 1851 to 1930. I use this resource to construct a number of measures of annual, district-to-district transport costs, which evolve as the railroad network expands.

My empirical approach proceeds in four steps, which follow the model's four predictions. In the first step (which tests prediction 1) I analyze interregional price differentials for eight different types of salt that are differentiated by region of origin. Each of these salt types was produced at a single source, but consumed throughout India. Prediction 1 therefore states that (for each different type of salt) the price difference between the source and any other district must be equal to the cost of trading salt between these locations. I use this result to measure the effect of the railroad network on bilateral trade costs, by modeling how the railroad network changed the optimal route between every pair of districts in every year. This is computationally feasible through the use of Dijkstra's shortest path algorithm, commonly used in transportation applications of graph

theory. I find that the railroad network caused the median bilateral trade cost between districts to fall by 73 percent. I estimate the parameters relating distance to trade costs separately for each mode of transportation.

The second step of my empirical analysis tests prediction 2—that bilateral trade flows between regions of India take on a gravity equation form. I find strong support for this prediction. I go on to show that the parameters estimated from the gravity equation—in conjunction with the trade cost parameters estimated in step 1—allow all of the model's parameters to be identified. I therefore estimate the entire model using data on salt prices and trade flows alone.

My third step tests prediction 3 using the responsiveness of agricultural prices (for each of 17 agricultural goods) to crop-specific rainfall shocks. I find that before a district is connected to the railroad network, an adverse rainfall shock in that district raises prices there; but after a railroad line penetrates the district, rainfall shocks there have no effect on local prices. This is consistent with the idea that railroads caused market integration in India. To probe this further I test the second part of prediction 3: that the extent to which a district's agricultural prices respond to rainfall shocks in neighboring districts depends on the cost of trading with these districts. Using the trade cost measures estimated in step 1, I find strong support for this prediction.

In the fourth step of my empirical analysis, guided by prediction 4, I estimate the total effect on a district's real agricultural income of being connected to the railroad network. I find that the arrival of a railroad line in a district raised real agricultural income there by a statistically and economically significant 18.2 percent. However, in accordance with prediction 4, the arrival of a railroad line in a neighboring district (controlling for own railroad access) reduced real agricultural income by (a statistically significant) 4.2 percent.

Finally, I use the estimated model parameters (obtained from steps 1 and 2 only) to compute a (unobserved) variable that I refer to as 'autarkiness'—the fraction of a region's expenditure that it buys from itself as opposed to other regions. According to the model, autarkiness is a sufficient statistic for all the effects of the entire railroad network on a district's real agricultural income. When I include the autarkiness term in the regressions from step 4, I find that the model's prediction (that the coefficient on this sufficient statistic measure should be one) cannot be rejected (its 95 percent confidence interval lies between 0.85 and 1.04). With this autarkiness term included, the coefficient on railroad access falls to a statistically insignificant 2.1 percent (from 18.2 percent without the autarkiness term included), implying that the model explains 88 percent of the observed impact of the railroad network. I interpret this as evidence that increased trade opportunities played an important role in explaining the impact of the railroad network in India.

This paper contributes to a growing literature that attempts to estimate the causal effects of a wide range of large infrastructure projects on economic welfare.[1] A distinguishing feature of my approach to evaluating the Indian railroad infrastructure project is the use of a model to assess quantitatively the role played by a particular mechanism in explaining the observed total effect. An attraction of this approach is that it is likely to improve external validity: an estimated and validated model can be used to inform quantitative predictions about the impact of future transportation infrastructure projects in different environments.

This paper also contributes to a rich literature on the role of railroads in the economic development of India, the United States, and many other countries.[2] This literature has been heavily influenced by the Fogel (1964) 'social savings' approach to calculating the benefits of railroads, even though this method has been criticized by many authors (for example, Fogel (1979), Fishlow (1965) and Williamson (1974)), for missing several potential effects that railroads may have caused.[3] My reduced-form estimates, by contrast, measure all of the effects of railroads—and unsurprisingly, my estimate of 18.2 percent is significantly larger than the Hurd (1983) social savings from Indian railroads of 9 percent of GDP. My approach is closest in spirit to Williamson (1974), who calibrates a multi-region, multi-good, dynamic model of the United States economy as the railroad network was expanded there. A distinguishing feature of my approach is that I compare the welfare estimates from my model (whose parameters are estimated from data not appearing in the welfare specification) to those obtained from a 'reduced-form' econometric approach that measures all of the welfare effects of railroads.

The remainder of this paper proceeds as follows. In the next section I outline my

[1]For example: Michaels (2008) estimates the effect of the United States Interstate Highway system on the skilled wage premium in rural US counties. Jensen (2007) evaluates the reduction in fish wastage and price volatility due to the construction of cellular phone towers in Kerala, India. Duflo and Pande (2007) use the average land slope in Indian districts to predict dam construction there, and estimate the effect of these dam projects on agricultural production and income. Dinkelman (2007) uses a similar identification strategy to estimate the effect of electrification on female labor force participation in South Africa.

[2]On India: Hurd (1975) documents greater price convergence among districts with railroads than among those without, and Andrabi and Kuehlwein (2005) find similar results among pairs of districts connected by railroads; Hurd (1983) estimates social savings; see also McAlpin (1974), McAlpin (1980), and Derbyshire (1985). On the United States, see Fogel (1964), Fishlow (1965), Williamson (1974) and the survey in Fishlow (2000). On other countries, see, for example, Metzer (1974) on Russia, Coatsworth (1979) on Mexico, Summerhill (2005) on Brazil, and the survey in O'Brien (1977).

[3]Social savings estimate the increase in national income that results from a technological improvement in the transportation sector, if the location of industries and economic activity stays the same, but the factors of production freed up in the transportation sector are instead used in their next-best substitute. For my analysis, the most relevant of these uncounted effects is that originally pointed out by Fishlow (1965): by reducing transport costs, railroads may have given rise to aggregate efficiency gains due to reallocations in the transportation-using sectors (even if the factors of production freed up in the transportation sector sit idle). This is the mechanism stressed in my comparative advantage trade model with iceberg transport costs.

theoretical framework and its predictions. Section 3 provides a historical background on the construction of India's railroad network. In section 4 I make clear how my theoretical model informs the econometric strategy I follow, as well as the instrumental variables and 'placebo' regressions I estimate. In section 5 I describe the new dataset that I have assembled on colonial India, and how it relates to the empirical strategy outlined in section 4. Section 6 presents my results, for each of the model's four predictions and the estimated structural model. Finally, section 7 concludes.

2 A Model of Trade in India

In this section I develop a general equilibrium model of trade among many regions in the presence of transport costs. The model serves two purposes. First, it delivers qualitative predictions that I use to guide my empirical evaluation of the construction of India's railroad network, which I interpret as reducing the cost of trading between regions of British India. Second, I estimate the model and use it to assess quantitatively how much of the total impact of the railroad network can be attributed to the role of increased opportunities to trade (the mechanism stressed in this model.)

2.1 Outline of the Model

2.1.1 Environment:

The economy consists of D regions (indexed by d); in my empirical application I take these regions to be both regions of India, and countries outside of India. There are K goods (indexed by k). Each good is available in a continuum of horizontally differentiated varieties (indexed by j) and the mass of this continuum is normalized to one. In my empirical application I work with data (on prices, allocations and trade flows) that refer to goods, and not individual varieties. While my empirical setting will consider 70 years of annual observations, for simplicity the model is static; I therefore suppress time subscripts until they are necessary.

2.1.2 Consumer Preferences:

Each region d is home to a mass (normalized to one) of identical agents, each of whom owns L_d units of land. Land is the only factor of production, is geographically immobile, and is supplied inelastically. Agents have Cobb-Douglas preferences over goods, and constant elasticity of substitution preferences over varieties within each good; that is,

their (log) utility function is

$$\ln U_d = \sum_{k=1}^{K} \left(\frac{\mu_k}{\varepsilon_k} \right) \ln \int_0^1 (Q_d^k(j))^{\varepsilon_k} dj, \tag{1}$$

where $Q_d^k(j)$ is the amount of variety j of good k consumed by a representative agent in region d, $\varepsilon_k = \frac{\sigma_k - 1}{\sigma_k}$ where σ_k is the (constant) elasticity of substitution, and the K Cobb-Douglas weights μ_k sum to one. Agents rent their land at the rate of r_d per unit, and use their income $r_d L_d$ to maximize utility from consumption.

2.1.3 Production and Market Structure:

Each variety j of the differentiated good k can be produced using a constant returns to scale production technology in which land is the only factor of production. Let $z_d^k(j)$ denote the amount of land required to produce a unit of variety j of good k in region d. I follow Eaton and Kortum (2002) in modeling $z_d^k(j)$ as the realization of a stochastic variable Z_d^k drawn from a Type-II extreme value distribution whose parameters vary across regions and goods; that is,

$$F_d^k(z) \doteq \Pr(Z_d^k \leq z) = \exp(A_d^k z^{-\theta_k}) \tag{2}$$

where $A_d^k \geq 0$ and $\theta_k > 0$. These random variables are drawn independently for each variety, good and region.[4] The parameter A_d^k increases the probability of high productivity draws of good k in region d, while θ_k captures (inversely) how variable the productivity of good k in any region is around its average. In my empirical setting I proxy for A_d^k using observable climatic variables that are known to be important determinants of a region's good-specific productivity. I treat θ_k as an unknown structural parameter to be estimated.

There are many competitive firms in region d with access to the above technology for

[4]The assumption of within-sector heterogeneity characterized by a continuous stochastic distribution of productivities is a standard feature in the literature on trade with heterogeneous firms (eg Melitz (2003).) It is common in that literature to assume that the productivity distribution is Pareto (to which the upper tail of a Type-II extreme value distribution converges), and that productivities are drawn independently across varieties (firms), goods and countries (eg Melitz and Ottaviano (2008), Chaney (2008) and Helpman and Rubinstein (2008).) As in the present paper, these assumptions are convenient for their ability to generate closed-form expressions for the model's endogenous variables. One advantage of the Type-II extreme value distribution used here is its potential microfoundations: because of the extremal types theorem (see, for example, Galambos (1978)), if producers draw ideas about production techniques in an iid manner from *any* continuous distribution, but choose to use only the maximum of the techniques they draw, then the distribution of technologies actually used in a population of producers will converge to a distribution in the extreme value family. Kortum (1997) makes this point formally in a model of innovation.

producing variety j of good k; consequently, firms make zero profits. These firms will therefore charge a mill (that is, pre-transport costs) price of $p^k_{dd}(j) = r_d/z^k_d(j)$, where r_d is the land rental rate in region d.

2.1.4 Opportunities to Trade:

Consumers can import varieties from other regions in order to take advantage of the favorable productivity draws potentially available there. But the movement of goods is subject to trade costs, which include transport costs, tariffs or other barriers to trade. These trade costs take the convenient and commonly used Samuelson (1954) 'iceberg' form. That is, in order for one unit of good k to arrive in region d, $T^k_{od} \geq 1$ units of the good must be produced and shipped in region o; trade is free when $T^k_{od} = 1$. (Throughout this paper I refer to trade flows between an *origin* region o and a *destination* region d. All bilateral variables, for example T^k_{od}, refer to quantities *from o to d*.) Because of free spatial arbitrage, trade costs satisfy the property that it is always (weakly) cheaper to ship directly from region o to region d, rather than via some third region m: that is, $T^k_{od} \leq T^k_{om}T^k_{md}$, for all regions o, d and m, and all goods k. Finally, I normalize $T^k_{oo} = 1$ for all regions o and goods k.[5] In my empirical setting I proxy for T^k_{od} with measures calculated from the observed transportation network, which incorporates all possible modes of transport between region o and region d. I allow these trade cost measures to vary across goods k, by relating them to observed characteristics such as weight and bulkiness. Trade costs drive a wedge between the price of an identical variety in two different regions. Let $p^k_{od}(j)$ denote the price of variety j of good k produced in region o, but shipped to region d for consumption there. Iceberg trade costs imply that any good in region d will cost T^k_{od} times more than in region o; that is, $p^k_{od}(j) = T^k_{od}p^k_{dd}(j) = r_dT^k_{od}/z^k_d(j)$.

2.2 Equilibrium Prices and Allocations

Consumers have preferences for all varieties j along the unit continuum of varieties of good k. But they are are indifferent about where a given variety is made—they simply buy from the region that can provide the variety at the lowest cost.[6] The first step in solving

[5] This is necessary because, like most attempts to infer trade costs from prices and trade flows (surveyed in Anderson and van Wincoop (2004)), my empirical methodology can only identify trade costs up to a proportionality constant. That is, T^k_{od} captures the additional costs of shipping good k from region o to region d, relative to the same shipment from region o to itself.

[6] This is in contrast to the Armington specialization assumption — where goods are differentiated by region of origin in consumers' utility functions— frequently used in the trade literature (see, for example, Anderson and van Wincoop (2003).) An attraction of the present framework is that specialization and the location of production are endogenous.

for the model's equilibrium is therefore to characterize the prices that consumers in a region d actually face, given that they will only buy a variety from the region (including their own) that can provide it at the lowest price.

The price of a variety sent from region o to region d, denoted by $p_{od}^k(j)$, is stochastic because it depends on the stochastic variable $z_d^k(j)$. Since $z_d^k(j)$ is drawn from the CDF in equation (2), $p_{od}^k(j)$ is the realization of a random variable P_{od}^k drawn from the CDF

$$G_{od}^k(p) \doteq \Pr(P_{od}^k \leq p) = 1 - \exp[-A_o^k (r_o T_{od}^k)^{-\theta_k} p^{\theta_k}]. \tag{3}$$

This is the price distribution for varieties (of good k) that could *potentially* be bought in region d from region o. The price distribution for the varieties that consumers in d will *actually* consume (denoted by $G_d^k(p)$) is the distribution of prices that are the lowest (among all D regions of the world):

$$\begin{aligned} G_d^k(p) &= 1 - \prod_{o=1}^{D} [1 - G_{od}^k(p)], \\ &= 1 - \exp(-\Phi_o^k p^{\theta_k}), \end{aligned}$$

where

$$\Phi_d^k = \sum_{o=1}^{D} A_o^k (r_o T_{od}^k)^{-\theta_k}. \tag{4}$$

Given this distribution of equilibrium prices in region d, it is straightforward to calculate any moment of the prices of interest. Two price moments are relevant to my empirical analysis. First, the expected value of the price of any variety j of good k found in region d in equilibrium is given by[7]

$$E[p_d^k(j)] \doteq p_d^k = \lambda_1^k (\Phi_d^k)^{-1/\theta_k}. \tag{5}$$

Second, the exact price index (given CES preferences) over all varieties of good k for consumers in region d is given by[8]

$$\widetilde{p}_d^k = \lambda_2^k (\Phi_d^k)^{-1/\theta_k}. \tag{6}$$

Given the price distribution in equation (3), Eaton and Kortum (2002) derive two

[7]Here, $\lambda_1^k \doteq \Gamma(1 + \frac{1}{\theta_k})$ is a constant that depends on the parameter θ_k only. $\Gamma(.)$ is the Gamma function defined by $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$.

[8]The exact price index for CES preferences is $\widetilde{p}_d^k \equiv \left[\int_0^1 (p_d^k(j))^{1-\sigma_k} dj\right]^{1/1-\sigma_k}$. The price index is only well-defined for $\sigma_k < 1 + \theta_k$, so I assume this throughout. Here, $\lambda_2^k \doteq \Gamma(\frac{\theta_k + 1 - \sigma_k}{\theta_k})^{1/(1-\sigma_k)}$ is a constant that depends on the parameters θ_k and σ_k only.

important properties of the trading equilibrium, which carry over to the model here. First, the price distribution of the varieties that any origin region actually sends to destination region d does not depend on the identity of the origin region.[9] If the price distribution of any shipped variety is independent of the shipping region, then the share of *expenditure* that consumers in region d allocate to varieties of good k from region o must be equal to the probability that region o supplies a variety of good k to region d. That is $X_{od}^k / X_d^k = \pi_{od}^k$, where X_{od}^k is total expenditure in region d on goods of type k from region o, $X_d^k = \sum_o X_{od}^k$ is total expenditure in region d on goods of type k, and π_{od}^k is the probability that region d sources any variety of good k from region o. Second, this probability π_{od}^k is given by[10]

$$\pi_{od}^k = \frac{X_{od}^k}{X_d^k} = \lambda_3^k \, A_o^k \, (r_o T_{od}^k)^{-\theta_k} \, (p_d^k)^{\theta_k}, \tag{7}$$

which makes use of the definition of the expected value of prices (ie p_d^k) in equation (5). This important result states that bilateral trade flows for any good k (measured inclusive of trade costs) from any region o to any region d take a form that is reminiscent of the 'gravity equation', which has generated an enormous body of empirical and theoretical work.[11]

Equation (7) characterized trade flows conditional on the endogenous land rental rate, r_o (and all regions' land rental rates, which appear in p_d^k via Φ_d^k.) It remains to solve for these land rents in equilibrium, by imposing the condition that each region's trade is balanced. Region o's trade balance equation requires that the total income received by land owners in region o (ie $r_o L_o$ by zero profits) must equal the total value of all goods made in region o and sent to every other region (including region o itself). That is:

$$r_o L_o = \sum_d \sum_k X_{od}^k = \sum_d \sum_k \pi_{od}^k \mu_k r_d L_d, \tag{8}$$

where the last equality uses the fact that, because of Cobb-Douglas preferences across

[9] The intuition for this is as follows. If the varieties that a region (say, m) was supplying to d were cheaper on average than those from other regions, consumers in region d would buy more varieties from m. To buy more varieties from m, given that they are already consuming the cheapest varieties that m has to offer, consumers in d would have to shift into relatively more expensive varieties from m. This would raise the average price of varieties shipped from m to d.

[10] Here, $\lambda_3^k = (\lambda_1^k)^{-\theta_k}$ is a parameter that depends only on the parameter θ_k.

[11] Empirical work on the gravity model is surveyed by Anderson and van Wincoop (2004); an important recent contribution is Helpman and Rubinstein (2008). Examples of theoretical work include Anderson (1979), Deardorff (1998) and Chaney (2008). That the present framework yields a gravity equation is unsurprising given list of sufficient conditions in Anderson and van Wincoop (2004) for any trade model to do so—namely, separability of production and consumption allocations from trade allocations, CES aggregation across goods, and trade costs proportional to the quantity of trade. All three of these are satisfied here.

goods k, expenditure in region d on good k (X_d^k) will be a fixed share μ_k of the total income in region d ($r_d L_d$). Each of the D regions has its own trade balance equation of this form. I take the rental rate in the first region (r_1) as the numeraire good, so the equilibrium[12] of the model is the set of $D-1$ unknown rental rates r_d that solves this system of $D-1$ (non-linear) equations.[13] Once the equilibrium values of r_d are known, all other endogenous variables—prices via equations (5) or (6), bilateral trade flows via equation (7), as well as output, consumption and land allocations for each good—are easily solved for.

2.3 Theoretical Predictions

In this section I state explicitly, five of the model's predictions. These predictions are presented in the order in which I test for them in my empirical analysis (outlined in section 3.)

Prediction 1: Price Differences Measure Trade Costs

In the presence of trade costs, the price of identical goods will differ across regions. In general, the cost of trading a good between two regions places only an upper bound on their price differential. However, for the special case where a good can only be produced in one region (call the region s, for 'source'), equation (5) predicts that the (log) price differential between s and any other region d will be equal to the (log) cost of trading the good between them. That is:

$$\ln p_d^s - \ln p_s^s = \ln T_{sd}^s, \tag{9}$$

where the goods label k is replaced by s to indicate that this equation is only true for goods that are made exclusively in region s. This is an important prediction of the model, but it should be noted that this prediction — essentially just free arbitrage over space, net of trade costs — is common to most models of spatial equilibrium.

Prediction 2: Bilateral Trade Flows Take Gravity Form

Equation (7) describes bilateral trade flows explicitly, but I re-state it here in logarithms for reference: (log) bilateral trade (in expenditure terms and including trade costs) of any good k from any region o to any other region d is given by

$$\ln X_{od}^k = \ln \lambda_k + \ln A_o^k - \theta_k \ln r_o - \theta_k \ln T_{od}^k + \theta_k \ln p_d^k + \ln X_d^k. \tag{10}$$

[12] Alvarez and Lucas (2007) find sufficient conditions for an equilibrium in this class of models to exist and be unique. Their conditions place mild restrictions on the model's parameters, which are likely to hold here.

[13] By Walras's Law, only $D-1$ of the trade balance equations are independent.

In my empirical setting I estimate this equation directly using interregional and international trade data in India.

Prediction 3: Agricultural Price Responsiveness to Rainfall

Unfortunately, the multiple general equilibrium interactions in the model are too complex to admit a closed-form solution for the effect of reduced trade costs on agricultural prices. To make progress in generating qualitative predictions (to guide my empirical analysis) I therefore assume a much simpler environment in which: there are only three regions (called X, Y and Z); there is only one good (so I will dispense with the k superscripts on all variables); the regions are symmetric in their exogenous characteristics (ie $L_o = L$ and $A_o = A$ for all regions o); and the three regions have symmetric trade costs with respect to each other. I consider the comparative statics from a local change around this symmetric equilibrium. It is possible to derive the following results (all derivations are in appendix A):

1. $\frac{dp_X}{dA_X} < 0$: Prices in a region (say, X) rise when that region experiences adverse supply shocks (a reduction in A_X).

2. $\frac{d^2p_X}{dA_X dT_{YX}} = \frac{d^2p_X}{dA_X dT_{ZX}} < 0$: The responsiveness of prices in a region (say, X) to supply shocks *in the same region* (ie $\frac{dp_X}{dA_X}$) is weaker (ie less negative) when the region has low trade costs to other regions.

3. $\frac{d^2p_X}{dA_Y dT_{YX}} = \frac{d^2p_X}{dA_Z dT_{ZX}} > 0$: The responsiveness of prices in a region (say, X) to supply shocks *in other any other region* (ie $\frac{dp_X}{dA_Y}$ or $\frac{dp_X}{dA_Z}$) is stronger (ie more negative) when the cost of trading between these two regions (ie T_{YX} or T_{ZX}) is low.

In my empirical setting I test for these predictions using data on the prices of 17 agricultural crops, and measures of crop-specific rainfall (which proxies for the productivity term A_o, in this largely rain-fed agricultural environment.)

Prediction 4: Real Income Effects of Railroads

As was the case for prediction 3, the effects of lower trade costs on nominal and real income are also not solvable in closed form in this model. I therefore consider a similar simplified environment to that in prediction 3 (ie, three symmetric regions). My focus here is on the effect of trade costs on real income (which, as prediction 5 shows explicitly, is equal to welfare in this model). To simplify notation, let W_o represent real income per unit land area (ie $W_o \doteq \frac{r_o}{\widetilde{P}_o}$, where $\widetilde{P}_o$ is the aggregate price index in region o). Then the following results hold around the symmetric equilibrium:

1. $\frac{dW_X}{dT_{YX}} = \frac{dW_X}{dT_{ZX}} < 0$: Real income in a region (say, X) rises when the cost of trading between that region and any other region (ie T_{XY} or T_{XZ}) falls.

2. $\frac{dW_X}{dT_{YZ}} > 0$: Real income in a region (say, X) falls when the cost of trading between the two other regions (ie T_{YZ}) falls.

These results suggest that a reduction in trade costs in one part of the network is not good for all regions. A railroad project that reduces trade costs between two regions will raise welfare in these two regions; but this project will reduce welfare in the third, excluded region whose trade costs were unaffected by the project. This negative effect on excluded regions arises because of two effects: first, the excluded region's trading partners' factor costs have increased, which raises the prices of the goods that they ship to the excluded region; and second, the excluded region loses some of its export market because its trading partners now have a cheaper supplier (in each other.) In my empirical setting I test for the two predictions in this section using data on real agricultural income, and the construction of railroad lines in selected districts and years.

<u>Prediction 5: Sufficient Statistic for Welfare</u>

Given the utility function in equation (1), the indirect utility function for a consumer in region o is

$$V_o = \frac{(r_o L_o)}{\prod_{k=1}^{K} (\widetilde{p}_o^k)^{\mu_k}} = \frac{r_o L_o}{\widetilde{P}_o}. \tag{11}$$

The numerator of this expression is nominal income in region o relative to the numeraire (ie relative to land rental rates in region 1), and the denominator is the exact price index across all goods k (each of which has an exact price index $\widetilde{p}_o^k$ given by equation (6)) denoted by $\widetilde{P}_o$. That is, under these preferences welfare is equal to real income. Using the bilateral trade equation (7) evaluated at $d = o$, (log) real income per unit of land (defined as W_o as in prediction 4) can be re-written as[14]

$$\ln W_o = \Omega + \sum_k \frac{\mu_k}{\theta_k} \ln A_o^k - \sum_k \frac{\mu_k}{\theta_k} \ln \pi_{oo}^k. \tag{12}$$

This result draws a distinction between two forces that govern the welfare of a region o: First, welfare rises in region o's own productivity levels (A_o^k), which is unsurprising. Second, welfare falls in region o's 'autarkiness' level (ie π_{oo}^k), or the fraction of region o's expenditure that region o buys from itself (which is 1 in autarky). But the second force makes the very strong prediction that autarkiness is a sufficient statistic for the impact on welfare in region o of *all* exogenous changes in the model (such as trade costs, other regions' productivity levels, or the relative sizes of regions) other than region o's own productivity. If railroads lowered trade costs in India then, according to this model, all of the welfare impact of these railroads on regions of India work their way through the

[14]Here, $\Omega \doteq \sum_k \frac{\mu_k}{\theta_k} \ln \lambda_2^k$ is a constant that depends on the full set of parameters $\{\mu_k\}$ and $\{\theta_k\}$.

regions' autarkiness terms.

3 Empirical Strategy

The theoretical predictions outlined in section 2.3 take a naturally recursive order, both for estimating the model's parameters, and for tracing through the impact of railroads on welfare in India. I follow this order here. First, I evaluate the extent to which the railroads reduced trade costs within India (using prediction 1). Second, I use prediction 2 to measure how much these reduced trade costs increased trade within British India (and also between regions of British India and the rest of the world.) Third, I estimate the extent to which lower trade costs reduced price volatility in British India, by estimating the dependence of a region's agricultural prices on local productivity shocks as trade costs fell (prediction 3). Fourth, I estimate how a region's real income is affected when the railroad network is extended to that region, and when it is instead extended to other nearby regions.

I implement this strategy using a newly collected district panel dataset from colonial India, with annual observations from 1861 to 1930. (The data is described in section 4, and in detail in Appendix B.) To relate the static model in section 2 to my dynamic empirical setting I take the simplest possible approach and assume that all of the goods in the model cannot be stored. This assumption means that the static model simply repeats every period, with independence of all decision-making across time periods. Throughout the remainder of the paper I therefore add the subscript 't' to all of the variables in the model, but I assume that all of the model parameters are fixed over time.

3.1 Estimating Trade Costs from Salt Prices

British policy in India allowed salt to be produced at a small number of designated sources only.[15] I have collected data on salt prices in Northern India, in which the prices of salt are differentiated by source. These features allow me to exploit prediction 1, which states that the price difference, for a given type of salt, between its source and any destination

[15]This policy was in place so that a production tax could be levied with the smallest administrative overhead. However, once salt left its production site, trade was largely free and distribution unregulated. The one exception to this *laissez-faire* distribution policy was the Salt Line, which sought to tax cheap salt from Southern India upon entering Northern India (for the period up to 1878 during which production taxes were much higher in the North.) This Line consisted of an impenetrable thorn hedge and was patrolled by an army of over 40,000 officers, and smuggling efforts are thought to have been largely unsuccessful (Moxham 2001). To the extent that smuggling was non-existent (or orthogonal to the railroads), this policy is not a concern for my empirical strategy because for all of the sources I use that lay south of the Line, the destination districts that I use lie to the north of the Line (and hence the extra tax charged at the Line affects all of these pairs equally.)

is a measure of the cost of trading salt between these two location. I use this prediction to estimate the determinants of trade costs. A natural empirical counterpart of equation (9) is

$$\ln p_{dt}^k - \ln p_{st}^k = \rho \ln T_{sdt}^k + \eta_{dt}^k, \tag{13}$$

where p_{dt}^s is the price of salt of type s in district d in year t (similarly, p_{st}^s is the price of salt s at its source), and T_{sdt}^s is the cost of trading type-s salt from source s to district d in year t. η_{dt}^s is an error term that captures any unobserved determinants of the price of type-s salt in district d and year t.

Two problems of unobserved variables prevent me from estimating equation 13 as it stands. First, I do not observe salt prices exactly at the point where they leave the source (ie p_{st}^k).[16] I therefore proxy for the unobserved p_{st}^s term with a source-by-year fixed effect (denoted by γ_t^s). In addition to proxying for the unobserved variable p_{st}^s, this fixed effect controls for any unobserved determinants of prices of each type of salt in each year that affect the price in all destination districts d equally.

Second, the bilateral trade costs term (T_{sdt}^s) is also unobserved, so I use prediction 1 to estimate the relationship between unobserved trade costs and observable features of the trading environment in colonial India.[17] I use two different proxies (explained below) for T_{sdt}^s that place the focus on how the transportation network *changes over time*. Both of these trade cost proxies are functions of geographic features of the transportation network, which affect all goods equally. I denote these geographic features of either of the two trade cost proxy variables by G_{sdt}. To capture the idea that goods with high weight-to-unit value ratios cost more to trade, I allow the trade costs proxy to vary log-linearly by weight-to-unit value (which I denote by W^s). That is, I model the unobservable T_{sdt}^s as:

$$\ln T_{sdt}^s = \alpha_{sd}^s + \phi_{sd}^s t + \delta W^s \ln G_{sdt} + \varpi_{sdt}^s. \tag{14}$$

This specification includes a source-by-destination-by-good fixed effect (α_{ds}^s) which controls for all of the permanent determinants of the cost of trading good s between s and d (such as the distance and elevation gain along the route from s to d, or ethno-linguistic differences between s and d that may hinder trade.[18] The specification also includes a

[16] My price data is at the district level and was recorded as the average price of goods over a number of markets in a district. However, most salt sources were extremely localized within their own districts so these district average prices will not capture p_{st}^s accurately. Nevertheless, I obtain similar results when using the average price in source s's district to proxy for p_{st}^s.

[17] As argued by Anderson and van Wincoop (2004), the problem of unobserved trade costs is common to all of the trade literature. Even when transportation costs are observed (for example, from shipping receipts as in Hummels (2007)) these may fail to capture other barriers to trade, such as the time goods spend in transit (Evans and Harrigan 2005). In lieu of direct measures of trade costs, a large literature has used a similar proxy variables strategy to the one I use here.

[18] Rauch (1999) surveys the literature on communication-based barriers to trade; see also Anderson

separate trend term ($\phi^s_{ds}t$) for each district-source; these non-parametric trends control for any trade costs between s and d that vary over time in a (log) constant way (such as technological progress in road, river, or sea transport). Finally, ϖ^s_{sdt} is an error term that captures any remaining unobserved determinants of trade costs.

I use two different measures for G_{sdt}, the geographic determinants of trade costs between source s and district d in any year t. The first measure I use is a simple *bilateral railroad dummy* variable which is equal to one in all years when it is possible to travel from the district containing source s to district d by railroad (and zero otherwise).[19] I denote this variable $RAIL_{sdt}$, and estimate the semi-log specification

$$\ln p^s_{dt} = \alpha^s_{sd} + \phi^s_{sd}t + \gamma_{st} + \delta W^s RAIL_{sdt} + \varepsilon^s_{dt}. \tag{15}$$

using least squares.

The second measure I use is an *optimal route* measure of transportation costs, which I denote OR_{sdt}. This measure attempts to model the cost of transporting goods between any two locations under the assumption that agents will always take the lowest cost route available to them. To do this I represent the transportation system of British India (in any given year) as a weighted graph, as is common in the transportation literature.[20] A graph consists of nodes, links between pairs of nodes nodes called vertices, and weights assigned to each vertex that represent the 'strength' of that vertex. In my representation nodes represent locations (such as the salt source s, and the destination district d), vertices represent the possibility of travel between two nodes (directly, without visiting another node) using one of the four potential modes of transport (railroad, road, river, or sea), and weights represent the cost of transporting goods along a vertex using a given mode of transport. I assume that these link weights are proportional to the length of the link, where the constant of proportionality depends on the mode of transport that the link represents. Denote these constants α^{rail}, α^{road}, α^{river}, and α^{coast}, where a higher value of α represents higher cost of transporting goods, per unit of distance and weight. Since the optimal route between any locations is only affected by *relative* mode costs, I normalize $\alpha^{rail} = 1$. The other three α parameters are unknown, but will be estimated below.[21] I therefore denote the optimal route proxy for trade costs as a function of the

and van Wincoop (2004).

[19] More detail on the construction of this and other variables can be found in Appendix B.

[20] See, for example, Black (2003) for a textbook treatment.

[21] Relative freight rates are available from a number of historical sources (Deloche 1994, Deloche 1995, Derbyshire 1985). However, as with overall trade costs, these relative trade cost components potentially include distance-related determinants of trade costs that are not included in observed freight rates. Nevertheless, I compare my estimates of $\boldsymbol{\alpha}$ to historical freight rates.

three unknown parameters (ie $\boldsymbol{\alpha} \doteq (\alpha^{road}, \alpha^{river}, \alpha^{coast})$), and estimate the specification:

$$\ln p^s_{dt} = \alpha^s_{sd} + \phi^s_{sd} t + \gamma_{st} + \delta W^s \ln OR_{sdt}(\boldsymbol{\alpha}) + \varepsilon^s_{dt}. \tag{16}$$

This specification is non-linear in the parameter $\boldsymbol{\alpha}$, so I use a non-linear estimation routine. Conditional on a value of $\boldsymbol{\alpha}$, this routine calculates (for each year t) the cost of transporting goods from each node s to each node d, along the lowest-cost route. To compute these lowest-cost routes I use the Dijkstra shortest- path algorithm, a standard network flow algorithm for these settings (Ahuja, Magnanti, and Orlin 1993). The non-linear estimation routine then searches over all values of $\boldsymbol{\alpha}$ (recomputing the lowest-cost routes at each step), to find the set of all parameters that minimizes the residuals in equation (16).[22]

I use data on retail prices of 8 types of salt, observed annually from 1861-1930, in each of 124 districts to estimate equations (15) and (16). My goal is then to estimate the parameters δ and $\boldsymbol{\alpha}$ in these two equations, in order to measure the extent to which the railroad network reduced the cost of trading salt in British India. In next section I describe how I use these estimates to make further progress on estimating all of the model's parameters.

3.2 A Gravity Equation for Bilateral Trade Flows

In this section I describe how I estimate equation (10), a bilateral gravity-style equation for the value of trade between any two regions (whether within India, or outside of India.) Substituting the general (log) trade costs proxy discussed in the previous section, $W^k \ln TC_{odt}$, into equation (10) yields

$$\ln X^k_{odt} = \ln \lambda_k + \alpha^k_{od} + \phi^k_{od} t + \ln A^k_{ot} - \theta_k \ln r_{ot} - \theta_k \delta W^k \ln TC_{odt} + \theta_k \ln p^k_{dt} + \ln X^k_{dt} + \varepsilon^k_{odt}. \tag{17}$$

I estimate two versions of this equation, each with a different goal in mind. The first version attempts to test prediction 2 in a qualitative sense, by simply testing whether the construction of India's railroad network significantly increased trade in India. I perform this test while controlling for as much unobserved heterogeneity as possible, by estimating an equation of the form

$$\ln X^k_{odt} = \alpha^k_{ot} + \beta^k_{dt} + \gamma_{odt} + \mu^k_{od} + \rho W^k \ln TC_{odt} + \varepsilon^k_{odt}. \tag{18}$$

[22] Further details are in Appendix C. To ease the computational burden of this method I first partial out the fixed effects and trend terms in equation (16). Then minimizing the residuals of this equation is equivalent to maximizing the correlation between the adjusted (ie with fixed effects and trends partialled out) $\ln p^s_{dt}$ and the adjusted $W^S \ln OR_{sdt}(\boldsymbol{\alpha})$ measure.

In this specification, the term α_{ot}^k is an origin-year-good fixed effect, β_{dt}^k is a destination-year-good fixed effect, μ_{od}^k is an origin-destination-good fixed effect, and γ_{odt} is an origin-destination-year fixed effect. The coefficient ρ on the trade costs term $\ln C_{odt} W^k$ is therefore estimated purely from time variation in the railroad network (as it affects the optimal route from o to d) that affects each good differently proportional to its weight. Prediction 2 is that the coefficient ρ will be negative — that railroads (which lower trade costs) increase trade.

My second approach to estimating the bilateral trade flows gravity equation (17) aims to estimate all the parameters of the model using the structural gravity equation predicted by the model. To do this, I include a simple proxy for the productivity terms A_{ot}^k, which are unobservable.[23] My proxy is based on crop-specific rainfall shocks because water is a critical determinant of agricultural productivity in India.[24] Each crop has its own annual timetable for sowing, growing and harvesting, and the amount of rain that falls during these periods in a given region and year is therefore a crop-specific measure of shocks to that region-year's land productivity. I call this crop-specific rainfall measure $RAIN_{ot}^k$, and include it in equation (17) as follows:

$$\ln X_{odt}^k = \beta_{dt}^k + \mu_{od}^k + \kappa RAIN_{ot}^k - \theta_k \delta W^k \ln TC_{odt} + \varepsilon_{odt}^k. \tag{19}$$

In its present form, the parameters θ_k and δ on $\ln C_{odt} W^k$ cannot be separately identified in this specification. But the method proposed in section 3.1 estimates a parameter $\hat{\delta}^s = \delta W^s$ (where W^s is the weight per unit value of salt.) I therefore substitute this estimate from the salt price regressions into equation (19) and estimate

$$\ln X_{odt}^k = \beta_{dt}^k + \mu_{od}^k + \kappa RAIN_{ot}^k - \theta_k \hat{\delta}^s \frac{W^k}{W^s} \ln TC_{odt} + \varepsilon_{odt}^k. \tag{20}$$

I estimate this equation across 17 agricultural goods k.[25] Doing so delivers the parameters θ_k (for each of the 17 agricultural goods k), and the parameter κ, which completes the estimation of the model.

[23] This is because they represent the location parameter on region o's *potential* productivity distribution in year t. The productivities actually used at any point will be a strict subset of this potential distribution—trade allows region o to produce only its highest productivity draws, and import varieties that it can only produce with low productivity.

[24] I am not able to include non-agricultural goods in the method described here, for want of potential determinants of productivity in those sectors. But Appendix D presents results that exploit the fact that the fixed effects α_{ot}^k in equation (18) are related to A_{ot}^k through $\alpha_{ot}^k = A_{ot}^k - \theta^k r_{ot}$. Using this procedure, estimates of the model for all 85 goods can be obtained.

[25] In this specification, and others like it in this paper, I correct the standard errors in this regression to account for the presence of a generated regressor. In this case the regressor $\hat{\delta}^s \frac{W^k}{W^s} \ln TC_{odt}$ is generated because $\hat{\delta}^s$ was estimated in an earlier regression.

3.3 Agricultural Price Responsiveness to Rainfall

Prediction 3 from my theoretical framework is that, as a district is connected to the railroad network, agricultural goods prices in that district will be less responsive to local rainfall shocks and more responsive to distant rainfall shocks in the districts that are on the railroad network. I test this prediction in a linear estimation framework, which uses the proxy measures for T^k_{odt} (trade costs, proxied for by the railroad dummy and optimal route measures) and A^k_{ot} (productivity levels, proxied for by crop-specific rainfall shocks). Specifically, I estimate specifications of the form

$$\begin{aligned} p^k_{dt} = \alpha^k_d + \beta^k_t \quad &+ \quad \gamma_{dt} + \delta_1 RAIN^k_{dt} + \delta_2 RAIL_{dt} \times RAIN^k_{dt} + \chi_1(\tfrac{1}{N_d}) \sum_{o \in N_d} RAIN^k_{ot} \\ &+ \quad \chi_2(\tfrac{1}{N_d}) \sum_{o \in N_d} TC^k_{odt} + \chi_3(\tfrac{1}{N_d}) \sum_{o \in N_d} RAIN^k_{ot} \times TC^k_{odt} + \varepsilon^k_{dt}. \end{aligned} \tag{21}$$

In this specification, $RAIL_{dt}$ is a dummy variable that is equal to one in all years t when the railroad network is present in district d. The effect of neighboring districts' rainfall is captured in the two summation terms; I restrict these terms so that they only sum over districts within district d's neigborhood N_d, which I take be all districts within a fixed radius of d. The coefficients δ_1 and χ_1 should be negative, indicating that rainfall is a determinant of productivity (which reduces prices.) Prediction 3 concerns the coefficients δ_2 and χ_3: δ_2 should be positive (prices in district d should be less responsive to rainfall in district d if district d is on the railroad network); and the coefficient χ_3 should also be positive (lower transport costs make prices in district d more responsive to rainfall shocks in neighboring districts to d.)

3.4 The Welfare Effects of Railroads

The final two predictions from my theoretical framework concern the welfare effect of the construction of the railroad network in India. Prediction 4 states that a district's welfare (as measured by real income, in accordance with the model) will increase when it is connected to the railroad network, but that its welfare will fall as (holding its own access constant) one of its neighbors is connected to the railroad network. I test this prediction by estimating a specification of the form

$$\ln\left(\frac{r_{ot}}{\widetilde{P}_{ot}}\right) = \alpha_o + \beta_t + \gamma RAIL_{ot} + \phi(\tfrac{1}{N_o}) \sum_{d \in N_o} RAIL_{ot} + \varepsilon_{ot}, \tag{22}$$

where $RAIL_{ot}$ is a dummy variable that is equal to one in all years t in which some part of district o is on the railroad network. The inclusion of fixed effects α_o and β_t means

that the effect of railroads is identified entirely from variation within districts over time, after accounting for common macro shocks affecting all districts. Prediction 4 states that the coefficient γ on district o's own railway access should be positive, but the coefficient ϕ on district o's neighbors' access should be negative. The total effect of the railroad network on district o is the sum of the own-access term and the neighbors'-access terms.

While my theoretical framework makes predictions about the signs of these coefficients, it should be stressed that many other theories would make similar predictions about these signs. I therefore view the coefficients γ and ϕ as measuring the total, or 'reduced-form' impact of the railroad network on a district.

3.5 Share of Welfare Effect Explained by New Trade Opportunities

The final step of my empirical strategy compares the 'reduced-form' impact of the railroad network on each district to the impact that is predicted by my theoretical framework. This exploits prediction 5, which states that once rainfall is controlled for in the appropriate manner, the welfare effect of the railroad network on district o will be captured in (an appropriate sum of) its 'autarkiness' terms π^k_{oot}. To test this prediction I estimate the specification suggested by equation (12), but I also include all of the terms from equation (22):

$$\begin{aligned} \ln\left(\frac{r_{ot}}{\widetilde{P}_{ot}}\right) = \alpha_o + \beta_t \quad &+ \quad \gamma RAIL_{ot} + \phi(\tfrac{1}{N_o}) \sum_{d \in N_o} RAIL_{ot} \\ &+ \quad \rho_1 \sum_k \tfrac{\mu_k}{\theta_k} \ln A^k_{ot} + \rho_2 \sum_k \tfrac{\mu_k}{\theta_k} \ln \pi^k_{oot} + \varepsilon_{ot}. \end{aligned} \tag{23}$$

In order to calculate the last two terms in this equation, I use estimates of the model's parameters that were estimated using bilateral trade data in the second stage of my estimation strategy. These parameters are θ_k, the parameter κ relating rainfall to the unobservable productivity terms A^k_{ot}, and the parameter δ relating the transportation network to the unobservable trade costs terms T^k_{dot}. (I estimate the parameters μ_k from observed expenditure shares for all of India, as the model suggests is appropriate.) It should be stressed that I *compute* the terms π^k_{oot} using these estimated parameters and

the exogenous variables of the model.[26] In summary, I estimate the following equation:

$$\ln\left(\frac{r_{ot}}{\widetilde{P}_{ot}}\right) = \alpha_o + \beta_t \quad + \quad \gamma RAIL_{ot} + \phi(\tfrac{1}{N_o}) \sum_{d \in N_o} RAIL_{ot}$$
$$+ \quad \rho_1 \sum_k \frac{\hat{\mu_k}}{\hat{\theta_k}} \hat{\kappa} RAIN^k_{ot} + \rho_2 \sum_k \frac{\hat{\mu_k}}{\hat{\theta_k}} \ln \hat{\pi}^k_{oot} + \varepsilon_{ot}. \tag{24}$$

where a 'hat' over a parameter or π^k_{oot} indicates that it was estimated in a previous regression. Prediction 5 now states that the coefficients γ and δ will be zero, and the coefficients ρ_1 and ρ_2 equal one and minus one respectively; that is, the autarkiness term is a sufficient statistics for the welfare impact of the railroad network.

4 Description of the Data

In order to evaluate the impact of the railroad network on economic welfare in colonial India, and to test the predictions in section 2.3, I have constructed a new panel dataset on 239 districts that span most of colonial India. The dataset tracks these districts annually from 1861-1930, a period during which 98 percent of British India's current railroad lines were opened. In this section I describe briefly each of the variables in this dataset, as they relate to the empirical strategy laid out in section 3; more detail on the construction and sources of each of these variables is provided in Appendix A.

The data on *salt prices* that I use are retail prices for eight different types of salt, differentiated by their source location. Data on differentiated salt types are not available from Southern India, so only 125 districts are represented in the salt data sample.

I construct a new GIS database on the *transportation network* in India to construct measures of transportation costs between districts. This database represents the location of every railroad and navigable river in India, as well as its coastline (for coastal shipping), in each year from 1850 to 1930. Because a consistent series of road maps is unavailable throughout this time period, (and because the density of informal roads was extremely high), I include road transport routes in this GIS database by assuming that roads connect any two points in India along the straight (geodesic) line. Finally, I use the geographic centroid of each district to measure its location, based on a GIS database on the location of district boundaries (in 1891) that I constructed.

The *bilateral trade flows* data I use represent the sum of all trading activity on the

[26] While it would be possible in principle to use trade data to observe π^k_{oot} in the data, this faces three limitations: first, as the model makes clear, π^k_{oot} is endogenous to the error term in equation (24), so an instrumental variables methodology would be necessary; and second, the only internal trade data available from colonial India are presented at a more aggregated level, and begin in a later year, than the data on all other variables in equation (24).

railroads, major rivers and coastal shipping routes. Trade flows by the other mode of transportation, roads, were not recorded within India, but were recorded as goods left India. The bilateral trade flows data are available between regions of India known as 'trade blocks', and between these trade blocks and each of 24 other countries outside of India. Trade blocks are larger than districts; my sample consists of 45 of them, as compared to 239 districts. My dataset tracks 85 commodities, which include the 17 agricultural commodities for which I have price data, and an aggregate salt category. The trade data are available from 1880 to 1930 only.

The data on *agricultural good prices* that I use measure the retail prices of 17 agricultural goods. These commodities were chosen for their data availability and importance; the 17 agricultural commodities in my sample account for 92 percent of the cropped area of India in the middle year of my dataset, 1896.

The *agricultural output* data I use are available by crop for each of the 17 crops for which I have price data, and are measured in physical units of quantity. I use these crop-level output bundles to construct a measure of total nominal output (for each district and year) by valuing the real output bundle at contemporary, local prices. Finally, I construct a measure of real income (defined as nominal output divided by a consumer price deflator) as advocated by Diewert and Morrison (1986), Feenstra (2003) and Kehoe and Ruhl (forthcoming).

I construct a *crop-specific rainfall shock* variable that is a measure of the total amount of rainfall (in a district and year) that fell during each crop's ideal sowing and growing windows. To do this I use daily rainfall data from 3614 GPS-coded meteorological stations with rain gauges, and average these daily amounts of rain within each district. I then calculate the total amount of rain (in millimeters) that fell within the crop- and district-specific sowing and growing windows documented in the 1961 *Indian Crop Calendar* (Government of India, 1961).[27]

5 Results

To be written.

[27]This report describes the technology of agricultural practice (related to scheduling of activities) in each district. This particular aspect of agricultural technology is unlikely to have changed between my sample period and 1961. The earliest *Crop Calendar* that I was able to access is from 1908, but this presents data at larger geographic areas than the district. However, when I calculate crop-specific sowing and growing rainfall amounts for these larger geographic areas the correlation between these and the (area-weighted average) 1961 district-level amounts is 0.78. This is high, especially given that the district-level data is aggregated, indicating that the timing of agricultural activities has not changed dramatically, from 1908 to 1961 at least.

References

Ahuja, R. K., T. L. Magnanti, and J. B. Orlin (1993): *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall.

Alvarez, F., and R. E. Lucas (2007): "General equilibrium analysis of the Eaton and Kortum model of international trade," *Journal of Monetary Economics*, 54(6), 1726–1768.

Anderson, J. E. (1979): "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69(1), 106–116.

Anderson, J. E., and E. van Wincoop (2003): "Gravity with Gravitas: a Solution to the Border Puzzle," *American Economic Review*, 93(1), 170–196.

——— (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.

Andrabi, T., and M. Kuehlwein (2005): "Railways and Market Integration in British India," mimeo Pomona.

Black, W. R. (2003): *Transportation: A Geographical Analysis*. The Guilford Press.

Blyn, G. (1966): *Agricultural Trends in India, 1891-1947: Output, Availability, and Productivity*. University of Pennsylvania Press.

Chaney, T. (2008): "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *American Economic Review*.

Coatsworth, J. H. (1979): "Indispensable Railroads in a Backward Economy: The Case of Mexico," *The Journal of Economic History*, 39 (4), 939–960.

Deardorff, A. (1998): "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?," in *The Regionalization of the World Economy*, ed. by J. A. Frankel. University of Chicago Press (for NBER).

Deloche, J. (1994): *Transport and Communications in India Prior to Steam Locomotion: Volume I: Land Transport (French Studies on South Asian Culture and Society, No VIII)*. Oxford University Press.

——— (1995): *Transport and Communication in India Prior to Steam Locomotion: Volume 2: Water Transport (French Studies in South Asian Culture and Society VII)*. Oxford University Press.

Derbyshire, I. (1985): "Opening Up the Interior: The Impact of Railways upon the North Indian Economy and Society, 1860-1914," Ph.D. thesis, University of Cambridge.

Diewert, W. E. (2008): "Index Numbers," in *The New Palgrave Dictionary of Economics (Second Edition)*, ed. by S. Durlauf, and L. Blume. Palgrave.

Diewert, W. E., and C. J. Morrison (1986): "Adjusting Output and Productivity Indexes for Changes in the Terms of Trade," *The Economic Journal*, 96(3), 659–679.

Dinkelman, T. (2007): "The Effects of Rural Electrification on Employment: New Evidence from South Africa," mimeo Princeton.

Duflo, E., and R. Pande (2007): "Dams," *Quarterly Journal of Economics*, 122(2), 601–646.

Eaton, J., and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–79.

Evans, C. J., and J. Harrigan (2005): "Distance, Time, and Specialization: Lean Retailing in General Equilibrium," *American Economic Review*, 95(1), 292–313.

Feenstra, R. C. (2003): *Advanced International Trade: Theory and Evidence*. Princeton University Press.

Fishlow, A. (1965): *Railroads and the Transformation of the Ante-Bellum Economy*. Harvard University Press.

——— (2000): "Internal Transportation in the Nineteenth and Early Twentieth Centuries," in *Cambridge Economic History of the United States (vol II)*, ed. by S. L. Engerman, and R. E. Gallman. Cambridge University Press.

Fogel, R. (1964): *Railways and American Economic Growth*. Johns Hopkins Press.

——— (1979): "Notes on the Social Saving Controversy," *The Journal of Economic History*, 39(1), 1–54.

Galambos, J. (1978): *The Asymptotic Theory of Extreme Order Statistics*. Wiley-Interscience.

Helpman, Elhanan, M. M., and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*.

Heston, A. (1973): "Official Yields per Acre in India, 1886-1947: Some Questions of Interpretation," *Indian Economic and Social History Review*, 10(40, 303–334.

Hummels, D. (2007): "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, Summer 2007, 131–154.

Hurd, J. (1975): "Railways and the expansion of markets in India, 1861-1921," *Explorations in Economic History*, 12(30, 263–288.

——— (1983): "Railways," in *Cambridge Economic History of India (volume II)*. Cambridge University Press.

Jensen, R. T. (2007): "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector," *Quarterly Journal of Economics*, 122(3).

Kehoe, T. J., and K. J. Ruhl (forthcoming): "Are Shocks to the Terms of Trade Shocks to Productivity?," *Review of Economic Dynamics*.

Kortum, S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65(6), 1389–1419.

McAlpin, M. B. (1974): "Railroads, Prices, and Peasant Rationality: India 1860-1900," *The Journal of Economic History*, 34(3), 662–684.

——— (1980): "The impact of trade on agricultural development: Bombay presidency, 1855-1920," *Explorations in Economic History*, 17(1), 26–47.

McCallum, J. (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review*, 85(3), 615–623.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1720.

Melitz, M. J., and G. P. Ottaviano (2008): "Market Size, Trade, and Productivity," *Review of Economic Studies*, 75(1), 295–324.

Metzer, J. (1974): "Railroad Development and Market Integration: The Case of Tsarist Russia," *The Journal of Economic History*, 34(3), 529–550.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill — Evidence from the Interstate Highway System," *Review of Economics and Statistics*.

Moxham, R. (2001): *The Great Hedge of India*. Constable and Robinson.

O'Brien, P. (1977): *The New Economic History of the Railways*. Routledge.

Rauch, J. (1999): "Networks Versus Markets in International Trade," *Journal of Economic Literature*, 48(1), 7–35.

Samuelson, P. (1954): "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments," *The Economic Journal*, 64(2), 264–289.

Schwartzberg, J. E. (ed.) (1978): *A Historical Atlas of South Asia*. University of Chicago Press.

Summerhill, W. R. (2005): "Big Social Savings in a Small Laggard Economy: Railroad-Led Growth in Brazil," *The Journal of Economic History*, 65(1), 72–102.

Williamson, J. G. (1974): *Late Nineteenth-Century American Development: A General Equilibrium History*. Cambridge University Press.

World Bank (2007): *Evaluation of World Bank Support to Transportation Infrastructure*. World Bank Publications.

A Data Appendix

This appendix provides information (supplementary to that in section 4 on the data that I use in this paper.

A.1 Sample of Districts

The data I use in this paper cover the areas of modern-day India, Pakistan and Bangladesh –that is, most of the area known as British India.[28] The majority of British India was under direct British control, and was divided into nine large, administrative units known as provinces.[29] Each province was further sub-divided into a total of 226 districts, which are the units of analysis that I track from 1861 to 1930.[30][31] However, data was not consistently collected in three of these districts so I leave them out of my sample. Areas not under direct British control were known as 'Princely States'. For administrative purposes these were grouped into divisions similar to the provinces and districts described above, so in princely state areas I use the lower administrative units as my units of analysis and refer to them as districts, following the *Indian Census Administrative Atlas* (2002). There were 251 of these districts in my sample area, but data collection in the princely states was extremely incomplete and I include only 16 districts from the princely state regions in my final sample. To summarize, the data I use in this paper track 239 geographic units of analysis that I refer to as districts, for as much of the period 1861 to 1930 as possible.

District borders change over the 1861-1930 period (though only 18 of the 239 districts in my sample suffer from this in any way. Using the *Indian Census Administrative Atlas* I have made adjustments for this so that comparable geographic units can be tracked over time. When two districts split after 1891 (or, analogously, merge before 1891) I aggregate their data in the appropriate manner (summing variables such as areas or output, and calculating area-weighted averages for variables such as yields or prices). Likewise, when two districts merge after 1891 (or, analogously, split before 1891) I assign post-1891 values proportionately (by land area) to all post-1891 observations. However, my results do not change in any appreciable manner when I simply drop the 18 districts affected by these corrections.

A.2 Salt Prices

I use data on eight different types of salt, those from: the Bombay sea salt sources near the city of Bombay (which were shipped to the districts in my sample via the city of Bombay), undifferentiated salt distributed via Calcutta (predominantly imported salt from the United Kingdom), the Didwana salt source in Punjab, the Kohat mines in Punjab, the Mandi mine in Punjab, the Salt Range mines in Punjab (principally the

[28] I omit the areas of modern-day Afghanistan, Myanmar and Sri Lanka because I was unable to obtain a consistent time-series of data for these regions of British India.

[29] The three largest provinces were officially known as 'Presidencies', though for simplicity I use 'province' to refer to all middle-tier administrative units. The nine provinces in my sample (using borders and definitions as in the 1891 census) are: Assam, Bengal, Berar, Bombay, Central Provinces, Madras, Northwestern Provinces, Punjab and Sind.

[30] The agricultural year in India runs from May to April, due to the timing of the summer monsoon (when the sowing of most crops begins). For this reason, many official statistics publications in India published annual observations relating to the agricultural, rather than the calendar, year. In these cases I take the year in question to be the earlier of the two calendar years inside the given agricultural year because this is when the bulk of agricultural production occurs.

[31] The district is the smallest administrative unit for which data on my variables of interest was consistently published in colonial India.

Mayo mine), the Sambhar Salt Lake in Rajputana, and the Sultanpur source in the Central India Agency. My sample includes Assam, Bengal Presidency, Berar, Central Provinces, Punjab and United Provinces. Official statistics in Southern India made no attempt to distinguish salt by region of origin because there were many salt sources and their products were not easily distinguishable. Publications with salt price data were produced by each province and salt department region.[32]

A.3 Trade Cost Proxies

I construct two proxies for the unobserved cost of trading between any two districts, in any year. Both of these proxies are based on a newly constructed GIS database on the Indian transportation network, from 1851 to 1930. The database covers four modes of transportation: railroads, roads, rivers and coastal shipping. I began with a GIS database that contains the locations of contemporary railroad, river and coast lines from the *Digital Chart of the World*; this is a declassified CIA cartographic resource known for its high spatial accuracy. I then divided the contemporary railroad lines in this database into segments approximately 20 km in length and coded each segment according to the year in which it was opened using the statements in several official publications.[33] For river transport I keep only those rivers (of the contemporary rivers reported in the *Digital Chart of the World*) that are reported in Schwartzberg (1978) as navigable in 1857. I assume that the location of navigable rivers in India did not change from 1857 to 1930.

The final component of the colonial India GIS database that I construct is the location of each district (as well as the location of each salt source that features in my analysis of salt price data.) To calculate these locations I construct a digital map of the district borders in India (as they existed in 1891), from which I calculate district centroids using the centroid feature in ArcGIS. I used the maps in the *Indian Census Administrative Atlas* and *Constable's Hand Atlas of India* (Bartholomew, 1893).

I then convert the GIS database of transportation lines (railroad, river and coast lines) and district/salt source locations into a graph of nodes and links, where the number of nodes and the sparsity of links was low enough for network algorithms to be feasibly operated on it using a desktop computer. To do this, I use the 'simplify' command in ArcGIS to reduce the transportation system to a computationally-manageable number of nodes (the network has 7651 nodes in total).[34] Because I was unable to find a consistent series of road maps (and because the density of informal roads was extremely high (Deloche, 1993)), I allow road transport to occur along the straight line between any two nodes

[32]These publications were: the provincial *Administration Reports*, listed in the agricultural prices data section below, for the provinces of Assam, Bengal, Bihar and Orissa, Central Provinces, Northwestern Provinces, and Punjab; the *Salt Report of Northern India*, which covered the provinces of Central Provinces, Northwestern Provinces and Punjab; and the *Bengal Salt Report*. I use salt data from 1861 to 1930.

[33]These publications are: *History of Indian Railways, Constructed and in Progress* (Railway Board of India, 1918 and 1966), the 1966 volume of which refers to railway lines in modern-day India only; to obtain years of opening for line segments in modern-day Pakistan and Bangladesh from 1919 to 1930 I use the annual *Report on Indian Railways* (Railway Board of India, various).

[34]A line in ArcGIS is a series of connected vertices connected by straight lines. The 'simplify' command removes vertices in such a way as to minimize the sum of squared distances between the original line and a the simplified line. The original *Digital Chart of the World* railway layer, for example, consists of approximately 33,000 vertices; I simplify the railway layer to one of only 5616 vertices.

on the network. This assumption, taken literally, would yield a network in which travel along each of the over 58 million links (ie one for each pair of nodes) was permitted. The shortest path between each of the nodes on such a dense network cannot be calculated using desktop computing technology, so I restrict many of these links to be non-existent; the result is that the 7651-by-7651 matrix representing the network can be stored as a sparse matrix, and analyzed using sparse matrix routines (that increase computation speed dramatically). I restrict all 'road' links between nodes that are more than 1000 km apart, apart from the nodes that represent districts whose direct links are unrestricted. The result is a network with 7651 nodes, 5616 of which represent the railroad network, 660 of which represent the navigable river network, 890 of which represent coastal shipping routes, 477 of which represent the centroids of the 477 districts in India (in 1891 borders), and 8 of which represent the locations of the sources of 8 different types of salt. Because the railroad nodes are coded with a year of opening indicator, this network can be restricted to represent the transportation network for any year from 1851 to 1930.

Finally, I use this network representation of the Indian transportation system to calculate the two trade cost proxies described in section 3.1. The first such proxy (the second introduced in 3.1, for reasons that will become clear) is a measure of the cost of traveling between each pair of districts (or, in the case of salt sources, the districts containing each of the each salt sources) in a year using the lowest-cost route along the network (available in that year). In order to compute the lowest-cost route along a network that features different modes of transport, each with its own per-unit-distance cost, it is necessary to specify the *relative* costs of using these different modes of transportation. I normalize the cost of using each mode such that the cost (per unit distance) of using the railroads is one, and then specify the other modes' (relative) per unit distance costs as unknown parameters, α^{road}, α^{river} and α^{coast}. Conditional on values of these three parameters, I use a standard algorithm from graph theory and transportation science (Dijkstra's algorithm) to calculate the shortest path between every pair of districts, along the transportation network available in each year from 1861 to 1930.[35] The resulting measure is in units of railroad-equivalent kilometers.[36]

The second proxy for trade costs is a simple dummy variable that indicates when it is possible to travel between two districts without leaving the railroad network. I use the digital map of district borders described above to place each of the 7651 nodes in the transportation network into its appropriate district. I then set α^{road}, α^{river} and α^{coast} to very high numbers (I use 1,000,000) and calculate the cost of travelling between each pair of nodes in each year. Finally, I calculate the smallest node-to-node distance among all the node pairs within a given district pair (the two nodes in a pair need to be in different districts). If this smallest node-to-node distance is lower than 1,000,000 then it must be the case that those two nodes can be connected by using railroad links (those whose cost per kilometer is just one), and hence there is a way to travel between the two nodes' districts using the railroad network alone.

[35] I implement this algorithm in Matlab; for the thickest network of nodes (that relating to 1930) it takes just over one minute to calculate the shortest path costs for every district pair (using an Intel Core 2 Duo CPU, 2.4 GHz.

[36] For example, suppose that the optimal route from A to B is to take the road, along the straight line between them, and this straight line distance is 100 km. Then, if road travel is four times more costly per kilometer than railroad travel, the optimal route trade cost proxy for the journey A to B is 400 railroad equivalent kilometers.

A.4 Bilateral Trade Flows

The data I use on bilateral trade flows were collected from a variety of different sources, each relating to a different mode of transportation. I describe each of these modes in turn, and then how they were combined into aggregate data on trade flows.

Data on railroad trade within India were published separately for each province. The geographic unit of analysis in these data is the 'trade block', which represents between four and five districts. Trade blocks split into smaller blocks over time, but I aggregated over these splits to maintain constant geographic units. The trade blocks were always drawn so as to include whole numbers of districts, and I used maps of the trade blocks to match each district to a block. All bilateral block-to-block intra-provincial trade flows were published, except that from a block to itself (which was always unreported.) Inter-provincial trade flows were published from each internal block to each external province (and vice versa), but not by trade block within the external province. I therefore create a full set of inter-provincial block-to-block flows by following an analogous procedure to that used to prepare bilateral trade data on provincial-state trade between Canada and the United States in McCallum (1995) and Anderson and van Wincoop (2003).[37] In order to match the internal block-level railroad trade data to international trade data (leaving via specified ports, as described below), I apply a similar proportionality method. This is possible because the railroad trade data differentiate railroad trade to/from principal ports (in each province) from trade bound for non-port consumption. All of the railroad trade flow data represent final shipments between two regions (even if a shipment changed railroad companies). Only if a shipment was taken off the railroad system (presumably for transformation) and re-shipped onwards would it be counted as two separate shipments. Railroad trade statistics were not published by the princely states themselves, but each province's external trade to/from each of the large princely states were published. I therefore treat each of these large princely states (Central India Agency, Hyderabad, Mysore, Rajputana and Travancore) as a single trade block. The railroad trade data were published (in the format described here) in various annual, provincial publications from 1880 onwards (with the exception of Madras Presidency, where they began in 1909.)[38]

Data on river-borne trade within India was published in a similar manner to the railroad trade data, but only for the Ganges and Brahmaputra river systems. The other major navigable river system in India on which significant amounts of trade occurred is the Indus, but no river-borne trade data were published there. The river-borne trade

[37] This is best explained by example. In the case of the exports from trade block X (in province A) to trade block Y in province B, I observe only the total exports from block X to all of province B. But using the inter-provincial trade published by province B, I observe the total imports of block Y from all of province B. I then infer the trade flows from block X to block Y as a share of block X's total trade to province A, where the share is the fraction of province B's total imports from province A that were received by block Y.

[38] These publications are: *Returns of the Rail and River-borne Trade of Assam* (Assam Revenue Department, various)*Returns of the Rail and River-borne Trade of Bengal* (Bengal Presidency, various); *Report on the trade carried by rail and river in Bihar and Orissa*, (Bihar and Orissa Office of the Director of Industries, various); *Report on the Rail-Borne trade of the Bombay Presidency, exclusive of Sind* (Bombay Presidency Customs Department, various); *Report on Inland Trade* (Central Provinces, various); *Review and Returns of the Rail-borne Trade of the Madras Presidency* (Madras Presidency Board of Revenue, various); and *Annual Report on the Inland Trade* (United Provinces of Agra and Oudh Department of Industries and Commerce, various).

data were included with the railroad trade publications, for two provinces (Assam and Bengal) from 1880 to 1920.[39] But these publications also document river-borne trade along the Ganges and Brahmaputra river systems beyond these two provinces.

Data on trade within India that occurred via coastal shipping were published by each of the coastal provinces (Bengal, Bombay, Madras and Sind) in a similar manner to the railroad trade data.[40] These data are presented at different geographic scales in each province. Madras presented imports and exports to/from each of the 20 largest ports in the province, and I aggregate these up to the railway trade block level. Trade flows leaving Madras were divided into those going to each of Bengal's, Bombay's and Sind's major ports and the sum of their minor ports. Bengal, Bombay and Sind used more geographically aggregated regions. Each presented flows between the major port in their own province (Calcutta, Bombay and Karachi respectively), the sum of all minor ports in their own province, and also to the sum of all ports in each of the other three provinces. I converted this coastal trade information into a full bilateral set of flows between any two coastal trade blocks by using a proportionality assumption similar to that described for the railroad trade data above.

Data on international trade leaving India are available by maritime shipping, and by roads. Each province published its own maritime international trade statistics, with each reporting the trade to and from its major and minor ports.[41] These international maritime trade flows were presented disaggregated into over 30 foreign countries, but to maintain consistent geographic units over time I aggregated them into 24 foreign regions. Foreign trade by land occurred between the provinces of Bengal, Northwest Provinces and Punjab and neighboring foreign countries (modern-day Nepal, China, Afghanistan and Bhutan). These trade flows were published by each of these provinces, disaggregated by the border post through which trade left or arrived.[42] I assign each of these border posts to the internal trade block in which it is located, and assume that all of the (extremely minimal) foreign land trade came to/from these blocks only.

Trade flows using all modes of transport discussed above were published disaggregated by goods. The railroad and river-borne trade data featured 85-100 goods (depending on the year and province), the coastal shipping data 200-400 goods, and the international

[39]These publications are: *Returns of the Rail and River-borne Trade of Assam* (Assam Revenue Department, various)*Returns of the Rail and River-borne Trade of Bengal* (Bengal Presidency, various); and *Report on the trade carried by rail and river in Bihar and Orissa*, (Bihar and Orissa Office of the Director of Industries, various). The new province of Bihar and Orissa was created out of part of the Bengal Presidency in 1911.

[40]The coastal trade data were published in: *Annual Statement of the Sea-borne Trade and Navigation of the Madras Presidency, etc*, (Madras Customs Department, various); *Annual Statement of the Trade and Navigation of the Presidency of Bombay*, (Bombay Presidency, various); *Report on the Maritime Trade* (Bihar and Orissa, various); and *Report on the Maritime Trade of Bengal* (Bengal Presidency Customs Department, various).

[41]The maritime international trade data were published in the same publications as those containing the coastal trade data, described above.

[42]The overland international trade flows were published in: *Annual Report on the Trans-frontier Trade of Bihar and Orissa with Nepal* (Bihar and Orissa Office of the Director of Industries, various); *Bengal Frontier Trade: Trade of Bengal with Nepal, Tibet, Sikkim and Bhutan* (Bengal Presidency, various); *Accounts Relating to the Trade by Land of British India with Foreign Countries* (India Commercial Intelligence Department, various); *Annual Report on the Foreign Trade of the United Provinces* (United Provinces of Agra and Oudh Department of Industries and Commerce, various); and *Report on the External Land Trade of the Punjab* (Punjab Revenue Department, various).

maritime shipping data well over 400 goods. In order to compare goods across these different levels of aggregation, I aggregated all data to the 85-good level. This was possible because a similar classification system was used in all provinces (and it did not change significantly over time), and because the disaggregate data was grouped into commodity groups roughly at the 85-good level.

I aggregate the trade data on each of the modes (for each good separately) into one trade dataset. This is possible because all of the trade data are in nominal units (of rupees per physical quantity). Wherever relevant, I treat the regions of modern-day Afghanistan, Myanmar and Sri Lanka as foreign countries, since they are outside of the region on which I have other data from India. All of the above trade data are available from at least 1861 onwards, to 1930 (and beyond), except for the railroad trade data. The railroad trade data only starts in a coherent manner in 1880, and was discontinued for budgetary reasons (in most provinces) in 1920. I therefore use bilateral trade data from 1880 to 1920 only.

A.5 Crop-specific Rainfall Shocks

A thick network of 3614 rain gauges (spread throughout India) at metorological stations recorded daily rainfall amounts from 1891-1930. The data for the region of modern-day India, from 1900 onwards, have been digitized by the Global Historical Climatology Network (Daily) project; the GHCN (daily) dataset also provides the latitude and longitude of each station. I digitized the data, and looked up all station latitudes and longitudes, in years and regions outside of the GHCN (daily) dataset using the original publication *Daily Rainfall for India in the year...* (Indian Meteorological Department, various). In the years 1865 to 1890, very little daily rainfall data was published in colonial India, but monthly data from 365 stations (spread throughout India) were published by each province.[43] I convert monthly station-level data to daily station-level data using a procedure that is common in the meteorological statistics literature: using daily data from 1891 to 1930, I estimate the district-specific relationship between the pattern of monthly rainfall in a year and the rainfall on any day of that year; I then use these estimated relationships to predict the rainfall on any day in a given district and year from 1865 to 1890, conditional on the pattern of monthly rainfall actually observed in that district and year.[44] Most districts have multiple stations reporting rainfall every day; in these cases I take a simple average over every day's reported values in each district; when only one station reports data on a given day in a district, I use that observation. When a given district-day has no reported rainfall observations I use an inverse distance-weighted average of that day's rainfall in the 5 closest reporting stations.

The 17 crops in my agricultural prices and output data have heterogeneous water requirements because each is sown and harvested at different times of the year. The calendar dates of these sowing and harvesting periods also vary from district to district,

[43]These publications included the *Administration Reports* for each province, described in the agricultural price data section above. I use additional data (to increase the number of stations) that was published in selected provinces' *Sanitary Reports*.

[44]While these daily rainfall predictions are likely to be imprecise, much of the imprecision will be averaged over when I construct crop-specific rainfall shocks, which are measures of the total rainfall in a given period (of length 55 to 123 days.)

because of differences in the timing of the onset of the south-west monsoon (which affects the sowing dates), and because local soil conditions will affect the length of the growing period (which therefor changes the harvest date, conditional on a sowing date). The 1961 *Indian Crop Calendar* reports the dates of district- and crop-specific sowing and harvest windows. I use these dates to convert a district's rainfall on every day of a year into crop-specific measures of total rainfall during the sowing and growing periods of each crop, for that district and year. I match the 1961 districts in this publication to 1891 districts using the correspondence tables in the *Indian Census Administrative Atlas* (2002); when two or more 1961 districts lie inside an 1891 district I take the average of their sowing and harvesting dates.

A.6 Agricultural Prices

The 17 agricultural goods in my sample are: bajra (pearl millet), barley, bengal gram (chickpeas), cotton, indigo, jowar (sorghum), kangni (Italian millet), linseed, maize, opium, ragi (finger millet), rape and mustard seed (which were treated as one crop in all official statistics), rice, sesamum, sugarcane, tur (pigeon pea) and wheat.

Price observations are typically obtained from either the 'headquarters' (ie capital) town in each district, or an average of market towns in each district. Observations were taken by district officers once per fortnight and then averaged to produce an average price level for each district and year. Throughout this paper I use the term 'good' (indexed by k) to refer to the products names in price (as well as output and trade) statistics. As in my theoretical framework, these goods are available in different varieties, but district officers were unqualified to distinguish these different varieties. I therefore assume that the price observations in the data represent a random sample of varieties chosen from each good. This implies that the price data observed for good k in a district d is equal to p_d^k (the expected value of the price, over randomly sampled varieties of good k that are actually sent to district d) in the model. Price statistics were published in a number of different sources.[45] I exploit all of these sources in order to collect the widest possible set of districts, commodities and years, and to average over any possible measurement error in collection.[46]

[45]These publications were: *Prices and Wages in India* (India Department of Statistics, 1901 and 1922), with price data from 1861 to 1921; *Administration Reports* from all provinces, with price data from 1868 to 1901; (Government of Assam Province, various; Government of Bengal Presidency, various; Government of Bombay, various; Government of Central Provinces, various; Government of Hyderabad Assigned Districts, various; Government of Madras Presidency, various; Government of Mysore State, various; Government of Punjab Province, various; Government of United Provinces of Agra and Oudh, various); *Statistical Atlas of Andhra State* (Andhra State Bureau of Economics and Statistics, 1956) with price data (for the Madras Presidency) from 1875 to 1930; the *Season and Crop Reports* from various provinces () with price data from 1902 to 1930; and the *Sanitary Reports* from various provinces () with price data on food grains from 1867 to 1898.

[46]While I average over any multiple observations within a district-crop-year, I have experimented with specifications that use non-averaged data; in these regressions, the coefficient on statistical publication indicator variables is never significantly different from zero, indicating that any variation across statistical publications is orthogonal to my regressors of interest (conditional on the fixed effects used).

A.7 Agricultural Output

Systematic data on output in colonial India are limited to the agricultural sector. I use data that present the area under each of 17 crops (the 17 for which price data are available), and the yield per acre for each of these crops, in each district and year.[47] I take the product of each area and yield pair to create a measure of real output for each crop, district and year. I then evaluate this bundle of 17 real output measures at the prices prevailing for these crops (from the agricultural price data described above), in each district and year, to create a measure of total nominal agricultural output for each district and year. Finally, I divide nominal output by a consumer price index to create a measure of real output. I use is the chain-weighted, ideal Fisher index as the consumer price index, which has a number of attractive properties (Diewert 2008). In order to compute this (or any commonly used) consumer price index I use district and year specific consumption weights from the internal trade data, computing consumption as output minus net exports. These consumption weights are only available at the trade block level, so I apply the same consumption weight to each district within the same trade block. While a large literature has discussed the potential for large measurement errors in these sources (for example, Blyn (1966), and Heston (1973)), this measurement error is unlikely to be correlated with the regressors I use in this paper.

[47] These data were published in *Agricultural Statistics of India* (India Department of Commercial Intelligence and Statistics, various) from 1884 to 1930. For the years 1870-1883 I use data on crop areas and yields in the provincial *Administration Reports*, as described in the agricultural prices data section above. Data on agricultural output was published in each province's *Administration Report* except for Punjab and Bengal. For supplementary data I use each province's *Season and Crops Report*.