Railroads and the Demise of Famine in Colonial India *

Robin Burgess
LSE and NBER

Dave Donaldson
MIT and NBER

September 2012

Abstract

Whether openness to trade can be expected to reduce or exacerbate the equilibrium exposure of real income to productivity shocks remains theoretically ambiguous and empirically unclear. In this paper we exploit the expansion of railroads across India between 1861 to 1930—a setting in which agricultural technologies were rain-fed and risky, and regional famines were commonplace—to examine whether real incomes became more or less sensitive to rainfall shocks as India's district economies were opened up to domestic and international trade. Consistent with the predictions of a Ricardian trade model with multiple regions we find that the expansion of railroads made local prices less responsive, local nominal incomes more responsive, and local real incomes less responsive to local productivity shocks. This suggests that the lowering of transportation costs via investments in transportation infrastructure played a key role in raising welfare by lessening the degree to which productivity shocks translated into real income volatility. We also find that mortality rates became significantly less responsive to rainfall shocks as districts were penetrated by railroads. This finding bolsters the view that growing trade openness helped protect Indian citizens from the negative impacts of productivity shocks and in reducing the incidence of famines.

*Correspondence: `r.burgess@lse.ac.uk` and `ddonald@mit.edu` We thank Richard Blundell, Chang-Tai Hsieh, and seminar participants at Bocconi University and the 2012 Nemmers Prize Conference (at Northwestern) for helpful comments. This document is an output from research funding by the UK Department for International Development (DFID) as part of the iiG, a research programme to study how to improve institutions for pro-poor growth in Africa and South-Asia. The views expressed are not necessarily those of DFID.

1 Introduction

In recent decades, regions within countries and countries within the world economy have become more integrated due to reductions in the costs of trading. A large empirical literature has aimed to estimate the impact of trade cost reductions on the *level* of real incomes both across countries (Frankel and Romer (1999), Feyrer (2009b) and Feyrer (2009a)) and within countries (Donaldson 2010). But in many developing countries, large segments of the population continue to derive income from risky production technologies—the case of rain-fed agricultural incomes is perhaps the most obvious. An important, but unanswered, empirical question therefore concerns whether trade cost reductions—for example, due to trade liberalization or investment in transportation infrastructure—will reduce or increase the *volatility* of real incomes, that is the extent to which productivity shocks affect output and prices.

Theory is ambiguous on this matter. Openness makes nominal incomes more responsive to production shocks (due to both increased specialization and dampened offsetting price movements), but consumer prices less responsive, such that the net effect on real incomes is unclear.[1]

To shed new empirical light on this issue we turn to the case of colonial-era India during the period of the construction of its vast railroad network, from 1861 to 1930 (see Donaldson (2010)). In this setting, a majority of Indian households were engaged in rain-fed agriculture, an activity that was likened to a 'gamble in monsoons.' Widespread famine was a perennial occurrence. The effect of openness on volatilty was therefore—as we shall see—a matter of life and death. A significant advantage of studying real income volatility in this agricultural setting is that the underlying source of productivity volatility, rainfall variation, is directly observed. We are therefore able to trace through the effects of an exogenous change in the amount of rainfall on a series of equilibrium variables in the economy: prices, output, trade flows, income, and the mortality rate (our proxy for per capita consumption).

The end goal is to improve our understanding of how production shocks map into real income and how trade openness alters this mapping. To do so we exploit the differential arrival of rail transport in each district. Railroads dramatically reduced trade costs and increased trade between Indian districts and between India and the outside world (Donaldson 2010). Districts that had been largely closed economies opened up as they were penetrated by railroads. We compare empirically the extent of equilibrium *responsiveness* of a given outcome variable in a district (say Y_d) to exgenous rainfall in that district (denoted by A_d)

[1]Newbery and Stiglitz (1981) present a wide range of models in which openness to trade can either increase or decrease real income volatility.

before and after a reduction in trade costs (denoted by τ_d) due to the arrival of a railroad line in that district. That is, we study the empirical analog of the cross-derivative

$$\frac{d}{d\tau_d}\left|\frac{dY_d}{dA_d}\right|,$$

where we define $\left|\frac{dY_d}{dA_d}\right|$ as the equilibrium responsiveness of outcome Y. Our empirical findings—which we believe follow a natural sequence in tracing through the effects of a productivity shock on economic well-being, and are ugided by a multi-region general equilibrium Ricardian model of trade due to Eaton and Kortum (2002)—can be summarized as follows:

1. *Openness reduces the responsiveness of local prices to local productivity shocks:* Rainfall has a large (in absolute value) effect on agricultural prices before the arrival of railroads in a region, but that the dependency of local prices on local rainfall falls to almost zero after the arrival of railroads in that region (even when focusing purely on rainfall variation across crops, within a district and year). This implies that railroads brought India's district economies close to the small open economy limit where local conditions have no effect on local prices.

2. *Openness increases the responsiveness of local nominal incomes to local productivity shocks:* In a closed economy, a negative productivity shock in one sector reduces the physical output from that sector, but local prices are likely to rise and thereby off-set the effect of the shock on nominal incomes. In a perfectly small and open economy, however, there can be no such price adjustment. The effect of an equivalent shock, therefore, on nominal incomes is larger in an open economy than in a closed economy. We find empirical support for this logic, in that railroads increase the extent to which local nominal agricultural incomes respond to local rainfall shocks.

3. *Openness reduces the responsiveness of local real incomes to local productivity shocks:* The net effect of railroads on the responsiveness of real incomes to productivity shocks will, in general, depend on the strength of the weakened price responsiveness (result 1) relative to that of the heightened nominal income responsiveness (result 2). We find that the first effect dominates the latter and hence that openness reduces real income responsiveness in our setting.

4. *Local net exports respond posititvely to local productivity shocks:* Consistent with result 1 concerning reduced price responsiveness due to trade openness we find an equivalent result concerning quantities: when a region experiences a negative productivity shock

in a product it imports more of that product. In addition, the total value of all agricultural net exports responds positively to productivity shocks, whereas non-agricultural exports respond negatively.

5. *Openness reduces the responsiveness of local mortality rates to local productivity shocks:* The extent to which reduced real income responsiveness passes through to reduced consumption responsiveness will depend on rural residents' abilities to borrow, save and obtain insurance. But what is clear is that if trade openness reduces real income responsiveness (as found in result 3) then it should also reduce consumption responsiveness. We lack data on consumption, so we cannot test this prediction directly. But in this low-income, poor-health environment it is likely that mortality rates are correlated with consumption levels and hence the mortality responsiveness to rainfall shocks should fall as regions open to trade. This is precisely what we find. Indeed, our results suggest that railroads brought the responsiveness of mortality to rainfall shocks from a very high level to one that is virtually non-existent. This is strong evidence consistent with the argument that railroads dramatically mitigated the the scope for famine in India.[2]

6. *The effect of openness on real income responsiveness can be well accounted for by a model-based sufficient statistic for real income:* In the Eaton-Kortum model of trade (and more generally, as in Arkolakis, Costinot, and Rodriguez-Clare (2012)), real income in any location and time period is a function of the location's exogenous productivity level and the location's endogenous 'trade share', ie the share of the location's expenditure that is produced in that location. In this model, therefore, the trade share is a sufficient statistic for the extent to which trade openness can reduce the responsiveness of real income to a local productivity shocks. Following a similar strategy to that in Donaldson (2010) we explore this empirical prediction by comparing the extent to which railroads reduced real income responsiveness (as in result 3) with and without an empirical control for this sufficient statistic. We find that the reduced-form effect of railroads on income responsiveness (result 3) becomes small and statistically

[2]This evidence is buttressed by the finding that railroads also reduced the extent to which a qualitative index of 'famine intensity' (based on official declarations compiled by Srivastava (1968)) that we reported in Burgess and Donaldson (2011). In contrast to that earlier work, the present paper aims to trace through the impacts of railroads on equilibrium volatility of real income levels, by understanding how railroads affected local price responsiveness, nominal income responsiveness, and finally real income responsiveness, to local (climatically-induced) productivity shocks. We also examine a quantitative measure of mortality (the crude death rate) which overcomes many of the drawbacks of Srivastava's qualitative work. The focus of the present work is therefore focused squarely on how openness can mitigate equilibrium economic volatility caused by productivity shocks in general, not on a relatively small number of discrete events that were officially or unofficially declared as 'famines'.

insignificant once the sufficient statistic is included, while the coefficient on the sufficient statistic is large and statitstically significant. This finding is consistent with an interpretation of our results in which railroads reduce real income volatility through the reduction in the costs of trading goods.

In short, these results suggest that the lowering of transportation costs via investments in transportation infrastructure played a key role in raising welfare by lessening the degree to which climatic shocks translated into income volatility. Mortality rates also become less responsive to rainfall shocks as districts are penetrated by railroads which bolsters the view that growing trade openness helped protect Indian citizens from the negative impacts of climatic shocks. The arrival of the railroads, in effect, made the lives of Indian households during the colonial period less risky due to their ability to export excess food output in times of plenty and import surfeit food output in times of scarcity.

Even in the modern era many areas of the developing world remain poorly integrated both domestically and internationally. And citizens in these remote regions of the developing world often remain exposed to the vicissitudes of weather for their livelihoods due to their dependence on subsistence agriculture. Our results suggest that in these settings the building of transportation infrastructure can significantly improve welfare by lessening the extent to which climatic shocks translate into real income volatility. This paper therefore documents the potential for improvements in a nation's transportation infrastructure to improve not only the level of real incomes (as shown in Donaldson (2010)) but also the volatility of real incomes.

Our results on mortality underline how critical trade openness may be to supporting rural livelihoods settings where climatic shocks are frequent. Our period (1861-1930) is one where famines were relatively frequent—the eleven officially declared famines that occurred in colonial India from 1861 to 1906 killed in the range of 15-35 million people (out of a population of around 150 million). Severe famine was a persistent feature of economic life in colonial India, but the fact that these calamities ceased—at least in peacetime—after 1906, just as railroad construction was at its zenith is suggestive that railroad expansion through the mechanisms examined in this paper may have played a role in protecting Indian citizens from the vagaries of climate.[3]

[3] A major exception to this statement is the 1943 famine in Bengal. But, as argued by O'Grada (2008), it is plausible that the war-time conditions that were commensurate with this calamity were an extreme case. O'Grada (2008) argues, in particular, that because Bengal's railroad lines and rolling stock were being diverted to the war effort, market integration withing Bengal suffered considerably. In addition, Burma, a major rice supplier to Bengal, was captured by Japan in early 1942 and all trade between the regions was suspended—though the plausibility of this claim is disputed by Sen (1981). Finally, the 1943 rice harvest fell victim to a rare infestation of brown spot disease, which Padmanabhan (1973) and Tauger (2003) argue was

This work relates to a nascent empirical literature on the effects of openness on volatility. di Giovanni and Levchenko (2009) explores the determinants of nominal output volatility in a cross-sectional, cross-country setting. A natural concern with such an exercise is that it is difficult to isolate exogenous variation in the extent of a country's openness to international trade; our within-country, panel setting, based on the plausibly exogenous arrival of railroads (exploiting the colonial motives for railroad placement that were largely militarily-oriented—see Donaldson (2010)) removes some of these concerns. Also related is work that, following the theoretical arguments in Rodrik (1997), explores the volatility of nominal wages in open economies. Hasan, Mitra, and Ramaswamy (2007) document empirical evidence from post-independence India that is consistent with higher trade openness leading to larger elasticities of labor demand, and Krishna and Senses (2009) find that nominal labor incomes are indeed relatively more volatile in the United States in sectors that are relatively more open to trade. Finally, as we do in this paper, recent work by Caselli, Koren, Lisicky, and Tenreyro (2011) uses an Eaton-Kortum model to explore how openness could affect real income volatility in a calibrated global economy. A distinguishing feature of our work, relative to the above papers, is that rather than studying equilibrium volatility of an outcome variable (such as real incomes), we study the response of an outcome variable to a exogenous, stochastic productivity (rainfall) shocks. The advantage of such an approach is that we can trace the equilibrium relationship between an underlying cause of volatility (in our case, climatic variation) and the volatility of an outcome variable of interest.

A second body of work to which this paper relates is a vast literature that aims to understand the economic and political causes of famine. (See, for example, Dreze (1988), Dyson (1991), McAlpin (1983) and Maharatna (1996) for the case of India, and O'Grada (2010), Ravallion (1997) and Sen (1981) more generally.) However, it is important to stress—as we describe more fully below—that India's colonial government took a staunchly *laissez-faire* approach to famine prevention throughout our time period. Our work, which relies on comparisons over time within India over our sample period, therefore has little relation to the interpretations of certain famines as events driven by political conflict (de Waal 1989, Dreze and Sen 1989) or due to the institutional features of a planned economy (Meng, Qian, and Yared 2011).

The remainder of this paper proceeds as follows. The next section describes the setting of the colonial era in which the empirical exercises of this paper are conducted; the emphasis here is on the elements of this setting that concern volatility, mortality and famines. Section 3

the extraneous circumstance that gave rise to this anomalous and anachronistic famine. Another potential exception occurred in the state of Maharashtra in 1972-73; O'Grada (2007) argues that this was a famine that killed 130,000 people.

outlines a general equilibrium trade theory that delivers predictions about the responsiveness of observable variables to rainfall shocks in India, and how this responsiveness is predicted to change when railroads arrive in India. Section 4 then reports on empirical results that are motivated by the model's predictions. Finally, Section 5 concludes.

2 Background and Data

In this section we describe some of the essential background features of colonial India's era of famines, and the economic conditions that gave rise to these extreme events. We also describe the data we have collected in order to shed new empirical light on the relationship between trade openness and real income volatility from this unique setting.[4]

2.1 Rainfall and Agricultural Incomes in India

Rainfall in India was extremely volatile from year to year and only 12 % of cultivated land was irrigated in 1885. The volatility of rainfall and the rain-fed nature of the vast majority of agricultural production gave rise to the common description of colonial Indian agriculture as a"gamble in monsoons."[5]

To shed light on this climatic volatility we use the measure of crop-specific rainfall (the amount of rainfall that fell in a district and year during which a given crop was under the soil) introduced in Donaldson (2010) and explained in more detail below. Table 1 (which contains summary statistics relating to the data used in this paper) documents the extent of this rainfall volatility, and demonstrates that there was no tendency for this climatic volatility to change significantly over time. Similarly, Figure 1 illustrates that even when aggregating rainfall up to province-level averages, there is significant volatility in all provinces, throughout our study period.

Like rainfall, prices, nominal agricultural incomes, and real agricultural incomes were also volatile over time within districts. This reflected the fact that rainfall was an essential input to agricultural production in this predominantly rain-fed environment. However, as shown in Table 1, there is some evidence that the volatility of these equilibrium economic variables changed over time in a way that rainfall did not. That is, there was a tendency for the volatility of prices and real income to decrease over the time period under study in this paper, but for the volatility of nominal incomes to rise. Below we explore the potential for

[4] Additional details on the construction of the dataset can be found in Donaldson (2010).

[5] See, for example, Gadgil, Rajeevan, and Francis (2007). The phrase is still used to refer to agriculture today—for example, in the state of Orissa's 2005 *Economic Survey*, a state in which only 56 percent of cultivated land is irrigated.

the arrival of railroads to have changed the mapping between an exogenous and stochastic production input, rainfall, and these economic variables.

2.2 Mortality in Colonial India

The agricultural income volatility described in the previous section occurred against a background of extremely low average agricultural incomes and subsistence living by many. In addition, the health infrastructure was poor, and only a very small minority of citizens had access to formalized health care of even the best Victorian standards (Arnold 1993).

In such a setting it is natural to expect high mortality rates and low life expectancies. We have collected the best available data on district-level death rates from 1870 onwards in order to examine how the death rate responds to rainfall variation throughout the period from 1870 to 1930, and how railroads changed this responsiveness. The mortality rate estimates stem from a (compulsory) vital events registration system that most of British India's provinces had in place by 1870.[6] Like the vital events data from many registration-based systems, the mortality rates used here are known to suffer from under-reporting, so the average death rate reported over time is almost surely an underestimate (Dyson 1991, Davis 1951). However, as Table 1 shows, even this underestimate is very high by both contemporary and historical standards.

Table 1 also documents the volatility of the mortality rate and its evolution over time. Under-reporting may have been particularly bad in times of crisis, such as famines, so the volatility is also likely to be understated. Still, there is considerable volatility in these registration data. In this low-income and low-health environment it is natural to expect real income volatility to be particularly damaging to human survival, and to give rise to volatile death rates. But notably, like the series for agricultural prices and real agricultural income, the volatility of the mortality rate is falling over time from 1870 to 1930.

The mortality data we use here are differentiated by cause of death. It is important not to overstate the precision with which these causes of death were classified—in most cases, mortality registers were completed by village headmen or watchmen, who were not qualified to make medical diagnoses (Maharatna 1996). The causes of death are also particularly coarse, with categories for 'fevers' (known now to be overwhelmingly malarial deaths (Davis 1951)), 'respiratory diseases', 'diarrhoeal diseases/bowel complaints', 'smallpox', 'plague', 'cholera', 'accidents' and 'other' only. The 'fevers' category accounted for 41 % of deaths in our sample period, and the second most common cause of death was 'diarrhoeal diseases/bowel complaints'. Maharatna (1996) argues that in times of famine, deaths due to all of these causes

[6] These data are reported in each province's annual *Sanitary Report*.

are likely to rise, with the exception of 'plague' and 'accidents'. Because these causes were also particularly easy to diagnose they provide a natural falsification variable for the exercise we pursue below, in that they should not respond to rainfall variation either before or after a district is accessible by railroad.

2.3 Colonial India's Famine Era

Famines were a recurrent blight on India prior to and throughout most of the colonial period—they were India's "late Victorian Holocaust" in the words of historian Mike Davis. There is little doubt that the principal cause of these famines was a shortage of rainfall, giving rise to crop failure.[7] In the low-income setting of colonial-era India, where almost three-quarters of citizens earned their incomes in the agricultural sector directly,[8] it is not surprising that rainfall shortages can lead to death, given the importance of rainfall in India's rain-fed agricultural environment.

In times of extreme mortality in a district, India's colonial government would officially declare the district to be in a 'famine'. The exact criteria for this declaration are unclear, potentially changing over time, and potentially endogenous to the arrival of railroads. For these reasons we deliberately do not use the official famine declarations in our analysis, and focus on the raw mortality data instead.

However, the official famine declarations, along with anecdotal sources, reveal four striking features of India's 11 officially declared famines between 1860 and 1930. First, they were responsible for enormous amounts of death—15 million people at extremely conservative estimates.[9] Naturally, rainfall shortages may have been responsible for considerable loss of life during the many years that lay beneath the extreme 'famine' threshold.

A second important feature of the officially declared famine districts and events, and even those that were widely agreed to be suffering from 'scarcity', was their regional concentration in any given year. Srivastava (1968) plots India's famines and 'lesser scarcities' (the latter according to his definition) from 1858 to 1918. While the location of the distressed regions occurred in different areas of the subcontinent from year to year, rarely did a given years's

[7] McAlpin (1979) argues: "There is not much dispute that India has been subject to periodic crop failures for centuries, nor that the proximate cause of these crop failures is the lack of adequate timely rainfall" (p. 143). This is echoed in the standard references on Indian famines such as Srivastava (1968) and Bhatia (1967).

[8] Among the employed adult population, 69 % earned their incomes in the agricultural sector in the 1901 census (and this share fell from only 73 % in 1872, the first census year in our sample, to 68 % in 1931). Of course, the rest of rural India (which comprised almost 85 % of Indian citizens in 1900) may also have been exposed to climatic variation through interlinkages between agriculture and non-agricultural activities.

[9] We calculate this figure from Appendix Table 5.2 of Visaria and Visaria (1983). Davis (2001) argues that this is a gross underestimate and suggests an upper estimate of 34 million deaths. In particular, Visaria and Visaria (1983) do not include any deaths from the severe famine of 1899-1900, for want of data.

famine or scarcity distress an area larger than the size of a single British Province—or approximately one fifteenth the area of colonial India. This limited spatial extent of famines and crop failures suggests a potential role for other regions, those not suffering from crop failure, to ship food to famine-stricken districts, if trade costs are low enough to allow for these trades to occur.

A third feature of the officially declared famine events is their diminishing frequency over time during the late colonial era (and the period under study here, 1861-1930, in particular). Only one, in 1906, occurred after 1899—and as discussed in the Introduction, to most observers the famine in 1906 was the last peace-time famine that India has seen.

As a final point about famines in colonial India it is important to consider how food shortages might map into death in times of famine. The exact trajectory of death during a famine—whether those in colonial India, modern sub-Saharan Africa or elsewhere—is an area of uncertainty and active research. Famines are often times of widespread starvation, accidental poisoning, epidemic disease, unsanitary temporary living conditions, dislocation from families and support networks, and violence. All of these are contributing factors to the aggregate death toll of a famine. In India, epidemic disease seems to have been particularly important, while dislocation and migration was comparatively less so (Dyson 1991). But the fundamental cause of death due to epidemic disease during a famine is still regarded as inadequate nutrition, which leaves an individual's immune system compromised and vulnerable to disease (Maharatna 1996). An implication of this is that the aggregate consumption-mortality elasticity is likely to be larger than the individual equivalent, because of the externalities of mortality due to disease.

2.4 Famine Prevention in India and the Role of Railroads

Despite the enormous human tragedy posed by famine and 'ordinary' rainfall variation, the British colonial government in India was reluctant to intervene. Stokes (1959) and McAlpin (1979) argue that the Indian civil service, educated in the *laissez-faire* tradition of Adam Smith, preferred to help market forces to prevent famine. In this regard, India's rapidly unfolding railroad network (shown in Figure 2) was thought sufficient to enable inter-regional trade that would limit the effects of local scarcity. An additional endeavor was the 1880 Famine Commission's recommendation discussed in Donaldson (2010)—to build new railroad lines in a manner that might alleviate famine. Donaldson (2010) demonstrated evidence consistent with a significant contribution of railroads to the reduction of trade costs in India. With this in mind, it is likely that food was able to move from famine-free districts to famine-stricken districts at low cost and great speed. Indeed, contemporary observers quoted in Johnson (1963), argued as much (p. 123):

> "The contribution of railways in moving supplies and meeting the critical shortages in areas affected by famine was indeed considerable. Quite apart from the economy in the costs of carriage, the knowledge that in a few days at most, more supplies would be arriving by railway trains, helped often to keep down the expensive rise in prices which increased demands produced. When, for example, a famine raged in northern India in 1884, railway stations in Coimbatore, hundreds of miles away, were crammed with grains for transportation and sale in north India. Even a hint of scarcity sufficed to attract movement of goods by railways from areas of relative abundance."

In the remainder of this paper we explore the potential for this movement of food brought about by railroads to mitigate the effect of rainfall shortages on death in colonial India. This investigation sheds light on the relationship between trade openness and real income volatility, and on the role played by railroads in ridding India of famine.

3 A Model of Railroads and Income Volatility

In this section we outline a theoretical framework in which multiple regions are able to trade with one another, at a cost, and explore the implications for real income volatility of reducing this trade cost. To simplify the presentation, and to ease connection with our empirical approach below, we consider a purely static model. This may seem a surprising modeling choice when studying volatility. But in this static model we consider how the level of an equilibrium variable (for example, the level of real income) responds to changes in the supply of a given productive input (for example, rainfall). If this productive input is stochastic (and uncorrelated with other shocks), then studying comparative statics relating to the responsiveness of the level of an equilibrium variable to the level of the productive input captures the key mechanism that would generate equilibrium volatility in a similar but dynamic model.

To shed light on the scope for trade openness to mitigate real income volatility—or equivalently, for railroads to mitigate the effects of productivity shocks on real incomes—we draw on the theoretical model of Ricardian trade introduced by Eaton and Kortum (2002). In particular we follow the multi-industry version of Eaton and Kortum (2002), as in, for example, Donaldson (2010) and Costinot, Donaldson, and Komunjer (2012). We first lay out this model in its full generality. Unfortunately, the multiple general equilibrium interactions in that model are too complex to admit a closed-form solution for the effect of reduced trade costs on agricultural prices and their responsiveness to productivity shocks. To make progress in generating qualitative predictions we therefore restrict attention, later in this section, to a simplified version of the model environment. Our empirical strategy, however, embraces the full generality of the model.

3.1 Model Environment and Equilibrium

What follows is a brief description of the model—see Appendix A for a detailed description.

The economy consists of D regions (indexed by either o or d depending on whether the region is the origin or destination of an exchange) and K commodities (indexed by k). Each available in a continuum of differentiated varieties indexed by j. Each region o is home to a representative agent who owns L_o units of land of inelastically supplied land. All agents have Cobb-Douglas preferences over commodities (k), with weight μ_k on each commodity, and constant elasticity of substitution preferences over varieties (j) within each commodity, with own-price elasticity of demand equal to σ_k.

Each variety j of the commodity k can be produced using a constant returns to scale production technology in which land is the only factor of production. Let $z_o^k(j)$ denote the amount of variety j of commodity k that can be produced with one unit of land in region o. We follow Eaton and Kortum (2002) in modeling $z_o^k(j)$ as the realization of a stochastic variable Z_o^k drawn from a Type-II extreme value distribution whose parameters vary across regions and commodities in the following manner

$$F_o^k(z) \doteq \Pr(Z_o^k \leq z) = \exp(-A_o^k z^{-\theta_k}), \tag{1}$$

where $A_o^k \geq 0$ and $\theta_k > 0$. These random variables are drawn independently for each variety, commodity and region. The exogenous parameter A_o^k increases the probability of high productivity draws and the exogenous parameter θ_k captures (inversely) how variable the (log) productivity of commodity k in any region is around its (log) average. There are many competitive firms in region o with access to the above technology; consequently, firms make zero profits. These firms will therefore charge a pre-trade costs (ie, 'free on board') price of $p_{oo}^k(j) = r_o/z_o^k(j)$, where r_o is the land rental rate in region o.

Trade of commodities over space incurs costs (which include transport costs and other barriers to trade). We assume that, in order for one unit of commodity k to arrive in region d, $\tau_{od}^k \geq 1$ units of the commodity must be produced and shipped in region o. (Free trade therefore corresponds to $\tau_{od}^k = 1$.) Trade costs drive a wedge between the price of an identical variety in two different regions. Let $p_{od}^k(j)$ denote the price of variety j of commodity k produced in region o, but shipped to region d for consumption there. Trade costs imply that any variety in region d will cost τ_{od}^k times more than it does in region o; that is, $p_{od}^k(j) = \tau_{od}^k\, p_{oo}^k(j) = r_o\, \tau_{od}^k / z_o^k(j)$.

Consumers have preferences for all varieties j along the continuum of varieties of commodity 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 (after accounting for trade costs). We therefore follow Eaton and Kortum (2002) and solve for the equilibrium prices that consumers in a region d actually pay, given that they will only buy a given variety from the cheapest source region (including their own). The expected value of the price (of any variety j) of commodity k in region d will be

$$E[p_d^k(j)] \doteq p_d^k = \lambda_1^k \left[\sum_{o=1}^{D} A_o^k (r_o \tau_{od}^k)^{-\theta_k} \right]^{-1/\theta_k}, \tag{2}$$

where λ_1^k is a constant. $\doteq \Gamma(1+\frac{1}{\theta_k})$. In our empirical application below we treat these expected prices as equal to the observed prices collected by statistical agencies.[10]

A second key result in this model describes the volume of trade between any two locations in any commodity. Letting X_{od}^k denote the total expenditure in region d on commodities of type k from region o, $X_d^k \doteq \sum_o X_{od}^k$ is total expenditure in region d on commodities of type k, and $\pi_{od}^k = \frac{X_{od}^k}{X_d^k}$, trade flows can be shown to satisfy:

$$\frac{X_{od}^k}{X_d^k} = \pi_{od}^k = \lambda_3^k\, A_o^k\, (r_o tau_{od}^k)^{-\theta_k}\, (p_d^k)^{\theta_k}, \tag{3}$$

where $\lambda_3^k = (\lambda_1^k)^{-\theta_k}$, and this equation makes use of the definition of the expected value of prices (ie, p_d^k) from equation (12). This expression for trade flows is often referred to as a gravity equation.

A final and important result in this model follows directly from equation (13). In this model (even in the fully general version of the model outlined above), despite the presence hundreds of interacting product and factor markets, the level of log real income (ie welfare)

[10] A second price moment that is of interest for welfare analysis is the exact price index over all varieties of commodity k for consumers in region d. Given CES preferences, this is $\widetilde{p}_d^k \doteq \left[\int_0^1 (p_d^k(j))^{1-\sigma_k} dj\right]^{1/1-\sigma_k}$, which is only well defined here for $\sigma_k < 1+\theta_k$ (a condition we assume throughout). The exact price index is given by $\widetilde{p}_d^k = \lambda_2^k p_d^k$, where $\lambda_2^k \doteq \frac{\gamma^k}{\lambda_1^k}$ and $\gamma^k \doteq [\Gamma(\frac{\theta_k+1-\sigma_k}{\theta_k})]^{1/(1-\sigma_k)}$. That is, if statistical agencies sampled varieties in proportion to their weights in the exact price index, as opposed to randomly as in the expected price formulation of equation (12), then this would not jeopardize our empirical procedure because the exact price index is proportional to expected prices.

in any region o can be written as follows:

$$\ln W_o = \Omega + \sum_k \frac{\mu_k}{\theta_k} \ln A_o^k - \sum_k \frac{\mu_k}{\theta_k} \ln \pi_{oo}^k, \tag{4}$$

where Ω is a constant and π_{oo}^k is the share of region o's expenditure on commodity k that is produced in region o which we refer to as the trade share. (That is, $\pi_{oo}^k = 1$ in autarky.) This result highlights that if railroads (ie reductions in trade costs, τ_{od}^k) change the mapping between productivity (A_o^k) and welfare (W_o), then they do so—according to this model—through the endogenous variable π_{oo}^k, the trade share.

Equation (13) characterizes trade flows conditional on the endogenous land rental rate, r_o (and all other regions' land rental rates, which appear in p_d^k). It remains to solve for these land rents in equilibrium, by imposing the condition that each region's trade is balanced up to an exogenous net transfer (which could be zero if trade were balanced) into region o from all other regions, B_o. Region o's trade balance equation requires that the total income received by land owners in region o ($r_o L_o$) must equal the total value of all commodities made in region o and sent to every other region (including region o itself). That is:

$$r_o L_o + B_o = \sum_d \sum_k X_{od}^k = \sum_d \sum_k \pi_{od}^k \, \mu_k \, r_d \, L_d. \tag{5}$$

Each of the D regions has its own trade balance equation of this form. The equilibrium of the model is the set of D-1 unknown rental rates r_d that solves this system of D-1 (non-linear) independent equations.

3.2 Six Predictions

We now outline six predictions that stem from the model presented above, in the order in which these predictions inform our empirical analysis below. All of these predictions concern *equilibrium responsiveness*, which we define as the extent to which an endogenous variable of interest in a location (for example, the level of real income at the location) changes as the location's productivity level changes. Further, we are not interested in responsiveness *per se*, but rather how responsiveness changes as the location becomes more open—that is, as the location is able to trade with other locations at a lower cost. In the notation of the model above, for any outcome variable Y_d in district d, we are therefore interested in comparative statics of the form:

$$\frac{d}{d\tau_{do}^k} \left| \frac{dY_d}{dA_d^k} \right|, \tag{6}$$

where τ_{do}^k represents the cost of trading between district d and some other district o, and A_d^k is the productivity level of district d (both of which terms could, in principle, be conditioned on commodity k). Our four main predictions below concern how openness (lower τ_{do}^k) affects equilibrium responsiveness (which we define as $\left|\frac{dY_d}{dA_d^k}\right|$) for a series of four outcome variables (Y_d): prices, nominal incomes, real incomes, and mortality rates.

In all cases below we consider a (considerably) simplified version of the general model outlined above. Our goal in doing so is to elucidate the workings of this model in as simple a setting as is necessary. Specifically, throughout this section we assume that there are only three regions (named 1, 2 and 3), that there is only one commodity (so we will dispense with the k superscripts on all variables), that the regions are symmetric in their exogenous characteristics (ie $L_o = L$ and $A_o = A$ for all regions o), and that the three regions have symmetric trade costs with respect to each other. We consider the comparative statics of a local change around this symmetric equilibrium, where it is straightforward to show that the model makes the following predictions.

Prediction 1: Openness Reduces Price Responsiveness

This first prediction concerns the extent to which local prices respond to local productivity shocks, and how this responsiveness changes as trade costs fall. For a local change around the symmetric equilibrium considered in this paper:

(a) $\frac{d}{dT_{12}}\left|\frac{dp_1}{dA_1}\right| > 0$: The responsiveness of prices in a region (say, 1) to a productivity shock *in the same region* (ie $\left|\frac{dp_1}{dA_1}\right| < 0$) is weaker (ie less negative) when the region has low trade costs to another region (say, 2).

(b) $\frac{d}{dT_{12}}\left|\frac{dp_1}{dA_2}\right| < 0$: The responsiveness of prices in a region (say, 1) to productivity shocks *in other any other region* (say region 2, so the price responsiveness of interest here is $\left|\frac{dp_1}{dA_2} < 0\right|$) is stronger (ie more negative) when the cost of trading between these two regions (ie T_{12}) is low.

Prediction 2: Openness Increases Nominal Income Responsiveness

This prediction concerns the effect of an exogenous change in productivity on a region's nominal income. Let r_o be the nominal income (the nominal land rental rate, per unit land area) of district o (relative to the numeraire good, taken to be the land rental rate of some reference region). Then around the three-region symmetric equilibrium:

- $\frac{d}{dT_{12}}\left|\frac{dr_1}{dA_1}\right| < 0$: The responsiveness of *nominal* income in a region (say, 1) to a productivity shock in that region (ie $\left|\frac{dr_1}{dA_1}\right| > 0$) is stronger when the cost of trading between this region and any other region (say, 2) falls.

This prediction can be thought of as following directly from Prediction 1; if prices become less reponsive to productivity shocks then nominal incomes (the product of prices and quantities) will become more responsive to productivity shocks because prices and quantities will move to offset one another.

Prediction 3: Openness Reduces Real Income Responsiveness

This prediction is analogous to Prediction 2, but concerns a region's real income rather than its nominal income. Let W_o be the real income of district o (per unit land area), which is equal to the nominal land rental rate divided by an appropriate consumer price index (ie $W_o \doteq \frac{r_o}{\widetilde{P_o}}$ where $\widetilde{P_o}$ is the price index). Then around the three-region symmetric equilibrium:

- $\frac{d}{dT_{12}}\left|\frac{dW_1}{dA_1}\right| > 0$: The responsiveness of *real* income in a region (say, 1) to a productivity shock in that region (ie $\left|\frac{dW_1}{dA_1}\right| > 0$) is stronger when the cost of trading between this region and any other region (say, 2) falls.

Prediction 4: Openness Increases Export and Import Responsiveness

Equation (13) makes it clear that in a gravity model such as this one, conditional on factor prices r_o and destination location effective demand $w_d L_d p_d^{\theta}$, a region's exports will rise if productivity increases. This prediction is true even unconditionally on factor prices and destination demand around a symmetric equilibrium in a three-region and two-commodity (ie $K = 2$) model:

(a) $\frac{d}{dT_{12}}\left|\frac{d(X_{12}^{k_1} - X_{21}^{k_1})}{dA_1^{k_1}}\right| > 0$: The responsiveness of net *exports* from a region (say, 1) in a sector (say, sector k_1) to a productivity shock in that region in that but not the other sector (say, sector k_2) (ie $\left|\frac{d(X_{12}^{k_1} - X_{21}^{k_1})}{dA_1^{k_1}}\right| > 0$) is stronger when the cost of trading between this region and any other region (say, 2) falls.

(b) $\frac{d}{dT_{12}^{k_2}}\left|\frac{d(X_{21}^{k_2} - X_{12}^{k_2})}{dA_1^{k_1}}\right| < 0$: The responsiveness of net *imports* from a region (say, 1) in the non-shocked sector (ie sector k_2) to a productivity shock in a sector (say, sector k_1) (ie $\left|\frac{d(X_{21}^{k_2} - X_{12}^{k_2})}{dA_1^{k_1}}\right| > 0$) is stronger when the cost of trading between this region and any other region (say, 2) falls.

Prediction 5: Openness Reduces Consumption (and hence Mortality) Responsiveness

In this static model, consumption equals income, and hence consumption responsiveness is equal to income responsiveness. Absent data on consumption, our empirical analysis below exploits the mortality rate as a proxy for per-capita consumption (though of course the mortality rate is a measure of economic welfare that is also of its own interest). This model does not specify how consumption affects the probability of dying. But it is trivial to introduce a consumption-mortality relationship that is monotonically increasing, as in, for example, Ravallion (1997). This relationship is typically assumed to be concave. And it is plausible that, in the current setting of colonial India, the level of average consumption lies in a region of the consumption-mortality relation in which the slope of the relationship is particularly high. No distinction is made in Ravallion (1997) between individual and aggregate consumption-mortality relationships, though, as argued earlier, it is plausible that the aggregate relationship is also monotonic and likely to be even stronger due to the scope for externalities (such as disease). With this sort of monotonic consumption-mortality relationship added to the model we would expect the mortality rate to respond to productivity shocks and trade costs in the same manner as real income above—but with opposite sign. That is, a positive productivity shock would reduce the mortality rate, but a reduction in trade costs would reduce the extent to which the mortality rate is sensitive to this positive productivity shock.

Prediction 6: Openness Does Not Affect Real Income Responsiveness, at Constant Trade Share

In this model (even in the fully general version of the model outlined in Section 3.1 above), as show in equation (4), the level of real income in region o can be written as a simple function of the productivity level in region o and the level of the trade share variable, π^k_{oo}, for region o. Put differently:

- $\frac{d}{dT_{od}} \left| \frac{dW_1}{dA^k_o} \right|_{\pi^k_{oo}=0}$: The responsiveness of real income in a region (say, region o) to a productivity shock in that region (ie $\left| \frac{dX_{12}}{dA_1} \right| > 0$) is no different when the cost of trading between this region and any other region (say, d) falls, if the trade share for region o in all commodities k (ie π^k_{oo}) is held constant.

In the remainder of this paper we compare these six predictions with data on Indian districts from 1861-1930, with the goal being to gain an improved understanding of how productivity shocks map into endogenous outcome variables and how openness to trade changes this mapping. We use rainfall variation as an exogenous shock to productivity (ie A^k_o in the model) and we use the arrival of railroads in a district as an exogenous change in the cost of trading goods beteen that district and other districts (ie τ^k_{od} in the model).

4 Empirical Method and Results

We now present a series of empirical results that are motivated by the theoretical analysis in the previous section, and that aims to shed new empirical light on the relationship between trade openness and real income volatility. These results proceed in six steps, which trace through the six predictions in Section 3 sequentially. We illustrate how openness to trade, brought about by railroads, altered the responsiveness to local productivity shocks of each of five local outcome variables (prices, nominal incomes, real incomes, trade flows, and mortality rates) and then describe how responsiveness to local productivity shocks .

4.1 Empirical Procedure

The six predictions in Section 3 all take a similar form: they describe the extent to which each of five local outcome variables (prices, nominal incomes, real incomes, trade flows, and mortality rates) respond to exogenous local productivity shocks, and how this equilibrium 'responsiveness' changes as the local economy in question becomes increasingly integrated with other economies. We use variation as an exogenous shifter of agricultural productivity. To study this changing equilibrium responsiveness empirically we estimate regressions of the following form:

$$\ln Y_{dt} = \alpha_d + \beta_t + \gamma_1 RAIL_{dt} + \gamma_2 RAIN_{dt} + \gamma_3 RAIL_{dt} \times RAIN_{dt} + \varepsilon_{dt}, \tag{7}$$

where d indexes the Indian district, t indexes the year. The dependent variable is Y_{dt}, an outcome variable of interest (such as a price, the level of nominal income, the level of real income, or the mortality rate) in district d and year t. The independent variables are $RAIL_{dt}$ (a dummy variable indicating that the district has been penetrated by a railroad line), $RAIN_{dt}$ (in most cases, just the level of annual rainfall in the district), and their interaction, $RAIL_{dt} \times RAIN_{dt}$. The terms α_d and β_t represent district- and year-specific fixed effects, respectively.

In this specification, equilibrium *responsiveness* (of the outcome variable Y_{dt} to an exogenous productivity shifter $RAIN_{dt}$) is captured empirically by the expression $\gamma_2 + \gamma_3 RAIL_{dt}$. This expression illustrates how equilibrium responsivenss to productivity shocks is not a fixed constant. By contrast, as described theoretically in Section 3 above, the extent of responsiveness in a locality (here, a district) depends on the level of openness of that locality (here, proxied empirically by the penetration of the railroad network, ie the variable $RAIL_{dt}$. In this sense, the main coefficient of interest in our study is γ_3, which captures the extent to which openness (as proxied by the arrival of railroads) alters a location's equilibrium responsiveness to productivity shocks. The first five predictions in Section 3 concern the sign of γ_3, whereas our sixth prediction is that $\gamma_3 = 0$ when an empirical proxy for the log trade share (ie for $\sum_k \frac{\mu_k}{\theta_k} \ln \pi^k_{oo}$, from equation (4)) is included in the regression from equation (7).

4.2 Empirical Results

We now discuss estimates of equation (7)—and similar specifications—for each of our the five outcome variables (prices, nominal incomes, real incomes, trade flows, and mortality rates) suggested by the six predictions in Section 3 above.

4.2.1 Prediction 1: Openness and Price responsiveness

Following prediction 1 of the model, we test the hypothesis that railroads reduced the responsiveness of local agricultural prices to local rainfall (an exogenous determinant of local productivity). In a small open economy (SOE), price responsiveness is zero since local prices are equal to the (exogenous) 'world' price level. However, as trade costs rise and an economy departs from the SOE limit, price responsiveness in that economy should rise (as in prediction 1). The extent of price responsiveness in a district is therefore a novel and powerful test of its openness to trade, which motivates the empirical exercise in this section.[11] Put another way, if openness is to affect the responsiveness of living standards to productivity shocks through conventional price theory mechanisms, a necessary condition is prediction 1—that openness changes price responsiveness. We examine the empirical relevance of this necessary condition here.

Prediction 1 actually has two parts: (a) when a district is connected to the railroad network, agricultural goods prices in that district will be less responsive to productivity shocks in that district; and (b), when a railroad line connects two districts, agricultural goods prices in a district will be more responsive to productivity shocks in the other district.

Doing full justice to the data (in which prices and rainfall shocks are available separately by crop) and to prediction 1 requires a departure from the general framework suggested by

[11] To our knowledge, this is a novel test for assessing a change in market integration. However, two papers are closely related. First, Shiue (2002) examines how the price correlation (over many years) between pairs of markets in 19th Century China is related to the weather correlation (over the same years) in these pairs, comparing this correlation in inland locations to that along rivers or the coast. Second, Keller and Shiue (2007) estimate formally, in the same setting, how the spatial dependence of weather shocks (on prices) varies between inland and water-accessible regions. Neither of these papers focuses on the responsiveness of local prices to local rainfall, nor on whether the spatial transmission of weather shocks is different along some transportation links (such as railroads) than along others (such as roads), in the manner we do here.

equation (7). We test prediction 1 by estimating the following specification:

$$\begin{aligned}\ln p^k_{dt} = \beta^k_d \;&+\; \beta^k_t + \beta_{dt} + \chi_1 RAIN^k_{dt} + \chi_2 RAIL_{dt} \times RAIN^k_{dt} \\ &+\; \chi_3 (\tfrac{1}{N_d}) \sum_{o \in \mathbf{N}_d} RAIN^k_{ot} \\ &+\; \chi_4 (\tfrac{1}{N_d}) \sum_{o \in \mathbf{N}_d} RAIN^k_{ot} \times RAIL_{odt} + \varepsilon^k_{dt}. \end{aligned} \qquad (8)$$

Here, p^k_{dt} represents the retail price of agricultural crop k in district d and year t. $RAIN^k_{dt}$ is the amount of crop-specific rainfall that fell in district d in year t (described in detail below). The variable $RAIL_{dt}$ is a dummy variable equal to one when the railroad network enters the boundary of district d, while the variable $RAIL_{odt}$ is a dummy variable equal to one when it is possible to travel from district o to district d using only the railroad network. Finally, the variable $RAIN^k_{ot}$ represents the amount of crop-specific rainfall in district o, where district o is a neighbor of district d—one of the the N_d districts ($o \neq d$) in district d's neighborhood $\mathbf{N}_d$ (taken to be all districts that lie even partially inside a 250 km radius of district d's centroid).[12] The summation terms in equation (8) are divided by the number of districts N_d in the neighborhood $\mathbf{N}_d$ to reflect an average effect.

We estimate equation (8) using fixed effects for each district-year (β_{dt}), which control for any unobservable variables affecting prices that are constant across crops within a district and year. This means that we identify price responsiveness through variation in how a given amount of annual rainfall in a district affects each of that district's crops differently. We also include fixed effects for each district-crop (β^k_d) to control for unobservables that permanently affect a district's productivity of a given crop (such as the district's soil type), and fixed effects for each crop-year (β^k_t) to control for country-wide shocks to the price of each crop.

To the extent that rainfall is a significant determinant of productivity (as was found to be the case revealed in trade flows, in Donaldson (2010)), the coefficients χ_1 and χ_3 will be negative. Prediction 1 (a) states that the coefficient χ_2 is positive (prices in district d are less responsive to rainfall in district d if district d is on the railroad network). And prediction 1 (b) states that the coefficient χ_4 is negative (lower transport costs should make prices in district d more responsive to rainfall shocks in neighboring districts to d). A positive coefficient χ_2 is consistent with railroads increasing the extent of market integration in India.

[12] While in principle the rainfall in any district o could affect prices in district d, in the Eaton and Kortum (2002)model these effects are likely to die out quickly over distance. In a partial equilibrium sense (that is, without allowing for the land rental rate r_o to adjust), this can be seen easily in equation (12). Here, each distant district's productivity term A^k_o affects local prices p^k_d in a manner proportional to $(T^k_{od})^{-\theta_k}$, where T^k_{od} is the trade cost and a typical estimate of θ_k is 3.8 (Donaldson 2010). We therefore restrict the effect of non-local rainfall on district d's prices to that in a short (250 km) range, though our results are insensitive to using smaller (eg 100 km) or larger (eg 500 km) ranges.

We estimate equation (8) using annual data on the retail price of 17 agricultural commodities, in 239 districts, from 1861-1930. These prices were collected by district officers who visited the 10-15 largest retail markets in each district once every two weeks. India-wide instructions were issued to each province to ensure that prices of each commodity were recorded in a consistent manner across the provinces. We construct the crop-specific rainfall variables ($RAIN_{dt}^{k}$, etc) as in (Donaldson 2010). Specifically, we use the *Indian Crop Calendar* and daily rainfall data for each district to compute the amount of rain that fell during the specific periods of the year during which each crop k was, according to the *Crop Calendar*, either sowing or growing each crop. Finally, we construct the railroad variables ($RAIL_{dt}$ and $RAIL_{odt}$) using the GIS database described in Section 2 above.

Table 2 presents results from OLS estimates of equation (8). Column 1 begins by regressing (log) agricultural prices in a district on the district's crop-specific local rainfall. The coefficient on local rainfall is negative and statistically significant, suggesting that rainfall has a positive impact on crop output, and this increase in supply transmits into local retail prices. This is indicative of imperfect market integration (ie, non-zero trade costs) in these agricultural commodities *on average* over the time period 1861-1930 in India. The coefficient estimate implies a large amount of price responsiveness on average over the period: a one standard deviation (ie 0.604 m) increase in a crop's crop-specific rainfall decreases that crop's prices by approximately 15 percent.

Column 2 of Table 2 then tests the first part of prediction 1: that a district's prices will be less responsive to local rainfall after the district is connected to the railroad network. In this specification the coefficient on local rainfall (χ_1 in equation (8)) represents price responsiveness before railroads penetrate a district. The estimated coefficient is negative, statistically significant, and demonstrates a great deal of price responsiveness in the pre-railroad era of each district. Further, in line with prediction 1, the coefficient χ_2 on rainfall interacted with a dummy for railroad access ($RAIL_{dt}$) is positive and statistically significant. The sum of the coefficients χ_1 and χ_2 represents the extent of price responsiveness after the district is brought into the railroad network. The estimated coefficients sum to -0.014 which implies that prices are still responsive to local rainfall, but in a dramatically reduced sense when compared to the coefficient of -0.428 that measures price responsiveness in the pre-rail era. However, we cannot reject the null of zero price responsiveness in the post-rail era. These findings suggests that the imperfect market integration from 1861-1930 found in column 1 reflects an average of two extreme regimes separated by the arrival of a railroad line in a district: a first regime of imperfect integration before the railroad arrives (where local supply shocks have large effects on local prices), and a second regime of near-perfect integration after the railroad arrives (where local supply shocks have a negligible effect on local prices).

Column 3 repeats the specification in column 1, but with the inclusion of average rainfall in neighboring districts. The effect of neighboring districts' rainfall on local prices is negative and statistically significant, which implies that, on average over the period from 1861-1930, neighboring districts' supply shocks affected local prices, as is consistent with some degree of market integration.

However, the estimates in column 4 demonstrate that, as was the case in column 2, the average effect in column 3 is masking the behavior of two different regimes. Column 4 estimates equation (8) in its entirety by including an interaction term between each neighboring district's rainfall and a dummy variable for whether that district is connected to the 'local' district by railroad (ie $RAIL_{odt}$). As is consistent with the second part of prediction 3, the coefficient on this interaction term is negative and statistically significant. Furthermore, the coefficient on neighboring districts' rainfall (which is now the effect of rainfall in districts not connected by railroad to the local district) is not significantly different from zero. Column 4 therefore suggests that local prices do respond to neighboring districts' supply shocks when those neighbors are connected to the local district by railroads; however, neighboring districts' supply shocks are irrelevant to local prices when there is no railroad connection.

To summarize the results from this section, we find that railroads played a strong role in facilitating market integration, as revealed by price responsiveness, among the 17 agricultural goods in our sample. This is consistent with both parts of prediction 1 of the model, and suggests that railroads significantly aided the trade of agricultural items across districts in colonial India, so much so that local scarcities brought about by local rainfall shortages were rapidly filled by supply from surrounding regions.

While this reduction in price responsiveness is useful as a metric of market integration or openness, it says nothing about the welfare consequences of heightened openness for the residents of a location. The results below probe such consequences in full detail. But, as described above, a reduction in price responsiveness due to openness (ie prediction 1) is an important necessary condition for what follows below.

4.2.2 Prediction 2: Openness and Nominal Income Responsiveness

The above results suggest that prior to the arrival of railroads in a district, agricultural prices were considerably responsive to adverse rainfall shocks in that district. However, after railroads arrive in a district, the responsiveness of local prices to local rainfall shocks virtually disappears. As discussed above, this is to be expected if railroads reduced trade costs and allowed a district's local prices to be determined by supply conditions in many regions, rather than just locally.

This finding implies that, after the arrival of railroads, when a locality receives a negative productivity shock in agriculture, local quantities of agricultural goods produced will fall, but there will be no offsetting price effect. As a result, the effect of these shocks on the value of output—that is, on nominal agricultural incomes—should become *more* responsive to these shocks after the arrival of railroads than before. In this section we explore the extent to which this logic—the logic behind Prediction 2—holds empirically in the case of colonial India.

The empirical strategy we follow to test Prediction 2 follows the general specification laid out in equation (7) above. Prediction 2 states that railroads will increase the responsiveness of nominal agricultural income in a district to its own rainfall shocks. Prediction 2 suggests that the coefficients γ_1 and γ_2 should both be positive (both railroad access and favorable rainfall raise nominal incomes), and that the coefficient γ_3 should also be positive, indicating that nominal agricultural income becomes more responsive to rainfall shocks once a district is open to trade through railroad access.

As a measure of nominal income (the dependent variable) we use the land rental rate, r_{ot}, which is the measure of nominal income in the model, as discussed above. Unfortunately, there are no available data on land rental rates in this setting, but in a perfectly competitive, constant-returns, one-factor setting such as the model used here, the land rental rate is equal to nominal agricultural output per unit land area. We use data on nominal agricultural output (the price-weighted sum of physical output quantities over 17 principal agricultural crops) from 1870-1930. We then divide this measure of nominal output by the total land area under cultivation in a district to create a measure of nominal agricultural incomes accruing to the representative owner of one unit of cultivable land.

Table 3 presents the results of this test of prediction 2. Column 1 confirms that rainfall is an important determinant of agricultural production, and therefore nominal agricultural income. This is in line with the above results on prices (where high rainfall was found to decrease prices in the absence of railroads, a result that is consistent with railroads reducing supply).

However, the results in column 2 demonstrate that rainfall is a stronger determinant of nominal agricultural income once a district gains railroad access than before. That is, the coefficient on rainfall ('rainfall in district') is smaller than in column 1 (and still statistically significant). This coefficient represents the responsiveness of nominal agricultural income to rainfall before the district is connected to the railroad network. By contrast, the effect of rainfall on nominal agricultural income after a district gains railroad access (represented by the sum of the coefficients on the 'rainfall in district' term and the interaction term between railroad access and rainfall) is larger, indicating heightened nominal income responsiveness.

Further, the pattern of all four coefficients in column 2 is in line with that predicted by prediction 2.[13]

The results in columns 1 and 2 of Table 3 are supportive of Prediction 2. They suggest that the move to trade openness due to railroads in India made nominal agricultural incomes more vulnerable to the vagaries of the monsoon rains. This may have been harmful to individuals who consume primarily non-agricultural goods. But for the majority of citizens in this low-income economy the heightened volatility of nominal incomes, while still harmful, will have been offset by the dampened volatility of the prices of the agricultural goods that they consume in large measure. That is, the net effect of railroads on the volatility of these poor consumers' real income streams will reflect a combination of the nominal income responsiveness results seen in this section and the price responsiveness results from the previous section. In the following section we go on to explore the net effect of these two offsetting changes in the responsiveness of rainfall shocks due to India's new railroad network.

4.2.3 Prediction 3: Openness and Real Income Responsiveness

We now turn to the set of empirical results motivated by Prediction 3, that the effect of local productivity shocks on local *real* agricultural income should be lower in an open economy than in a closed economy. This is in contrast to the results of the previous section, where *nominal* agricultural incomes were found to become more responsive once districts opened to rail trade. We see the results on real income here as the most important in the paper.

As discussed above, the effect of openness on the responsiveness of real incomes to productivity shocks is in principle ambiguous. It is natural to expect openness to reduce the responsiveness of prices to productivity shocks—as expected in Prediction 1 and found empirically in Section 4.2.1 above. Given this, it is then natural to expect trade openness to increase the responsiveness of nominal incomes to productivity shocks—as expected in Prediction 2 and found empirically in Section 4.2.2 above. However, the net effect of openness on the responsiveness of real incomes—the ratio of nominal incomes to a consumer-based price index—to productivity shocks will then depend on the relative strength of the counterveiling nominal income and price index effects. Prediction 3 was unambiguous about this

[13] It is possible that while railroad access increased the responsiveness of a district's nominal income to its own rainfall, railroad connections could also have altered the responsiveness to neighboring districts' rainfall (as we found in the case of prices, in Table 2 above). We have estimated a specification similar to equation (7) but with an extension to include dependence on neighboring districts' rainfall and an interaction term for neighboring districts that are bilaterally connected by rail to the district of observation. However, the coefficients on these two additional terms (neighbors' rainfall and neighbors' rainfall for railroad connected neighbors) are small and not statistically different from zero (jointly or individually) so we have not pursued this further. This result is perhaps unsurprising given the weak price spillover effects estimated in Table 2.

net effect, predicting that real income responsiveness should be reduced by trade openness. But in a wider class of models it is likely that this result could be overturned.

In this section we therefore explore what the data from colonial India have to say about the relationship between trade openness and real income responsiveness to productivity shocks. That is, we examine the extent to which railroads altered the mapping from a given rainfall shock to the level of real agricultural income. As discussed in section 2.4, the 1880 Famine Commission recommended railroad construction in exactly the hope that this mapping would change for the better.

To test prediction 3 we estimate a specification that is identical to that in the previous section (ie equation (7)) but we instead use *real* agricultural income as the dependent variable. Prediction 3 suggests that the coefficients γ_1 and γ_2 will take the same signs as in previous sections (both positive). However, the coefficient γ_3 is now expected to be negative according to Prediction 3, indicating that real agricultural income is now less responsive to rainfall shocks once a district is able to trade by the use of railroads.

The data used here are identical to those used in the previous section, except that the dependent variable is now real rather than nominal agricultural income. We construct a measure of real agricultural income by dividing nominal agricultural income by a consumption-based price index, defined over the 17 main agricultural crops for which price data are available. To construct consumption weights we follow the procedure used in Donaldson (2010). This procedure uses inter-regional trade data (available only for 'trade blocks', regions that comprise approximately five districts each) and output data (at the district level) to estimate consumption patterns at the trade block level, and then assigns each trade block's consumption weights to all of the districts in the given trade block).

The results from this test of Prediction 3 are presented in columns 3 and 4 of Table 3. These columns are the analogues of columns 1 and 2 respectively, but are based on the dependent variable of real agricultural income rather than nominal agricultural income.

Column 3 of Table 3 demonstrates that, on average, rainfall played a considerable role in determining the level of real agricultural income in colonial India. The results in column 4, however, illustrate how railroads mitigate the effect of agricultural productivity shocks on real agricultural income. That is, in line with Prediction 3, the coefficient on the interaction between rainfall and railroads is negative—which implies that an equivalent negative rainfall shock did less damage to real agricultural incomes after the arrival of railroads than before. Of course, a symmetric implication of the results from this symmetric specification is that a given *positive* shock did less to *raise* real income after the arrival of railroads as well. This result suggests that transportation infrastructure projects like India's railroads can bring about significant real income insurance, at least for a representative agent. In the next and final section of this paper we investigate whether this real income insurance appears to have given rise to consumption insurance too.

4.2.4 Prediction 4: Openness and Export Responsiveness

A natural implication of the results corresponding to Prediction 1—namely, that price responsiveness fall with openness—is that openness allows trade flows so as to counter-act the local quantity surpluses or deficits created by rainfall variation. Put differently, it is natural to expect that net exports respond positively to productivity shocks and that this responsiveness rises with trade openness. To maintain trade balance, net imports should be expected to rise similarly in the non-agricultural sectors not subject to rainfall variation. We explore the empirical validity of these two predictions here, and additionally explore whether trade remains balanced (in the short-run) in the presence of productivity shocks due to rainfall.

These results draw on the best available data on intra-national trade within colonial India in our time period, that on the flow of goods on railroads, rivers and coastal shipping routes from 1880-1921 collected by Donaldson (2010). These data track trade flows among 45 trade blocks that span the Indian sub-continent; naturally, these 45 trade blocks are smaller than any of the 239 districts used as the geographic unit of observation throughout the rest of this paper. An attractive feature of these data is that they delineate trade in each of 78 commodities, 17 of which are agricultural and the remainder of which are not. As suggested by Prediction 4 we look for differential effects across agricultural goods (whose production is rain-fed) and non-agricultural goods (whose production process is, presumably, less tied to rainfall).

Table 4 presents the empirical results based on this data, corresponding to Prediction 4. Following the same pattern of analysis as in the previous sub-sections above, Columns (1) and (2) explore the responsiveness to rainfall of net exports (that is, a region's exports minus its imports) in agricultural goods. Column (1) shows that, as one might expect given the above results on production, net exports in agricultural goods respond positively and statistically significantly to the amount of rainfall available. And column (2) shows how this responsiveness increases as a trade block has greater railroad penetration.

Columns (3) and (4) of Table 4 demonstrate that similar results hold for the second component of Prediction 4, namely that net imports of non-agricultural goods are also postively responsive to rainfall variation on average (column (3)) and more so with heightened railroad penetration (column (4)). This result is less mechanical than that in columns (3) and (4) as it works through the general equilibrium nature of production and consumption in an open economy—if net exports rise in a sector due to an exogenous shock to productivity in that sector then either net imports in other sectors have to rise or trade will be (more) imbalanced (ie in net surplus).

The result in columns (1) through (4) suggest that net exports in agricultural goods and net imports in non-agricultural goods move in off-setting ways in the response to productivity shocks in the agricultural sector. A natural question is whether these off-setting responses are of similar magnitude such that net trade surpluses (ie net exports minus net imports) are largely unaffected by rainfall shocks. In a one-period model such as that we develop above, trade is always balanced (or perhaps, equivalently, in a state of exogenous imbalance). But in a dynamic model it is possible that a region that enjoys a productivity shock will export more (on net) without necessarily receiving more net imports in return—the trade surplus could be financed through lending more (or borrowing less). Our results in column (5) and (6) of Table 4 suggest that regional trade imbalances were remarkably constant—at least in the face of exogenous rainfall shocks—in this time period. That is, the coefficient on rainfall in column (5) is small and statistically insignificant.

4.2.5 Prediction 5: Openness and Mortality Responsiveness

The results above have demonstrated that, in colonial India, railroads had an economically and statistically significant effect on reducing real income volatility, by reducing the extent to which a given productivity shock affected the real value of output in terms of what a consumer could buy with this output.

A natural follow-up to this investigation into real income volatility would lead to real consumption volatility; unfortunately, however, there exist no district-level consumption series from this time period in colonial India. But it is likely, in this setting, that a reduction in the volatility of real income may have passed through into consumption volatility, as most citizens had no access to formal insurance or banking facilities. For example, Roy (2001) describes how even even the wealthiest members of society in colonial India (outside of major cities) resorted to money-boxes and jewelery as the only means to save. Rosenzweig and Binswanger (1993) and Rosenzweig and Wolpin (1993) document limited access to insurance even in richer, post-Independence India. The gains from reduced consumption volatility may have been even more important to poor consumers due to subsistence concerns (or if risk aversion decreases with income more generally).

In this section we present results relating to the mortality rate—both because the mortality rate should provide a window on consumption in this low-income setting, and because the results relating to mortality are, we believe, important in their own right. They are especially relevant in the current setting where consumption volatility was apparently so severe as to give rise to eleven famines in just 46 years. In a sense, the results below relating to the mortality rate put a human face on earlier results which documented how lower trade costs reduced real income volatility.

We estimate a specification that follows the general procedure of equation (7), but with the log total (all-causes) mortality rate as the dependent variable rather than log real agricultural income. Prediction 5 suggests that the pattern of coefficient signs should be the same (but flipped, since mortality should track consumption inversely) as in Prediction 3. Importantly, Prediction 5 states that $\gamma_3 > 0$, or that openness (due to railroad access) should reduce the equilibrium exposure of local mortality rates to local productivity shocks (due to exogenous rainfall variation).

We use data on the total mortality rate for each district and year for which this measure is available post-1870. While the raw data list multiple separate causes of death, we first aggregate over this information to create a measure of the number of deaths due to all causes. We then divide this figure by district population (interpolated exponentially between census decadal years) to construct a mortality rate per 1,000 inhabitants.

The results from this test of Prediction 5 are reported in Table 5, columns (1) and (2). These two columns are analogous to those in columns (3) and (4) of Table 3, but use the log mortality rate as the dependent variable rather than log real income. A similar pattern of coefficient estimates—but with signs inverted since we expect mortality and income to co-move negatively—exists in columns (1) and (2) of Table 5 as in columns (3) and (4) of Table 3, suggesting that the death rate tracks real income reasonably closely.

The results in column (2) of Table 5 are particularly striking. They suggest that while local rainfall variation played a considerable role in determining the local death rate prior to the arrival of railroads in a district, after the railroad's arrival there is virtually no effect of local rainfall on local death. This suggests an important but novel insurance role played by railroads in this setting, and is, to the best of our knowledge, a first test of the claim that railroads did much to rid India of famine, as was hoped by the 1880 Famine Commission.

To explore this finding in more detail we next disaggregate the death reate into causes of death that Maharatna (1996) has argued do not increase in times of famine (ie times of low rainfall)—namely, 'plague', 'small-pox' and 'accidents'—and causes of death that may increase in times of famine—namely, 'fevers', 'respiratory diseases', 'diarrhoeal diseases/bowel complaints', 'cholera', 'accidents' and 'other'. Column (3) presents results corresponding to the first set (what Maharatna (1996) argues are the non-famine causes of death) and column (4) presents results corresponding to the second set (what Maharatna (1996) argues are the famine causes of death). The results corroborate our interpretation of our results throughout this paper. Neither railroad penetration nor rainfall amounts (nor indeed their interaction) appear to have a bearing on the mortality of Indian residents in the types of causes of death that may be expected to have little to do with rainfall variation. By contrast, column (4) suggests that our main results in column (2) are clearly being driven by exactly the causes of death that Maharatna (1996) argues one should see respond to rainfall variation and that, presumbably, freer movement of people and goods via railroads could help to mitigate.

4.2.6 Prediction 6: A Sufficient Statistic for the Impact of Openness on Responsiveness

The findings corresponding to Predictions 1-5 above are qualitatively consistent with the model we outlined in section 3. However, in this section we explore not just the qualitative consistency of our model with the data from this setting but also the extent of quantitative agreement between the two. This is a simple exercise thanks to the extremely strong nature of Prediction 6—that railroads should have no effect on equilibrium real income responsiveness while controlling for the trade share. Put another way, in this model, the only manner in which railroads can affect real income (and hence real income responsiveness) is via the trade share, which is a sufficient statistic for real income in the model.

To implement this test we follow Donaldson (2010) and compute the trade share, ie π^k_{oot} for each district, year and commodity. We do so using the full general equilibrium structure of the model. This is deliberate: not only is data on the trade share not directly available (only imperfect proxies for the extent to which a district trades with itself are available) but, as the model makes clear, the trade share is an endogenous variable (ie it is determined simultaneously in the model with the level of real income) so computing the predicted trade share (ie predicted on the basis of the empriical counterparts of the exogenous variables, trade costs and rainfall, in the model) and using this in our test obviates concerns of simultaneity bias. For details of our estimation and computation procedure see Donaldson (2010).

Table 6 presents our empirical results corresponding to Prediction 6. Column (1) replicates column (4) from Table (3)—that is, it reports estimates of the interaction between railroads and rainfall, a coefficient that is negative, in line with Prediction 3. Column (2) reports a regression that is identical to that in column (1) but for the inclusion of an additional variable, the trade share (as computed in the model, as described above). As suggested by Prediction 6, the coefficient on the rainfall-railroad interaction term (ie γ_3 in equation (7)), falls from a value that was large and statistically significant in column (1) to one that is small and statistically insignificant in column (2). That is, controlling for the trade share dramatically reduces the explanatory power of a railroad penetration dummy to account for real income responsiveness. Columns (3) and (4) show that a similar finding holds in the case of how railroads appear to reduce the responsiveness of mortality to rainfall—that is, again, and in line with Prediction 6, this is quantitatively well accounted for by the model-based sufficient statistic.

5 Concluding Remarks

By exploiting the change in trade costs brought about by railroads—which was large and differentiated over regions and time—as well as the relatively exogenous impact of district rainfall shocks as an observable underlying source of volatility, this paper has aimed to credibly establish whether openness increases or decreases volatility in a poor, agricultural setting. Our goal has been to move beyond reduced-form analysis of the impact of openness and volatility and trace out the economic effects of climatic shocks on equilibrium living standards as they differ across relatively closed and open trading environments. Our findings suggest that railroads strongly reduced the responsiveness of prices, and increased the responsiveness of nominal incomes, to productivity shocks. But in effect the price stabilizing effect of increased openness dominates the destabilizing effect on nominal incomes thus dampening overall real income volatility. Our results are corroborated by similar findings on the responsiveness of trade flows and mortality (an empirical proxy for consumption in a setting such as this). And our results are well accounted for quantitatively by an Eaton-Kortum model of trade in the sense that controlling a model-based sufficient statistic for any potential railroad-driven change in economic responsiveness appears to substantially reduce the empriical effect of railroads on both real income and mortality responsiveness. However, while our findings appear to be (both qualitatively and quantitatively) consistent with this class of models—as well as with a wider explanation of our findings that would be grounded in a reduction in income or mortality responsiveness driven by the cheaper ability to move goods—alternative interpretations for our findings are possible. For example, it is possible that railroads enabled the cheaper movement of people, of factors of production, or of transfers (though our finding about relatively static trade balances suggests that such transfers may not have been particularly responsive to rainfall variation). In ongoing work we aim to use additional sources of data to probe these alternative explanations.

Overall these results suggest that openness can play an important role in reducing the exposure of rural citizens' livelihoods to the riskiness of their environments. Further, this set of results documents benefits of openness and transportation infrastructure that go beyond the traditional effects of higher income levels found in Donaldson (2010) and work such as Frankel and Romer (1999).

The paper thus uncovers a central role of infrastructure investments that reduce mobility costs in protecting the welfare of rural citizens against climatic shocks. As many parts of the developing world remain poorly integrated both domestically and internationally trade facilitation seems just as relevant today as it did during the colonial period in India. Our findings point to the fact that the value of opennness in these remote, agriculture dependent locations lies not just with its role in raising real incomes but also in the key role it can play in reducing the volatility of those income streams.

References

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

Arnold, D. (1993): *Colonizing the Body: State Medicine and Epidemic Disease in 19th-century India*. Berkeley: University of California Press.

Bhatia, B. M. (1967): *Famines in India*. Delhi: Konark Publishers.

Burgess, R., and D. Donaldson (2011): "Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era," *American Economic Review Papers and Proceedings*.

Caselli, F., M. Koren, M. Lisicky, and S. Tenreyro (2011): "Diversification Through Trade," LSE working paper.

Costinot, A., D. Donaldson, and I. Komunjer (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79(2), MIT working paper.

Davis, K. (1951): *The Population of India and Pakistan*. Princeton: Princeton University Press.

Davis, M. (2001): *Late Victorian Holocausts: El Nino Famines and the Making of the Third World*. New York: Verso.

de Waal, A. (1989): *Famine that Kills*. Oxford: The Clarendon Press.

di Giovanni, J., and A. Levchenko (2009): "Trade Openness and Volatility," *Review of Economics and Statistics*.

Donaldson, D. (2010): "Railroads of the Raj: Estimating the Economic Impact of Transportation Infrastructure," mimeo MIT.

Dreze, J. (1988): "Famine prevention in India," WIDER working paper number 45.

Dreze, J., and A. Sen (1989): *Hunger and Public Action*. Oxford: Oxford University Press.

Dyson, T. (1991): "On the Demography of Famines: Part I," *Population Studies*, 45(2), 279–297.

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

Feyrer, J. (2009a): "Distance, Trade, and Income—The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment," NBER working paper 15557.

——— (2009b): "Trade and Income—Exploiting Time Series in Geography," NBER working paper 14910.

Frankel, J. A., and D. Romer (1999): "Does Trade Cause Growth?," *American Economic Review*, 89(3), 379–399.

Gadgil, S., M. Rajeevan, and P. A. Francis (2007): "Monsoon variability: Links to major oscillations over the equatorial Pacific and Indian oceans," *Current Science*, 93 (2), 182–194.

Hasan, R., D. Mitra, and K. Ramaswamy (2007): "Trade Reforms, Labor Regulations, and Labor-Demand Elasticities: Empirical Evidence from India," *Review of Economics and Statistics*, 89(3), 466–481.

Johnson, J. (1963): *The Economics of Indian Rail Transport*. Bombay: Allied Publishers.

Keller, W., and C. H. Shiue (2007): "The Origins of Spatial Interaction," *Journal of Econometrics*, 140(1), 304–332.

Krishna, P., and M. Z. Senses (2009): "International Trade and Labor Income Risk in the United States," NBER working paper 14992.

Maharatna, A. (1996): *The Demography of Famines: An Indian Historical Perspective*. Delhi: Oxford University Press.

McAlpin, M. B. (1979): "Dearth, Famine, and Risk: The Changing Impact of Crop Failures in Western India, 1870-1920," *Journal of Economic History*, 39(1), 143–157.

——— (1983): *Subject to Famine: Food Crises and Economic Change in Western India, 1860-1920*. Princeton: Princeton University Press.

Meng, X., N. Qian, and P. Yared (2011): "The Institutional Causes of Famine In China 1959-61," NBER working paper 16361.

Newbery, D. M., and J. E. Stiglitz (1981): *The Theory of Commodity Price Stabilization : a Study in the Economics of Risk*. Oxford: Clarendon Press.

O'Grada, C. (2007): "Making Famine History," *Journal of Economic Literature.*

——— (2008): "Sufficiency, sufficiency and sufficiency: revisiting the Bengal Famine of 1943-44," mimeo University College Dublin.

——— (2010): *Famine: A Short History.* Princeton: Princeton University Press.

Padmanabhan, S. Y. (1973): "The Great Bengal Famine," *Annual Review of Phytopathology*, 11, 11–24.

Ravallion, M. (1997): "Famines and Economics," *Journal of Economic Literature*, 35(3), 1205–1242.

Rodrik, D. (1997): *Has Globalization Gone Too Far?* Institute for International Economics.

Rosenzweig, M. R., and H. Binswanger (1993): "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *Economic Journal*, 103, 56–78.

Rosenzweig, M. R., and K. I. Wolpin (1993): "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India," *Journal of Political Economy*, 101, 223–244.

Roy, T. (2001): *The Economic History of India, 1857-1947.* Oxford: Oxford University Press.

Sen, A. (1981): *Poverty and Famines: An Essay on Entitlement and Deprivation.* Oxford: Oxford University Press.

Shiue, C. H. (2002): "Transport Costs and the Geography of Arbitrage in Eighteenth-Century China," *American Economic Review*, 92(5), 1406–1419.

Srivastava, H. (1968): *The History of Indian Famines and Development of Famine Policy, 1858-1918.* Sri Ram Mehra.

Stokes, E. (1959): *The English Utilitarians and India.* Oxford University Press.

Tauger, M. (2003): "Entitlement, Shortage and the 1943 Bengal Famine: Another Look," *The Journal of Peasant Studies*, 31, 45–72.

Visaria, L., and P. Visaria (1983): "Population (1757-1947)," in *The Cambridge Economic History of India, Volume 2*, ed. by D. Kumar. Cambridge: Cambridge University Press.

A Theoretical Appendix

In this Appendix we provide a more detailed description of the model from Section 3 of the main text.

The economy consists of D regions (indexed by either o or d) and K commodities (indexed by k). Each available in a continuum (with mass normalized to one) of horizontally differentiated varieties (indexed by j). Our empirical work below refers to data on prices and output of commodities (aggregates over sets of varieties), not individual varieties.

Consumer Preferences:

Each region o is home to a mass (normalized to one) of identical agents, each of whom owns L_o units of land. Land is geographically immobile and supplied inelastically. Agents have Cobb-Douglas preferences over commodities (k) and constant elasticity of substitution preferences over varieties (j) within each commodity; that is, their (log) utility function is

$$\ln U_o = \sum_{k=1}^{K} \left(\frac{\mu_k}{\varepsilon_k}\right) \ln \int_0^1 (C_o^k(j))^{\varepsilon_k} dj, \tag{9}$$

where $C_o^k(j)$ is consumption, $\varepsilon_k \doteq \frac{\sigma_k - 1}{\sigma_k}$ (where σ_k is the constant elasticity of substitution), and $\sum_k \mu_k = 1$. Agents rent out their land at the rate of r_o per unit and use their income $r_o L_o$ to maximize utility from consumption.

Production and Market Structure:

Each variety j of the commodity k can be produced using a constant returns to scale production technology in which land is the only factor of production. Let $z_o^k(j)$ denote the amount of variety j of commodity k that can be produced with one unit of land in region o. We follow Eaton and Kortum (2002) in modeling $z_o^k(j)$ as the realization of a stochastic variable Z_o^k drawn from a Type-II extreme value distribution whose parameters vary across regions and commodities in the following manner

$$F_o^k(z) \doteq \Pr(Z_o^k \leq z) = \exp(-A_o^k z^{-\theta_k}), \tag{10}$$

where $A_o^k \geq 0$ and $\theta_k > 0$. These random variables are drawn independently for each variety, commodity and region. The exogenous parameter A_o^k increases the probability of high productivity draws and the exogenous parameter θ_k captures (inversely) how variable the (log) productivity of commodity k in any region is around its (log) average.

There are many competitive firms in region o with access to the above technology; consequently, firms make zero profits. These firms will therefore charge a pre-trade costs (ie, 'free on board') price of $p^k_{oo}(j) = r_o/z^k_o(j)$, where r_o is the land rental rate in region o.

Opportunities to Trade:

Without opportunities to trade, consumers in region d must consume even their region's worst draws from the productivity distribution in equation (10). The ability to trade breaks this production-consumption link. This allows consumers to import varieties from other regions in order to take advantage of the favorable productivity draws available there, and allows producers to produce more of the varieties for which they received the best productivity draws. These two mechanisms constitute the gains from trade in this model.

However, there is a limit to trade because the movement of goods is subject to trade costs (which include transport costs and other barriers to trade). These trade costs take the convenient and commonly used 'iceberg' form. That is, in order for one unit of commodity k to arrive in region d, $T^k_{od} \geq 1$ units of the commodity must be produced and shipped in region o; trade is free when $T^k_{od} = 1$. (Throughout this paper we refer to trade flows between an origin region o and a destination region d; all bilateral variables, such as T^k_{od}, refer to quantities from o to d.) Trade costs are assumed to 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, $T^k_{od} \leq T^k_{om} T^k_{md}$. Finally, we normalize such that $T^k_{oo} = 1$.

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 commodity k produced in region o, but shipped to region d for consumption there. Trade costs imply that any variety in region d will cost T^k_{od} times more than it does in region o; that is, $p^k_{od}(j) = T^k_{od}\, p^k_{oo}(j) = r_o\, T^k_{od}/z^k_o(j)$.

Equilibrium Prices and Allocations:

Consumers have preferences for all varieties j along the continuum of varieties of commodity 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 (after accounting for trade costs). We therefore solve for the equilibrium prices that consumers in a region d actually pay, given that they will only buy a given variety from the cheapest source region (including their own).

The price of a variety sent from region o to region d, denoted by $p^k_{od}(j)$, is stochastic because it depends on the stochastic variable $z^k_o(j)$. Since $z^k_o(j)$ is drawn from the CDF in equation (10), $p^k_{od}(j)$ is the realization of a random variable P^k_{od} drawn from the CDF

$$G^k_{od}(p) \doteq \Pr(P^k_{od} \leq p) = 1 - \exp[-A^k_o (r_o T^k_{od})^{-\theta_k} p^{\theta_k}]. \tag{11}$$

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

$$\begin{aligned} G_d^k(p) &= 1 - \prod_{o=1}^{D}[1 - G_{od}^k(p)], \\ &= 1 - \exp\left(-\left[\sum_{o=1}^{D} A_o^k (r_o T_{od}^k)^{-\theta_k}\right] p^{\theta_k}\right). \end{aligned}$$

Given this distribution of the actual prices paid by consumers in region d, it is straightforward to calculate any moment of the prices of interest. The price moment that is relevant for our empirical analysis is the expected value of the equilibrium price of any variety j of commodity k found in region d, which is given by

$$E[p_d^k(j)] \doteq p_d^k = \lambda_1^k \left[\sum_{o=1}^{D} A_o^k (r_o T_{od}^k)^{-\theta_k}\right]^{-1/\theta_k}, \tag{12}$$

where $\lambda_1^k \doteq \Gamma(1 + \frac{1}{\theta_k})$.[14] In our empirical application below we treat these expected prices as equal to the observed prices collected by statistical agencies.[15]

Given the price distribution in equation (11), Eaton and Kortum (2002) derive two important properties of the trading equilibrium that carry over to the model here. First, the price distribution of the varieties that any given origin actually sends to destination d (ie, the distribution of prices for which this origin is region d's cheapest supplier) is the same for all origin regions. This implies that the share of expenditure that consumers in region d allocate to varieties from region o must be equal to the probability that region o supplies a variety to region d (because the price per variety, conditional on the variety being supplied to d, does not depend on the origin). That is $X_{od}^k / X_d^k = \pi_{od}^k$, where X_{od}^k is total expenditure in region d on commodities of type k from region o, $X_d^k \doteq \sum_o X_{od}^k$ is total expenditure in region d on commodities of type k, and π_{od}^k is the probability that region d sources any variety of

[14] $\Gamma(.)$ is the Gamma function defined by $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$.

[15] A second price moment that is of interest for welfare analysis is the exact price index over all varieties of commodity k for consumers in region d. Given CES preferences, this is $\widetilde{p}_d^k \doteq \left[\int_0^1 (p_d^k(j))^{1-\sigma_k} dj\right]^{1/1-\sigma_k}$, which is only well defined here for $\sigma_k < 1 + \theta_k$ (a condition we assume throughout). The exact price index is given by $\widetilde{p}_d^k = \lambda_2^k p_d^k$, where $\lambda_2^k \doteq \frac{\gamma^k}{\lambda_1^k}$ and $\gamma^k \doteq [\Gamma(\frac{\theta_k + 1 - \sigma_k}{\theta_k})]^{1/(1-\sigma_k)}$. That is, if statistical agencies sampled varieties in proportion to their weights in the exact price index, as opposed to randomly as in the expected price formulation of equation (12), then this would not jeopardize our empirical procedure because the exact price index is proportional to expected prices.

commodity k from region o. Second, this probability π_{od}^k is given by

$$\frac{X_{od}^k}{X_d^k} = \pi_{od}^k = \lambda_3^k \, A_o^k \, (r_o T_{od}^k)^{-\theta_k} \, (p_d^k)^{\theta_k}, \tag{13}$$

where $\lambda_3^k = (\lambda_1^k)^{-\theta_k}$, and this equation makes use of the definition of the expected value of prices (ie, p_d^k) from equation (12).

Equation (13) characterizes trade flows conditional on the endogenous land rental rate, r_o (and all other regions' land rental rates, which appear in p_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 ($r_o L_o$) must equal the total value of all commodities made in region o and sent to every other region (including region o itself). That is:

$$r_o L_o = \sum_d \sum_k X_{od}^k = \sum_d \sum_k \pi_{od}^k \, \mu_k \, r_d \, L_d, \tag{14}$$

where the last equality uses the fact that (with Cobb-Douglas preferences) expenditure in region d on commodity k (X_d^k) will be a fixed share μ_k of the total income in region d (ie, of $r_d L_d$). Each of the D regions has its own trade balance equation of this form. The equilibrium of the model is the set of D-1 unknown rental rates r_d that solves this system of D-1 (non-linear) independent equations.

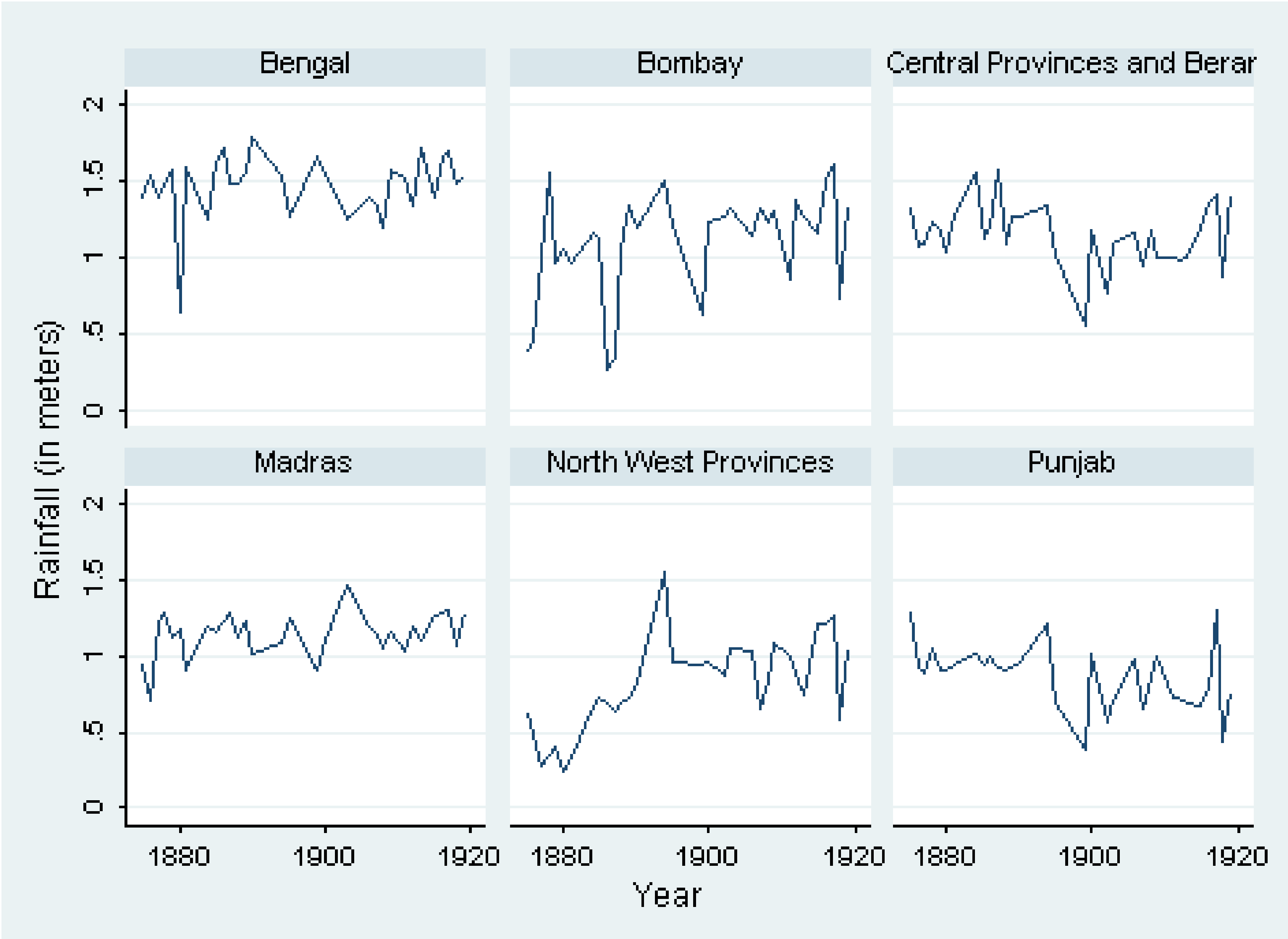

Figure 1: Annual Rainfall by Indian Province, 1875-1919: This figure plots the average amount of annual rainfall by province, averaging over the British districts within each province.

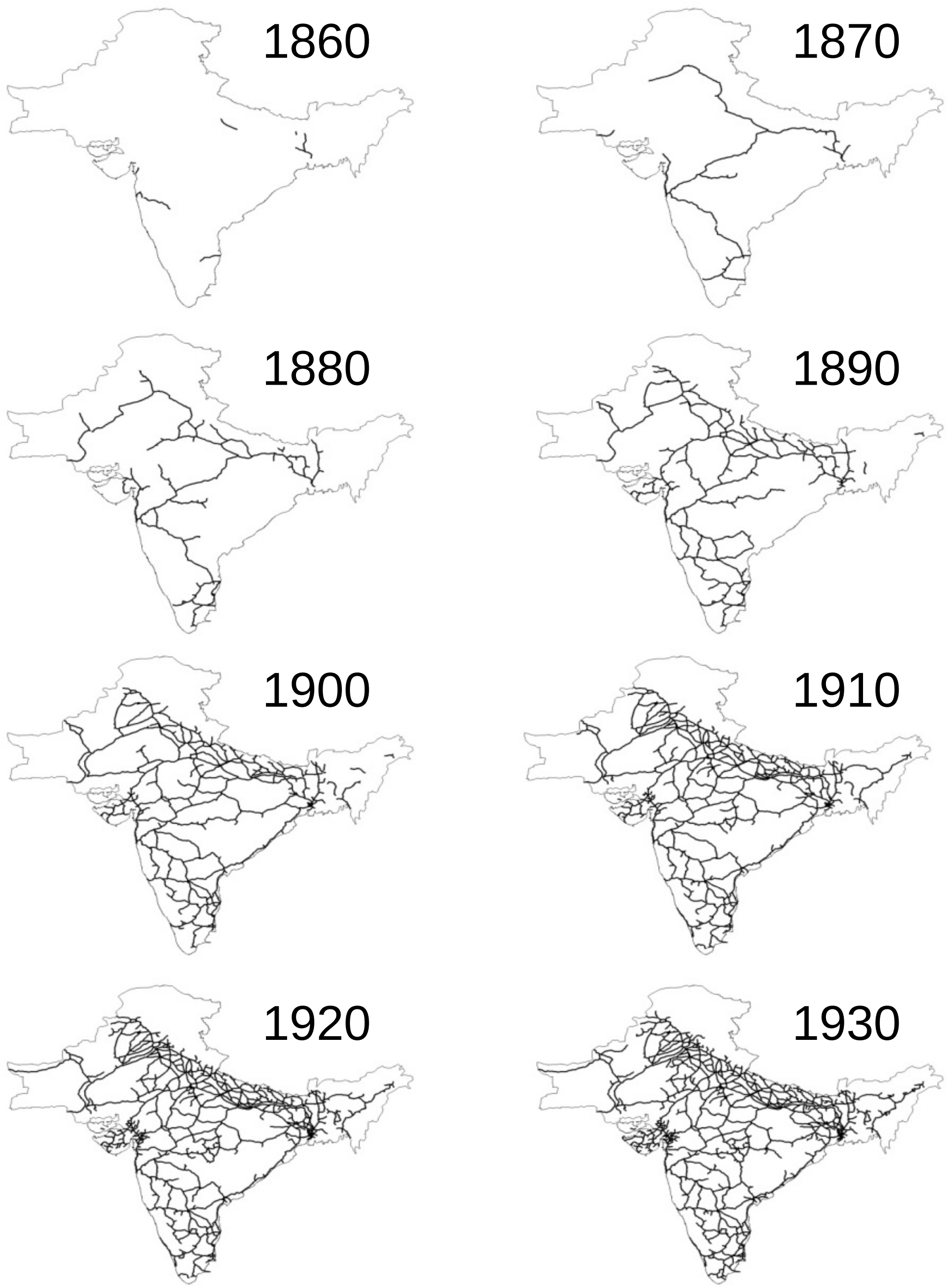

Figure 2: The evolution of Indias railroad network, 1860-1930: These figures display the decadal evolution of the railroad network (railroads depicted with thick lines) in colonial India (the outline of which is depicted with thin lines). The first railroad lines were laid in 1853. This figure is based on a GIS database in which each (approximately) 20 km long railroad segment is coded with a year of opening variable.

	Number of observations	Beginning of available data	End of available data
Crop-specific rainfall shock, coefficient of variation over past 5 years	68,821	0.520 (0.417)	0.541 (0.480)
Agricultural prices, averaged over all crops, coefficient of variation over past 5 years	13,384	0.108 (0.114)	0.024 (0.031)
Nominal agricultural output, coefficient of variation over past 5 years	13,384	0.110 (0.085)	0.135 (0.106)
Real agricultural income, coefficient of variation over past 5 years	13,384	0.06 (0.04)	0.04 (0.03)
Mortality rate due to all causes of death, per 1,000 population	14,672	31.8 (24.0)	35.4 (21.2)
Mortality rate due to all causes of death, coefficient of variation over past 5 years	13,006	0.221 (0.236)	0.116 (0.127)

Table 1: Summary Statistics: Notes: Values are sample means over all observations for the year and question, with standard deviations in parentheses. Beginning and end of available data are: 1870 and 1930 for real and nominal agricultural incomes, and mortality rates; and 1861 and 1930 for agricultural prices.

Dependent variable: Log agricultural price	(1)	(2)	(3)	(4)
Local rainfall in sowing and growing periods	-0.256	-0.428	-0.215	-0.402
	(0.102)**	(0.184)***	(0.105)**	(0.125)***
(Local rainfall) x (Railroad in district)		0.414		0.375
		(0.195)**		(0.184)**
Rainfall in neighboring districts			-0.051	-0.021
			(0.023)**	(0.018)
Average: (Rainfall in neighboring districts) x				
(Connected to neighboring districts by railroad)				-0.082
				(0.036)***
Crop x Year fixed effects	YES	YES	YES	YES
Crop x District fixed effects	YES	YES	YES	YES
District x Year fixed effects	YES	YES	YES	YES
Observations	73,000	73,000	73,000	73,000
R-squared	0.891	0.892	0.894	0.899

Table 2: The Effect of Railroads on Agricultural Price Responsiveness: Notes: OLS Regressions estimating equation (8) using data on 17 agricultural crops, from 239 districts in India, annually from 1861 to 1930. 'Local rainfall in sowing and growing period' (abrev. 'local rainfall') refers to the amount of rainfall (measured in meters) in the district in question that fell during crop- and district- specific sowing and harvesting dates. 'Railroad in district' is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. 'Rainfall in neighboring districts' is the variable 'local rainfall' averaged over all districts within a 250 km radius of the district in question. 'Connected to neighboring district' is a dummy variable that is equal to one if the district in question is connected by a railroad line to each neighboring district within 250 km. Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1 % level; ** indicates 5 % level; and * indicates 10 % level.

Dependent variable:	Log nominal agricultural income per acre		Log real agricultural income per acre	
	(1)	(2)	(3)	(4)
Rainfall in district	1.410	0.931	1.248	2.021
	(0.632)***	(0.414)***	(0.430)***	(0.839)**
Railroad in district	0.241	-0.368	0.186	1.196
	(0.114)**	(0.148)**	(0.085)**	(0.482)***
(Railroad in district) x (Rainfall in district)		0.865		-1.444
		(0.381)**		(0.496)***
District fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Observations	14,340	14,340	14,340	14,340
R-squared	0.771	0.776	0.767	0.772

Table 3: The Effect of Railroads on Nominal and Real Income Responsiveness Notes: OLS Regressions estimating equation (7) using agricultural income (as the variable Y_{dt} in equation (7)) constructed from crop-level data on 17 principal agricultural crops, from 239 districts in India, annually from 1870 to 1930. 'Railroad in district' is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. 'Rainfall in district' is the amount of rainfall that fell in the district-year (in meters). Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1 % level; ** indicates 5 % level; and * indicates 10 % level.

Dependent variable:	Log net exports of agricultural goods		Log net imports of non-agricultural goods		Log trade surplus	
	(1)	(2)	(3)	(4)	(5)	(6)
Average rainfall in trade block	0.467	0.184	0.361	0.144	0.048	0.043
	(0.186)***	(0.099)*	(0.149)**	(0.096)	(0.095)	(0.101)
Share of districts (in trade block) with railroad	0.871	0.501	0.815	0.531	0.071	0.059
	(0.217)***	(0.119)**	(0.199)***	(0.225)***	(0.104)	(0.095)
(Share of districts with railroad) x (Average rainfall)		0.512		0.394		0.009
		(0.202)***		(0.219)*		(0.099)
Trade block fixed effects	YES	YES	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES	YES	YES
Observations	1,845	1,845	1,845	1,845	1,845	1,845
R-squared	0.639	0.657	0.610	0.623	0.451	0.467

Table 4: The Effect of Railroads on Export and Import Responsiveness Notes: OLS Regressions estimating equation (7) using total trade flows (as the variable Y_{dt} in equation (7)) from 45 trade blocks in India, annually from 1880 to 1921. The dependent variable in columns (1) and (2) is the net exports (ie exports minus imports, summed across all destinations and origins) of agricultural commodities (where agricultural commodities are the 17 principal agricultural commodities used to compute agricutural income in Table 3). The dependent variable in columns (3) and (4) is the net imports (ie imports minus exports, summed across all origins and destinations) of non-agricultural commodities (all traded commodities other than those classified as agricultural). The dependent variable in columns (5) and (6) is the trade surplus (ie the sum of net exports and net imports, across all commodities). 'Share of districts (in trade block) with railroad' is the average of a dummy variable whose value is one if any part of a district in question is penetrated by a railroad line, averaged across the (typically 4-5) districts in each trade block. 'Average rainfall in trade block' is the average (across districts in the trade block) amount of rainfall that fell in the trade block-year (in meters). Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1 % level; ** indicates 5 % level; and * indicates 10 % level.

Dependent variable:	Log total mortality rate		Log mortality due to plague, smallpox or accidents	Log mortality due to all other causes
	(1)	(2)	(3)	(4)
Rainfall in district	-0.064	-0.118	-0.012	-0.123
	(0.032)**	(0.052)**	(0.052)	(0.050)**
Railroad in district	-0.080	-0.144	0.024	-0.156
	(0.061)	(0.078)*	(0.065)	(0.079)*
(Railroad in district) x (Rainfall in district)		0.098	0.006	0.104
		(0.044)**	(0.021)	(0.041)**
District fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Observations	13,512	13,512	13,512	13,512
R-squared	0.642	0.649	0.434	0.716

Table 5: The Effect of Railroads on Mortality Responsiveness Notes: OLS Regressions estimating equation (7) using mortality rates (as the variable Y_{dt} in equation (7)) from 239 districts in India, annually from 1870 to 1930. The dependent variable in columns (1) and (2) is the all-causes mortality rate (annual deaths divided by estimated mid-year population). The dependent variable in column (3) is the mortality rate computed using only the causes-of-death, 'plague', 'small-pox' and 'accidents'. The dependent variable in column (4) is the mortality rate computed using all causes of death other than those used in column (3), namely 'fevers', 'respiratory diseases', 'diarrhoeal diseases/bowel complaints', 'cholera', 'accidents' and 'other'. 'Railroad in district' is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. 'Rainfall in district' is the amount of rainfall that fell in the district-year (in meters). Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1 % level; ** indicates 5 % level; and * indicates 10 % level.

Dependent variable:	Log real agricultural income per acre		Log total mortality rate	
	(1)	(2)	(3)	(4)
Rainfall in district	2.021	1.058	-0.118	-0.103
	(0.839)**	(0.484)**	(0.052)**	(0.055)*
Railroad in district	1.196	0.022	-0.144	-0.034
	(0.482)***	(0.095)	(0.078)*	(0.091)
(Railroad in district) x (Rainfall in district)	-1.444	0.038	0.098	0.014
	(0.496)***	(0.061)	(0.044)**	(0.082)
Trade share, as computed in model		-0.897		0.105
		(0.159)***		(0.036)***
District fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Observations	14,340	14,340	13,512	13,512
R-squared	0.772	0.786	0.649	0.662

Table 6: The Effect of Railroads on Responsiveness, Controlling for a Model-Based Sufficient Statistic Notes: OLS Regressions estimating equation (7) using either real agricultural income (columns (1) and (2)) or the all-causes mortality rate (columns (3) and (4)) as the variable Y_{dt} in equation (7), from 239 districts in India, annually from 1870 to 1930. 'Railroad in district' is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. 'Rainfall in district' is the amount of rainfall that fell in the district-year (in meters). 'Trade share, as computed in model' is the share of the district in question's expenditure that it produces itself, averaged appropriately (ie as in the model) across 17 agricultural goods; this variable is computed using the full general equilibrium structure of the estimated model (estimated in Donaldson (2010)) using only the empirical counterparts of the model's exogenous variables (ie trade costs based on railroads, and rainfall). Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses. *** indicates statistically significantly different from zero at the 1 % level; ** indicates 5 % level; and * indicates 10 % level.