RAILWAYS AND PRICE CONVERGENCE IN BRITISH INDIA

June 2008

Tahir Andrabi*

tandrabi@pomona.edu

Michael Kuehlwein*
mkuehlwein@pomona.edu

Abstract

The period 1861 to 1920 witnessed sharp price convergence in British Indian wheat and rice markets. Previous authors have assumed that since a vast railway system was built during that time, railways were probably the main cause of that convergence. But tests examining relative price differences between individual districts provide surprisingly weak support for that hypothesis. Railways clearly mattered, but even controlling for the effects on nearby districts, railways seem capable of explaining only about 20% of the decline in grain price dispersion. The modest estimated effect fundamentally derives from the fact that prices were converging during this time almost as rapidly between districts without railways as with railways. Our fixed effect estimation using yearly dummies points out that the pattern of price convergence between a given district pair was no different before they got railways than after they got the rail. One reason for this may be that India was already a well established, partially integrated economy at the time of the railroad expansion. Some lines were built along existing trade routes where there was already significant commerce between districts. The impact of railroads on districts not obviously linked by trade was shown to be about five times larger than the impact on previously well linked districts. There was also considerable heterogeneity of the railway effect depending on whether districts were already close to railways or not.

* Department of Economics, Pomona College, Claremont, CA 91711. We would like to thank Jishnu Das, Asim Ijaz Khwaja, Steve Marks and seminar participants at USC, the Pacific Coast Development Conference 2007, WEA Annual meetings 2005, and the NEUDC 2005 for valuable comments. All remaining errors are ours.
RAILWAYS AND PRICE CONVERGENCE IN BRITISH INDIA

A person living two thousand years ago would have recognized the main features of human existence [in India] right up to the middle almost of the 19th century… Then came the great change heralded by the telegraph and the railways which has progressively altered the very texture of life.1

Theory suggests that a unified goods market should generate important benefits for an economy. There is a comparative advantage effect that comes from greater regional specialization. There can be economies of scale. Larger markets may increase investment opportunities, leading to higher rates of capital formation. The ability to sell one’s product in a bigger market should also create incentives for innovation and technical progress.

One important dimension of market integration is how close prices are throughout the economy for the same goods. Declining price dispersion should be an indicator of greater market integration. Transportation costs, in turn, are an important determinant of price dispersion. Falling transport costs, therefore, should lead to less price dispersion and better integrated markets. That is why it is particularly instructive to study markets in the nineteenth century, as that period witnessed the emergence of extensive transportation networks in many parts of the world.

There is a rich literature on how the introduction of steamships, railways, the construction of canals and better roads during this time transformed transportation and brought prices closer together in many countries (e.g. O’Rourke and Williamson 1999).

---

The building of railroads, in particular, has been identified as crucial to these processes of price convergence and market integration.

This was certainly true for British India. During the latter half of the nineteenth century, the country built the fourth most extensive railway system in the world. By 1910 India had over 30,000 miles of track, just behind Germany and Russia.\(^2\) The period 1860 to 1920 was also marked by markedly declining price dispersion across districts in grain markets. Several historians contend that India’s railway system caused the fall in price dispersion and thereby unified the Indian economy.\(^3\) However, many of these assertions rest on the simple fact that price convergence and railway construction occurred at the same time. There have been very few attempts to go beyond this and rigorously test whether railroads were primarily responsible for this convergence.

We examine the effect of railway introduction on price dispersion in annual retail wheat and rice prices in over 150 Indian districts from 1860-1920. Our main contribution is to rigorously estimate the effect of railways on price dispersion. We examine relative prices for each district pair in a multiple regression analysis with fixed effects. The fixed effects control for constant characteristics of each pair such as the distance between them, physical features such as being on the Ganges, and historical relationships. In addition, we utilize yearly dummy variables to capture secular trends such as other transportation, communication or political changes happening in India that might have affected all districts. Once we include fixed effects, we are able to measure the average effect on relative grain prices for each pair before it got the railroad versus after.

\(^2\) The US, however, was way ahead. Railway density is perhaps more important and it was lower in India than in the US and Europe, but it was still higher than in South America, Africa, the rest of Asia, Canada, Russia, and Australia. (Prakash 1999, p. 165)

\(^3\) See Hurd (1975) for example.
We find that the effect of railways is statistically significant but small. Railways can explain about 10% of the decline in the relative price gap that we observe over these 60 years. There was a large drop in grain price dispersion across India during these years, but this drop was happening even before districts got connected by railroads. Ignoring this trend, therefore, produces misleadingly large effects of railways.

We extend our analysis in three directions to investigate the small magnitude of the railway effect. First, we check to see if the effect starts small in our sample period but grows large. We find no evidence of that. Secondly, we could be underestimating this effect if railways also impacted districts that were nearby. So we allow railways to affect prices in districts that lie close to districts that got railways. We find that there is a significant effect on nearby districts, and that allowing for it roughly doubles our railways estimate. Still, railways are able to explain only about 20% of the observed decline in price dispersion.

Third, we test whether railways had a small effect because they were imposed on an already integrated economy. We provide statistical evidence that Indian grain markets were integrated as early as the 1860s. We then look at particular routes such as the northern corridor from Delhi to Calcutta that had well established communication and transportation links such as the Ganges river and the Grand Trunk (GT) road. It is to be expected that the effect of railways on price dispersion would be less along that route. Similarly, the railway effect should be larger if it connected districts that were previously not well linked. The decision to link Bombay to the northern region provides us with a useful natural experiment. Interestingly, Bombay was not directly connected to either Delhi or Calcutta. Instead the connecting line was built through Jubbulpore in central
India to Allahabad in the United Provinces where it intersected midway the Delhi-Calcutta line. We find that among the districts along the Bombay-Allahabad line the effect of railways was five to six times larger than along the Delhi-Calcutta line. We conduct a similar exercise for districts in southern India, with its access to roads and the ocean, and find that railways had little effect there too.

One major issue in our research is how to control for the endogeneity of track placement. A possible solution is to examine the effect of railways on lines built for non-commercial purposes, such as military considerations. When we try that we still measure a modest railway effect, though the standard errors are large. Even with our various specifications, it is likely that we have some endogeneity-related bias. We discuss this issue and argue that any remaining bias is likely to inflate our railway estimates. We also explain why the distance between each district and a straight line between major cities is probably not a valid instrument in this context.

Our Fogelesque conclusion, therefore, is that railways played a surprisingly modest role in grain price convergence in British India. It is important to note, however, that we found considerable heterogeneity in the magnitude of this effect, depending on whether districts were already close to railways or lay on routes which had pre-existing transportation and communication networks.

The plan for the rest of the paper is as follows. The next section is a brief survey of the literature on changes in transportation and price convergence. Then we discuss the development of the Indian railway system and details of their wheat and rice markets. We lay out the theory and data used in the analysis and present the basic results as well as extensions investigating the small magnitude of the railway effect. A conclusion contains
speculation about what might have caused prices to converge if not railways and suggests extensions to our research.

RELATED LITERATURE

A large literature exists on commodity price convergence across nations, but more relevant for the purposes of our paper is the research on convergence within nations. Berry (1943, 106) reports that with the introduction of steamships in America, price differences for lard, pork, and wheat flour between New Orleans and Cincinnati fell by over 70% between 1816 and 1860. The gap between New York and Cincinnati prices declined even more. Williamson (1974, 259) provides evidence on shrinking grain price differences between the US Midwest and East from 1870-1910. Slaughter (2001) discovers the same pattern among 10 goods prices for 6 large US cities during the period 1820-1860 and notes these were years in which an extensive system of canals and railroads was being constructed. Metzer (1974) documents a narrowing of inter-provincial price differences in wheat and rye in Russia after 1870 and attributes it to the nation’s expanding railway system. O’Rourke and Williamson (1999) highlight the dramatic decline in the Bavarian/Prussian price gap in wheat and oats from 1854-1904, a time when Germany was establishing its rail system. Dobado and Marrero (2005) offer evidence that railroads accelerated the rate of interstate convergence in corn prices in Porfirian Mexico from 1885 to 1908.

A few papers on this issue focus on India. Hurd (1975) analyzes annual retail wheat and rice data from 188 districts in British India and reports that the coefficient of variation dropped by 60% between 1861 and 1920, the same time that India was building
its railways. In addition, he shows that price dispersion was consistently lower for
districts with railways. In a study of Bengal rice prices, Mukherjee (1980) calculates a
similar decline in dispersion between 1855 and 1912, and also finds less dispersion
among subdivisions that had early linkages to railways. Finally, McAlpin (1974)
document a sharp convergence in Indian cotton prices during this same period.
However, she finds that there were no concomitant changes in acreage devoted to either
cotton or food grains, which one would expect if railways impacted relative prices.4

Taken together, the literature demonstrates that there was substantial convergence in
grain and other commodity prices within several important economies in the 1800’s and
early 1900’s, and argues that the introduction of railroads was probably the main cause of
these changes. Many of these arguments, however, rest on the simple fact that price
convergence and railway construction occurred at the same time. But there have been no
serious attempts to measure if railways were the main force driving this process.

THE INDIAN RAILWAY SYSTEM

While discussions about railways started with Lord Dalhousie’s arrival in India as
Governor General in 1848, the impetus to develop a railway system came in the 1850s
with the passage of the parliamentary act following Dalhousie’s famous 1853 minute.
That same year saw the first commercial train introduced in Bombay by the Great Indian
Peninsula Railways (GIPR). Service near Calcutta with the East Indian Railways (EIR)
commenced the following year, and the Madras and Southern Mahratta Railway

4 More indirectly, Collins (1999) points out that trade theory indicates that commodity price equalization in
late nineteenth century India driven by transportation improvements could lead to wage convergence, of
which he finds little evidence. However he also references Samuelson (1948), who argues that the theory
rests on numerous strong assumptions.
(M&SM) opened a line in Madras in 1856. The construction and management of the railways was given to private companies initially, but with a financial guarantee and tight regulation by the government. Railways grew from the main ports towards the inland areas (see Figure 1). The EIR, for instance, was created to connect Calcutta to north India. The completion of the Yamuna bridge in 1866 resulted in an uninterrupted flow of rail service from Calcutta to Delhi. In March 1870 a major north-south connection was made as the GIPR reached Jubbulpore from Itarsi over the Thull Ghat, linking up with EIR track from Allahabad and establishing a link between Bombay and Calcutta. In 1871 the GIPR line extended over the Bhore Ghat to reach Raichur, connecting Bombay with Madras. By 1873 all 20 of India’s largest cities had rail service. The last major connection in the system occurred in 1878 between Multan and Karachi.5

Track mileage grew rapidly during and after these years. Track mileage was only 838 miles in 1860 but grew to 4,771 miles by 1870, 24,751 miles by 1900, and 35,199 miles by 1920.6 Though the system eventually covered most of the peninsula, initially it ran mainly from east to west and was concentrated in the north. In 1871, for instance, EIR track totaled 1350 miles, GIPR 875 miles, M&SM 680 miles, the Sind, Punjab, and Delhi railway 400 miles, the Bombay, Baroda and Central India railway 300 miles, the East Bengal Railway 115 miles, and the Great Southern Railway 170 miles. Significant construction of branch lines in central and southern India had to wait until the 1870’s and later.

---

5 Trains still, though, had to cross the Indus River at Sukkur by steam ferry eight cars at a time until the Landsdowne Bridge opened in 1888. See Administration Report on Railways in India 1911 for all the details.

6 Hurd 1975, p. 277.
Railways were built, in part, for economic gain. Trunk routes, wrote Dalhousie, “should be chosen according to the extent of political and commercial advantages which they are calculated to afford.”\(^7\) One of their main economic benefits to Britain was establishing a dependable supply of cheap food and other basic commodities. Minutes from planning meetings clearly indicate that lines were designed to reach major cotton and rice producing regions.\(^8\) They were also successful in that regard, helping to boost the value of Indian exports from £21 million in 1853 to £84 million by 1883 and £333 million by 1920. Britain was consistently the largest importer of Indian goods. Nonetheless, many of the lines were unprofitable. In 1900, 70% of the lines earned a return less than the 5% rate guaranteed to investors and had to be subsidized by the British government.\(^9\)

Military considerations were also a factor in track placement, especially in the northwest where lines were constructed to facilitate troop and supply movements to the frontier. Official strategic lines were few and relatively late in coming, but they included those stretching to the Bolan Pass near Quetta, the Khyber Pass beyond Peshawar, and Zahidan (now in Iran). Several other lines officially designated commercial, however, had significant military purposes such as those to Lahore, Rawalpindi, Multan, and Dera Ismael Khan with their important cantonments.

Famine lines designed to protect regions from disastrous crop failures were another salient part of the system. Indeed, in the early 1850’s the prevention of famines was advanced in Britain as one of the arguments for building an Indian railway system. It wasn’t until after the Famine Commission Report of 1880, however, that famine

---

\(^7\) Quoted in MacPherson 1955, p. 178.
\(^8\) Ibid.
\(^9\) Hurd 1983, p. 742.
considerations became important in the construction of new lines.10 Even then, famine line construction was limited. The Commission recommended that 5000 miles of line be built, but the Afghan War of 1878-80 convinced the British that those funds were needed elsewhere. The Indian Midland and Bengal-Nagpur Railways were, however, established in the 1880s in part to address famine concerns in Orissa and central India.11

There were thus a variety of factors influencing why and where railways were built in India. This will be relevant when we discuss strategies to obtain unbiased estimates of the impact of railways on price differences.

**INDIAN WHEAT AND RICE MARKETS**

Wheat production and consumption during this period were concentrated in northern India, principally the Punjab and United Provinces (see Figure 2). In 1919 those two states comprised 70% of wheat acreage in the country. Most wheat was consumed internally. In 1895, only 5% of wheat produced was exported, though that fraction rose to 15% by 1918. One-twelfth of the population in India was estimated to live on wheat.12

Rice production and consumption occurred more along the east coast including Bengal, Bihar and Orissa, and Madras. In 1919 those states made up 70% of cultivated rice acreage. The fraction of the population consuming rice was estimated in 1883 to vary from two-thirds in Bengal and Assam to only 5% in the Punjab.13 Most production was again for domestic purposes: only 7% of the 1895 rice crop was exported. That fraction remained fairly constant through the end of our sample. For the country as a

---

10 MacPherson 1955, p. 177.
11 Bell 1894, pp. 30 and 50.
12 Connell 1885, p. 238.
13 Hunter 1886, p. 487.
whole, rice was the more important staple; production and consumption levels were regularly three times that of wheat. Whether rice was the primary food grain in the country is unclear. Some sources claim that millet was more important (e.g. Hunter 1886), but by 1895 rice acreage was double that of millet acreage.

Statistics towards the end of our sample also suggest that railways transported a significant fraction of wheat and rice production. Almost three million tons of wheat were carried by 29 separate railways in 1910. There could be some double-counting in that figure since wheat could be transferred from one railway to another. But mitigating this problem is the fact that 19 of those railways carried tiny amounts of wheat (3% combined), one railway - the North-Western State - carried 44% of the total, and most of the major railways were connected to ports. If we ignore the possibility of double-counting, these numbers indicate that 29% of the wheat crop in 1910 was moved by rail. The estimate for rice is 14%. These percentages would seem high enough to expect railways to affect grain price dispersion within the country.

We don’t know how much of the wheat exported in 1910 was carried by rail. But if we make the extreme assumption that all of it was, that would imply that wheat transported by rail for domestic consumption was still roughly double the amount transported for export that year. For rice the breakdown is half for domestic consumption, half for export. These statistics are telling because they contradict the assumption that railways served solely to supply the British with cheap commodities. Even if that was their original purpose, railways served the domestic market in important ways too.

Average prices by state are displayed in Figure 3 and are generally consistent with the production levels discussed above. For wheat, prices in the high production states were the lowest in both decades. Similarly, mean rice prices were lowest in the Bengal and Bihar/Orissa states (though Madras is an exception). There is also some evidence of price convergence over time. This is clearest in the rice data, where the gap between Bombay and Bengal prices shrank from 167% to 22%. It is harder to see in the wheat data, largely because of the continued high relative prices in Coorg, Mysore, and Nizam’s Territories. Averages from those three states, however, should be interpreted with caution since they only contain one, two, and two districts respectively. Excluding them, the highest initial price differential was Bombay and United Provinces, and it fell from 98% to 31%.

Falling transportation costs provide a ready explanation for this convergence. Hurd (1975, p. 266) estimates that railway costs of transporting grain during these years were 80% lower than cart costs, though Derbyshire (1987, p. 525) suggests a lower figure of 55% on metal roads. Boat rates were more competitive. An 1873 government report estimated a cost advantage of 20% for railways (Mukherjee 1980, p. 198), but other sources claim steamboats and downstream riverboats were as much as 55% cheaper through the 1870’s. Of course, access to this form of transportation was limited geographically to areas near waterways and seasonally to periods when water levels were neither too high (monsoon) nor too low. There were also huge time savings with railways.\textsuperscript{15} Whereas pack bullocks traveled 6-8 miles a day, bullock carts covered 13-20 miles a day, and upstream riverboats could sometimes complete only 10 miles a day,

\textsuperscript{15} Coleman (2005, 21) provides evidence that in the late nineteenth century the big advantage of railroads in the US corn market was their speed, not cost; Great Lake steamships were in fact cheaper.
railroads could cover 400.\textsuperscript{16} Railroads also better protected cargo from damage and spoilage. Railroads, then, had many advantages over alternative means of transportation in British India. How much they contributed to price convergence, though, remains an open question.

THEORY

In a friction-less world of complete information, arbitrage should ensure that homogeneous commodities sell for one price. In the real world, however, the Law of One Price (LOP) may not hold for several reasons including incomplete information and search costs, transportation costs, trade barriers, menu costs, and local monopolies engaging in differential pricing in separate markets. There is evidence that transportation costs in India prior to railroads were high. Estimates supplied by the Famine Commission of 1880 suggest that freight charges for goods shipped by cart ranged from 15-50\% of the price of wheat per 100 miles carried, depending upon road conditions (Hurd 1975).

The effect of proportional transport costs on the LOP has been extensively studied (e.g. Taylor 2001). In its simplest “iceberg” form, assume a fraction $f$ of the good melts in transit. Let $P_A$ be the price of wheat in market $A$ and $P_B$ be the price in market $B$. Then it is profitable to ship wheat from market $A$ to market $B$ if $(1-f) P_B > P_A \rightarrow 1-f > P_A/P_B$. Likewise shipments in the reverse direction would only occur if $(1-f) P_A > P_B \rightarrow P_A/P_B > 1/(1-f)$. Thus there is a band within which relative wheat prices can fluctuate given by $(1-f) < P_A/P_B < 1/(1-f)$. Outside that band arbitrage should push relative prices

back into it. The simple logic of this model is that by reducing transportation costs, railroads should narrow this band and reduce price dispersion.

Because we have actual price data instead of price indexes we can directly measure the impact of factors on price convergence. The approach we use is similar to that employed by Engel, Rogers, and Wang (2003) in their study examining the determinants of commodity price differences across US and Canadian cities. Our dependent variable is the absolute value of the log price difference between districts: $|\log p_{i,t} - \log p_{j,t}|$ where $\log p_{i,t}$ is the log of the wheat price in district $i$ at year $t$. This variable approximately measures the percentage gap between prices in the two districts.

Since the introduction of railways varies both across districts pairs and over time, we can isolate the effect of the railways variable by using a dummy variable for each district pair. Thus our base econometric specification uses district-pair fixed effects to capture all time invariant factors specific to each district pair (distance between them, location on a river, etc.):

$$LPD_{j,k,t} = \alpha + \beta_1 BothRail_{j,k,t} + \sum_j \sum_k \beta_{jk} D_{j,k} + \mu_{j,k,t}$$

where $LPD_{j,k,t}$ is the (absolute value of the) log price differential at time $t$ between districts $j$ and $k$, $\alpha$ is the constant, $BothRail$ is a dummy equal to one if both districts have railways that year and $D_{j,k}$ is a dummy variable for each district pair $j,k$.17 We later add time variables to this specification.

---

17 We tried random effects estimation too, but a Hausman specification test indicated that random effects was inconsistent, so we used fixed effects.
DATA

Annual retail wheat and rice price data by district were obtained from the 1896 and 1922 issues of *Prices and Wages in India* and are expressed in rupees per ser (2.057 lbs). They were based on fortnightly purchases at district headquarters. The data run from 1861 to 1920, though some districts don’t have price data for all those years. There are slightly more districts with rice data than wheat data: 183 versus 159. All of the major cities in India are contained in our sample except for Madras, which only has rice data. A few districts had so few observations that they were omitted.

The quality of the price data appears to be high. In 1886 the Department of Finance and Commerce of the Government of India cross-checked the retail price data with the books of merchants and grain dealers and with other Administration reports and found that the data were trustworthy. Hurd (1975) expresses confidence in the data and argues that accurate retail price data should have been easy for bureaucrats to collect from local bazaars. McAlpin affirms that the data are internally consistent and concludes that they were “probably among the more reliable statistics available for historical work on India.” (1983, p. 900) Nonetheless it is important to note that they do not necessarily reflect prices in the countryside away from district headquarters. We also don’t know if prices were sampled every two weeks throughout the year, which could affect their quality.

Railway opening dates by city were found in the 1947 issue of the *History of Indian Railways*. The opening date for railways used in our analysis was the first year a railway entered a district. We obtained district boundary information from historical maps and the 1911 *Encyclopedia Britannica*. By 1861, there were 16 districts in our sample with
railways. By 1920, 156 of our 159 districts had railways. The growth in the number of districts with railways over time is displayed in Figure 4.

Data on how close districts were to the coast and the Ganges were obtained from historical and contemporary maps of India. Distances between districts were based on the longitude and latitude of an important city (almost always the district headquarters) in each district. Most came from www.mapsofindia.com/lat_long. The Haversine formula was employed to estimate the distance between them. City population data came from the Indian Census of 1872 found in the Statistical Abstract Relating to British India from 1867/8 to 1876/7. Finally, the list of districts that operated as quasi-independent Princely (Native) States during this period came from the same Statistical Abstract supplemented by the 1911 Encyclopedia Brittanica.18

Figure 4 also displays a measure of wheat and rice price dispersion over our sample period: the annual coefficient of variation. Despite some fluctuations, the trends are clearly downward. The measure falls almost 60% between the early 1860s and the early 1900s. Grain prices, therefore, moved significantly closer together during this period of railway expansion. This convergence is consistent with Hurd’s (1975) analysis and suggests that the introduction of railways may have been the principal reason for it.

One concern is that prices may have been administered by either local or national officials and therefore did not reflect market prices. We could find no evidence for that in the data or literature. Figure 5 displays annual prices for the largest city in India – Calcutta - and two other randomly chosen districts. They are fairly representative of our

18 Districts from the province of Berar were not counted as Princely States because starting in 1860 they were held in trust by the British Administration. Districts from the province of Mysore were counted as Princely States because, although in the 1870’s they were temporarily under British Administration, in 1881 they were restored to the Raja.
sample. Wheat prices appear flexible, fluctuating significantly from one year to the next. The average standard deviation of district wheat prices is 0.030, large relative to the average mean of 0.083. There were very few cases where nominal district prices were constant from one year to the next, which could be evidence of administered prices. Authors such as Hurd and McAlpin who worked with these data before believed that they reflected market forces and were therefore useful for measuring the impact of events such as the coming of railways.

With complete data our wheat sample would contain 753,660 ((159x158)/2 district pairs times 60 years) annual price differences, but missing data reduce it to 682,132 observations. Similarly, missing data for rice reduce that sample size from 999,180 to 918,180 observations. The sample summary statistics are presented in Table 1. Starting with the wheat data, the mean log price differential is 0.26 or roughly 26%. It ranges from 0 to more than 2.5, so some prices differed by as much as a factor of 13. The average distance between districts is close to 500 miles. Only small percentages of our district pairs are both from big cities, along the Ganges, on the GT road, or by the coast. About 18% had one district from a Princely state and one from British India, and 55% were connected by railways. The final two variables will be discussed later in the paper. The summary statistics for our rice sample are broadly similar.

RESULTS

Table 2 displays our first set of results. The coefficients on the district pair dummies are suppressed. The first column is with no time controls. For both grains we see a large

---

19 Nominal prices were constant for three years in a row in only 19 cases out of 8939 in our sample. Our results did not change when those districts were omitted from the sample.
and significant negative effect of railways on the difference in log prices. With mean log price differences in wheat falling from 0.50 in the early 1860s to 0.15 just before 1910, the introduction of railways is estimated to be responsible for about 45% (0.16/0.35) of that decline. For rice the average log price differential falls from 0.55 in 1864 to 0.16 in 1920, so railways can explain 51% (0.20/0.39) of that decrease. Both estimates clearly support the hypothesis that railways greatly facilitated price convergence in British India.

However, the estimated effect of railways could be inflated by a spurious correlation between the trend of falling price dispersion and the rising number of railways. To test that we added yearly dummy variables to our regression. The results are in column 2. There is a dramatic change in our railways coefficient: the estimate falls by about five to six times for both wheat and rice, though both remain highly significant. Column 3 runs the same regressions with linear and quadratic time trends instead of yearly dummies. The estimates are virtually identical. Both regressions, therefore, suggest that the introduction of railways still has an important impact on price dispersion, but the magnitude of the impact is now considerably small. Railways are estimated to reduce wheat price differences by 10% of the total fall in price dispersion over the sample period, and a similar percentage for rice.

Some intuition on why these estimated effects are small comes in Figure 6. In this diagram mean log price differences are calculated each year for districts connected by railways and districts not connected. It is analogous to Hurd’s (1975) diagram displaying a consistently higher coefficient of variation for districts without railroads. One must carefully interpret these graphs. There is no attempt to control for other variables that should be correlated with price dispersion such as distance, population, coastal access,
etc. It could be the case, for instance, that connected districts were closer to each other than unconnected districts, which may have been more likely to come from remote parts of the country, and that that is why price dispersion is lower among them. The composition of these two groups also changes over time, as district pairs got railways and jumped to the connected group.

These caveats notwithstanding, we obtain similar results to Hurd in that price dispersion for districts connected by rails is lower than for districts not connected. However, the gap is fairly small for most years, in many cases 0.05 or less. By the middle of our sample period the gap is close to zero. It widens in the last decade for wheat, but the number of districts without railways after 1910 is so few (7) that those statistics should not be given much weight. The other noteworthy feature of the graph is that price dispersion falls significantly among district pairs without railways. This suggests that something independent of railways was bringing prices closer together during these years.

A more rigorous way to measure the impacts of railways on price dispersion is to control for fixed effects and examine how dispersion changes before and after railways were introduced. Figure 7 displays those estimates. It comes from regressing price dispersion on dummies measuring the number of years (up to 15) before and after railways linked districts. For wheat, price dispersion falls only modestly the year railways connect two districts: a couple of percentage points. It also falls by about the same amount each of the two years before railways appeared. There is clearly a downward trend in price dispersion after railways are introduced, but dispersion was trending downwards before as well. The introduction of railways does not seem to affect
the trend much. The effect on rice price dispersion is even smaller: there is virtually no immediate effect. Again, it is hard to distinguish the post-railway trend from the pre-railway trend.

Based on the estimated effect of railways from Table 2, one can calculate the impact that an increasing number of districts receiving railways should have on price dispersion over our sample period. The effect should start small, as only a small number of district pairs got railways early, but should grow as more districts got linked. Figure 8 displays this estimated effect for both grains starting at peak price dispersion in the early 1860s. It is a visual representation of the results from Table 2 taking into account the fraction of the sample affected by railways. It is a rather striking picture of the modest estimated effect of railways on price convergence.

**Why is the Railway Effect so Small?**

These results raise an obvious question: why is the estimated railway effect so small? Given the advantages of railways, how could their impact on price differentials be so limited? There are several potential explanations for these results. A simple one is that it just took railways time to affect prices and markets. As David (1990) has shown, it can take a while for firms and workers to understand a new general purpose technology and to take advantage of it. The effect, therefore, may start small but eventually grow large. This could especially be true if improvements in railroading, such as more efficient locomotives, gradually increased its price advantage over alternative means of transportation. Rail freight charges, in fact, fell by 50% between 1881 and 1916-17.\(^\text{20}\)

\(^{20}\) Hurd 1983, p. 754.
A straightforward way to test this is to allow the railway effect to vary by year and look for a trend. Figure 9 has those results. Surprisingly there is no overall downward trend which one would expect if railways had a bigger negative effect in later years. From 1870 onward the effect appears fairly constant. One reason for this could be that alternative forms of conveyance were also becoming cheaper during these years. Steamboat charges, for example, fell two-thirds between the 1840/50s and 1870s. The average of these effects across the sample period is close to our 0.03 estimates from Table 2. Allowing the railway effect to change over time, therefore, does not seem to help solve our puzzle.

Another possibility is that our model omits relevant variables. The model focuses on the effects of establishing a direct railway link between districts. It ignores, however, the possibility that there may be effects on nearby districts that don’t have railways. If grain was traded between neighboring districts, prices between those districts could be linked, so when one district got a railway the impact on its price could be transmitted to the neighboring district. By not allowing for that in our model we could be missing an important transmission mechanism for railways.

To test this hypothesis we included two additional variables in our regressions: OneRailOneClose and BothClose. OneRailOneClose means that one district has a railway and the other does not, but its district headquarters is within 100 miles of a district headquarters with a railway. BothClose means that neither district has a railway but both are within 100 miles of districts that do. The missing dummy is where neither

21 Derbyshire 1987, p.525.
22 Dobado and Marrero (2005), using a different methodology, find an increasing effect of railways over time in Mexico.
23 Hurd (1975, p. 275) conjectures that this is why the coefficient of variation fell for non-railway districts in his sample too.
district has rail and neither is close to a railway. The expectation is that OneRailOneClose might have an effect on price convergence, but it should be weaker than BothRail. For similar reasons, the effect of BothClose should be weaker still.

The results are contained in Table 3. Because the results with time trends were so similar to those with yearly dummies, we only report the latter. All three railway variables enter significantly, and our expectations about the relative magnitude of their effects are confirmed. Having a railway directly link two districts does more to bring prices together than if one of them is merely close to a railway, which has a larger effect than if both of them are just close to rail lines. The size of the BothRail effect (relative to a pair where neither district has a railway nor neither was close to a railway also) is now roughly twice as large as it was in our earlier fixed effects regressions. Apparently, one of the reasons our earlier estimates were smaller is that before some districts got connected by railways, they were already close to railways. In fact, 42% of our wheat pairs were in the OneRailOneClose category before both districts got railways, and 70% of our rice pairs did. So even before a number of district pairs were directly linked by rail, railways were already reducing the price gap between them. Being directly linked by rail, then, would only have the marginal effect of switching from OneRailOneClose to BothRail, which is smaller. This doubling of our parameter estimate is significant. It suggests that now railways are capable of explaining about 20% of the price convergence we observe over our sample period. That is no longer such a small percentage, but it still leaves 80% unexplained.
Existing Transportation Networks

We can extend our discussion of the small magnitude of the railway effect by placing the introduction of the railway system in the context of the existing transportation networks in the Indian economy at that time. The key point is that the railway system was, at least in some significant areas, overlaid on an already functioning transportation network and a somewhat integrated economy: “It is true that in India railways were not opening up virgin territory and did not have to create their own traffic.” (MacPherson 1955 181). Support for this hypothesis comes from time series tests, described in an appendix, that show that wheat and rice markets were already integrated by the 1860s, before most railroads were built.24

The notion that railways were sometimes imposed on existing trade networks was certainly the case in the north between Calcutta and Delhi. Yang (1998) describes the interconnections along the Calcutta to Delhi corridor, particularly in Bihar. He discusses in some detail the principal pre-existing channels of transportation in this region via rivers and roads. The Ganges flowing through North India was widely used for transportation; large cities that had significant commercial linkages had grown around the banks of the river. Interestingly the development of railways also followed the existing commercial linkages. In Bihar, the shortest and most direct route between Calcutta and Benares was through Gaya, but the development of the line was along the Ganges on a northerly arc through Patna.

24 This contrasts with the claim that railways created an integrated market for foodgrains in India (e.g. Hurd 1975 p. 275)
This region also possessed notable road transportation with the principal linkage being the GT road connecting Delhi to Calcutta. There is some discussion about whether the road network was used principally for foot traffic or freight traffic. But for our purposes, a qualitatively similar result should hold for the road network as with the Ganges. Combining the effects of roads and the Ganges would then imply that the introduction of railways should have a relatively small effect along the Delhi-Calcutta route. In the same vein, one should expect a region such as southern India, with its relatively dense network of roads and near access to the ocean, to experience a smaller fall in price dispersion once railways appeared.

An interesting counterexample was the decision not to connect Bombay directly to either Calcutta or Delhi by rail. Instead, the British built a line from Bombay to Allahabad, midway between Calcutta and Delhi, through Jubbulpore (see Figure 1). Since this was a less traversed commercial route, one should expect to see a greater than average decrease in price dispersion among districts along this route.

Table 4 strongly confirms this hypothesis. The effect of introducing railways along the Delhi-Calcutta route (or within 50 miles of a straight line between both cities) or within the South Indian region is less than the average effect for all of India. In contrast, the effect of railways on districts along or near the Bombay-Allahabad line is five to six times greater than that of the Delhi-Calcutta line. To the extent, then, that railways were built on top of existing trade routes in an integrated economy, their effect on price differentials is definitely muted.

---

25 The GT road extended beyond Delhi to Peshawar.
Endogeneity

The endogeneity of the timing of railways is clearly an important issue in our analysis. As already discussed, when and where railroads were built was not random. Remember our regressions control for district pair fixed effects. Thus any missing time invariant variables that affected the price differential between each district pair would be controlled for in that specification. This should eliminate, for instance, any bias arising from district pairs with close prices systematically receiving railways before pairs with distant prices.

With respect to the remaining endogeneity, it should be noted that the process used by the British in choosing where to build railways could be cryptic. For instance, some of the most abundant cotton lands were entirely bypassed by railroads until close to the end of the century (Hurd 1983 p. 742). Nonetheless, if there is omitted variable bias, it seems most plausible to us that it would inflate our railway estimates — that districts that got railways earlier had more of a propensity for price convergence. This could arise if districts were chosen for their commercial promise and that promise also made them quicker to respond to other pro-market integration changes occurring in the economy, such as the elimination of internal tolls. So correction for this bias would probably only strengthen our claim that railways had a relatively small effect on price convergence.

One possible way around this bias is to examine the impact of railways built for non-commercial purposes. Perhaps the clearest example of that are the military lines. Unfortunately they turned out to be a tiny percentage of our sample: only Quetta, the Upper Sind Frontier, and Dera Ismael Khan were located on official strategic lines. This subsample was too small to do any meaningful analysis. So we added six districts from
the north-west that had cantonments. The results from this group are displayed in Table 5. The railway point estimates are similar to those found earlier, but unfortunately, only one of them is significant at the 10% level. So it is hard to get precise estimates of the effect of railroads in non-commercial cases.

Another strategy involves identifying a good instrument for the timing of railways. It argues that many railroads were initially built along straight lines between major cities. Thus towns on a straight line between major cities were “exogenously” connected to the railway system in the sense that they got railways independently of their own characteristics. Because the chances of finding towns exactly on a straight line are low, one can use the shortest distance from the town to the straight line as an instrument for when that district got a railway.

There exist two problems with this instrumentation strategy in our context. First, we already mentioned two prominent cases in the development of the Indian railway system where the straight-line assumption of this instrument was violated. To connect Calcutta to Delhi the planners placed the trunk line along a northern route next to the Ganges rather than along a “straight line” between both big cities. Also, no trunk line was initially built between either Bombay and Delhi or Bombay and Calcutta, but instead went from Bombay to Allahabad, a much smaller city. For these reasons, the distance to

---

26 They were Jhelum, Lahore, Peshawar, Rawalpindi, Sialkot, and Multan.  
27 To confound matters, several of these cantonments were on the GT road.  
29 Michaels (forthcoming) is a good example of using such an instrumentation strategy in the context of highway development.  
30 “Although the shortest route would have been to follow a direct line between Calcutta and Banaras, a course roughly paralleling the path of the Grand Trunk Road, the initial alignment followed the course of the Ganges, that is, roughly the line of the old Mughal trunk route. From Banaras the East India Railway proceeded along the north of Shahabad and Patna, then continued on to Monghyr and Bhagalpur en route to Calcutta. And like the Banaras-Patna-Calcutta road along the Ganges its major stops were, from west to east, Buxar, Arrah, Danapur, Patna-Bankipur, Fatwa, Bakhtiyarpur, Barh, and Mokameh, before entering Monghyr.” Yang (1999), pp. 46
a straight line between major cities proved to be a very weak instrument in our sample for when districts got railways. Second, this instrument probably violates the “exclusion restriction” whereby the instrument should only affect the dependent variable through the endogenous variable in question. The problem is that early roads, such as the GT road, also tended to lie along straight lines between major cities. Without controlling for those roads, this instrument could artificially capture their effects.

OTHER FACTORS AFFECTING RELATIVE PRICES

One unsatisfying aspect of fixed effect estimation is that the fixed effect is a bit of a black box. Knowing what is in it might be valuable. For instance, we might want to know what the effect of distance is on price differences between districts. Estimating such an effect could be a way to check our results and also to put the effects of railways in context. To explore these possibilities, we drop the fixed effect and add district pair information that might be relevant to price dispersion.

Our new specification is then:

\[
LPD_{j,k,t} = \alpha + \beta_1 \text{BothRail}_{j,k,t} + \beta_2 \text{LnDistance}_{j,k} + \beta_3 \text{BothLargeCity}_{j,k} + \\
\beta_4 \text{BothonGanges}_{j,k} + \beta_5 \text{BothonGTRoad} + \beta_6 \text{BothonCoast}_{j,k} + \\
\beta_7 \text{SameState}_{j,k} + \beta_8 \text{PrincelyState}_{j,k} + \mu_{j,k,t}
\]

\(LnDistance\) is the log of the distance between the district headquarters, \(\text{BothLargeCity}\) is a dummy equal to one if both districts have a city with more than 50,000 people in them according to the 1872 census, \(\text{BothonGanges}\) is a dummy equal to one when both districts
are within 10 miles of the Ganges River, \textit{BothonCoast} is equal to one if both districts are within 10 miles of the coast, and \textit{SameState} equals one when both districts lie in the same state. \textit{PrincelyState} is a dummy equal to 1 if one of the districts is in a Princely State and the other is in British India.

As shown in the literature, price dispersion should increase nonlinearly with distance. Larger cities may have more sophisticated markets and advanced communication and transportation linkages that would reduce price dispersion. Being near the Ganges, on the GT road, or along the coast implies a richer set of communication and transportation options which should lower dispersion. Since we are controlling for distance, the variable \textit{SameState} helps to capture other effects, including institutional and cultural factors, specific to that state.

During the 1840s and 50s, a number of important Indian states were absorbed into the British Empire including the Punjab (1849), Sambalpur (1849), Bhagat (1850), Udaipur (1852), Jhansi (1853), Tanjore (1853), Nagpur (1854), and Oudh (1856). These acquisitions brought most of India under the same administrative rule with a similar set of institutions: legal, monetary, educational, even linguistic over time. That should have facilitated commerce between districts within British India. The exceptions were the independent Princely States, which is why we included dummy variables for them in our regressions.

Table 6 presents the results for both samples after controlling for time effects with yearly dummies.\(^{31}\) Examining wheat first, it is still the case that BothRail has the largest estimated effect, followed by OneRailOneClose and BothClose. The BothRail coefficient is slightly larger than with fixed effects, though it is useful to remember that

\(^{31}\) Again, the results using time trends were very similar.
our fixed effects estimates should be more accurate. We are probably not controlling for all the time invariant factors that influence price dispersion, and any omitted variables could easily be correlated with our right-hand side variables. Both districts being on the Ganges, both on the GT road, or both on the coast reduces dispersion by approximately 5 percentage points or half the amount of a direct railway link. Distance boosts price dispersion: a distance of 500 miles (close to the average in the sample) boosts price dispersion by about 50 percentage points. Both districts having large cities is statistically significant, but has a tiny impact on price dispersion. It is also positive, which is counterintuitive. Being from the same state has virtually no economic effect. Finally, price dispersion between districts in British India and Princely States is positive and statistically significant, on the order of a quarter of the size of the railway effect. There seems to be, therefore, a border effect similar to what Engel and Rogers (1996) found between Canada and the US.32

The results from the rice sample are a bit different. The estimated coefficients for railways and distance shrink, but are still highly significant. But both districts having large cities now significantly reduces price dispersion, and the effect is about a third as large as being linked by railways. The effects of being on the Ganges or coast fall dramatically and the GT road no longer has a significant effect on price differences. Being in the same state now becomes economically important, with an impact roughly 40% as large as railways. And the price gap between princely states and British India quadruples in size and is now even slightly larger than our railway estimate.

32 However, Gorodnichenko and Tesar (2005) warn that border effects as measured by Engel and Rogers can be spuriously produced by country heterogeneity. Furthermore we don’t deal with the difficult question of why these borders exist where they are.
Overall, these results are encouraging. Most of our variables enter significantly and have the expected sign. They indicate that railways mattered more than the possibility of transportation via river, road, or sea, which makes sense given the advantages of railways over those other modes of transportation. The results for SameState and OnePrincelyState imply that institutional and cultural factors could also play an important role in grain price dispersion, potentially as much as railways. Again, though, it is useful to remember that these results are more suggestive than definitive, given that we don’t fully control for fixed-effects here.

CONCLUSION

The period 1861 to 1920 witnessed sharp price convergence in British Indian wheat and rice markets. Previous authors have assumed that since a vast railway system was built during that time, railways were probably the main cause of that convergence. But tests examining relative price differences between individual districts provide surprisingly weak support for that hypothesis. Railways clearly mattered, but even controlling for the effects on nearby districts, railways seem capable of explaining only about 20% of the decline in grain price dispersion. The modest estimated effect fundamentally derives from the fact that prices were converging during this time almost as rapidly between districts without railways as with railways. Putting our fixed effect estimation results a different way: the pattern of price convergence between a given district pair was no different before they got railways than after they got the rail.

One reason for this may be that India was already a well established, partially integrated economy at the time of the railroad expansion. Some lines were built along
existing trade routes where there was already significant commerce between districts.

The impact of railroads on districts not obviously linked by trade was shown to be about five times larger than the impact on well linked districts. There was also considerable heterogeneity of the railway effect depending on whether districts were already close to railways or not.

There were also a number of significant changes in technology and infrastructure occurring in India before and during this time which could have contributed to price homogeneity:

It should first be noted, however, that before railways were introduced into the United Provinces in 1860, there were already signs of expanding commerce, population growth and cultivated acreage extensions, particularly from the mid-1840s. Market connections with the recently annexed Central Provinces and the Punjab were improving, bullock carts were replacing pack animals, canals were being introduced into the western Doab and steamboats on to the Ganges below Allahabad.” (Derbyshire, p. 524).

An extensive telegraph system was built at the same time railways were introduced and a uniform postal service was put in place.33 Roads were constructed and paved with metal. Clearly there were many changes in transportation, communication, and agriculture in India during this time which could have impacted price dispersion.

Another possibility is that businesspeople became more commercially oriented during this period. Entrepreneurship is hard to measure, but it plausible that the dramatic rise in exports fostered greater awareness of profit opportunities in trade. Better information sharing throughout the country could also have alerted producers to large price differences of which they could take advantage. There is evidence that wheat and rice

33 Dalhousie referred to railways, the telegraph, and uniform postage as the “three great engines of social improvement” instituted under his leadership. (Ghosh 97). Lew and Cater (2004) provide evidence that the diffusion of the telegraph was an important factor in the growth of world trade in the late nineteenth century.
started moving throughout the country in response to price differentials, and that during good harvests farmers began to export surpluses they previously would have stored.34

Finally, one should perhaps not underestimate the impact of institutional reforms occurring at the time.35 As listed earlier, a number of important states were added to the country at the start of our sample period, providing them with a single legal and administrative framework. The Education Despatch of 1854 created a plan for uniform education in the country. After 1857 peace existed inside the country and law and order was strengthened.36 Internal tolls were abolished.37 A single currency was used in the country and a common language was beginning to be taught.38 In short, numerous impediments to commerce were being swept aside at this time. Reforms such as these in other countries have been linked to the growth of markets: “Market integration also increased substantially within Germany after political unification in 1871.” (O’Rourke and Williamson, 43) In our sample, the large gap in rice prices between British India and independent states suggests that these factors may matter. This ties in with recent research on the primacy of institutions in economic development (e.g. Acemoglu and Johnson 2005)

What this contributes to the debate about the impact of British colonialism on India is less clear. Because market integration is assumed to benefit economies, if the railways the British built were less responsible for price convergence than previously assumed, the assumed benefits of colonialism would shrink. On the hand, if the observed price

34 See McAlpin (1979, pp. 148 and 150).
35 Jacks (2006) found that commodity market integration in 19th and 20th century Europe and North America was due more to regime and policy changes than technological progress in transportation.
36 Morris D. Morris (1963 p. 609) argued that previous instability in India inhibited trade and investment.
37 Derbyshire 523.
38 A single currency, the Madras rupee, was established in 1835, but Indian states were only able to take advantage of that when they became part of British India (Adams and West, 1979, p. 55)
convergence was due to other British initiatives such as canal construction or the elimination of tolls, it wouldn’t matter as much that railways were less important. All we really know is that during this period grain prices converged dramatically. What was primarily responsible for that, however, is still uncertain.

One could extend this analysis in a number of important ways. The most obvious is to obtain precise measures of institutional change, which would allow one to test for their effects on price convergence. It would also be interesting to delve deeper into why the impact of railways on prices appears to be largest in the 1860’s and to taper off after that instead of the reverse. Keller and Shiue (2004) look at the predictive power of market integration in the 17th century in China on current economic outcomes. Banerjee and Iyer (2003) also look at the impact of historical institutions and events such as land revenue systems on current outcomes in India. So another extension of this research would be to estimate the persistent effects of price differentials in 19th century India on economic outcomes there today. Ultimately this work may help shed light on the relative contributions of institutions and technology to economic development.
Figure 1: Railways and Major Indian Cities: 1872
Figure 2: Major Wheat and Rice Growing States in British India
Figure 3A: Mean Wheat Prices by State

Figure 3B: Mean Rice Prices by State
Figure 4A: Coefficient of Variation of Wheat Prices and the Number of Districts with Railways (out of 159 districts)

Figure 4B: Coefficient of Variation of Rice Prices and the Number of Districts with Railways (out of 183 districts)
Figure 5A: Wheat and Rice Prices in Calcutta

Figure 5B: Wheat and Rice Prices in Sambalpur

Figure 5C: Wheat and Rice Prices in Bulandshahr
Figure 6A: Annual Mean Log Wheat Price Dispersion for District Pairs Connected and Not Connected by Railways

Figure 6B: Annual Mean Log Rice Price Dispersion for District Pairs Connected and Not Connected by Railways
Figure 7A: Wheat Price Dispersion Before and after Railroads

Figure 7B: Rice Price Dispersion Before and after Railroads
Figure 8A: Estimated Decline in Wheat Price Dispersion due to Railways

Based on fixed effects regression with yearly dummies

Figure 8B: Estimated Decline in Rice Price Dispersion Due to Railways

Based on fixed effects regression with yearly dummies
### Table 1

**Summary Statistics: Wheat Sample**

District Pairs from 159 Districts: 1861-1920*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>DifferenceLogPrices</td>
<td>682,132</td>
<td>0.257</td>
<td>0.238</td>
<td>0</td>
<td>2.581</td>
</tr>
<tr>
<td>LnDistance</td>
<td>682,132</td>
<td>6.164</td>
<td>0.383</td>
<td>1.949</td>
<td>7.464</td>
</tr>
<tr>
<td>BothBigCities</td>
<td>682,132</td>
<td>0.061</td>
<td>0.239</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothonGanges</td>
<td>682,132</td>
<td>0.023</td>
<td>0.150</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothonGTRoad</td>
<td>682,132</td>
<td>0.025</td>
<td>0.156</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothonCoast</td>
<td>682,132</td>
<td>0.007</td>
<td>0.081</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SameState</td>
<td>682,132</td>
<td>0.128</td>
<td>0.334</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OnePrincelyState</td>
<td>682,132</td>
<td>0.179</td>
<td>0.383</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothRail</td>
<td>682,132</td>
<td>0.549</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OneRailOneClose</td>
<td>682,132</td>
<td>0.211</td>
<td>0.408</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothClose</td>
<td>682,132</td>
<td>0.061</td>
<td>0.239</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

* Some pairs missing data for some years

---

### Summary Statistics: Rice Sample

District Pairs from 183 Districts: 1861-1920*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>DifferenceLogPrices</td>
<td>918,180</td>
<td>0.290</td>
<td>0.264</td>
<td>0</td>
<td>2.939</td>
</tr>
<tr>
<td>LogDistance</td>
<td>918,180</td>
<td>6.286</td>
<td>0.692</td>
<td>1.948</td>
<td>7.491</td>
</tr>
<tr>
<td>BothBigCities</td>
<td>918,180</td>
<td>0.061</td>
<td>0.239</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothonGanges</td>
<td>918,180</td>
<td>0.018</td>
<td>0.132</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothonGTRoad</td>
<td>918,180</td>
<td>0.019</td>
<td>0.137</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothonCoast</td>
<td>918,180</td>
<td>0.007</td>
<td>0.084</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SameState</td>
<td>918,180</td>
<td>0.111</td>
<td>0.314</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OnePrincelyState</td>
<td>918,180</td>
<td>0.154</td>
<td>0.361</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothRail</td>
<td>918,180</td>
<td>0.558</td>
<td>0.496</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OneRailOneClose</td>
<td>918,180</td>
<td>0.192</td>
<td>0.394</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BothClose</td>
<td>918,180</td>
<td>0.035</td>
<td>0.184</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

* Some pairs missing data for some years
### Table 2
Log Price Dispersion Regressions with District-Pair Fixed Effects

<table>
<thead>
<tr>
<th>Sample</th>
<th>Variable</th>
<th>No time effects</th>
<th>Year dummies</th>
<th>Time trends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>BothRail</td>
<td>-0.156</td>
<td>-0.034</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(272.1)***</td>
<td>(47.2)***</td>
<td>(45.3)***</td>
</tr>
<tr>
<td></td>
<td>Observations</td>
<td>682,132</td>
<td>682,132</td>
<td>682,132</td>
</tr>
<tr>
<td></td>
<td>Adj. R-squared</td>
<td>0.42</td>
<td>0.50</td>
<td>0.47</td>
</tr>
<tr>
<td>Rice</td>
<td>BothRail</td>
<td>-0.200</td>
<td>-0.032</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(366.9)***</td>
<td>(47.9)***</td>
<td>(51.1)***</td>
</tr>
<tr>
<td></td>
<td>Observations</td>
<td>918,180</td>
<td>918,180</td>
<td>918,180</td>
</tr>
<tr>
<td></td>
<td>Adj. R-squared</td>
<td>0.41</td>
<td>0.50</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Robust t statistics in parentheses; *** significant at 1%; time trends include a linear and a quadratic trend.

### Table 3
Log Price Dispersion with District-Pair Fixed Effects Including Close Railway Effects

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wheat Prices</th>
<th>Rice Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>BothRail</td>
<td>-0.077</td>
<td>-0.068</td>
</tr>
<tr>
<td></td>
<td>(67.5)***</td>
<td>(69.4)***</td>
</tr>
<tr>
<td>OneRailOneClose</td>
<td>-0.057</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td>(51.2)***</td>
<td>(52.2)***</td>
</tr>
<tr>
<td>BothClose</td>
<td>-0.036</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>(24.4)***</td>
<td>(23.4)***</td>
</tr>
<tr>
<td>Observations</td>
<td>682,132</td>
<td>918,180</td>
</tr>
<tr>
<td>Adj. R-Squared</td>
<td>0.50</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Robust t statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; regressions include year dummies; close: within 100 miles of a railway.
Table 4
Log Price Dispersion Regressions with District-Pair Fixed Effects
Wheat Subsamples: Delhi-Calcutta (DC) & Bombay-Allahabad (BA) Lines
and Southern India

<table>
<thead>
<tr>
<th></th>
<th>On actual DC railway line</th>
<th>Within 50 miles of DC straight line</th>
<th>On actual BA railway line</th>
<th>Within 50 miles of BA straight line</th>
<th>Within Southern India</th>
</tr>
</thead>
<tbody>
<tr>
<td>BothRail</td>
<td>-0.028</td>
<td>-0.013</td>
<td>-0.136</td>
<td>-0.084</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(1.3)</td>
<td>(5.7)***</td>
<td>(4.7)***</td>
<td>(10.7)***</td>
<td>(4.1)***</td>
</tr>
<tr>
<td>Observations</td>
<td>4,249</td>
<td>27,821</td>
<td>1,680</td>
<td>3,949</td>
<td>12,797</td>
</tr>
<tr>
<td>Adj. R-squared</td>
<td>0.50</td>
<td>0.45</td>
<td>0.60</td>
<td>0.58</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Robust t statistics in parentheses; *** significant at 1%; regressions include year dummies

Table 5
Log Price Dispersion Regressions with District-Pair Fixed Effects
Military Subsample

<table>
<thead>
<tr>
<th></th>
<th>Wheat Prices</th>
<th>Rice Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>BothRail</td>
<td>-0.022</td>
<td>-0.047</td>
</tr>
<tr>
<td></td>
<td>(1.1)</td>
<td>(1.8)*</td>
</tr>
<tr>
<td>Observations</td>
<td>1,928</td>
<td>1,938</td>
</tr>
<tr>
<td>Adj. R-squared</td>
<td>0.43</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Robust t statistics in parentheses; * significant at 10%; regressions include year dummies
Table 6
Log Price Dispersion Regressions without District-Pair Fixed Effects

<table>
<thead>
<tr>
<th>Sample:</th>
<th>Wheat</th>
<th>Rice</th>
</tr>
</thead>
<tbody>
<tr>
<td>BothRail</td>
<td>-0.104</td>
<td>-0.076</td>
</tr>
<tr>
<td></td>
<td>(92.6)***</td>
<td>(74.6)***</td>
</tr>
<tr>
<td>OneRailOneClose</td>
<td>-0.074</td>
<td>-0.056</td>
</tr>
<tr>
<td></td>
<td>(64.2)***</td>
<td>(51.9)***</td>
</tr>
<tr>
<td>BothClose</td>
<td>-0.067</td>
<td>-0.047</td>
</tr>
<tr>
<td></td>
<td>(44.03)***</td>
<td>(28.1)***</td>
</tr>
<tr>
<td>LnDistance</td>
<td>0.088</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>(201.4)***</td>
<td>(151.8)***</td>
</tr>
<tr>
<td>BothLargeCities</td>
<td>0.005</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(5.0)***</td>
<td>(25.7)***</td>
</tr>
<tr>
<td>BothonGanges</td>
<td>-0.052</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(44.7)***</td>
<td>(9.1)***</td>
</tr>
<tr>
<td>BothonGTRoad</td>
<td>-0.045</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(37.9)***</td>
<td>(0.3)</td>
</tr>
<tr>
<td>BothonCoast</td>
<td>-0.061</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(24.4)***</td>
<td>(9.4)***</td>
</tr>
<tr>
<td>SameState</td>
<td>-0.002</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>(2.6)***</td>
<td>(38.7)***</td>
</tr>
<tr>
<td>OnePrincelyState</td>
<td>0.021</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>(30.4)***</td>
<td>(118.9)***</td>
</tr>
</tbody>
</table>

Observations 682,132 918,180
Adj. R-squared 0.26 0.24

Robust t statistics in parentheses; *** significant at 1%; close: within 100 miles of a railway; regressions include year dummies
APPENDIX: TIME SERIES MARKET INTEGRATION TESTS

There is a common test in the time series literature for market integration (e.g. Parsley and Wei 1996). It involves testing if relative prices across districts converge after moving apart. Because one is dealing with changes over time, one has to allow for random walks (unit roots) in the data. Two estimation methods that do so are the Levin-Lin and Im-Pesaran-Shin panel unit root tests. Panel data estimation significantly increases the power of the test. The simple specification is:

\[
\Delta q_{i,t} = \alpha_i + \beta q_{i,t-1} + \epsilon_{i,t}
\]

where \( q_{i,t} \) is the ratio of the log grain price in district \( i \) to the mean log grain price at time \( t \), and \( \Delta q_{i,t} = q_{i,t} - q_{i,t-1} \). The null is \( \beta = 0 \), or relative prices follow a random walk and are as likely to diverge as converge. The alternative, \( \beta < 0 \), implies price convergence (a high relative price is followed by a fall in that price) and market integration. The Levin-Lin test assumes a common \( \beta \) across districts, while Im-Pesaran-Shin allows the \( \beta \)'s to vary (so no single \( \hat{\beta} \) is available). Here are our results based on complete samples from 1861-1920:

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\beta} )</th>
<th>Probability ( \beta = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Levin-Lin Method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat Prices: 1861-1920</td>
<td>-0.50</td>
<td>0.000</td>
</tr>
<tr>
<td>Wheat Prices: 1861-1870</td>
<td>-0.67</td>
<td>0.000</td>
</tr>
<tr>
<td>Rice Prices: 1861-1920</td>
<td>-0.37</td>
<td>0.000</td>
</tr>
<tr>
<td>Rice Prices: 1861-1870</td>
<td>-0.55</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Im-Pesaran-Shin Method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat Prices: 1861-1920</td>
<td>various</td>
<td>0.000</td>
</tr>
<tr>
<td>Wheat Prices: 1861-1870</td>
<td>various</td>
<td>0.000</td>
</tr>
<tr>
<td>Rice Prices: 1861-1920</td>
<td>various</td>
<td>0.000</td>
</tr>
<tr>
<td>Rice Prices: 1861-1870</td>
<td>various</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Clearly one can reject the hypothesis that $\beta=0$. So there is strong evidence that grain markets were integrated during this time. More importantly, one can easily reject the hypothesis that $\beta=0$ for the first decade of our sample: 1861-1870. That means that even before railroads became widespread in India, district grain markets were linked to each other.
REFERENCES


*Administration Report on the Railways in India for Calendar Year 1910.* London: His Majesty’s Stationery Office, 1911.


*Statistical Abstract Relating to British India from 1867/8 to 1876/7*, London: Her Majesty's Stationary Office, 1878.

