NBER WORKING PAPER SERIES

THE ORIGINS OF SPATIAL INTERACTION

Wolfgang Keller
Carol H. Shiue

Working Paper 10069
http://www.nber.org/papers/w10069

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2003

We thank Luc Anselin, Richard Baldwin, Tim Conley, Dan Hamermesh, Gordon Hanson, and Preston McAfee for suggestions. All errors are our own. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

The Origins of Spatial Interaction
Wolfgang Keller and Carol H. Shiue
NBER Working Paper No. 10069
October 2003
JEL No. N5, N7, F1, O1, R1

ABSTRACT

Geography shapes economic outcomes in a major way. This paper uses spatial empirical methods to detect and analyze trade patterns in a historical data set on Chinese rice prices. Our results suggest that spatial features were important for the expansion of interregional trade. Geography dictates, first, over what distances trade was possible in different regions, because the costs of ship transport were considerably below those for land transport. Spatial features also influence the direction in which a trading network is expanding. Moreover, our analysis captures the impact of new trade routes both within and outside the trading areas. We also discuss the long-run implications this might have.

Wolfgang Keller
International Monetary Fund
Research Department
700 19th Street, NW
Washington, DC 20431
and NBER
keller@eco.utexas.edu

Carol H. Shiue
The World Bank
Development Research Group
1818 H Street, NW
Washington, DC 20433
shiue@eco.utexas.edu

1 Introduction

Geography exerts a major influence in many economic areas. Trade declines with geographic distance, and per-capita incomes vary with climatic conditions, for instance (Anderson and van Wincoop 2003, and Sachs 2003, respectively). But does geography—mineral deposits, soil quality, or rivers, say—immediately determine how rich people are, or is the impact of geography on the choice sets of economic agents more indirect?[1] Does geography influence also long-run outcomes as some have claimed (Diamond 1997)? And are there strong reasons to believe that the influence of geography will be lower in the future than it has been in the past?

In an attempt to shed new light on these questions, we study how geography guides the evolution of interregional trade patterns in 18th century China. Recent research has highlighted the general importance of interregional trade during this period (Chen, Wang, and Hu 1999, Shiue 2002). Geographic factors, in particular local climate and access to relatively low-cost ship transport, were important determinants of interregional trade (Perkins 1969, Chuan and Kraus 1975, Wang 1989, and Shiue 2002). Analyzing the evolution of this trade empirically, we contend, can provide valuable insights on comparative economic development in China and elsewhere.

This paper studies interregional trade by examining the spatial pattern of rice price differences in 121 Chinese prefectural markets between the years 1742 to 1795. It puts the paper into the well-established literature that uses price data to look at trade (Engel and Rogers 1996, O' Rourke and Williamson 2000, and Slaughter 2001). At the same time, we emphasize more the geographic features underlying these price differences by using information on the actual geography of China (climate, topography) and Geographic Information System-based spatial empirical methods (Cliff and Ord 1981, Anselin 1988).[2]

[1]Cronon (1991) has called this first-nature and second-nature geography, respectively.

[2]See e.g., Case (1991), Kelejian and Robinson (1992), and Anselin, Varga, and Acs (1997) for other recent work

Our results suggest that spatial features have shaped the expansion of interregional trade. First, geography dictates the distance over which trade was possible in different regions. Second, because in addition to distance our methods track the spatial position of one region relative to all others, we can also capture the impact that the emergence of new trade routes has in different regions, and the direction a trading network is most likely to expand.[3] Overall, we think that using spatial methods gives a much clearer picture of the evolution of interregional trade.[4]

Geography not only affects the costs of interregional trade, but also the autarky prices in different regions. Specifically, the (relative) price of rice under autarky is related to the relative abundance of arable land (Heckscher 1919 and Ohlin 1924), and trade will tend to equalize prices and factor returns across regions (e.g., O'Rourke, Taylor, and Williamson 1996, Richardson 1995). We suspect that this effect was present in 18th century China too, although there is little systematic data that we can use to confirm this point. The main focus of this paper is the extent to which geography affected trade costs, and thus trading possibilities, and whether these differences could trigger different development paths across regions.

The question is fundamental to related work on institutions and technological change. For example, trade might facilitate technological innovations through learning, which could give some regions earlier access to new transport technologies than other regions.[5] In addition, trade may be associated with institutional innovations, for instance information sharing within a group of traders as a monitoring device in the absence of enforceable contracts (Greif 1989).[6] Once formed, such a group

---

that has used spatial empirical methods.

[3] Quah (2002) and Quah and Simpson (2003) emphasize the importance of capturing relative spatial position as well. These authors, as well as Hanson (2001) also suggest that the recent literature on agglomeration driven by scale economies might benefit from incorporating more spatial elements.

[4] O'Rourke and Williamson also note that spatial analysis helps to understand changes in market integration in New World and European economies of the 19th and early 20th century (2000, 47).

[5] Along these lines, Mokyr emphasizes the importance of trade *primarily* insofar as it contributes to the diffusion of technological knowledge (1990, 134); Keller (2003) discusses some of the more recent evidence.

[6] North and Thomas also note that trade goes hand-in-hand with institutional innovation (1973, 12), while Williamson emphasizes that ex-ante mutually beneficial trade often shapes ex-post institutional outcomes (2000, 599-600).

of traders can serve as a network that lowers the costs of looking for trading opportunities outside the current trade area.[7] Last but not least, the strategic expansion of trade can have lasting economic and political consequences. A case in point might be the abandonment of long-distance trade by China in the early 15th century, versus the state-supported trade of the British and Dutch East India companies in Asia a couple of centuries later.[8] Additional examples of how initial geographic advantages can be self-enforcing may be found.[9]

The remainder of the paper is as follows. Section 2 describes the characteristics and sources of the data and gives summary statistics. Section 3 examines the spatial autocorrelation of prices both globally as well as locally. The spatial econometric results are discussed in section 4, while section 5 provides some concluding discussion. Additional background on the data is given in the appendix.

2 Data

2.1 Characteristics and sources

The geographic area studied in this paper consists of 10 out of a total of 18 core provinces of China; these provinces are Anhwei, Fuijan, Guangdong, Guangxi, Guizhou, Hubei, Hunan, Jiangsu, Jiangxi, and Zhejiang. The area is situated in the center and south-east of the country, and includes some of the most agriculturally fertile areas, some of the most developed areas, as well as some poorer areas. The ten provinces were selected on the basis that they all produced rice as a major grain crop in the period under analysis. Figure 1 shows the sample area within the borders of contemporary China. The provinces have retained a basic correspondence to their historical geographic positions.

The 10 provinces in the sample are made up of 121 largely contiguous prefectures; a limited

[7] Rauch (2001) discusses the importance of social networks in international trade.

[8] See e.g. Findlay and O'Rourke (2002).

[9] At the same time, Krugman (1991) has shown that a region's geographic destiny might be altered. Davis and Weinstein (2002) have provided evidence on this question.

number of independent administrative units are not included in the sample. There are ten to fourteen prefectures in a province, and the sample includes the full set of the prefectures in each province that we can identify on historical maps. Figure 2 gives a map of these prefectures and the boundaries of the ten provinces that are the focus of the spatial analysis.

This paper uses weather and price data from 1742-1795 from the 121 prefectures. Systematic rainfall recording began as early as the Tang Dynasty (618-907 A.D.), and from at least the 17th century, during the reign of the Qing Dynasty (1644 to 1911), the collection of rainfall and weather reports at the county level had become standard government practice (Wilkinson 1969). In addition, the reporting of prices of the major grains and different grades of rice was also required at a minimal frequency of once a month. The prices recorded were selling prices of grains in each of the city markets, given in the standard government accounting unit of taels (silver currency) per bushel. It is generally believed that the grain prices closely correspond to free market prices (Chuan and Kraus 1975, Shiue 2003 forthcoming).

The price reports were sent to the prefectural level, where the highest and the lowest price observed in the city market of every county in a particular month were recorded and compared with the highest and lowest observed in the previous month. These reports are provided in the *Gongzhong zhupi zouzhe, nongye lei, liangjia qingdan* [Grain Price Lists in the Agricultural Section of the Vermilion Rescripts in the Palace Archives]. The series we use consists of the 2nd and 8th lunar months from 1742-1795.[10]

Historical weather data, from the State Meteorological Society (1981), gives weather data throughout China for each year for 120 "stations" (a regional designation equal to one or two prefectures). The sample we construct is created by using the information from the table and from pinpointing the location of prefectures on the weather maps. The variable is a discrete indicator of the degree

[10] Because of missing values, we estimate parts of the data; see the appendix for more on this.

of "wetness and aridity", from floods, droughts, monsoons, or rainfall. Bad weather ranks are 1 and 5 (exceptional drought and flood), fair weather ranks are 2 and 4 (limited drought and flood), and good weather is rank 3 (favorable conditions). From this data, we have constructed three different weather variables (see appendix for details).

2.2 Summary statistics

The location of the prefectures can be described using Geographic Information System (GIS) data, which we will employ extensively below. Table 1 reports the prefectures' longitude and latitude, by province. The North-South axis is spanned by Jiangsu and Guangdong provinces (latitudes of about 32 and 23, respectively), whereas the West-East range is given by Guizhou and Zhejiang (longitudes of about 107 and 120, respectively). The area of analysis is approximately 1600 kilometers by 1100 kilometers (km). In our analysis, the longitude and latitude of the capital city of each prefecture is used to measure Euclidean distance between prefectures. This distance ranges from about 10 to 1730 km in our sample.

We focus primarily on the middle price, computed as the average between highest and lowest price in each prefecture in each month.[11] Table 2 presents summary statistics on the mid-, lowest and highest prefectural price by province. There are a total of 13,068 observations, corresponding to two monthly (2nd and 8th month) observations for 54 years and 121 prefectures. There is a substantial amount of variation, with the mid-price ranging from 0.600 to 3.250 taels and a standard deviation of 0.334. Across provinces, the mean ranges from a low of 0.972 in Guizhou to a high of 1.794 in Jiangsu. In general prices in inland regions are lower than on the coast.

Table 3 shows how the weather, a key determinant of the quality of the harvest and hence agricultural output, varies across regions. The Bad Weather variable indicates that the prefectures

[11] We compute this average for regions that differ substantially in size (Figure 2). The implied spatial averaging does not affect our results in a major way; using either the low or the high price leads to similar results.

of Fujian province experience exceptional floods or droughts (mostly floods, related to monsoons) in about 17% of the years, for instance, whereas among the prefectures in Guangxi this occurs in only about 4% of the years. This highlights the relatively low weather variability in the inland areas. The table also indicates that for the sample as a whole exceptional drought and flood years account for about 10% of all cases (last line).

Our analysis focuses on how the scope and direction of interregional trade has changed over time. To this end, it is useful to examine the changes in prices and weather over this period. Tables 4 and 5 provide the major trends by reporting summary statistics for three 18-year periods, namely 1742/59, 1760/77, and 1778/95. Table 4 shows that, overall, prices have risen from about 1.33 to 1.48, or 0.2% per year on average, and the rate of price increase is somewhat higher in the later years. An important fact for our analysis is that with the exception of Guizhou province, all provinces experienced a slight price increase over time. The table also suggests that the 2nd and 8th month prices behave not too differently, which allows us to focus largely on the 8th month prices without losing much.

As the ease of arbitrage across markets within a region increases, the variation of prices in any one market is expected to decline. Table 4 reports the coefficient of variation, defined as the standard deviation over the mean, both by province and for the sample as a whole. Within-province price variation tends to decline over time (70% of the cases), but for the sample area as a whole, the variation in price increases (from 0.214 to 0.248 for the 8th month prices; last row). To the extent that these changes in price variation are due to the emergence of interregional grain trade, its geographic scope seems to be limited: trade appears to increase among the prefectures within a single province, but there does not seem to be strong enough forces of arbitrage that would bring about one and the same price at a national level.

Lastly, Table 5 shows how the weather changed over the sample period. In general, weather is an exogenous variable that is also expected to be random, even though a certain region may be more (or less) susceptible to a harsh climate. In the weather data for this sample, the percentage of years of Bad Weather, e.g., does not vary much overall (from 10.1 over 9.2 to 11.1 percent; at the bottom).

We now turn to the empirical analysis.

3 Measures of spatial autocorrelation

Two of the most widely used measures of spatial autocorrelation are Moran's I and Geary's c statistics (Moran 1950, and Geary 1954, respectively). Let $\{\delta_{ij}\}$ be a connection matrix in which $\delta_{ij} = 1$ if the ith and jth prefecture are joined, and $\delta_{ij} = 0$ otherwise. Frequently, jurisdictions are considered to be joined if they share a boundary. This can be generalized by considering matrices for different degrees of connectedness, or, different spatial lags. Let k denote the spatial lag, $k = 1, ..., K$, with $\{\delta_{ij}^{(k)}\}$ the corresponding connection matrices. In our analysis, the spatial lags are based on distance.

We use the prefectures' location in terms of longitude and latitude to compute the Euclidean distance D_{ij} between any pair of prefectures i and j. Then, let $\{\delta_{ij}^{(1)}\}$ be the connection matrix for spatial lag $k = 1$, where $\delta_{ij}^{(1)} = 1$ if $D_{ij} \leq 200$ km, and $\delta_{ij}^{(1)} = 0$ otherwise. Similarly, let $\{\delta_{ij}^{(2)}\}$ be the spatial lag ($k = 2$) matrix, with $\delta_{ij}^{(2)} = 1$ if $200 < D_{ij} \leq 400$ km, and $\delta_{ij}^{(2)} = 0$ otherwise. Define also the following spatial lags by successive 200 km distance bands, up to $k = 9$, where $\delta_{ij}^{(9)} = 1$ if 1600 km $< D_{ij} \leq 1800$ km, and $\delta_{ij}^{(9)} = 0$ otherwise. Let p_i denote the log of the mid-price of prefecture i, $i = 1, ..., N$.

Moran's I statistic for spatial lag k is defined as

$$I_k = \frac{N}{2J_k} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{ij}^{(k)} z_i z_j}{\sum_{i=1}^{N} z_i^2}, i \neq j, \tag{1}$$

where $z_i = p_i - \overline{p}$, $\overline{p}$ is the average price, $\overline{p} = N^{-1} \sum_{i=1}^{N} p_i$, and J_k is the number of nonzero values of $\delta_{ij}^{(k)}$. Correspondingly, Geary's c statistic for spatial lag k is defined as

$$c_k = \frac{(N-1)\sum_{i=1}^{N}\sum_{j=1}^{N}\delta_{ij}^{(k)}(p_i - p_j)^2}{4J_k\sum_{i=1}^{N} z_i^2}, i \neq j. \tag{2}$$

Both the Moran and the Geary statistic measure the covariance of prices in connecting prefectures relative to the variance of the price across prefectures. Under the null hypothesis that the z_i are identically and independently distributed normal variates, the expected value of the Moran statistic is $E[I_k] = -(N-1)^{-1}$. For Geary's c_k statistic, the expected value under these assumptions is $E[c_k] = 1$. Inference is based on the result that the I_k and c_k statistics are asymptotically normal under the null.[12] In the following we examine how spatial autocorrelation varies across different spatial lags.

3.1 Spatial autocorrelation and distance

Table 6 shows Geary's c statistic for different groups of prefectures and subperiods by distance band. The latter ranges from (0,2] to (14,16]. For the full sample of 121 prefectures, the first column shows that Geary's c rises monotonically from a value of 0.289 for the (0,2] band to 2.47 for the (14,16] band.[13] Under the null hypothesis of no autocorrelation, the expected value is one. Therefore there is evidence for positive autocorrelation for distances up to 800 km and evidence for negative autocorrelation for distances above 1,000 km.

Intuitively, this result can be best understood by considering two opposing pictures of spatial patterns: clusters versus the checkerboard pattern. If similar prices are clustered together, this will produce a positive covariance among connected regions, and hence positive spatial autocorrelation.

[12] See, e.g., Cliff and Ord (1981), who also derive higher moments of the Moran and Geary statistics.

[13] These are means across 54 years; the table also shows the standard error of these means in parentheses. The latter tend to be relatively small, but increases if the number of prefectural pairs for a given distance is small.

However, if prices vary a lot in different markets for a certain spatial lag, there is negative spatial autocorrelation—a checkerboard pattern. In the case of rice markets in 18th century China, higher spatial autocorrelation for shorter distances is plausible: since transport technology was not very advanced, there should be a diminished possibility of trading grain over long distances relative to shorter distances. This means that trade will tend to connect markets for relatively short distances before it does so over longer distances.

Negative autocorrelation over longer distances might be due to several factors. Perhaps most importantly, prices tend to be increasingly dissimilar for geographic locations that are further apart because of differences in soil, weather, and other reasons. Thus, even if one finds negative spatial autocorrelation over long distances, it could still be the case that long distance trade contributes to price covariation if in the absence of trade the spatial autocorrelation at this distance would be even lower. At the same time, positive autocorrelation over short distances does not necessarily imply trade—it is merely consistent with it. The most important alternative reason of why prices might be spatially autocorrelated is probably common weather patterns.

How important is spatial autocorrelation for weather relative to prices? Geary's c for weather suggests that the variable is spatially autocorrelated, but the effect is less pronounced than that for prices. For example, for the 200 to 400 km band, Geary's c is 0.46 for prices and 0.85 for weather (recall that the expected c under the null hypothesis of no spatial autocorrelation equals one) . While part of the correlation in price might be attributable to common weather shocks, prices are clearly autocorrelated even after we take weather into account. Moreover, for very long distances, the spatial autocorrelation for prices is lower than for weather. That also suggests that the results for prices capture more than weather effects.

We have obtained similar results using Moran's I statistic.[14] This is the case in general for our

[14]For instance, the correlation between Geary's c_k and Moran's I_k for all 121 prefectures across all distance bands

sample and so we present only the Geary's c statistic.

3.2 Spatial correlation in different regions

This section considers how the geographic features of the cross-sectional heterogeneity across prefectures might affect spatial patterns in price. Because transport costs of grain over land was much higher than that of ship transport for the same distance in 18th century China (Evans, 1984), a major distinction can be made between areas where grain transport by ship was possible and those where it was not possible (only overland transport, by e.g. porter or cart). First, we divide the sample into those prefectures that lie directly on or near to the coast—relatively low-cost ship transport—and those that are located more inland. The Coastal sample is defined as the prefectures in the provinces of Anhwei, Fuijan, Guangdong, Jiangsu, and Zhejiang (59 prefectures), whereas the Inland prefectures are those in Guizhou, Guangxi, Hubei, Hunan, and Jiangxi (62 prefectures).

Second, there is evidence for the importance of the Yangzi River, China's longest navigable waterway, for interregional transport in China, both historically as well as today. We therefore define a set of prefectures located near the Yangzi River. In the 18th century, the Yangzi was navigable by sizable watercraft for at least 1,000 kilometers upriver from its mouth near the city of Shanghai (Worcester 1971). The Yangzi's path is outlined in Figure 2. This Yangzi River sample consists of 21 prefectures, and has maximum distance band (8,10].[15]

The Geary's c statistics for these Inland, Coastal, and Yangzi River samples are shown in Table 6, and spatial correlograms are presented in Figure 3. For each of these groups, the Geary statistic increases with distance. Across all samples, there is consistent evidence of price clustering at distances of 0 to about 400 km, little or no evidence of autocorrelation at distances between 400 km and 800

and the three subperiods given in Table 6 is with -0.98 close to -1.

[15] These are prefectures 1, 4, 5, 6, 12, 64, 65, 66, 69, 78, 79, 87, 88, 89, 90, 91, 93, 94, 96, 106, and 107; see Figure 2. Note that the Yangzi sample includes several of the prefectures that are also in the Coastal sample.

km, and negative autocorrelation for distances above 800 km.

The figure shows that the level of spatial autocorrelation differs markedly across regional samples with distance, confirming that the full sample of prefectures contain cross-sectional heterogeneity in spatial patterns. The Geary's c statistic rises with the steepest incline for the Yangzi River sample. By contrast, the Coastal sample displays a relatively flat incline with progressively higher distance bands, and the inland areas show moderate increases with distance. The patterns suggests that it is in the Yangzi River areas where we would find the greatest degree of local spatial clustering, as well as the greatest degree of overall heterogeneity in the covariation between markets within the sample. This may occur if strong linkages between major trading partners coexist with areas that are not as well connected. The Coastal sample displays less short-distance clustering, but as distance increases, the covariation between markets does not decline quite as rapidly, suggesting a degree of homogeneity in the spatial pattern of markets in this sample that does not appear in the Yangzi River sample. Spatial clustering also occur in the Inland sample, especially at short distances, but compared to the Yangzi River sample the Inland area also displays relatively small changes in the strength of clustering as distance increases.

We note also the presence of spatial non-stationarity appearing in the data. Prices in the interior of China are lower than on the coast, in part due to lower population (per unit of arable land) in inland regions. This East-West gradient affects the value of our autocorrelation statistic for the Yangzi sample in particular, as the Yangzi prefectures are exactly located along this gradient.[16] In contrast, there is more directional variation in the other samples, which limits the effect. Below we will use local spatial autocorrelation techniques in an attempt to obtain a more detailed picture.

[16] The numerator of Geary's c sums over the squared price differential for a given distance band, $(p_i - p_j)^2$, see equation (2) above. In the Yangzi sample, due to the East-West price gradient, this means that for long distances c captures large price differences, between one very low- and one very high-price market; Moran's I is affected similarly.

3.3 Changes in spatial autocorrelation over time

Our sample period of more than half a century allows us to consider changes over time. To this end, we have computed the mean of the Geary statistics for the four groups of prefectures—All, Inland, Coastal, and Yangzi—separately for three 18-year subperiods: the years 1742 to 1759, 1760 to 1777, and finally the years 1778 to 1795. The statistics are presented in Table 6, columns two to four.

For the sample as a whole, spatial autocorrelation is increasing over time for distances of 0 to 800 km while it is decreasing above 800 km.[17] Similarly, for Inland prefectures clustering is increasing for shorter and decreasing for longer distances, except that the cutoff distance is around 600, not 800 km. This suggests that it is necessary to keep track of distance brackets in analyzing changes over time.

If initially there is no interregional trade at any distance, there will be relatively little price clustering at any distance (controlling for the effect of weather and other factors). Once trade over short distances starts to emerge, central market locations will appear: this leads to more clustering within regional markets. At the same time, the formation of scattered central markets within different regions will lead to what will appear to be less clustering across a sample of *interregional* markets. For our case, since interregional trade does not yet connect markets that are more than 800 km apart for the sample as a whole, the cutoff value of about 800 km gives the distance for which trade would lead to a strengthening of within-region clusters.

The degree of spatial autocorrelation changes quite differently among the Inland prefectures compared to the Yangzi River prefectures, see Figures 4a and 4b. The top part shows that over time, there is more clustering in Inland prefectures for distances below 600 km, and less clustering for distances above that. For the Yangzi prefectures, Figure 4b shows that for the sample period as

[17] The tables report standard errors for these estimates; formal tests regarding the statistical significance of differences in these estimates are available from the authors upon request.

a whole exactly the opposite is true: less clustering (at least weakly) for distances up to 600 km and more clustering for distances above 600 km. This difference underscores the possibility that Inland and Yangzi prefectures were at different stages of interregional market development. For Inland prefectures, trade was only starting to connect markets at the relatively short distances, whereas for Yangzi River prefectures, trade was already, and increasingly, linking prices in markets that were much further apart.

If the increasing feasibility of trade over time means that prices tend to covary more (higher spatial autocorrelation)—even though this occurs at different distances for different groups—then what could account for the reduction in price covariation for distances of less than 600 km among the Yangzi prefectures? One possibility is that over time, it becomes more efficient to trade with a different set of partners. Assume for simplicity that there are two Yangzi prefectures that are 600 km apart, with no other prefecture between them. Suppose that initially shipping costs are too high for trade to occur between them. Further, assume that overland trade costs are low enough for there to be trade for distances up to 200 km away from the river. This would result in some covariation of prices in the river prefectures with their respective local hinterland markets (up to 200 km), but no covariation of prices over longer distance, between the two river prefectures. Over time, if the technology for river transport improves faster than that for land transport, it could be efficient for the Yangzi regions to switch their primary trade partner from the respective hinterland to more distant markets along the Yangzi river. The result would be a weakening of price covariation for short distances, and an increase in price covariation for longer distances, as depicted in Figure 4b.

We conclude the section with the following local spatial autocorrelation analysis.

3.4 Local Spatial Autocorrelation

This section uses a variant of Geary's c to identify the prefectures that have the greatest relative spatial clustering. Geary's c_i, due to Anselin (1995), is used to calculate a local spatial association coefficient for each individual locality i instead of for a certain distance class. For distance band k and locality i, this coefficient is defined as

$$c_i^k = \frac{N \ \sum_{j=1}^{N} \delta_{ij}^{(k)} \left(p_i - p_j\right)^2}{\sum_{q=1}^{N} \left(p_q - \overline{p}\right)^2}. \tag{3}$$

As the numerator indicates, this local measure is based only on a single summation, for each locality i, whereas Geary's global coefficient sums over localities i and j (see equation (2)). Two separate distances bands were used in this procedure on the price of rice: the 0 to 300 km and the 300 to 600 km bands (denoted (0,3] and (3,6], respectively). The Geary's c_i coefficients were then separately ranked and the prefectures falling in the lowest 25% of all prefectures (i.e., the 30 prefectures with the highest local spatial autocorrelation) were plotted on a map of China. The locations of the waterways were in no way incorporated in the calculation of Geary's c_i.

The results of the ranking of the prefectures with greatest local distance autocorrelation in price (shaded squares) show that the strongest clusters of trade originate along the Yangzi River and its main tributaries: the Yuan River, Gan River, and Huai River. In addition, there are also indications of local clustering in the southern provinces.[18] The association between the waterways and spatial market clusters is visible in Figure 5, which superimposes the location of the prefectures on the coastline and major lakes and rivers in China for distance band (0,3].[19]

Figure 6 shows the ranking at a further distance band, (3,6], for each individual locality. Com-

[18] These are prefectures 29, 56, and 57, respectively, in Figure 2.

[19] The location of rivers and coastal boundaries shown in this map comes from China Historical GIS (2002).

pared to distance band (0,3], the local clustering found in the southern provinces is relatively weaker, and they no longer rank among the 25% most spatially autocorrelated markets at that distance. Twelve prefectures appear in both maps; these prefectures are for the most part located along the Yangzi River, but there are also a number of prefectures directly on the Yangzi River at distance (0,3] that are not strongly autocorrelated with the (3,6] band. These are likely to be prefectures that had relatively strong local connections with nearby markets, but not with far away ones. In the (3,6] band, we also observe strong autocorrelations among prefectures that are located somewhat further from the Yangzi River, for instance, at markets more distant from the main artery, yet still on a tributary. It is likely that these are the prefectures most closely linked to the Yangzi River trade at a longer distance.

We now turn to a regression framework that allows to analyze this further.

4 Regression results

We start out with the linear regression model

$$y = X\beta + \varepsilon, \tag{4}$$

where y, the dependent variable, is $N \times 1$, X is a $N \times K$ matrix of exogenous variables, and ε is a $N \times 1$ error term distributed as $\varepsilon \sim NID\left(0, \sigma^2\right)$. Here y is the (log) mid price and X consists of a constant and the weather variable. Under the stated conditions, ordinary least squares is the best linear unbiased estimator, and we will report it as a baseline. Given the size differences of the prefectures, we also present Huber (1967) and White (1980) heteroskedasticity-consistent standard errors.

Our results on spatial autocorrelation above strongly suggest that the residuals of (4) are spatially dependent. One approach would be to test this assumption by applying a formal test for spatial dependence.[20] A second approach is to adjust the estimated covariance of regression (4) for spatial dependence. Conley's (1999) nonparametric approach can be viewed as the spatial counterpart of the Newey and West (1987) heteroskedasticity and autocorrelation consistent time series covariance estimation. His covariance estimator uses weighted averages of sample autocovariances that are computed from subsets of observation pairs falling within a given distance band. We compute these standard errors for a number of different distance bands.[21] Relative to OLS, these methods affect only the computation of the standard errors.

We then move to regression models that incorporate the spatial autocorrelation into the regressive structure. The two most influential models are the spatial error dependence and the spatial lag dependence model. The former is given by

$$y = X\beta + \varepsilon \tag{5}$$

with

$$\varepsilon = \lambda W \varepsilon + u \tag{6}$$

where λ is the spatial autoregressive coefficient, $u \sim NID(0, \sigma^2)$, and W is $N \times N$ matrix of known spatial weights w_{ij}. These weights correspond to the connection matrix $\{\delta_{ij}\}$, defined above, in that they capture the spatial structure. If $\lambda \neq 0$, ignoring the spatial dependence means OLS is inefficient

[20] In addition to a Moran or Geary test on the residuals of (4), a number of other tests have been proposed (e.g., Kelejian and Robinson 1992, Anselin and Bera 1998, Baltagi and Li 2001). See also Kelejian and Prucha (2001) on the relationship of different tests and further results.

[21] As emphasized by Anselin (2001), the asymptotics of Conley's (1999) covariance estimator relies on the data generation process being spatially stationary, and this might not hold in the present context (see the discussion in section 3.2). Therefore care is required in interpreting these results.

but remains unbiased. The spatial lag dependence model is given by

$$y = \rho W y + X\beta + \varepsilon \tag{7}$$

where ρ is the spatially autoregressive parameter and $\varepsilon \sim NID\left(0, \sigma^2\right)$. If $\rho \neq 0$, leaving out the term $\rho W y$ from equation (7) and running least squares gives biased and inconsistent results. We will present results from maximum likelihood estimation of both models below.

Our spatial weights are based on distance. We considered a variety of specifications, including one- and two-window distance bands and exponential specifications. For one-window distance bands, $w_{ij} = 1$ if D_{ij} is less than some maximum $D_{\max}$, and $w_{ij} = 0$ otherwise, and different spatial structures are captured by varying $D_{\max}$.[22] An example of a two-window weighing matrix, $W(0, 3, 6)$, might specify a weight of one for distances between $(0, 3]$, a weight of one-half between $(3, 6]$, and zero for distances above six. Exponential weights are of the form $w_{ij} = \exp(-\theta D_{ij})$, where a higher value of θ leads to a more rapid decline in the size of the weights as distance increases.

In a limited grid search in terms of likelihood for a good weighting matrix, the exponential specification with parameter $\theta = 1.4$ tended to perform best.[23] Among the distance band specifications, the one-window specification with $D_{\max} = 3$ and $D_{\max} = 6$ (corresponding to distances of 300 and 600 kilometers, respectively) performed best. The results for these three matrices are shown in the following table.

[22] The own weight, w_{ii}, is set to zero $\forall i$ in all weighting matrices.

[23] With distance measured in units of hundreds of kilometers, this leads to weights from 0 to 0.85, with a mean of 0.017. Also note that all estimations use row-standardized weights that sum to one.

4.1 Spatial regression with all prefectures, 1742-1795

Table 7 shows the mean of the estimates for the models above across the 54 sample years, 1742 to 1795.[24] First, the OLS results might suggest that a prefecture's weather has a significant positive effect on the local price.[25] The point estimate of 0.08 (s.e. of 0.03) suggests that moderately bad weather, either above normal wet or dry, raises the price by about eight percent, whereas exceptional floods and droughts are associated with a 16% higher price. In terms of fit, the OLS regression has an R^2 of about 0.10, and the log likelihood is given by 19.463 (last line). The second column shows Huber (1967)-White (1980) heteroskedasticity-consistent standard errors. They are similar to the usual standard errors.

Conley's (1999) nonparametric spatial standard errors are presented in Table 7 for distance bands $(0, 3]$ and $(0, 6]$, respectively (denoted "Spatial corr. adjusted s.e.'s"). They are about 60% larger than the non-spatial standard errors. That spatial standard errors are larger than conventional ones is plausible, because if there is spatial dependence, the effective size of the sample is reduced relative to one with independent observations (e.g., Cressie 1993).[26] Now one cannot reject the null that local weather has no effect on price at the 5% significance level. A reduction in the effect of local weather on price once spatial dependence in prices is incorporated is consistent with trade being the cause of spatial dependence.

The results for the spatial error and spatial lag dependence models are in the lower part of Table 7. They are for the three different weighting matrices discussed above: distance bands (0,3] and (0,6], and $w_{ij} = \exp(-\theta D_{ij})$, with $\theta = -1.4$. First, notice that the spatial models fit much better

[24] Also Tables 8 and 9 show the mean estimate across 54 years, while Table 10 shows mean estimates from 18-year subperiods. We have omitted the standard errors of these means in the interest of space; they are available upon request.

[25] Among the three weather variables, we focus on Weather Deviation, the extent to which the weather differed from medium dryness, the generally best harvest conditions. The other two weather variables give similar results.

[26] At the same time, Conley (1999) demonstrates that his spatial standard errors are not necessarily larger than conventional standard errors.

than the models that omit the spatial structure, with a log-likelihood ranging from about 60 to 85, versus about 20 before. Second, this improvement is clearly due to the spatial structure. The coefficient of the spatial error model, λ, lies between 0.9 and 0.95, with a standard error of about 0.05. The test statistic of the LM test—see e.g. Anselin and Bera (1998), Baltagi and Li (2001)—is χ^2 distributed with one degree of freedom, lies between 239 and 495, with a p-value of zero. The results for the spatial lag model are similarly strong. Lastly, evidence on the influence of weather is now further weakened: its point estimate is between 0.02 and 0.04, compared to 0.08 before, and it is generally not significant at standard levels. This is an example where accounting for spatial structure qualitatively changes the inferences.

Comparing the results for the six models in terms of fit, one sees that the exponential weighing matrix performs better than the other two. Moreover, the spatial lag model has always a somewhat better fit than the corresponding error model. The error model is sometimes seen as appropriate when the spatial autocorrelation in the data is the outcome of numerous factors. This is in contrast to the spatial lag model, which isolates one factor (Haining 1990). Our finding that the spatial lag model is always preferred to the corresponding spatial error model suggests that controlling for local weather, the price in one locality is predominantly related to prices in other regions, but not much else. This is consistent with interregional trade leading to a link in prices and that there is no other major reason that underlies the spatial autocorrelation in the data.

4.2 Coastal versus Inland prefectures

Table 8 shows the regression results for the Inland and Coastal samples. In the OLS regression (upper left corner), the point estimate for weather is higher in the Inland than in the Coastal sample, but in neither equation does the coefficient enter significantly at standard levels. The different constants, 0.146 (Inland) and 0.448 (Coastal), confirm that on average prices are higher in Coastal areas. As was the case for the whole sample, the Huber-White standard errors are quite similar to the conventional ones, whereas the spatially adjusted standard errors are higher.

For the spatial models, the results are shown in the lower part of Table 8. We focus now on the (0,3] window and the exponential weighting matrix ($\theta = -1.4$, as before). First, the log likelihood values indicate that the spatial models lead to a substantial increase in fit, although it is not quite as large as for the sample as a whole. The evidence for spatial dependence is strong, with the LM test indicating that the H_0 of no dependence can always be rejected at a 5% level. For the (0,3] weighting matrix, the spatial autoregressive parameter for the Inland sample is around 0.83, whereas for the Coastal sample it is about 0.63. Correspondingly, one cannot reject the H_0 of no dependence at a 1% level for the Coastal sample. This is consistent with less clustering in the 0 to 300 km range along the coast.

Interestingly, the difference in terms of estimated spatial dependence narrows between the Inland and Coastal samples when the exponential weighting matrix is employed (lower right of the table): the p-values are about 0.1% for both samples. The exponential matrix gives some weight to prefectures in the medium to long-distance range while the 0 to 300 kilometers matrix does not. It is possible that for the coastal prefectures there is stronger evidence for spatial dependence with exponential weights because some trade relations are already in the middle to long-distance range.

4.3 Yangzi River versus Non-River prefectures

Table 9 highlights the importance of river transport by isolating the Yangzi River prefectures. First, in the OLS specification weather has an influence on the price only in non-Yangzi regions. The Huber-White adjustment makes only a small difference, which is in line with our earlier results. However, here the spatially adjusted standard errors are larger than the conventional ones only for the Non-Yangzi prefectures (upper part on the right). This may be indirect evidence for less spatial dependence in the Yangzi sample than in the Non-Yangzi sample.[27] With spatial standard errors, the effect of weather on price tends to vanish at standard levels of significance even for the Non-Yangzi sample.

The specifications for the spatial regression models in the lower part of Table 9 help to clarify these issues. Overall, there is evidence for spatial dependence in both subsamples. At the same time, it is weaker for the Yangzi prefectures. While for the Non-Yangzi prefectures, the LM test p-value is zero irrespective of model and spatial weights, for the Yangzi prefectures one can reject the H_0 of no dependence at the 10% level, but never at the 1% level. This suggests that prices are less strongly clustered along the Yangzi River. Also note that the evidence for spatial clustering in the Yangzi sample is slightly lower using the 0 to 300 km spatial weights than for the exponential weights (p-values of 5.8% and 2.3%, respectively). The 0 to 300 km window appears to be too small for Yangzi River prefectures. In this sense, these LM tests confirm that trade along the Yangzi had linked markets over substantial distances by then.

4.4 Spatial dependence over time

Table 10 show how spatial dependence among Chinese rice prices has changed over the sample period. There is strong evidence for positive spatial dependence, and for the sample as a whole the clustering has somewhat increased over time: the spatial autoregressive parameter ρ is estimated to be 0.856 during the years 1742-59, 0.904 for 1760-77, and 0.917 during the years 1778-95.[28] This increase in positive spatial autocorrelation is stronger among the Inland prefectures: initially, ρ is estimated to be 0.71 (LM test p-value of 0.3%), and towards the end of the sample period ρ has risen to 0.89 and

[27] Other possibilities include problems coming from the relatively small size of the Yangzi sample ($N = 21$), or non-stationarity of the data.

[28] Are these ρ estimates significantly different at standard levels? The standard error of the mean ρ of 0.856 for 1742-59 is 0.014, and that for the mean ρ of 0.917 for 1778-95 is 0.007. Thus, at standard levels, the subperiod estimates here are different, although this is not always the case. The full set of standard errors, with tests results regarding their difference, is available from the authors upon request.

the LM test's p-value is zero. For the set of Coastal prefectures, there is no clear pattern of spatial dependence over time.

The pattern for the Yangzi River prefectures differs substantially from the sample as a whole. Instead of an increase in spatial autocorrelation, the parameter estimate of ρ decreases from about 0.8 to 0.63, with most of the change occurring after the year of 1777. Moreover, during the last eighteen years of the sample there is no strong evidence for price clusters among Yangzi River prefectures (p-value of LM test of 6.8%).

Overall, a picture emerges according to which the level and change in spatial interaction varies systematically with opportunities to interregional trade among these prefectures. For some prefectures, the intensification of local trade tightens price clusters over short distances; these seem to be primarily the inland regions of China during the late 18th century. Regions on the coast and along rivers have good local opportunities to trade as well, but they have also medium- to long-distance trade opportunities. This has shifted the frontier of what constitutes the local economy for these economies, and price clusters at short distances have begun to disappear.

We now turn to a concluding discussion.

5 Conclusion

This paper has shown how estimators of spatial autocorrelation may be employed with price data to provide evidence on interregional trade patterns. The geographic locations of the most locally integrated markets in our sample are found to lie along the Yangzi River and its tributaries, a result that is consistent with historical accounts that have emphasized the importance of physical geography in the emergence of trade. The results also suggest that accounting for spatial structure in contexts where spatial effects play a major role alters the inferences on economic outcomes. For our sample, we were able to obtain information on local weather shocks, a variable that could lead to spurious price correlation because weather, like trade, is a geographically localized variable. We find that the spatial models effectively removes the significance of weather in the estimation results.

The spatial patterns in the data indicate that markets which are most likely to be integrated over longer distances are not necessarily also the same locations which are most integrated with nearby markets. The difference appears to depend on the location of a market with respect to its most efficient trade route.

When data for long periods is available, there are clear benefits to allowing for temporal as well as spatial changes in the model. Over time, we find that markets appear to substitute between local and more distant markets. It is likely that differences in the rate of improvement of different transport technologies, as well as its location with respect to the preferred trade route, determine whether a certain market begins to engage in long-distance trade. When long-distance trade does emerge, our results suggest that its linkages with more nearby markets may diminish.

In future work, it would be interesting to combine this analysis with direct evidence on how trade routes evolved in China over time. Another direction for future research is to examine the long-run consequences of market access and interregional trade in China. From what we know about regional differences in income per capita in China in the 18th century, income appears to be correlated with market access and trade. Notably, the Chinese regions that are rich today (the Yangzi Delta including Shanghai, the area around Guangdong and Hong Kong, as well as the coastal areas of Fujian) include many that were relatively rich already a couple of centuries ago. The impact of geography would seem to have lasting effects.

A Price data

During the era of the Qing emperors (years 1644 to 1911), we have relatively good price records for the 18th century. In these years an empire-wide price reporting system was in place, serving the purpose of an early warning system to prevent famines. The fact that these data were put to practical uses suggests accuracy would have been important and, to the extent that is was possible, enforced. In addition, a system of unannounced checks and audits of the price data was in place. The quality of the data is generally considered good compared to other historical price records.

Our focus is on rice, the most important type of grain in China. This leads to the exclusion of parts of China to the North and West where also other crops, for instance barley, play a major role. With about 60% of the Chinese economy–a population of about 120 million people, about 20% of the world population at the time–, our sample is still sizable. For our sample period from 1742 to 1795, about 24% of the data is missing in the original source. The percentage of missing data is similar across prefectures, and there is no evidence to suggest that the missing data is systematically related to known prefectural characteristics. We have estimated the missing data using the methods developed by Gomez and Maravall (1997). The main findings of this study are not affected by the estimation of missing data.

The original source of the price data is *Gongzhong zhupi zouzhe, nongye lei, liangjia qingdan* [Grain Price Lists in the Agricultural Section of the Vermilion Rescripts in the Palace Archives]. Today it is located in the Number One Historical Archives in Beijing. The data was originally collected by C. H. Shiue, see Shiue (2002) for additional details.

B Weather data

The historical weather data comes from the State Meteorological Society (1981). The weather scale is defined as follows by the compilers of these maps: "Level 1 represents years in which there have been exceptional rainfall, leading to major floods, typhoons, water related disasters, and the destruction of all crops. Level 2 rain encompasses cases where there is heavy rainfall, but limited in scope and/or resulting in only minor flooding. Level 3 weather is the most favorable weather. Level 4 indicates minor droughts of limited consequence, while level 5 denotes the years of exceptional drought, lasting two or more seasons of the year, and leading to major harvest failures." For this paper, the above rankings are used to compute three weather variables: (1) Dryness: This variable gives the weather levels as given in the source. Dryness takes values of 1, 2, 3, 4, and 5 (1 being least and 5 being most dry). (2) Weather deviation: This variable is defined as $|Dryness - 3|$, taking values of 0, 1, and 2. (3) Bad weather: This variable equals one if Dryness is 1 or 5, and zero otherwise. Summary statistics for these variables by province are given in Table 3.

References

[1] Anderson, James, and Eric van Wincoop (2003), "Trade costs", mimeo, Boston College, September.

[2] Anselin, Luc (2002), "Under the Hood. Issues in the Specification and Interpretation of Spatial Regression Models", working paper, University of Illinois, Urbana-Champaign, Urbana, IL.

[3] Anselin, Luc (2001), "Spatial Econometrics", chapter 14 in *A Companion to Theoretical Econometrics*, ed. Badi Baltagi, Blackwell Publishers.

[4] Anselin, Luc (1995). "Local Indicators of Spatial Association–LISA", *Geographical Analysis*, 27: 93-115.

[5] Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

[6] Anselin, Luc, and Anil Bera (1998), "Spatial dependence in linear regression models with an introduction to spatial econometrics", in A. Ullah and D. Giles (eds.), *Handbook of Applied Economic Statistics*, New York, Marcel Dekker.

[7] Anselin, Luc, Attila Varga, and Zoltan Acs (1997), "Local Geographic Spillovers between University Research and High Technology Innovators", *Journal of Urban Economics* 42: 422-448.

[8] Baltagi, Badi, and Dong Li (2001), "Double Length Artificial Regressions for Testing Spatial Dependence", *Econometric Reviews* 20: 31-40.

[9] Case, Anne C. (1991) "Spatial Patterns in Household Demand." *Econometrica*, 59(4): pp. 953-979.

[10] Chen, Zen-Yi, Yeh-Chien Wang, and T.H. Hu (1999), "A Time Series Analysis of Rice Prices in Suzhou during the Eighteenth Century", *Academia Economic Papers* Vol.27, No.3, 311-334.

[11] China Historical GIS (2002), China Historical GIS Dataset, Version 1.0: Qing Dynasty 1820, April 2002.

[12] Chuan, Han-sheng, and Richard Kraus (1975). *Mid-Ch'ing Rice Markets and Trade, and Essay in Price History.* Cambridge: East Asian Research Center, Harvard University Press.

[13] Cliff A. and J. Keith Ord (1981). *Spatial Processes. Models and Applications*, London: Pion Limited.

[14] Conley, Timothy (1999), "GMM estimation with cross sectional dependence", *Journal of Econometrics* 92: 1-45.

[15] Cressie, Noel (1993), *Statistics for Spatial Data*, New York: Wiley.

[16] Cronon, William (1991), *Nature's Metropolis: Chicago and the Great West*, New York: Norton.

[17] Davis, Don, and David Weinstein (2002), "Bones, Bombs, and Break Points: The Geography of Economic Activity", *American Economic Review* 92: 1269-1289.

[18] Diamond, Jared (1997), *Guns, Germs, and Steel: The Fate of Human Societies*. New York: Norton.

[19] Engel, Charles, and John H. Rogers (1996), "How Wide is the Border?", *American Economic Review* 86: 1112-1125.

[20] Evans, Laurence (1984), "Junks, Rice, and Empire: Civil Logistics and the Mandate of Heaven", *Historical Reflections 11(3):* pp. 271-313.

[21] Findlay, Ronald, and Kevin O'Rourke (2002), "Commodity Market Integration, 1500-2000", in M. Bordo, Alan Taylor, and Jeffrey Williamson (eds.), *Globalization in Historical Perspective*, The University of Chicago Press for the NBER.

[22] Geary, R. F. (1954), "The contiguity ratio and statistical mapping", *The Incorporated Statistician* 5: 115-145.

[23] Gomez, Victor, and Agustin Maravall (1997), *Time Series Regression with ARIMA Noise, Missing Observations and Outliers*, Banco de Espana, Madrid.

[24] Greif, Avner (1989), "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders", *Journal of Economic History* 49: 857-882.

[25] Haining, Robert (1990), *Spatial data analysis in the social and environmental sciences*, Cambridge: Cambridge University Press.

[26] Hanson, Gordon (2001), "Scale Economies and the Geographic Concentration of Industry", *Journal of Economic Geography* 1: 255-276.

[27] Heckscher, Eli (1919), "The Effect of Foreign Trade on the Distribution of Income", first published in Swedish in *Ekonomisk Tidskrift* 21, 497-512; reprinted in *Heckscher-Ohlin Trade Theory*, H. Flam and M.J. Flanders (eds.), MIT Press 1991.

[28] Huber, P. J. (1967), "The behavior of maximum likelihood estimates under non-standard conditions", in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA: University of California Press.

[29] Kelejian, Harry, and Ingmar Prucha (2001), "On the distribution of the Moran I test statistic with applications", *Journal of Econometrics* 104: 219-257.

[30] Kelejian, Harry, and Dennis Robinson (1992), "Spatial autocorrelation. A new computationally simple test with an application to per capita county police expenditures", *Regional Science and Urban Economics* 22: 317-331.

[31] Keller, Wolfgang (2003), "International Technology Diffusion", IMF Working Paper, forthcoming.

[32] Krugman, Paul (1991), "History versus Expectations", *The Quarterly Journal of Economics* 106: 651-667.

[33] Mokyr, Joel (1990), *The Lever of Riches*, Oxford: Oxford University Press.

[34] Moran, P.A.P. (1950), "Notes on Continuous Stochastic Processes." *Biometrika*, 37: pp. 17-23.

[35] Newey, Whitney, and Ken West (1987), "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica* 55: 703-708.

[36] North, Douglass, and Robert Thomas (1973), *The Rise of the Western World*, Cambridge: Cambridge University Press.

[37] Ohlin, Bertil (1924), "The Theory of Trade", first published in Swedish as *Handelns Teori,* Stockholm; reprinted in *Heckscher-Ohlin Trade Theory*, H. Flam and M.J. Flanders (eds.), MIT Press 1991.

[38] O'Rourke, Kevin, Alan Taylor, and Jeffrey Williamson (1996), "Factor price convergence in the late nineteenth century", *International Economic Review* 37: 499-530.

[39] O'Rourke, Kevin, and Jeffrey Williamson (2000), *Globalization in History. The Evolution of a Nineteenth Century Atlantic Economy*, MIT Press.

[40] Perkins, Dwight H. (1969). *Agricultural Development in China, 1368-1968.* Chicago: Aldine.

[41] Quah, Danny (2002), "Spatial agglomeration dynamics", *American Economic Review (Papers and Proceedings)* 92: 247-252.

[42] Quah, Danny, and Helen Simpson (2003), "Spatial Cluster Empirics", mimeo, London School of Economics and Institute for Fiscal Studies, June.

[43] Rauch, James (2001), "Business and Social Networks in International Trade", *Journal of Economic Literature* 39: 1177-1203.

[44] Richardson, J. David (1995), "Income inequality and trade: How to think and what to conclude", *Journal of Economic Perspectives* 9, no.3 (Summer): 33-55.

[45] Sachs, Jeffrey (2003), "Institutions don't Rule: Direct Effects of Geography on Per-Capita Income", NBER Working Paper # 9490, February.

[46] Shiue, Carol H. "Local Granaries and Central Government Disaster Relief: Moral Hazard and Intergovernmental Finance in Eighteenth and Nineteenth Century China." Forthcoming in the *Journal of Economic History.*

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

[48] Slaughter, Matthew (2001), "Does Trade Liberalization Converge Factor Prices? Evidence from the Antebellum Transportation Revolution", *Journal of International Trade and Economic Development* 3: 339-362.

[49] State Meteorological Society (1981), *Zhongguo jin wubai nien hanlao fenbu tuji,* [Collected maps of droughts and floods in China in the past five hundred years], Beijing: Ditu chuban she.

[50] Wang, Yeh-chien (1989). “Food Supply and Grain Prices in the Yangtze Delta in the Eighteenth Century”, in *Proceedings of the Second Conference on Modern Chinese Economic History,* 423-262. Taibei: Institute of Economics, Academia Sinica.

[51] White, Hal (1980), “A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity”, *Econometrica* 48: 817-830.

[52] Wilkinson, Endymion (1969) “The Nature of Chinese Grain Price Quotations, 1600-1900,” Transactions of the International Conference of Orientalists in Japan/Kokusai Toho Gakusha Kaigi Kiyo, 14: pp. 54-65.

[53] Williamson, Oliver (2000), ”The New Institutional Economics: Taking Stock, Looking Ahead”, *Journal of Economic Literature* 38: 595-613.

[54] Worcester, G.R.G. (1971), *The Junks and Sampans of the Yangzi.* Annapolis: U.S. Naval Institute.

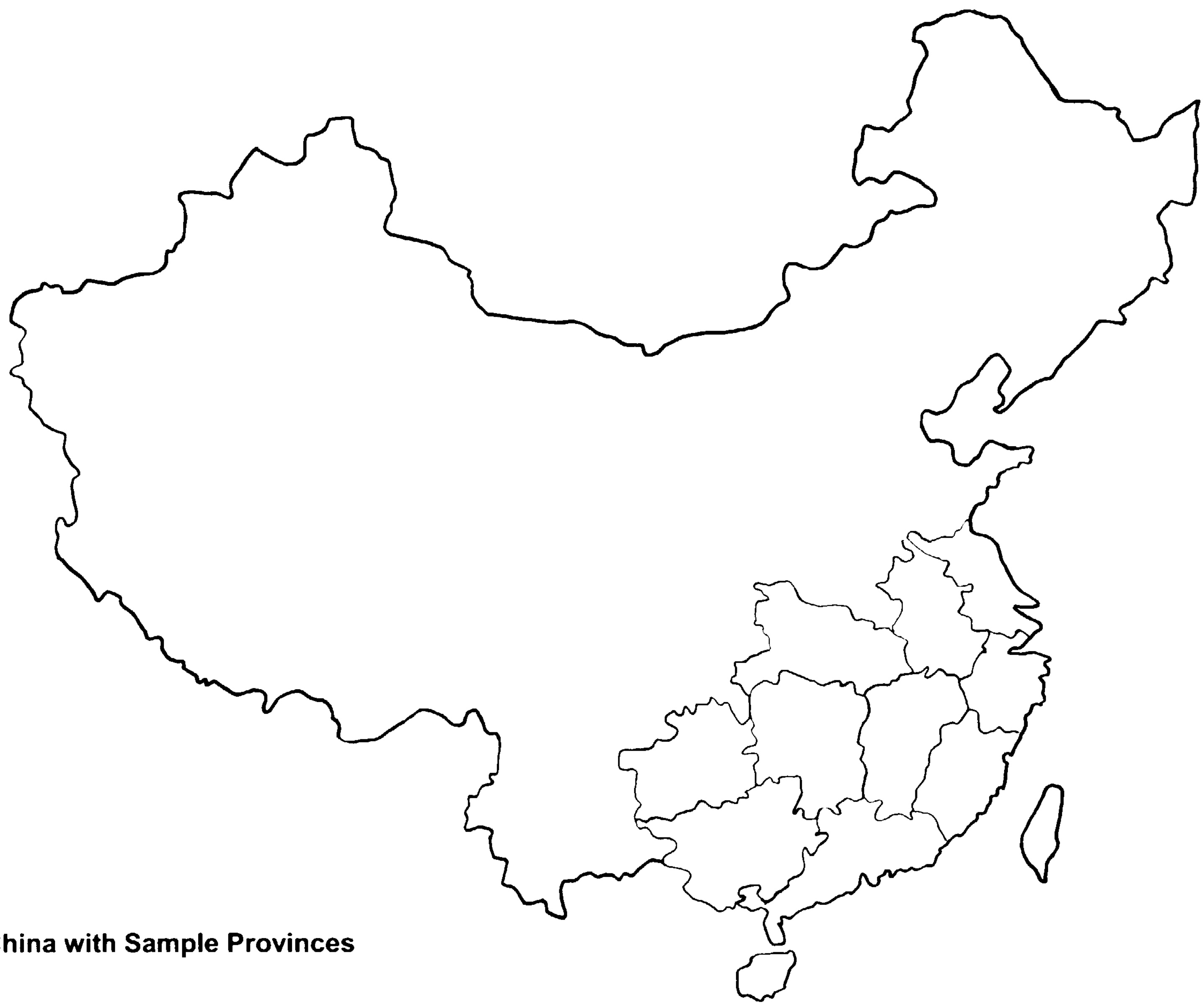
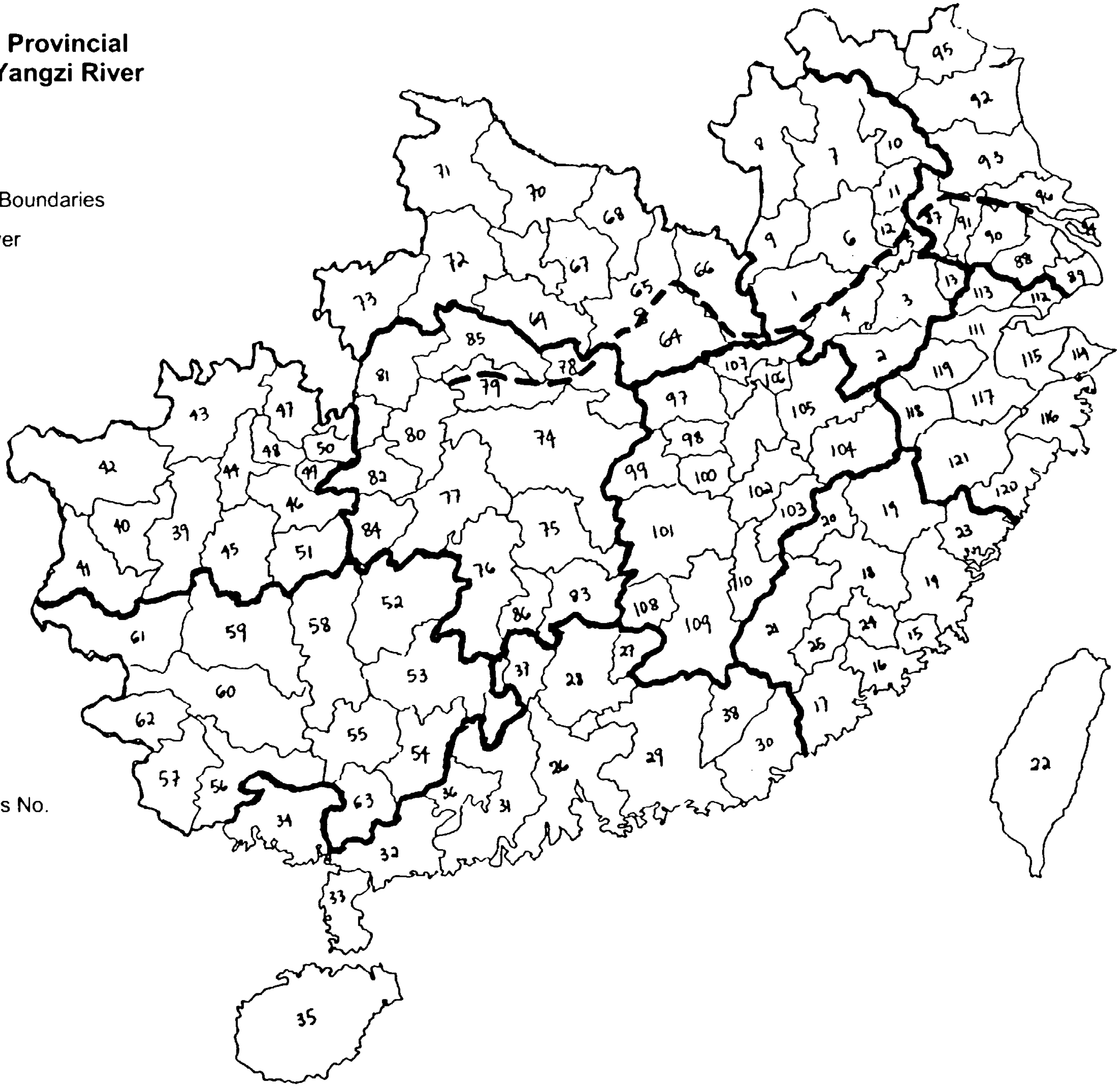


Figure 1
Map of China with Sample Provinces

Figure 2
Chinese Prefectures, Provincial Boundaries and the Yangzi River



———	Provincial Boundaries
- - - - - -	Yangzi River

Province	Prefectures No.
Anhwei	1 to13
Fuijan	14 to 25
Guangdong	26 to 38
Guizhou	39 to 51
Guangxi	52 to 63
Hubei	64 to 73
Hunan	74 to 86
Jiangsu	37 to 96
Jiangxi	97 to 110
Zhejiang	111 to 121

Figure 3. Comparison of Spatial Autocorrelation in Yangzi River, Coastal, and Inland Prefectures, 1742-1795

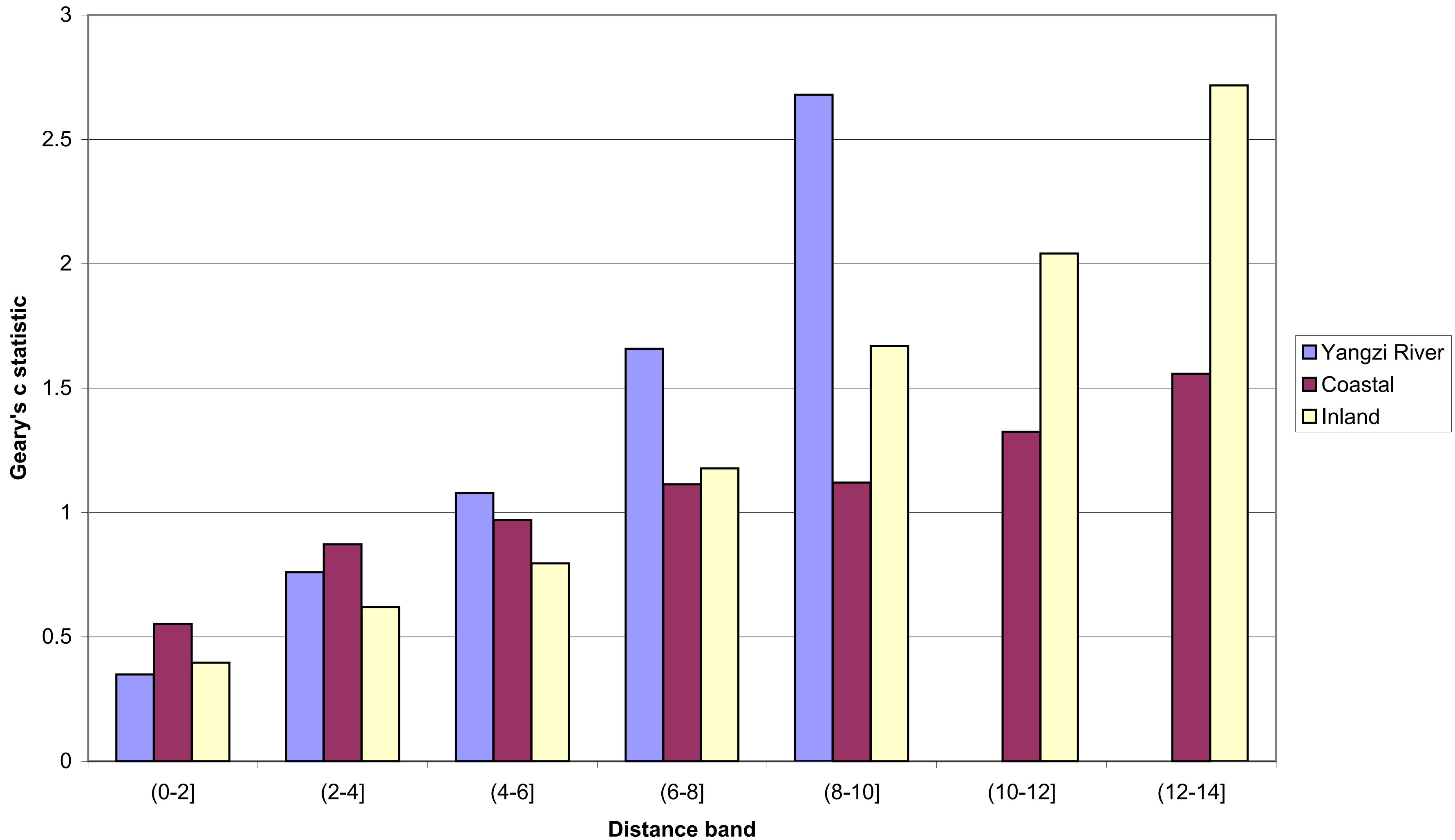


Figure 4a. Changes in Autocorrelation over Time among Inland Prefectures

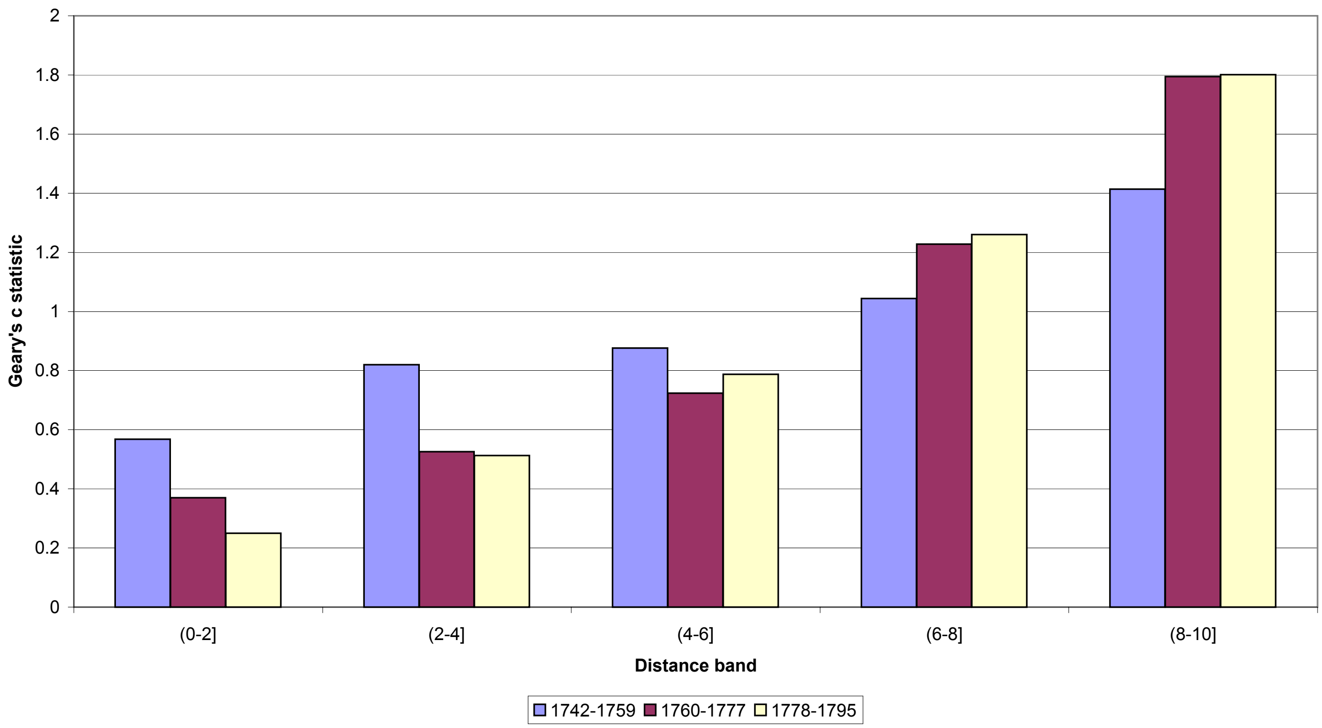
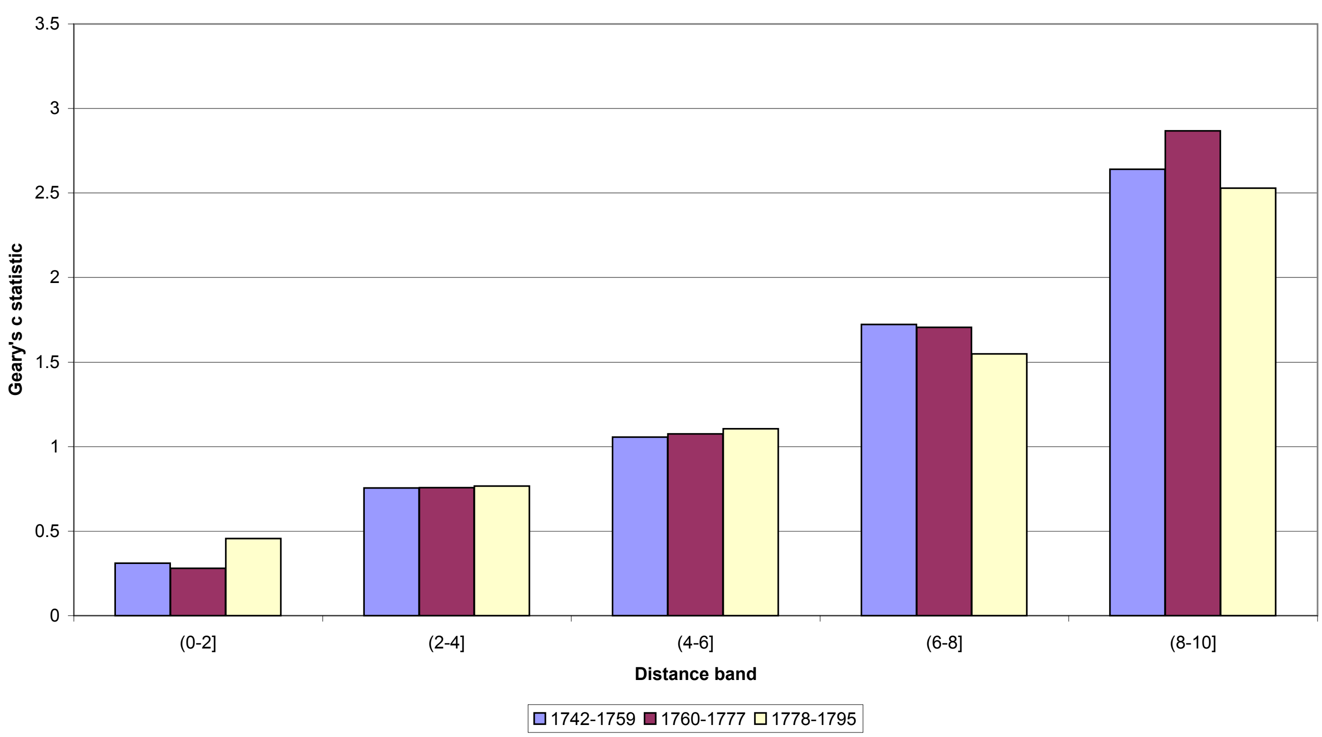


Figure 4b. Changes in Autocorrelation over Time among Yangzi River Prefectures



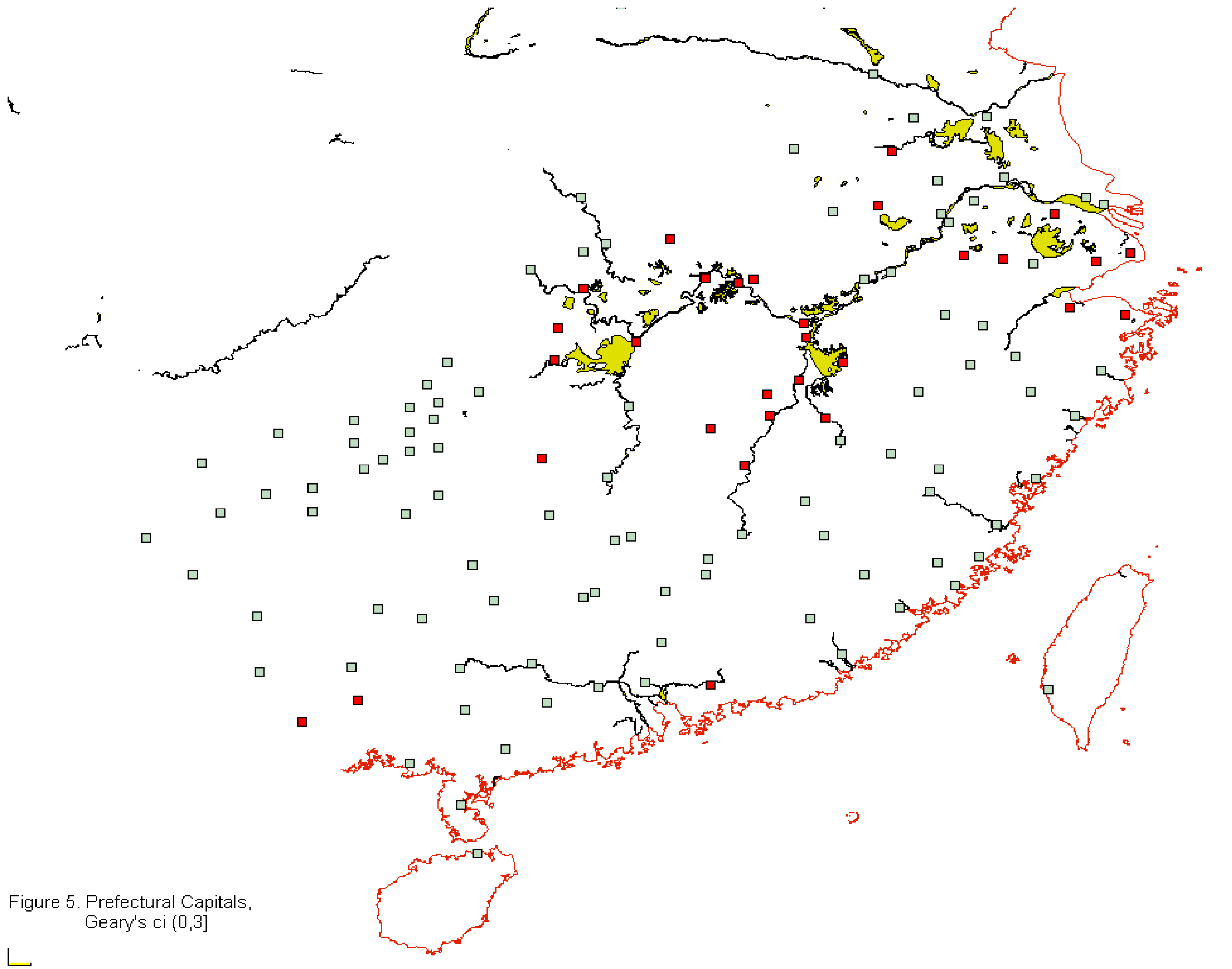


Figure 5. Prefectural Capitals,
Geary's ci (0,3]

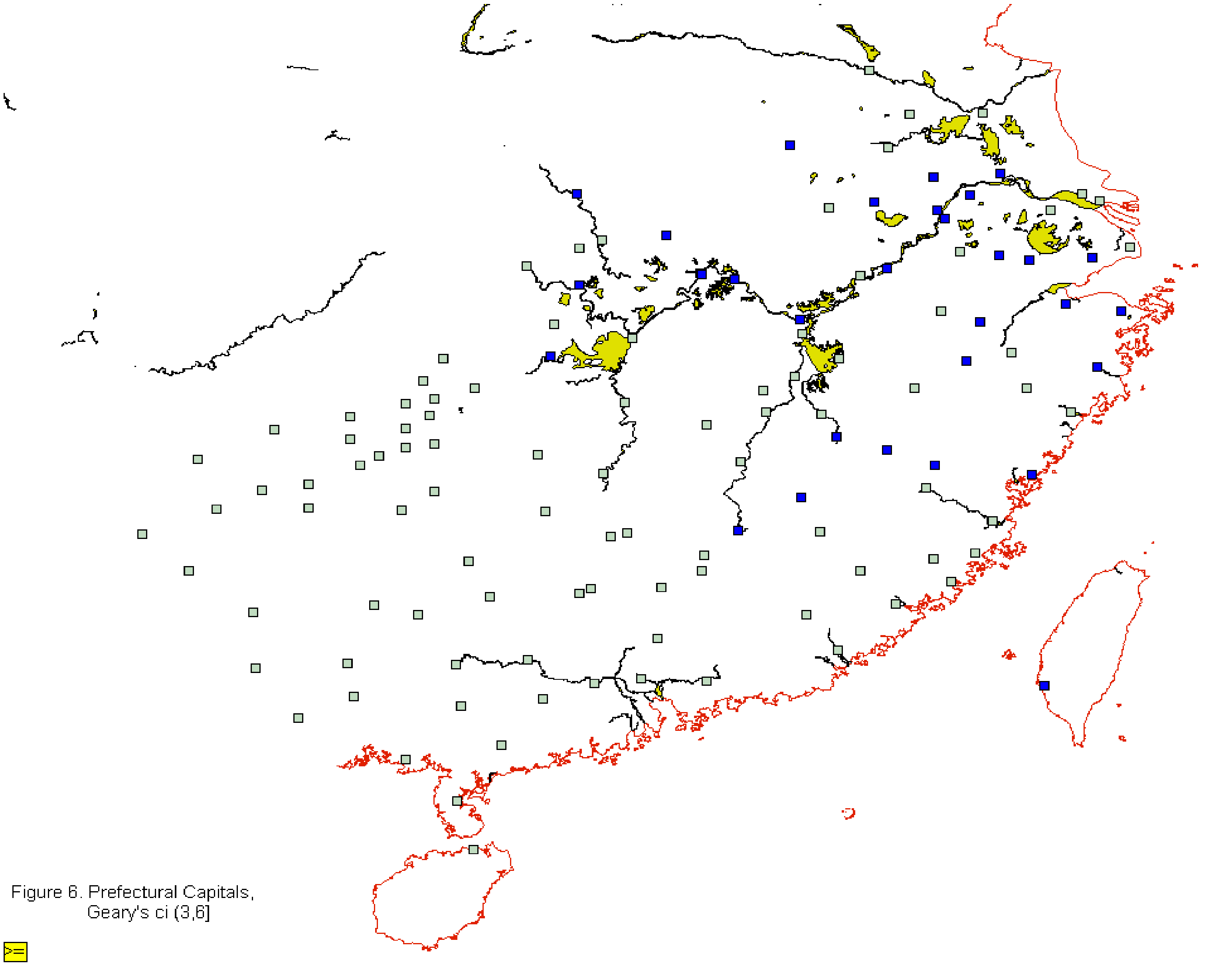


Figure 6. Prefectural Capitals,
Geary's ci (3,6]

>=

Table 1

Geographic Location: Longitudes and Latitudes

Province	Number of Prefectures	Coordinates	Mean	Standard Deviation
Anhwei	13	longitude	117.9	1.0
		latitude	31.6	1.0
Fujian	12	longitude	118.4	1.3
		latitude	25.8	1.1
Guangdong	13	longitude	112.4	2.4
		latitude	23.0	1.6
Guangxi	12	longitude	108.7	1.6
		latitude	23.5	1.0
Guizhou	13	longitude	107.5	1.2
		latitude	26.8	0.8
Hubei	10	longitude	112.6	1.8
		latitude	31.2	0.9
Hunan	13	longitude	111.4	1.3
		latitude	27.6	1.4
Jiangsu	10	longitude	120.0	0.9
		latitude	32.2	1.0
Jiangxi	14	longitude	115.9	1.2
		latitude	27.9	1.2
Zhejiang	11	longitude	120.3	0.7
		latitude	29.4	1.0

Table 2

Price Data by Province*

Province	Price Variable	Mean	Standard Deviation	Min	Max	N
Anhwei	Middle	1.553	0.277	0.950	2.890	1404
	Lowest	1.408	0.252	0.790	2.600	1404
	Highest	1.698	0.335	1.050	3.400	1404
Fujian	Middle	1.667	0.262	0.945	3.250	1296
	Lowest	1.507	0.280	0.730	3.200	1296
	Highest	1.827	0.279	1.100	3.300	1296
Guangdong	Middle	1.498	0.284	0.830	2.515	1404
	Lowest	1.315	0.278	0.650	2.240	1404
	Highest	1.681	0.326	0.920	2.840	1404
Guangxi	Middle	1.117	0.162	0.665	1.705	1296
	Lowest	0.999	0.157	0.590	1.600	1296
	Highest	1.234	0.194	0.740	2.000	1296
Guizhou	Middle	0.972	0.172	0.600	1.555	1404
	Lowest	0.835	0.172	0.420	1.450	1404
	Highest	1.109	0.213	0.670	2.000	1404
Hubei	Middle	1.330	0.235	0.780	2.800	1080
	Lowest	1.156	0.203	0.660	2.100	1080
	Highest	1.504	0.297	0.900	3.600	1080
Hunan	Middle	1.190	0.156	0.855	2.095	1404
	Lowest	1.090	0.160	0.670	2.030	1404
	Highest	1.290	0.185	0.910	2.660	1404
Jiangsu	Middle	1.794	0.323	1.125	3.250	1080
	Lowest	1.639	0.316	0.950	3.000	1080
	Highest	1.950	0.353	1.200	3.500	1080
Jiangxi	Middle	1.413	0.189	0.945	2.475	1512
	Lowest	1.294	0.184	0.780	2.300	1512
	Highest	1.533	0.218	1.000	2.710	1512
Zhejiang	Middle	1.564	0.202	1.010	2.715	1188
	Lowest	1.414	0.200	0.830	2.350	1188
	Highest	1.713	0.226	1.120	3.100	1188
Total	Middle	1.400	0.334	0.600	3.250	13068
	Lowest	1.257	0.322	0.420	3.200	13068
	Highest	1.543	0.370	0.670	3.600	13068

* Price of middle-quality rice in taels per bushel

Table 3

Weather by province

province	**weather variable***	**mean**	**standard deviation**	**N**
Anhwei	Dryness	2.677	0.940	702
	Wdeviation	0.739	0.664	702
	Bad Weather	0.124	0.330	702
Fujian	Dryness	2.748	1.089	648
	Wdeviation	0.906	0.654	648
	Bad Weather	0.171	0.377	648
Guangdong	Dryness	2.946	1.054	702
	Wdeviation	0.906	0.542	702
	Bad Weather	0.104	0.305	702
Guangxi	Dryness	2.966	0.680	648
	Wdeviation	0.392	0.556	648
	Bad Weather	0.035	0.185	648
Guizhou	Dryness	3.006	0.740	702
	Wdeviation	0.405	0.619	702
	Bad Weather	0.071	0.257	702
Hubei	Dryness	2.850	0.887	540
	Wdeviation	0.557	0.706	540
	Bad Weather	0.126	0.332	540
Hunan	Dryness	2.839	0.805	702
	Wdeviation	0.503	0.649	702
	Bad Weather	0.085	0.280	702
Jiangsu	Dryness	2.865	0.904	540
	Wdeviation	0.639	0.654	540
	Bad Weather	0.098	0.298	540
Jiangxi	Dryness	2.696	0.853	756
	Wdeviation	0.640	0.641	756
	Bad Weather	0.090	0.286	756
Zhejiang	Dryness	2.842	0.910	594
	Wdeviation	0.623	0.681	594
	Bad Weather	0.114	0.319	594
Total	Dryness	2.841	0.900	6534
	Wdeviation	0.633	0.659	6534
	Bad Weather	0.101	0.302	6534

* Weather is recorded as a discrete variable taking the values 1 (floods; least dry), 2,3 (normal),4, and 5 (droughts)
This is the variable Dryness; Wdeviation = |Dryness-3|, the deviation from normal (good) weather
Bad Weather indicates severe floods or droughts; it is 1 for Dryness=1 or Dryness=5, zero otherwise

Table 4

Prices over time
Mid-price

Province	Statistic	(i) 1742/59 month 2	(i) 1742/59 month 8	(ii) 1760/77 month 2	(ii) 1760/77 month 8	(iii) 1778/95 month 2	(iii) 1778/95 month 8
Anhwei	mean	1.472	1.405	1.500	1.500	1.717	1.724
	sd *	0.269	0.236	0.214	0.208	0.269	0.285
	sd/mean	0.183	0.168	0.143	0.139	0.157	0.165
Fujian	mean	1.577	1.564	1.636	1.660	1.764	1.801
	sd *	0.289	0.257	0.181	0.170	0.285	0.272
	sd/mean	0.183	0.164	0.111	0.102	0.162	0.151
Guangdong	mean	1.455	1.412	1.490	1.503	1.566	1.564
	sd *	0.321	0.277	0.285	0.283	0.258	0.246
	sd/mean	0.221	0.196	0.191	0.188	0.165	0.157
Guangxi	mean	1.134	1.138	1.021	1.039	1.171	1.197
	sd *	0.209	0.176	0.099	0.096	0.139	0.140
	sd/mean	0.185	0.155	0.097	0.093	0.119	0.117
Guizhou	mean	0.983	1.016	1.012	1.010	0.908	0.903
	sd *	0.179	0.174	0.173	0.177	0.142	0.142
	sd/mean	0.182	0.171	0.171	0.175	0.156	0.158
Hubei	mean	1.210	1.204	1.296	1.296	1.483	1.492
	sd *	0.183	0.160	0.136	0.122	0.286	0.276
	sd/mean	0.151	0.133	0.105	0.094	0.193	0.185
Hunan	mean	1.156	1.151	1.139	1.147	1.270	1.276
	sd *	0.150	0.126	0.102	0.103	0.197	0.168
	sd/mean	0.130	0.109	0.090	0.090	0.155	0.132
Jiangsu	mean	1.693	1.671	1.778	1.785	1.915	1.924
	sd *	0.338	0.292	0.252	0.241	0.346	0.359
	sd/mean	0.200	0.175	0.142	0.135	0.181	0.187
Jiangxi	mean	1.332	1.276	1.421	1.382	1.551	1.517
	sd *	0.212	0.150	0.129	0.111	0.184	0.170
	sd/mean	0.159	0.118	0.091	0.080	0.119	0.112
Zhejiang	mean	1.526	1.468	1.600	1.588	1.595	1.607
	sd *	0.257	0.162	0.195	0.174	0.184	0.190
	sd/mean	0.169	0.110	0.122	0.110	0.115	0.118
Total	mean	1.346	1.322	1.380	1.381	1.483	1.489
	sd *	0.325	0.283	0.309	0.304	0.369	0.369
	sd/mean	0.242	0.214	0.224	0.221	0.249	0.248

* sd = standard deviation

Table 5

Weather Trends

		(i) 1742/59		(ii) 1760/77		(iii) 1778/95	
Province	**Statistic**	**Dryness**	**Bad Weather**	**Dryness**	**Bad Weather**	**Dryness**	**Bad Weather**
Anhwei	mean	2.470	0.162	2.739	0.107	2.821	0.103
	sd *	0.941	0.370	0.943	0.310	0.904	0.304
Fujian	mean	3.000	0.218	2.593	0.093	2.653	0.204
	sd *	1.189	0.414	0.858	0.291	1.151	0.404
Guangdong	mean	3.043	0.090	2.829	0.107	2.966	0.115
	sd *	1.072	0.286	1.009	0.310	1.076	0.320
Guangxi	mean	3.111	0.019	2.787	0.065	3.000	0.023
	sd *	0.592	0.135	0.760	0.247	0.640	0.151
Guizhou	mean	3.000	0.047	3.124	0.068	2.893	0.098
	sd *	0.718	0.212	0.679	0.253	0.803	0.298
Hubei	mean	2.761	0.094	2.872	0.089	2.917	0.194
	sd *	0.794	0.293	0.770	0.285	1.067	0.397
Hunan	mean	2.812	0.090	2.825	0.047	2.880	0.120
	sd *	0.791	0.286	0.655	0.212	0.946	0.325
Jiangsu	mean	2.767	0.039	2.944	0.172	2.883	0.083
	sd *	0.799	0.194	1.018	0.379	0.880	0.277
Jiangxi	mean	2.675	0.111	2.615	0.071	2.798	0.087
	sd *	0.926	0.315	0.817	0.258	0.805	0.283
Zhejiang	mean	3.045	0.131	2.657	0.121	2.823	0.091
	sd *	0.994	0.339	0.839	0.327	0.851	0.288
Total	mean	2.866	0.101	2.796	0.092	2.862	0.111
	sd *	0.920	0.301	0.853	0.289	0.924	0.314

* sd = standard deviation

Table 6
Geary's c statistics by regional groups and sub-periods, mean (standard error of mean)

distance	(i) 1742-1795	(ii) 1742-1759	(iii) 1760-1777	(iv) 1778-1795
(0-2], full sample	0.289 (0.012)	0.368 (0.022)	0.263 (0.013)	0.237 (0.012)
Yangzi River	0.349 (0.020)	0.311 (0.021)	0.281 (0.026)	0.456 (0.040)
Coastal	0.552 (0.018)	0.585 (0.026)	0.481 (0.028)	0.591 (0.032)
Inland	0.396 (0.022)	0.568 (0.030)	0.370 (0.021)	0.250 (0.010)
(2-4], full sample	0.463 (0.017)	0.565 (0.031)	0.416 (0.021)	0.408 (0.018)
Yangzi River	0.760 (0.027)	0.756 (0.050)	0.757 (0.041)	0.767 (0.050)
Coastal	0.873 (0.021)	0.917 (0.039)	0.852 (0.042)	0.849 (0.029)
Inland	0.620 (0.026)	0.820 (0.041)	0.526 (0.027)	0.513 (0.020)
(4-6], full sample	0.641 (0.017)	0.714 (0.030)	0.596 (0.029)	0.613 (0.020)
Yangzi River	1.079 (0.026)	1.056 (0.054)	1.075 (0.033)	1.106 (0.048)
Coastal	0.971 (0.020)	0.956 (0.036)	0.880 (0.030)	1.076 (0.024)
Inland	0.796 (0.015)	0.876 (0.027)	0.724 (0.018)	0.787 (0.018)
(6-8], full sample	0.864 (0.012)	0.907 (0.022)	0.852 (0.024)	0.835 (0.010)
Yangzi River	1.658 (0.046)	1.722 (0.067)	1.705 (0.078)	1.548 (0.084)
Coastal	1.113 (0.035)	1.090 (0.073)	1.110 (0.060)	1.140 (0.046)
Inland	1.177 (0.016)	1.044 (0.018)	1.228 (0.020)	1.260 (0.013)
(8-10], full sample	1.161 (0.012)	1.092 (0.020)	1.201 (0.017)	1.189 (0.013)
Yangzi River	2.679 (0.087)	2.641 (0.120)	2.867 (0.125)	2.529 (0.192)
Coastal	1.120 (0.031)	1.110 (0.051)	1.215 (0.062)	1.036 (0.038)
Inland	1.670 (0.036)	1.414 (0.059)	1.794 (0.042)	1.801 (0.035)
(10-12], full sample	1.602 (0.021)	1.476 (0.035)	1.639 (0.028)	1.692 (0.023)
Yangzi River	N/A	N/A	N/A	N/A
Coastal	1.324 (0.054)	1.286 (0.107)	1.448 (0.105)	1.239 (0.063)
Inland	2.041 (0.055)	1.717 (0.102)	2.240 (0.069)	2.166 (0.064)
(12-14], full sample	2.062 (0.040)	1.890 (0.070)	2.051 (0.073)	2.245 (0.036)
Yangzi River	N/A	N/A	N/A	N/A
Coastal	1.557 (0.090)	1.623 (0.167)	1.622 (0.195)	1.427 (0.094)
Inland	2.717 (0.102)	2.338 (0.230)	3.032 (0.130)	2.782 (0.112)
(14-16], full sample	2.472 (0.059)	2.382 (0.117)	2.540 (0.119)	2.494 (0.059)
Yangzi River	N/A	N/A	N/A	N/A
Coastal	1.898 (0.107)	1.823 (0.172)	2.102 (0.235)	1.770 (0.136)
Inland	N/A	N/A	N/A	N/A

Table 7

All 121 prefectures, 1742-95
Dependent variable: log mid price, month 8

	OLS	Huber-White standard errors	Spatial corr. adjusted s.e.'s (0,3]	Spatial corr. adjusted s.e.'s (0,6]
Constant	0.264 **	0.264 **	0.264 **	0.264 **
s.e.	0.027	0.028	0.055	0.074
Weather	0.080	0.080	0.080 *	0.080 *
s.e.	0.031	0.030	0.048	0.053
lnL	19.463	19.463	19.463	19.463

	Spatial error (0,3]	Spatial lag (0,3]	Spatial error (0,6]	Spatial lag (0,3]	Spatial error (exp)	Spatial lag (exp)
Constant	0.300 *	0.016	0.385	-0.002	0.318 **	0.022
s.e.#	0.157	0.022	0.373	0.023	0.121	0.020
Weather	0.020	0.024	0.036	0.043 **	0.017	0.019
s.e.#	0.022	0.018	0.022	0.021	0.022	0.017
lambda	0.915 **		0.954 **		0.906	
s.e.#	0.049		0.043		0.045	
rho		0.900 **		0.945 **		0.892 **
s.e.#		0.053		0.048		0.047
LM test	320.192	328.131	495.353	508.979	239.071	249.600
p-val	0.000	0.000	0.000	0.000	0.000	0.000
lnL	75.936	76.469	59.425	60.536	85.152	85.630

* (**) significant at 10 (5)% level
\# robust standard errors

Table 8: Inland versus Coastal prefectures, 1742-95

Dependent variable: log mid price, month 8
Inland sample: N = 62, coastal sample: N = 59

	OLS		Huber-White standard errors		Spatial corr. adjusted s.e.'s (0,3]		Spatial corr. adjusted s.e.'s (0,6]	
	Inland	**Coastal**	**Inland**	**Coastal**	**Inland**	**Coastal**	**Inland**	**Coastal**
Constant	0.146 **	0.448 **	0.146 **	0.448 **	0.146 **	0.448 **	0.146 **	0.448 **
s.e.	0.029	0.031	0.028	0.029	0.046	0.038	0.056	0.038
Weather	0.044	0.009	0.044	0.009	0.044	0.009	0.044	0.009
s.e.	0.043	0.032	0.035	0.030	0.048	0.036	0.050	0.036
lnL	24.638	31.136	24.638	31.136	24.638	31.136	24.638	31.136

	Spatial error (0,3]		Spatial lag (0,3]		Spatial error (exp)		Spatial lag (exp)	
	Inland	**Coastal**	**Inland**	**Coastal**	**Inland**	**Coastal**	**Inland**	**Coastal**
Constant	0.145	0.441 **	0.020	0.160 **	0.176 *	0.440 **	0.028	0.117 **
s.e.#	0.135	0.060	0.023	0.077	0.095	0.066	0.021	0.058
Weather	0.008	0.007	0.012	0.006	0.005	0.005	0.010	0.004
s.e.#	0.028	0.030	0.024	0.026	0.032	0.029	0.024	0.024
lambda	0.843 **	0.637 **			0.819 **	0.735 **		
s.e.#	0.081	0.163			0.077	0.117		
rho			0.828 **	0.626 **			0.804 **	0.724 **
s.e.#			0.083	0.158			0.078	0.116
LM test	92.849	32.336	98.487	33.395	73.122	35.821	78.058	37.599
p-val	0.006	0.030	0.005	0.023	0.001	0.002	0.001	0.001
lnL	46.815	38.823	47.083	38.799	49.076	43.869	49.272	43.948

* (**) significant at 10 (5)% level
# robust standard errors

Table 9

Yangzi River versus Non-Yangzi prefectures
Dependent variable: log mid price, month 8
Yangzi sample: N=21, NonYangzi: N=100

	OLS		Huber-White standard errors		Spatial corr adjusted s.e.'s (0,3]		Spatial corr adjusted s.e.'s (0,6]	
	Yangzi	**NonYangzi**	**Yangzi**	**NonYangzi**	**Yangzi**	**NonYangzi**	**Yangzi**	**NonYangzi**
Constant	0.387 **	0.238 **	0.387 **	0.238 **	0.387 **	0.238 **	0.387 **	0.238 **
s.e.	0.051	0.030	0.047	0.031	0.060	0.055	0.062	0.074
Weather	0.020	0.091 **	0.020	0.091 **	0.020	0.091 *	0.020	0.091 *
s.e.	0.057	0.035	0.050	0.034	0.051	0.051	0.047	0.055
lnL	12.663	15.107	12.663	15.107	12.663	15.107	12.663	15.107

	Spatial error (0,3]		Spatial lag (0,3]		Spatial error (exp)		Spatial lag (exp)	
	Yangzi	**NonYangzi**	**Yangzi**	**NonYangzi**	**Yangzi**	**NonYangzi**	**Yangzi**	**NonYangzi**
Constant	0.400 **	0.289 **	0.112	0.020	0.381 **	0.292 **	0.096 *	0.027
s.e.#	0.114	0.142	0.078	0.024	0.110	0.107	0.059	0.022
Weather	0.011	0.023	0.011	0.029	0.013	0.021	0.011	0.024
s.e.#	0.044	0.025	0.039	0.021	0.040	0.025	0.033	0.020
lambda	0.693 **	0.892 **			0.767 **	0.878 **		
s.e.#	0.236	0.056			0.138	0.051		
rho			0.700 **	0.872 **			0.740 **	0.860 **
s.e.#			0.195	0.060			0.138	0.053
LM test	11.928	195.883	14.294	207.111	14.158	155.151	16.521	165.726
p-val	0.117	0.000	0.058	0.000	0.054	0.000	0.023	0.000
lnL	17.374	58.512	17.674	59.137	20.407	65.059	20.587	65.648

* (**) significant at 10 (5)% level
\# robust standard errors

Table 10: Regression Results by Subperiod

Spatial Lag Dependence Model*
Dependent variable: log mid price, month 8

	All Prefectures			Inland Prefectures		
	1742-59	**1760-77**	**1778-95**	**1742-59**	**1760-77**	**1778-95**
Constant	0.027	0.019	0.021	0.037 *	0.025	0.022
s.e.#	0.020	0.019	0.021	0.023	0.020	0.021
Weather	0.020	0.019	0.019	0.010	0.012	0.007
s.e.#	0.017	0.016	0.017	0.022	0.024	0.025
rho	0.856 **	0.904 **	0.917 **	0.710 **	0.810 **	0.892 **
s.e.#	0.057	0.043	0.041	0.107	0.079	0.049
LM test	187.195	278.514	283.092	41.436	80.633	112.105
p-val	0.000	0.000	0.000	0.003	0.000	0.000
lnL	84.476	89.903	82.512	46.293	54.181	47.340
N	121	121	121	62	62	62

	Coastal Prefectures			Yangzi River Prefectures		
	1742-59	**1760-77**	**1778-95**	**1742-59**	**1760-77**	**1778-95**
Constant	0.111 **	0.096 *	0.142 **	0.064 *	0.065	0.158 *
s.e.#	0.054	0.053	0.067	0.043	0.050	0.085
Weather	0.003	0.007	0.001	0.000	0.012	0.022
s.e.#	0.025	0.021	0.025	0.031	0.028	0.038
rho	0.691 **	0.770 **	0.711 **	0.791 **	0.800 **	0.630 **
s.e.#	0.121	0.103	0.123	0.121	0.120	0.174
LM test	33.447	48.940	30.409	19.460	18.749	11.355
p-val	0.002	0.000	0.000	0.000	0.001	0.068
lnL	42.385	49.264	40.195	20.501	22.030	19.228
N	59	59	59	21	21	21

* Spatial weights: exponential with parameter theta = -1.4
# s.e. : robust standard errors
* (**) significant at 10 (5)% level