The Origins of Spatial Interaction

Evidence from Chinese Rice Markets, 1742-1795*

Wolfgang Keller† Carol H. Shiue‡

University of Texas University of Texas

September 2004

Abstract

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.

Keywords: Interregional trade, geography, spatial econometrics, trade costs, Geographic Information System coding, spatial autocorrelation, China, river transport, coastal transport

* We thank seminar participants at Penn State University, Texas A&M, and the NBER International Trade and Investment Fall 2003 meeting at Stanford University, as well as Luc Anselin, Richard Baldwin, Badi Baltagi, Tim Conley, Daniel Hamermesh, Gordon Hanson, and Preston McAfee for suggestions. We have also benefited from the comments of two anonymous referees. All errors are our own.

†Department of Economics, BRB 3.152, Austin, TX 78712; email: keller@eco.utexas.edu. Also affiliated with the National Bureau of Economic Research and the Centre for Economic Policy Research.

‡Department of Economics, BRB 2.126, Austin, TX 78712; email: 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, see Anderson and van Wincoop 2004, 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 will the influence of geography 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. 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 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 2002, 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 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.

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 determined by 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). 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 to consider the role of spatial dependence when drawing inferences about price formation and trade in geographic space.

The question is fundamental to related work on the extent to which geography affected trade costs, and thus trading possibilities, and whether these differences could trigger different development paths across regions. For example, trade might facilitate technological innovations through learning, which could give some regions earlier access to new transport technologies than other regions.[4] In addition, trade may be associated with institutional innovations and information sharing within a network of traders.[5]

The remainder of the paper is as follows. Section 2 describes the characteristics and sources of

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] Along these lines, Mokyr (1990, 134) emphasizes the importance of trade primarily insofar as it contributes to the diffusion of technological knowledge; Keller (2004) discusses some of the more recent evidence.

[5] North and Thomas (1973, 12) note that trade goes hand-in-hand with institutional innovation, while Williamson (2000, 599-600) emphasizes that ex-ante mutually beneficial trade often shapes ex-post institutional outcomes. Information sharing among traders may also act as a monitoring device in the absence of legally enforceable contracts (Greif 1989).

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 presented in section 4, and section 5 provides the concluding discussion. Additional background on the data is given in the appendices.

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, Fujian, 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 that we can identify on historical maps. There are between ten to fourteen prefectures in each province, and on the next-lower administrative level, about eighty counties; a limited number of independent administrative units are not included in the sample. In comparison with the contemporaneous United States of America, the province is most closely analogous to the state (the average U.S. state has about sixty-five counties). Figure 2 gives a map of the prefectures and the boundaries of the ten provinces that are the focus of the analysis.

This paper uses weather and price data for 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). 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 accounting unit of taels (silver currency) per bushel. It is generally believed that the grain prices closely correspond to market prices (Chuan and Kraus 1975).

The county level price reports were sent to the prefectural level, where the highest and the lowest price observed in a particular lunar month were recorded and compared with the highest and lowest observed in the previous month. These reports are provided in the *Gongzhong liangjiadan* [Grain Price Lists in the Palace Archives of the Number One Historical Archives in Beijing]. The series we use consists of the 2nd and 8th months from 1742-1795.[6] The recorded price is for mid-quality rice in all prefectures; there is no indication that there are in fact substantial quality differences across prefectures.

Historical weather data is from the State Meteorological Society (1981), which gives information on weather for each year for locations throughout China. The sample we construct is created by pinpointing the location of prefectures on the weather maps in this source. The variable denoted *dryness* is a discrete indicator of the degree 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 the data on *dryness*, we have also constructed two additional weather variables, *wdev*, which measures the deviation from good weather ($wdev = |dryness - 3|$), and *bad_weather* ($dryness = 1$ or 5). Appendix B provides additional details on the weather data.

[6]Because of missing values, we estimate parts of the data; see Appendix A.

2.2 Summary statistics

The location of the prefectures can be described using Geographic Information System (GIS) data, which we employ 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.

Table 2 presents summary statistics on the mid-, lowest and highest prefectural price by province.[7] 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 on average from a minimum of 0.722 to a maximum of 2.306 taels across the 54 years, and a standard deviation of 0.310.[8] 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. On average, prices are lowest at $dryness = 3$ (i.e., normal conditions). The variable *bad_weather* indicates that the prefectures 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. Inland

[7] The souce contains high and low prices. The mid-price is constructed. We also report results based on the low and high prices instead of mid-prices below. See Appendix A for additional details.

[8] This, as well as the other statistics presented in Tables 2 to 6, is based on the assumption of independence. Note, for example, that this will underestimate the true variability in the presence of positive spatial or temporal autocorrelation. We estimate the extent of spatial as well as temporal autocorrelation in section 4.

areas tend towards relatively low weather variability. The table also indicates that for the sample as a whole exceptional drought and flood occurs in about 10% of all years (last row).

Tables 4 and 5 provide the major temporal trends by reporting averages of the annual price and weather statistics over three 18-year periods, namely, 1742/59, 1760/77, and 1778/95. Over the entire period, Table 4 shows that prices in the entire sample rose from about 1.33 to 1.48, or an average of 0.2% per year, and the rate of price increase is somewhat higher in the later years. 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 loss of generality.

As a measure of price dispersion, Table 4 reports the coefficient of variation across all prefectures, as well as across prefectures within a given province. Over time, the change in price dispersion within a given province is mixed, and for the sample area as a whole, the price dispersion increases over the sample period (from 0.198 to 0.238 for the 8th month prices; last row). This suggests that for the sample as a whole, there is price divergence, whereas within individual regions prices may be converging.[9] To the extent that these price trends are due to interregional grain trade, its geographic scope seems to be limited: integration increases within certain provinces, but the force of arbitrage does not appear to be sufficiently strong to 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

[9] Another test for convergence is the following simple regression: $\Delta p_i = \beta_0 + \beta_1 p_i^0 + \varepsilon_i$, where Δp_i is the growth of the price in prefecture i between 1742 and 1795, and p_i^0 the log initial price, that is, in year 1742. A negative estimate of β_1 means that prices tend to grow slower in prefectures where they are relatively high to begin with. This is consistent with convergence. Here, we estimate β_1 at 0.114, insignificantly different from zero at standard levels, which also suggests that there is no general tendency of price convergence. The Huber (1967)-White (1980) robust standard error of β_1 is equal to 0.117. If the growth rate of the weather variable *wdev* is added to the regression, the coefficient on weather is 0.066 (s.e. 0.024), while the parameter on initial price (β_1) remains insignificant; these results use the 8th month price, with N = 121 prefectures.

of *bad_weather*, e.g., does not vary much overall (from 10.1 over 9.2 to 11.1 percent; at the bottom).

3 Patterns of spatial autocorrelation

In this section, we examine the spatial dimension with a number of spatial autocorrelation measures. Among the most widely used measures are Moran's I and Geary's c statistics (Moran 1950, and Geary 1954, respectively). These statistics provide evidence on the sample as a whole. We will also employ local measures of spatial autocorrelation, which give a spatial association coefficient for a particular locality $i,\ i = 1, ..., N$. The local measures have been introduced by Anselin (1995).

Let $\{w_{ij}\}$ be a connection matrix in which $w_{ij} = 1$ if the ith and jth prefecture are spatially connected, and $w_{ij} = 0$ otherwise. Spatial connectedness is defined in terms of the prefectures sharing a boundary, for example, or by some threshold bilateral geographic distance. Different degrees of connectedness are modeled by allowing for multiple connection matrices, one for each spatial lag, $k = 1, ..., K$, where $\{w_{ij}^{(k)}\}$ are the corresponding connection matrices.

For a given year, $t = 1742, ..., 1795$, Moran's I statistic for spatial lag k is defined as

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

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

$$c_k = \frac{(N-1) \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}^{(k)} \left(z_i - z_j\right)^2}{4J_k \sum_{i=1}^{N} z_i^2}, i \neq j. \tag{2}$$

Both the Moran and the Geary statistic measure the covariance of prices in connecting prefectures relative to the variance of the price across prefectures.[10] 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 $E[c_k] = 1$. It is apparent from the definitions of the statistics that if prices in connected prefectures tend to be relatively similar (dissimilar)—i.e., positive (negative) spatial autocorrelation—then Moran's I_k will tend to be greater (smaller) than zero, and Geary's c_k will tend to be below (above) one. One advantage of the Geary's statistic is that the expected value of Geary's c_k under the null hypothesis is independent of the sample size N, while the expected value of Moran's I_k is not. The variances of I_k and c_k are shown in Appendix C. Inference is based on the result that under the null hypothesis, the I_k and c_k statistics are distributed asymptotically normal.

The corresponding local measures compute spatial autocorrelation for each location i. In particular, the local Geary statistic is given by

$$c_i^k = \frac{N \sum_{j=1}^{N} w_{ij}^{(k)} (z_i - z_j)^2}{\sum_{q=1}^{N} z_q^2}, i \neq j. \tag{3}$$

Compared to the global Geary statistic, the major difference is that the local measure is based on a single summation, whereas Geary's global coefficient sums over both i and j (compare equation 3 with equation 2).

These spatial autocorrelation statistics require normally distributed variates with constant variance. To assess whether our data satisfies this, we have, first, employed the Shapiro and Wilk (1965) test for normality. At a 5% significance level, one cannot reject the null hypothesis that the mid-

[10] It is useful to think of the counterparts of both statistics in terms of temporal autocorrelation. Moran's I test was originally developed as a two-dimensional analog of the test of significance of the serial correlation coefficient in univariate time series; in the case of one dimension, it reduces to the familiar serial correlation coefficient. Geary's c is related to Durbin and Watson's d statistic (Durbin and Watson 1950, 1951); see Anselin and Bera (1998, 265-268), and Cliff and Ord (1981, 13-14), for more discussion.

price variable is normally distributed.[11] Regarding the assumption of constant variance, a major concern is that different prefecture sizes lead to heteroskedasticity. Although there are a number of spatial autocorrelation measures that adjust for such differences, for instance Oden's (1995) Moran I adjusted for population density, they tend to require data that we do not have in this case (including prefecture-level population data, as well as an equivalent to disease rates). To see how far off the assumption of constant variance might be in our data, we have compared the standard deviation of the price for several groups of relatively small versus relatively large prefectures. At a 5% level, statistical tests typically reject that the variances are the same, while at a 1% level, the assumption of constant variance in the relatively small and relatively large prefectures is not rejected.[12] This suggests that while prefectural size may be a source of heteroskedasticity, its influence should be limited.[13]

3.1 Local Spatial Autocorrelation

As a first cut at assessing spatial autocorrelation in our sample, we compute the local Geary statistic for the mid-price for two separate connection matrices $\{w_{ij}^{(k)}\}$. In the first matrix, for any prefecture i, $w_{ij}^{(1)} = 1$ if the Euclidian distance between prefectures i and j is less than or equal to 300 kilometers, and zero otherwise. Analogously, for the second matrix, $w_{ij}^{(2)} = 1$ if and only if the distance is greater than 300 but no more than 600 kilometers, and the two cases are denoted by (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; see Figure 3. The locations of the waterways were in no way incorporated in the calculation of Geary's c_i.

[11] Similar results are obtained for the the highest and lowest price.

[12] This result holds for the (log) mid-price for the five smallest versus largest prefectures, as well as for the ten smallest versus largest prefectures. Similar results are obtained for the lowest and highest price in a prefecture.

[13] We will return to the issue of heterogeneity across prefectures in section 4.2.

The results of the ranking of the prefectures with greatest local autocorrelation in price (shaded squares) suggest that spatial patterns are determined by the location of transport routes. The strongest clusters are 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.[14] The association between the waterways and price clusters is visible in Figure 3.[15]

These results on the prevalence of spatial price clusters are consistent with historical accounts that place much emphasis on the importance of waterway transport for grain trade (Evans 1984). Transport costs are related to the prevalence of trade because the lower these costs are, the more likely it is that the pre-trade price gap is small enough such that traders can arbitrage across locations, pulling prices together. It therefore appears that an analysis of spatial autocorrelation patterns can provide information on trade patterns.

Figure 4 shows the ranking at a further distance band, (3,6], for each individual locality. Compared 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 for 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 and coastal trade at a longer distance.

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

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

This local analysis suggests that spatial autocorrelation patterns vary systematically with access to water transport as well as other geographic characteristics. The results are consistent with differences in terms of interregional trade—its prevalence, and how far trading networks reached—across different regions in China. This will be taken into account when forming subgroups of prefectures in the next section.

3.2 Global Spatial Autocorrelation

Table 6 shows Geary's c statistic, given in equation (2), for different groups of prefectures and subperiods for differnet spatial lags, k, ranging from distances of 0 to 200 km (denoted by (0,2]) to distances of 1400 to 1600 km (denoted by (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.[16] There is evidence for positive autocorrelation for distances up to 800 km and evidence for negative autocorrelation for distances above 1,000 km.[17]

Positive autocorrelation for shorter distances is a plausible finding, because given that transport costs were increasing in distance, trade will tend to connect markets for relatively short distances before it does so over longer distances. But are there differences across different regions of 18th century China? First, we divide the sample into those prefectures that lie directly on or near to the coast—regions with better access to 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).

[16] In addition to the average of Geary statistics for each of 54 years, the table also shows the standard error of these means in parentheses.

[17] We have obtained similar results using Moran's I statistic. For instance, the correlation between Geary's c_k and Moran's I_k for all 121 prefectures across all distance bands and the three subperiods given in Table 6 is with -0.98 close to -1. Given that, in addition, Geary's c statistic has the advantage relative to Moran's I that the expected value of the former does not depend on N, we present only Geary results in this paper.

Second, the local Geary analysis above suggests that the Yangzi River, China's longest navigable waterway, matters for interregional transport in China. 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 sample consists of 21 prefectures.[18]

Geary's c statistics for these Inland, Coastal, and Yangzi River samples are shown in Table 6, column (i). For each of these groups, the Geary statistic increases with distance. The level of spatial autocorrelation also differs markedly across regional samples. The relatively low Geary for the Yangzi River prefectures at short distances (for (0,2], it is 0.349) confirms what we found in Figure 3 above. The Coastal sample displays less short-distance clustering, but as distance increases, the covariation of prices across markets does not decline quite as rapidly, suggesting a degree of homogeneity in the spatial pattern of markets in this sample that differs from 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 displays relatively small changes in the strength of clustering as distance increases.

Table 6 also reports the average 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 (columns ii-iv). Differences in spatial patterns across regions persist across the duration of the entire sample period. However, there are also some temporal changes, and these occur at different rates for different distances in some of the regions. It is possible that the rate of change of integration at different distances are interelated, as might happen, for example, if increasing long distance trade diverts trade that had been present at short or medium distances.

[18] Yangzi River prefectures are 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.

However, positive spatial autocorrelation can result not only from trade. Weather shocks, for instance, tend to vary less for shorter than they vary for larger distances. Weather, as well as possible spatial trends in prices (leading to spatial non-stationarity), may also explain some of the findings of negative autocorrelation for larger distances in Table 6. We will address these issues in the following section.

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 *wdev*. Under the stated conditions, ordinary least squares is the best linear unbiased estimator, and we report it as a baseline.[19] Given the size differences of the prefectures, we also present Huber (1967) and White (1980) heteroskedasticity-consistent standard errors.

The previous section strongly suggest that the residuals of (4) are spatially dependent. One approach that we take is 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

[19] Reported results are averages from cross-sectional regressions across all years, or within the relevant sub-period, but for convenience we generally suppress t, the subscript for time.

[20] In addition to a Moran or Geary test on the residuals of (4), a number of other tests have been proposed, see Kelejian and Robinson (1992), Anselin, Bera, Florax, and Yoon (1996), Anselin and Bera (1998), and Baltagi and Li (2001). See also Kelejian and Prucha (2001) on the relationship of different tests and further results.

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. Relative to OLS, this method affect only the computation of the standard errors.

The two most influential models that incorporate spatial autocorrelation into the regressive structure 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 $\tilde{w}_{ij}$. These weights correspond to the connection matrix $\{w_{ij}\}$, defined above, in that they capture the spatial structure of the sample. If $\lambda \neq 0$, ignoring the spatial dependence means OLS is inefficient 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 (MLE) of both models below.

These regression techniques require a number of key assumptions to be valid. In particular, the properties of the MLE estimation of the two spatial regression models rely on normally distributed variates with constant variance. In addition, the asymptotics of Conley's (1999) covariance estimator relies on the data generation process being spatially stationary. In the absence of these conditions, in general not very much is known on the properties of these estimators. We will therefore examine whether the assumptions underlying the estimators are satisfied in our case, and how sensitive the results appear to be to departures from those assumptions.

Our spatial weights $\tilde{w}_{ij}$ are based on distance (denoted D_{ij}). We considered a variety of specifications, including one- and two-window distance bands and exponential specifications. For one-window distance bands, $\tilde{w}_{ij} = 1$ if D_{ij} is less than some maximum $D_{\max}$, and $\tilde{w}_{ij} = 0$ otherwise, and different spatial structures are captured by varying $D_{\max}$.[21] An example of a two-window weighting 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 $\tilde{w}_{ij} = \exp(-\theta D_{ij})$, 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.[22] 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 Table 7.

[21] The own weight, $\tilde{w}_{ii}$, is set to zero $\forall i$ in all weighting matrices.

[22] 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.[23] First, the OLS results might suggest that a prefecture's weather has a significant positive effect on the local price.[24] The point estimate of 0.08 (s.e. of 0.03) suggests that moderately bad weather ($wdev = 1$) raises the price by about eight percent, whereas exceptional floods and droughts ($wdev = 2$) 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 row). 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 (see Cressie 1993, 14-15, as well as Anselin and Bera 1998, 247-248). Now one cannot reject the null that local weather is not correlated with 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 $\tilde{w}_{ij} = \exp(-\theta D_{ij})$, with $\theta = 1.4$. First, the spatial models fit much better than the models that omit the spatial structure, with a log-likelihood ranging from about 60 to 85, versus about 20

[23] We have omitted the standard errors of these means in the table in the interest of space.

[24] Among the three weather variables, we focus on weather deviation (*wdev*), the extent to which the weather differed from medium dryness, generally most favorable to harvests. The other two weather variables, *dryness* and *bad_weather*, give similar results.

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 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.

Table 7 also reports results based on the adjusted LM tests developed by Anselin, Bera, Florax, and Yoon (1996). These authors develop tests of the $H_0 : \lambda = 0$ in the presence of the nuisance parameter ρ, and conversely, of $H_0 : \rho = 0$ without restrictions on λ. Clearly, in our case there appears to be evidence for both λ and ρ to be different from zero, so a test that does not rely on one parameter to equal zero when testing the H_0 that the other is zero is preferred. Moreover, these results may also help to establish which is the better model. These adjusted LM test statistics in Table 7 range from 3.12 to 19.66 for the spatial error model, and from 13.85 to 33.66 for the spatial lag model. These values are substantially below those of the corresponding standard LM test discussed above. The adjusted LM statistic is also χ^2 distributed with one degree of freedom, with a 1% (10%) critical value of 6.64 (2.71). This means that in five out of six specifications shown in Table 7, one can reject the null of no spatial dependence at a 1% level, and for the sixth specification, the null can be rejected at a 10% level.

How does the presence of spatial dependence change our inferences on the influence of weather on the local price? Table 7 indicates that the evidence for a strong effect of weather shocks on local prices is now further weakened. The point estimate for the weather variable lies 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 specifications in terms of likelihood, one sees that the exponential weighting matrix performs better than the other two. This is a robust finding, and we will in the following show only results for the model with exponential spatial weights ($\tilde{w}_{ij} = \exp(-\theta D_{ij})$, with

$\theta = 1.4$). In addition, the spatial lag model has always a somewhat better fit than the corresponding error model. Consistent with that, the evidence for spatial dependence as measured by the adjusted LM test statistic is stronger for the spatial lag than for the spatial 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). For the preferred weighting matrix ($\tilde{w}_{ij} = \exp(-\theta D_{ij})$, with $\theta = 1.4$), there is evidence for dependence at the 5% significance level using the spatial lag model, but not when using the spatial error model. This 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 price linkages and there being no other major reason that underlies the spatial autocorrelation in the data.

4.2 Specification and Robustness

Table 8 presents additional results that allow us to discuss a number of issues. First of all, there is the question of omitted variables. By focusing on contemporaneous spatial dependence, our analysis thus far has ignored the effects that might arise due to temporal dependence, and specifically, the previous periods' prices. For instance, in the presence of interregional trade, it may matter substantially for the current price in location i whether the prices in the markets that neighbor location i have been relatively high or low in the previous period. Moreover, location i's own price in the previous period may also have an influence if there is serial correlation. If that were the case, the omission of these variables from our specification might lead to an overestimate of the degree of spatial dependence as measured by estimates of λ and ρ. Therefore, we extend our analysis to include the spatially lagged previous-period prices, denoted by $Wy(-1)$, as well as the prefectures' own previous period prices, denoted by $y(-1)$.[25]

The results are shown in Table 8. Specification (1) repeats the preferred specification of Table 7 for convenience (note that 53 years of data, rather than 54, are used in this specification). In the second column, the spatial lag of the previous period prices has been added to the specification. This variable enters with a coefficient of 0.547, significant at standard levels, and its inclusion raises the log-likelihood from about 86 to 94. If instead the lagged own price, $y(-1)$, is included, this also comes in with a positive coefficient, but the improvement in fit is larger (see (3)). The inclusion of either variable leads to a reduction in the estimated value of ρ, from about 0.90 to about 0.40, suggesting there is less evidence for spatial dependence. At the same time, as measured by the adjusted LM test statistic, there remains evidence for spatial dependence (p-value < 0.02), and the effect from weather is still insignificant.

When both $Wy(-1)$ and $y(-1)$ are included jointly, the coefficient on the spatially-lagged previous year price turns negative. In contrast, the coefficient on the temporally lagged own price remains positive, in fact it goes up somewhat, which may be a sign of collinearity between $Wy(-1)$ and $y(-1)$. Thus, there is evidence for serial correlation of prices, whereas the relation of price with spatially-lagged previous prices appears to be mixed.[26] Overall, the inclusion of temporally lagged price variables leaves the qualitiative finding unchanged: ρ is estimated to be about 0.64, with an adjusted LM test statistic of about 6.2 (p-value around 0.025). However, it is clear that without the $Wy(-1)$ and $y(-1)$ variables, the evidence for spatial dependence would have been overestimated.

A second issue is the assumption of normality. In particular, this assumption could be violated because the weather variable *wdev* is discrete and takes on only three values (0, 1, and 2). In specifications (5) to (7) we experiment with the variable *dryness* and dryness squared (denoted

[25] The spatial weights matrix W is the same that is applied to the contemporaneous prices (exponential in distance with $\theta = 1.4$).

[26] In the corresponding spatial error model, the coefficient on $Wy(-1)$ tends to be positive but insignificant when $y(-1)$ is included.

dryness2) instead of the variable *wdev*, finding that the results are similar.[27] Omitting the weather variable from the specification gives also similar results (compare specifications (6) and (3)). That non-normality due to the discreteness of the weather variable does not critically the results may be due to the fact that weather has only a very weak effect once spatial dependence is accounted for. In addition, a number of further specification checks, not reported here due to space reasons, suggest that the possible violation of the normality assumption does not critically affect our results on spatial dependence.

The remaining specifications (8) to (12) highlight a number of other issues. As noted in Appendix A, we use time-series techniques proposed by Gomez and Maravall (1997) to estimate the missing price data for this analysis. A simple alternative to that is to estimate any missing observation for location i by its mean across all years for which data is not missing. This is the data underlying specification (8). One notes that the fit is, as expected, worse than in the comparable specification (5), but the inferences regarding spatial dependence are largely the same.

Another concern is that heterogeneity (for instance in terms of size) across prefectures might induce heteroskedasticity and other problems in the estimation. In this respect, we note that reported in Table 8 to Table 10 are Huber-White heteroskedasticity-consistent standard errors. Moreover, we have also experimented with non-parametric techniques by computing bootstrap standard errors for a number of specifications. The bootstrapped standard errors are typically larger, and at times substantially larger than the Huber-White robust standard errors. At the same time, the bootstrapped standard errors are not so large as to affect our inferences.

The difference in the size of the prefectures may mean that relying on the computed mid-price in a prefecture is problematic. Specifications (9) and (10) of Table 8 show the results obtained with using the lowest and highest price in a prefecture, instead of the mid-price. The fit is somewhat

[27] We found more evidence for normality of *dryness* than *wdev* based on the test by Shapiro and Wilk (1965).

worse compared to specification (5), and the estimate of ρ drops to around 0.54 from 0.62, but the evidence for spatial dependence remains quite strong (the adjusted LM test p-values are less than 0.04).

The last two specifications in Table 8 provide some information on the possible effects of spatial non-stationarity on the results. In section 2 above it was noted that prices in our sample tend to be higher on the coast than in the interior regions of China, which is at least in part related to the relatively high land-to-labor ratio in the interior of China. This spatial price gradient—increasing from West to East, and from North to South—could mask some of our findings in terms of spatial dependence. In order to see to what degree this seems to be the case, we use two alternative price variables from which the basic price gradient has been purged in a first step. The price variable in specification (11) is the residual of a regression of the log mid-price on three regional fixed effects—one group of prefectures in the Northeast, another group on the coast near the Yangzi delta, and a third consisting of all other prefectures—while the price variable in (12) is the regression residual based on four groups in the first step.[28] Comparing the regression results in specifications (11) and (12) with the baseline specification (5), the differences seem to be quite small. This suggests that at least for the spatial regression results where we condition on weather as well as other variables, spatial non-stationarity, if present, does not appear to have a strong effect on our results.

We now turn to spatial regression results for some of the groups of prefectures discussed in section 3 above.

[28] In specification (11), the groups include (1) prefectures in Guangxi and Guizhou, (2) prefectures in Anhwei, Jiangsu, and Zhejiang, and (3) prefectures in Guangdong, Fujian, Jiangxi, Hubei, and Hunan. For specification (12), the previous group (3) is subdivided into (3') Guangdong, Fujian, and Jiangxi, and group (4) Hubei and Hunan prefectures. Our approach here is a very simple version of the widely-used trend-surface analysis; see, e.g., Cliff and Ord (1981, 222-228).

4.3 Spatial dependence in different regions and over time

Different regions Table 9 compares results for the spatial lag model with preferred specifications for two pairs of prefectures that were identified in section 3 above: first, the Inland versus the Coastal prefectures, and second, the Yangzi River versus the Non-Yangzi River prefectures. For each sample, two specifications are shown, one with and another without including the previous period price variables ($Wy(-1)$ and $y(-1)$).

There are several results that are common across all regions in China. First of all, weather shocks do not have a significant influence on the local price. Thus, controlling for spatial dependence makes a qualitative difference for inferences about the role of weather for variation in local prices. Second, the correlation of current to own previous period price ($y(-1)$) is positive, whereas the correlation of current to spatially lagged previous period prices is non-positive.[29]

The evidence for spatial dependence varies, however. For one, it is weaker once the temporally lagged price variables are included; for the Inland prefectures, for instance, the estimate of ρ falls from about 0.8 to 0.45. A special case seems to be the Yangzi River group of prefectures, where the estimate of ρ falls from about 0.73 to essentially zero. That we can reject the null hypothesis of spatial dependence for the Yangzi River prefectures is confirmed by both of the LM test statistics. Note, however, that the adjusted LM test statistic is considerably less affected than the standard LM test by whether $Wy(-1)$ and $y(-1)$ are included; both for the Coastal as well as the Non-Yangzi sample, the adjusted LM test statistic is quite similar for the case with and without the spatially lagged previous period prices in the regression. This underscores the importance of addressing the

[29] Note that the influence of last period's (own) price is stronger in Inland and Non-Yangzi prefectures than in Coastal and Yangzi River prefectures: the estimates on $y(-1)$ in the former are about 0.79, while for Coastal it is 0.73, and for Yangzi prefectures it is 0.63. The relatively strong intertemporal effect for Inland and Non-Yangzi River prefectures is consistent with grain storage being a more important mechanism of consumption smoothing in these prefectures, compared to the Coastal and Yangzi prefectures; this is in line with the findings in Shiue (2002). See also the analysis of grain storage in Shiue (2004).

nuisance parameter problem that apparently plagues the standard LM test statistic.

Overall, if we measure the evidence for spatial dependence by the adjusted LM statistic in the specification with the spatially lagged previous prices, the results are roughly in line with our analysis of spatial autocorrelation above, as well as with direct historic evidence. The adjusted LM statistic ranking is Inland $\approx$ Non-Yangzi River $>$ Coastal $>$ Yangzi River. The spatial weights matrix that we use (exponential in distance with $\theta = 1.4$) emphasizes relatively short distances. This ranking thus reflects the fact that the evidence for short-distance price clusters is strongest for Inland prefectures. These prefectures are less likely to have access to waterway transport, and due to the higher transport costs of land trade, rice trade can be expected to take place only for relatively short distances.

In contrast, the Coastal prefectures have some access to waterway transport, and the Yangzi River prefectures probably enjoyed the lowest transportation costs. Our failure to find evidence for spatial dependence in the Yangzi River sample does not mean there was an absence of trade pulling prices in different markets together. Rather, it suggests that trade also occurred among these prefectures at distances beyond the short-distance radius.

Spatial dependence over time Table 10 presents results for four samples separately for three subperiods (1743/59, 1760/77, and 1778/95). As shown in the top left, in the All Prefectures sample, the influence of weather shocks on local prices is never significant. There is evidence for spatial dependence in all three subperiods, even though it becomes slightly weaker over time (as measured by the adjusted LM test statistic). This result seems to be primarily driven by the Coastal prefectures, where the adjusted LM statistic falls from about 5.8 to 4.3. The estimate of ρ rises for the Yangzi River prefectures over time, although all values are not significantly different from zero at standard levels.

Overall, these results largely confirm earlier results that emphasize the heterogeneity of different regions in China, and in particular with respect to their access to waterway transport.

5 Conclusion

This paper has shown how estimators of spatial dependence 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 to trade. 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.

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 remove the significance of weather in the estimation results.

Patterns of spatial dependence may change over time, and when data for long periods are available, there are clear benefits to allowing for temporal as well as spatial changes in the model. In this paper, we have taken into account the lagged variables of the cross-sectional markets, but this could be taken further. In particular, the panel structure of the data, where there are repeated temporal observations on the same locations, can be used to help control for the presence of cross-sectional heterogeneity among the spatial units of observation. Another important direction for future work would be to estimate longer-horizon dynamics involving lags of more than one year.

This analysis can also be combined with direct evidence on how trade routes evolved in China over time to the present day 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 (Keller and Shiue 2004). 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, prices were reported approximately monthly from all parts of the empire as part of an early warning system of food shortages. The fact that these data were put to practical uses suggests accuracy would have been important. To the extent that is was possible, accuracy was enforced through a system of unannounced checks and audits. The quality of the data is generally considered good compared to other historical price records.

Our focus is on rice, a predominate type of grain for the central and southern parts of China. This leads to the exclusion of parts of China to the North and West where other types of grain, for instance barley and wheat, were more common. The sample covers about 60% of the Chinese economy—a sizeable area with a population of about 120 million people, or about 20% of the world population at the time. 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 several different approaches. For the most part, the results shown above are based on the time-series methods developed by Gomez and Maravall (1997). In addition, Table 8 shows results based on a simple cross-sectional data interpolation, which leads to similar results. We also find qualitatively the same results if we focus on years that are not missing in the original source.

For each prefecture, we have information on the highest and lowest price in a given month. These two prices yield the mid-price, which is the focus of our analysis; it is computed as the average between the highest and lowest price in a given prefecture. This raises a number of issues. First of all, the prefectures differ substantially in size (see Figure 2), which may give rise to heteroskedasticity.

Specifically, a given prefecture's lowest and highest price, respectively, might for different months be from different (county-level) markets. If these differences depend on prefectural size, the assumption of constant variance of prices across prefectures would be violated. Our preferred techniques are therefore those that are heteroskedasticity-consistent, and more generally, we discuss what the influence of heteroskedasticity might be on our results (see section 4.2 above).

The original source of the price data is *Gongzhong liangjiadan* [Grain Price Lists in the Palace Archives]. Today it is located in the Number One Historical Archives in Beijing. The data were 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 is equal to the weather levels given in the source. The variable *dryness* takes values of 1, 2, 3, 4, and 5 (1 being least and 5 being most dry). (2) The variable weather deviation: This variable is defined as $wdev = |Dryness - 3|$, taking values of 0, 1, and 2. (3) The variable *bad_weather* variable is as follows:

$$bad\_weather = \begin{cases} 1 \text{ if } dryness = 1 \text{ or } 5 \\ 0 \text{ otherwise.} \end{cases}$$

Summary statistics for these variables by province are given in Table 3.

C Variances of Geary's c and Moran's I statistics

Under the null hypothesis that z is an identically and independently distributed normal variate, the variance of Geary's c is given by

$$var(c) = \frac{(2S_1 + S_2)(N-1) - 4S_0^2}{2(N+1)S_0^2}, \tag{8}$$

where

$$S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}, \qquad S_1 = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (w_{ij} + w_{ji})^2, \tag{9}$$

and

$$S_2 = \sum_{i=1}^{N} (w_{i\cdot} + w_{\cdot i})^2, \quad \text{with} \quad w_{i\cdot} = \sum_{j=1}^{N} w_{ij} \quad \text{and} \quad w_{\cdot i} = \sum_{j=1}^{N} w_{ji}. \tag{10}$$

The variance of Moran's I is, under the null hypothesis that z is an identically and independently distributed normal variate, given by

$$var(I) = \frac{N^2 S_1 - N S_2 + 3 S_0^2}{S_0^2 \left(N^2 - 1\right)}. \tag{11}$$

See Cliff and Ord (1981, 20).

References

[1] Anderson, James, and Eric van Wincoop (2004), "Trade costs", NBER Working Paper # 10480.

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

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

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

[5] 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.

[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, Anil Bera, Raymond Florax, and Mann Yoon (1996), "Simple diagnostic tests for spatial dependence", *Regional Science and Urban Economics* 26: 77-104.

[8] 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.

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

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

[11] 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.

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

[13] 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.

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

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

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

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

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

[19] Durbin, J., and G. Watson (1950), "Testing for serial correlation in least squares regression I", *Biometrika* 37: 409-428.

[20] Durbin, J., and G. Watson (1951), "Testing for serial correlation in least squares regression II", *Biometrika* 38: 159-178.

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

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

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

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

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

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

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

[28] 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.

[29] 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.

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

[31] 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.

[32] Keller, Wolfgang (2004), "International Technology Diffusion", forthcoming, *Journal of Economic Literature*.

[33] Keller, Wolfgang and Carol H. Shiue (2004), "Market Integration and Economic Development: A Long-run Comparison", forthcoming, *Review of Development Economics*.

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

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

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

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

[38] Oden, Neal L. (1995), "Adjusting Moran's I for population density", *Statistics in Medicine* 14: 17-26.

[39] 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.

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

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

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

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

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

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

[46] Shapiro, S., and M. Wilk (1965), "An analysis of variance test for normality (complete samples)", *Biometrika* 52: 591-611.

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

[48] Shiue, Carol H. (2004) "Local Granaries and Central Government Disaster Relief: Moral Hazard and Intergovernmental Finance in Eighteenth and Nineteenth Century China", *Journal of Economic History* 64(1): 101-125.

[49] 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.

[50] 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.

[51] 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.

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

[53] 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.

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

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

Table 1

Geographic Location: Longitudes and Latitudes

province	number of prefectures	coordinate	mean	standard devation
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 by province
Averages from 54 annual cross-sectional statistics (includes both months 2 and 8)

province	**price variable**	**mean**	**standard deviation**	**min**	**max**	**obs.**
Anhwei	Mid-price	1.553	0.189	1.245	1.991	1404
	Lowest	1.408	0.168	1.121	1.765	1404
	Highest	1.698	0.252	1.331	2.289	1404
Fujian	Mid-price	1.667	0.200	1.340	2.079	1296
	Lowest	1.507	0.228	1.132	1.958	1296
	Highest	1.827	0.215	1.464	2.296	1296
Guangdong	Mid-price	1.498	0.233	1.103	1.985	1404
	Lowest	1.315	0.232	0.898	1.798	1404
	Highest	1.681	0.274	1.214	2.218	1404
Guangxi	Mid-price	1.117	0.109	0.930	1.336	1296
	Lowest	0.999	0.109	0.826	1.230	1296
	Highest	1.234	0.136	0.997	1.506	1296
Guizhou	Mid-price	0.972	0.149	0.727	1.310	1404
	Lowest	0.835	0.150	0.597	1.127	1404
	Highest	1.109	0.192	0.815	1.548	1404
Hubei	Mid-price	1.330	0.148	1.044	1.601	1080
	Lowest	1.156	0.144	0.884	1.402	1080
	Highest	1.504	0.186	1.177	1.853	1080
Hunan	Mid-price	1.190	0.114	1.011	1.440	1404
	Lowest	1.090	0.116	0.886	1.341	1404
	Highest	1.290	0.139	1.071	1.600	1404
Jiangsu	Mid-price	1.794	0.208	1.474	2.206	1080
	Lowest	1.639	0.209	1.299	2.042	1080
	Highest	1.950	0.238	1.599	2.426	1080
Jiangxi	Mid-price	1.413	0.099	1.233	1.620	1512
	Lowest	1.294	0.113	1.077	1.524	1512
	Highest	1.533	0.120	1.310	1.792	1512
Zhejiang	Mid-price	1.564	0.162	1.254	1.824	1188
	Lowest	1.414	0.167	1.095	1.690	1188
	Highest	1.713	0.180	1.364	2.011	1188
Total	Mid-price	1.400	0.310	0.722	2.306	13068
	Lowest	1.257	0.300	0.591	2.121	13068
	Highest	1.543	0.345	0.807	2.589	13068

Table 3: Weather by province

Averages from 54 annual cross-sectional statistics

province	weather variable*	mean	standard deviation	obs.
Anhwei	dryness	2.677	0.621	702
	wdev	0.739	0.512	702
	bad weather	0.124	0.174	702
Fujian	dryness	2.748	0.686	648
	wdev	0.906	0.493	648
	bad weather	0.171	0.251	648
Guangdong	dryness	2.946	0.668	702
	wdev	0.906	0.361	702
	bad weather	0.104	0.146	702
Guangxi	dryness	2.966	0.468	648
	wdev	0.392	0.434	648
	bad weather	0.035	0.060	648
Guizhou	dryness	3.006	0.235	702
	wdev	0.405	0.219	702
	bad weather	0.071	0.071	702
Hubei	dryness	2.850	0.527	540
	wdev	0.557	0.474	540
	bad weather	0.126	0.143	540
Hunan	dryness	2.839	0.504	702
	wdev	0.503	0.417	702
	bad weather	0.085	0.127	702
Jiangsu	dryness	2.865	0.556	540
	wdev	0.639	0.456	540
	bad weather	0.098	0.108	540
Jiangxi	dryness	2.696	0.461	756
	wdev	0.640	0.361	756
	bad weather	0.090	0.102	756
Zhejiang	dryness	2.842	0.630	594
	wdev	0.623	0.483	594
	bad weather	0.114	0.156	594
Total	dryness	2.841	0.818	6534
	wdev	0.633	0.629	6534
	bad weather	0.101	0.274	6534

* See Appendix B for the definition of the three weather variables

Table 4: Prices over time

Mid-price

		1742/59		1760/77		1778/95	
province	**statistic***	**month 2**	**month 8**	**month 2**	**month 8**	**month 2**	**month 8**
Anhwei	mean	1.472	1.405	1.500	1.500	1.717	1.724
	sd	0.179	0.161	0.188	0.181	0.187	0.196
	sd/mean	0.121	0.114	0.125	0.121	0.109	0.114
Fujian	mean	1.577	1.564	1.636	1.660	1.764	1.801
	sd	0.199	0.168	0.162	0.142	0.259	0.243
	sd/mean	0.126	0.107	0.099	0.085	0.147	0.135
Guangdong	mean	1.455	1.412	1.490	1.503	1.566	1.564
	sd	0.229	0.202	0.252	0.236	0.241	0.229
	sd/mean	0.158	0.143	0.169	0.157	0.154	0.146
Guangxi	mean	1.134	1.138	1.021	1.039	1.171	1.197
	sd	0.155	0.127	0.082	0.077	0.101	0.095
	sd/mean	0.136	0.112	0.081	0.074	0.086	0.079
Guizhou	mean	0.983	1.016	1.012	1.010	0.908	0.903
	sd	0.156	0.163	0.138	0.143	0.146	0.147
	sd/mean	0.159	0.160	0.136	0.141	0.161	0.163
Hubei	mean	1.210	1.204	1.296	1.296	1.483	1.492
	sd	0.133	0.119	0.118	0.112	0.185	0.183
	sd/mean	0.110	0.099	0.091	0.086	0.125	0.123
Hunan	mean	1.156	1.151	1.139	1.147	1.270	1.276
	sd	0.116	0.113	0.094	0.093	0.111	0.109
	sd/mean	0.100	0.098	0.082	0.081	0.087	0.086
Jiangsu	mean	1.693	1.671	1.778	1.785	1.915	1.924
	sd	0.165	0.205	0.141	0.162	0.239	0.248
	sd/mean	0.098	0.123	0.079	0.091	0.125	0.129
Jiangxi	mean	1.332	1.276	1.421	1.382	1.551	1.517
	sd	0.112	0.100	0.077	0.070	0.096	0.089
	sd/mean	0.084	0.078	0.054	0.050	0.062	0.059
Zhejiang	mean	1.526	1.468	1.600	1.588	1.595	1.607
	sd	0.157	0.144	0.152	0.141	0.170	0.175
	sd/mean	0.103	0.098	0.095	0.089	0.107	0.109
Total	mean	1.346	1.322	1.380	1.381	1.483	1.489
	sd	0.287	0.261	0.300	0.297	0.352	0.354
	sd/mean	0.214	0.198	0.217	0.215	0.237	0.238

* Mean and standard deviation: averages across 18 annual cross-sectional statistics
Coefficient of variation (= sd/mean): average(sd)/average(mean)

Table 5: Weather over time

province	statistic*	1742/59 dryness	1742/59 bad weather	1760/77 dryness	1760/77 bad weather	1778/95 dryness	1778/95 bad weather
Anhwei	mean	2.470	0.162	2.739	0.107	2.821	0.103
	sd	0.713	0.231	0.591	0.182	0.561	0.110
Fujian	mean	3.000	0.218	2.593	0.093	2.653	0.204
	sd	0.686	0.306	0.636	0.159	0.737	0.289
Guangdong	mean	3.043	0.090	2.829	0.107	2.966	0.115
	sd	0.794	0.145	0.676	0.171	0.533	0.123
Guangxi	mean	3.111	0.019	2.787	0.065	3.000	0.023
	sd	0.484	0.053	0.503	0.100	0.418	0.028
Guizhou	mean	3.000	0.047	3.124	0.068	2.893	0.098
	sd	0.254	0.079	0.208	0.052	0.243	0.082
Hubei	mean	2.761	0.094	2.872	0.089	2.917	0.194
	sd	0.562	0.157	0.420	0.108	0.601	0.163
Hunan	mean	2.812	0.090	2.825	0.047	2.880	0.120
	sd	0.563	0.180	0.396	0.084	0.552	0.117
Jiangsu	mean	2.767	0.039	2.944	0.172	2.883	0.083
	sd	0.611	0.045	0.568	0.215	0.489	0.063
Jiangxi	mean	2.675	0.111	2.615	0.071	2.798	0.087
	sd	0.576	0.141	0.406	0.040	0.400	0.126
Zhejiang	mean	3.045	0.131	2.657	0.121	2.823	0.091
	sd	0.519	0.121	0.697	0.196	0.674	0.149
Total	mean	2.866	0.101	2.796	0.092	2.862	0.111
	sd	0.865	0.281	0.782	0.255	0.807	0.287

*Averages from 54 annual cross-sectional statistics

Table 6: Geary's c statistic for different prefectures and by subperiod*

distance	prefectures	1742-1795	1742-1759	1760-1777	1778-1795
0- 200 km	All prefectures	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)
200 - 400 km	All prefectures	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)
400 - 600 km	All prefectures	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)
600 - 800 km	All prefectures	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)
800 - 1000 km	All prefectures	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)
1000 - 1200 km	All prefectures	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)
1200 - 1400 km	All prefectures	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)
1400 - 1600 km	All prefectures	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

* Averages from 54 (column (i)) or 18 (columns (ii)-(iv)) annual cross-sectional statistics; in parentheses: standard error of mean

Table 7: Regression results for All Prefectures, 1742-1795

Dependent variable: log mid price; N = 121; averages from 54 annual cross-sectional regressions

	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
wdev	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 (0,6]	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
wdev	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
Adjusted LM test ++	9.146	17.234	19.663	33.657	3.116	13.847
lnL	75.936	76.469	59.425	60.536	85.152	85.630

* (**) significant at 10 (5)% level

Huber-White robust standard errors

++ Chi-squared distributed, with 1 degree of freedom; critical values: 2.706, 3.841, and 6.635 at the 10%, 5%, and 1% level, respectively

Table 8: Specification and Robustness

Dependent variable: log price, N = 121; exponential weights (theta = -1.4); averages from 53 cross-sectional regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)☺	(9) ◊	(10) ○	(11) ☼	(12) ◘
wdev	0.020	0.017	0.013	0.010								
s.e.#	0.017	0.017	0.012	0.011								
dryness					-0.030		-0.033	-0.037	-0.025	-0.039	-0.026	-0.024
s.e.#					0.043		0.045	0.049	0.049	0.053	0.043	0.043
dryness2					0.006		0.007	0.007	0.005	0.007	0.005	0.005
s.e.#					0.008		0.008	0.009	0.009	0.009	0.008	0.008
Wy(-1)		0.547**		-0.415**	-0.394**			-0.296**	-0.274*	-0.323**	-0.426**	-0.447**
s.e.#		0.170		0.134	0.137			0.141	0.153	0.146	0.140	0.143
y(-1)			0.683**	0.771**	0.771**	0.677**	0.691**	0.661**	0.731**	0.765**	0.790**	0.789**
s.e.#			0.058	0.062	0.062	0.057	0.057	0.074	0.067	0.066	0.061	0.062
rho	0.893**	0.469**	0.338**	0.636**	0.618**	0.357**	0.332**	0.628**	0.546**	0.543**	0.613**	0.614**
s.e.#	0.047	0.159	0.068	0.111	0.114	0.067	0.068	0.113	0.127	0.125	0.115	0.116
LM test ++	251.389	15.672	31.619	52.810	45.996	36.521	29.964	50.771	31.755	33.133	49.855	48.109
Adjusted LM test ++	13.847	6.378	8.444	6.183	7.099	8.152	9.374	6.995	4.624	6.640	6.539	6.373
lnL	85.956	94.032	152.870	159.111	160.434	150.848	154.843	138.254	138.227	142.764	160.920	161.015

Huber-White robust standard error; *(**): significant at a 10%(5%) level
++ Chi-squared distributed with 1 degree of freedom; critical values: 2.706, 3.841, and 6.635 at the 10%, 5% and 1% level, respectively
☺ Alternative interpolation of missing data; see section 4.2
◊ Log lowest price, instead of log mid-price
○ Log highest price, instead of log mid-price
☼ Residual of log mid-price regression on fixed effects for 3 groups of prefectures; see section 4.2
◘ Residual of log mid-price regression on fixed effects for 4 groups of prefectures; see section 4.2

Table 9: Regression results for different prefectures
Dependent variable: log mid-price; averages from 53 cross-sectional regressions

	Inland		Coastal		Non-Yangzi River		Yangzi River	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
dryness	-0.024	-0.022	0.005	-0.016	-0.075	-0.029	-0.077	-0.054
s.e.#	0.115	0.073	0.099	0.063	0.084	0.049	0.156	0.112
dryness2	0.005	0.005	0.000	0.003	0.013	0.006	0.015	0.010
s.e.#	0.020	0.012	0.018	0.011	0.015	0.009	0.027	0.020
Wy(-1)		-0.239		-0.225		-0.358**		0.337
s.e.#		0.215		0.209		0.148		0.346
y(-1)		0.789**		0.733**		0.791**		0.632**
s.e.#		0.085		0.091		0.065		0.163
rho	0.799**	0.447**	0.711**	0.467**	0.864**	0.561**	0.727**	0.045
s.e.#	0.080	0.173	0.120	0.166	0.053	0.124	0.142	0.288
LM test ++	72.672	14.195	34.156	14.921	162.589	30.492	14.687	1.912
Adjusted LM test ++	8.868	5.324	4.461	4.889	5.203	5.250	5.774	2.240
lnL	50.637	92.980	44.709	77.025	66.475	133.809	21.233	32.168
N	62	62	59	59	100	100	21	21

Huber-White robust standard error; *(**): significant at a 10%(5%) level
++ Chi-squared distributed with 1 degree of freedom; critical values: 2.706, 3.841, and 6.635 at the 10%, 5% and 1% level, respectively

Table 10: Spatial dependence over time
Dependent variable: log mid-price; averages from annual cross-sectional regressions

	All Prefectures			Inland Prefectures			Coastal Prefectures			Yangzi River Prefectures		
	1743/59	**1760/77**	**1778/95**	**1743/59**	**1760/77**	**1778/95**	**1743/59**	**1760/77**	**1778/95**	**1743/59**	**1760/77**	**1778/95**
dryness	-0.032	-0.039	-0.019	-0.021	-0.044	-0.003	-0.022	-0.014	-0.011	-0.002	-0.091	-0.072
s.e.#	0.046	0.049	0.036	0.071	0.079	0.068	0.072	0.063	0.054	0.105	0.102	0.130
dryness2	0.006	0.008	0.003	0.004	0.008	0.003	0.006	0.003	0.002	0.002	0.017	0.011
s.e.#	0.008	0.009	0.006	0.012	0.015	0.011	0.013	0.012	0.009	0.018	0.018	0.024
Wy(-1)	-0.317**	-0.438**	-0.421**	-0.132	-0.318*	-0.269	-0.236	-0.235	-0.206	0.507	0.398	0.089
s.e.#	0.143	0.127	0.141	0.249	0.189	0.205	0.212	0.189	0.225	0.347	0.335	0.355
y(-1)	0.689**	0.794**	0.825**	0.722**	0.770**	0.867**	0.656**	0.735**	0.805**	0.560**	0.675**	0.663**
s.e.#	0.072	0.065	0.049	0.096	0.099	0.062	0.104	0.092	0.077	0.171	0.168	0.148
rho	0.626**	0.636**	0.594**	0.419**	0.537**	0.392**	0.499**	0.491**	0.412**	-0.011	-0.043	0.200
s.e.#	0.113	0.107	0.122	0.184	0.148	0.186	0.157	0.160	0.179	0.290	0.289	0.285
LM test ++	55.081	49.577	33.836	14.030	19.028	10.085	21.218	15.071	8.823	2.006	1.517	2.235
Adjusted LM test ++	7.801	7.425	6.109	4.652	7.117	4.164	5.735	4.673	4.306	1.736	1.702	3.255
lnL	132.407	159.879	187.459	75.277	91.540	110.913	63.126	77.879	89.298	30.820	33.447	32.242

Huber-White robust standard error; *(**): significant at a 10%(5%) level
++ Chi-squared distributed with 1 degree of freedom; critical values: 2.706, 3.841, and 6.635 at the 10%, 5% and 1% level, respectively

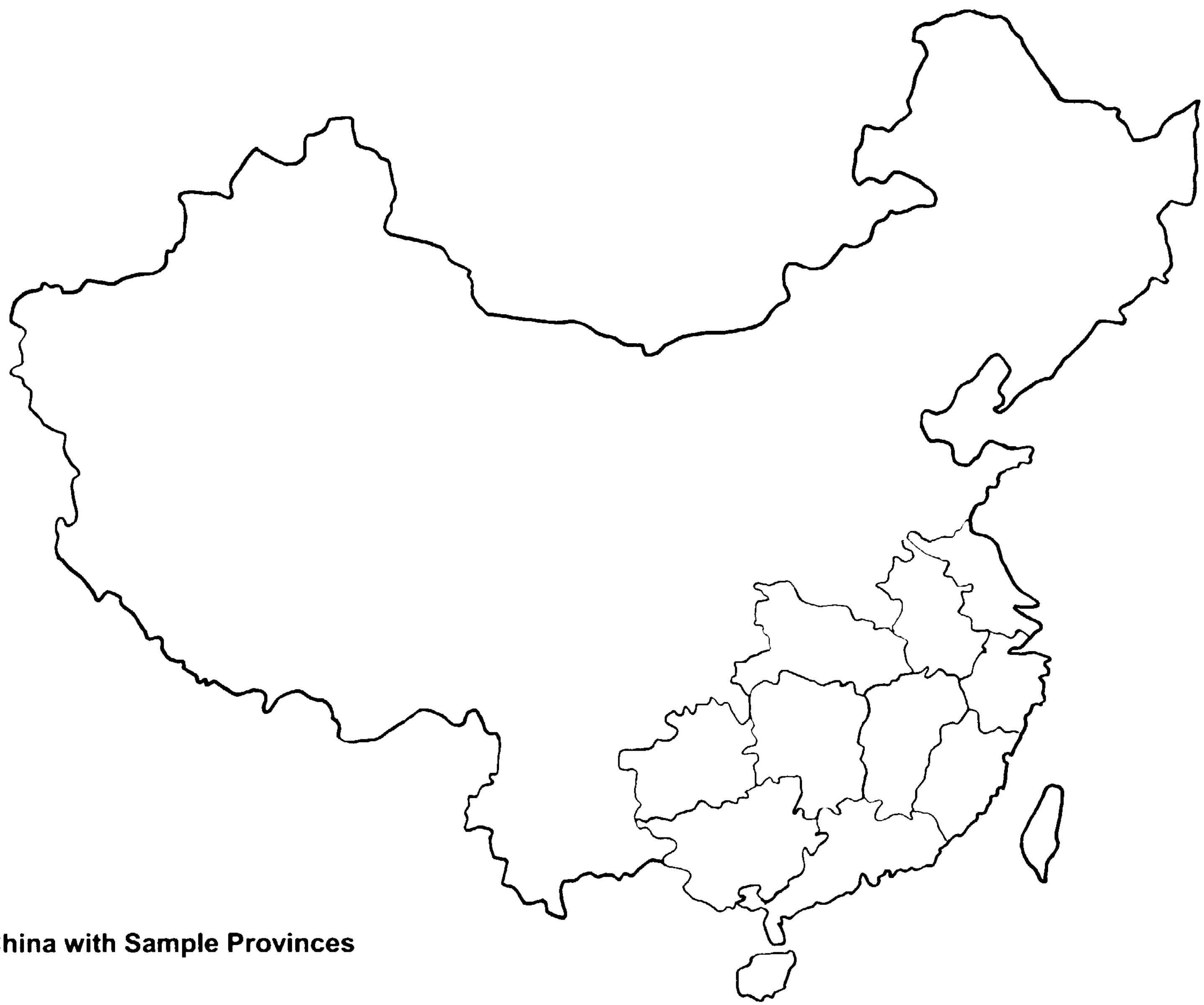
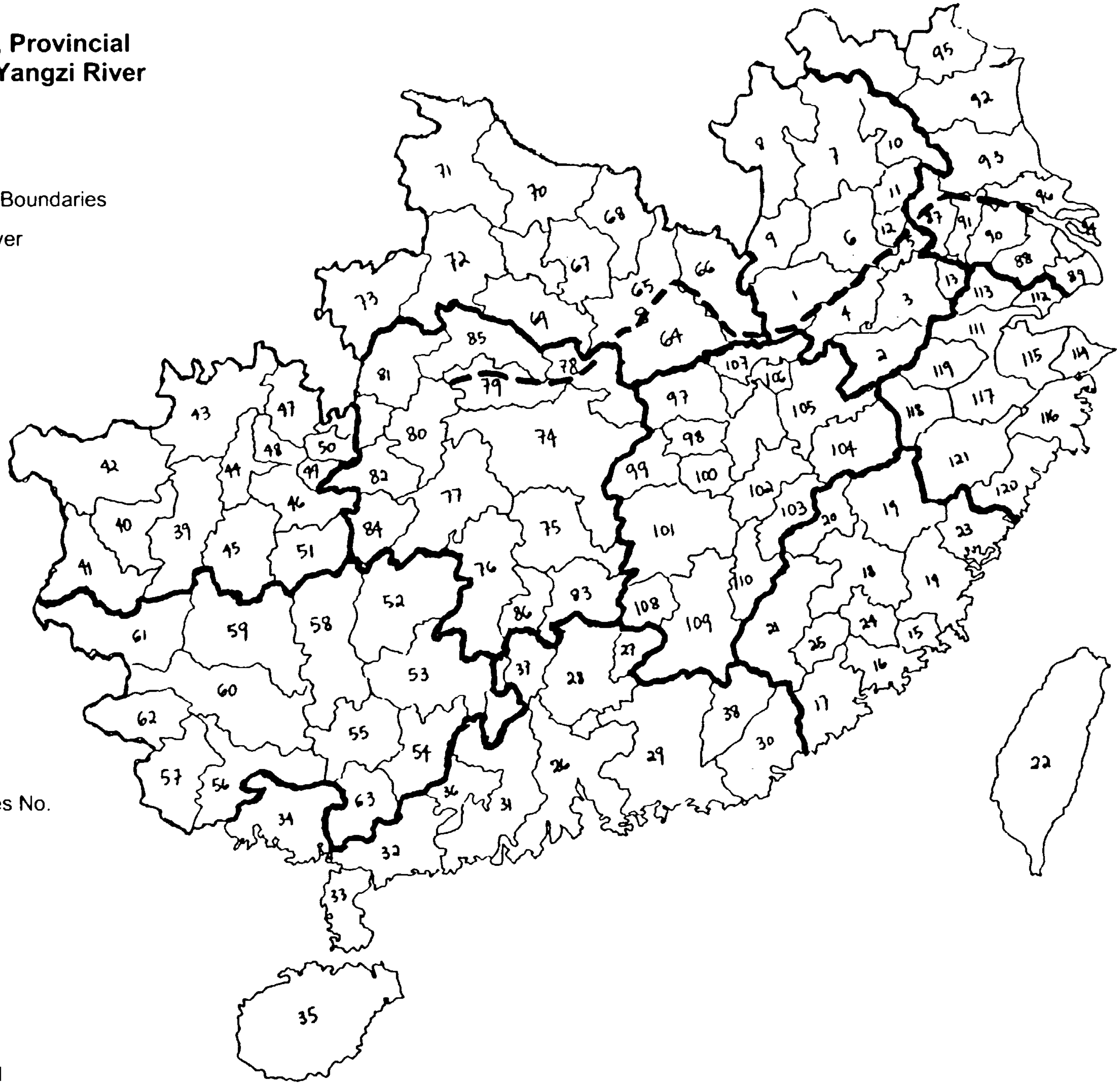


Figure 1
Map of China with Sample Provinces

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

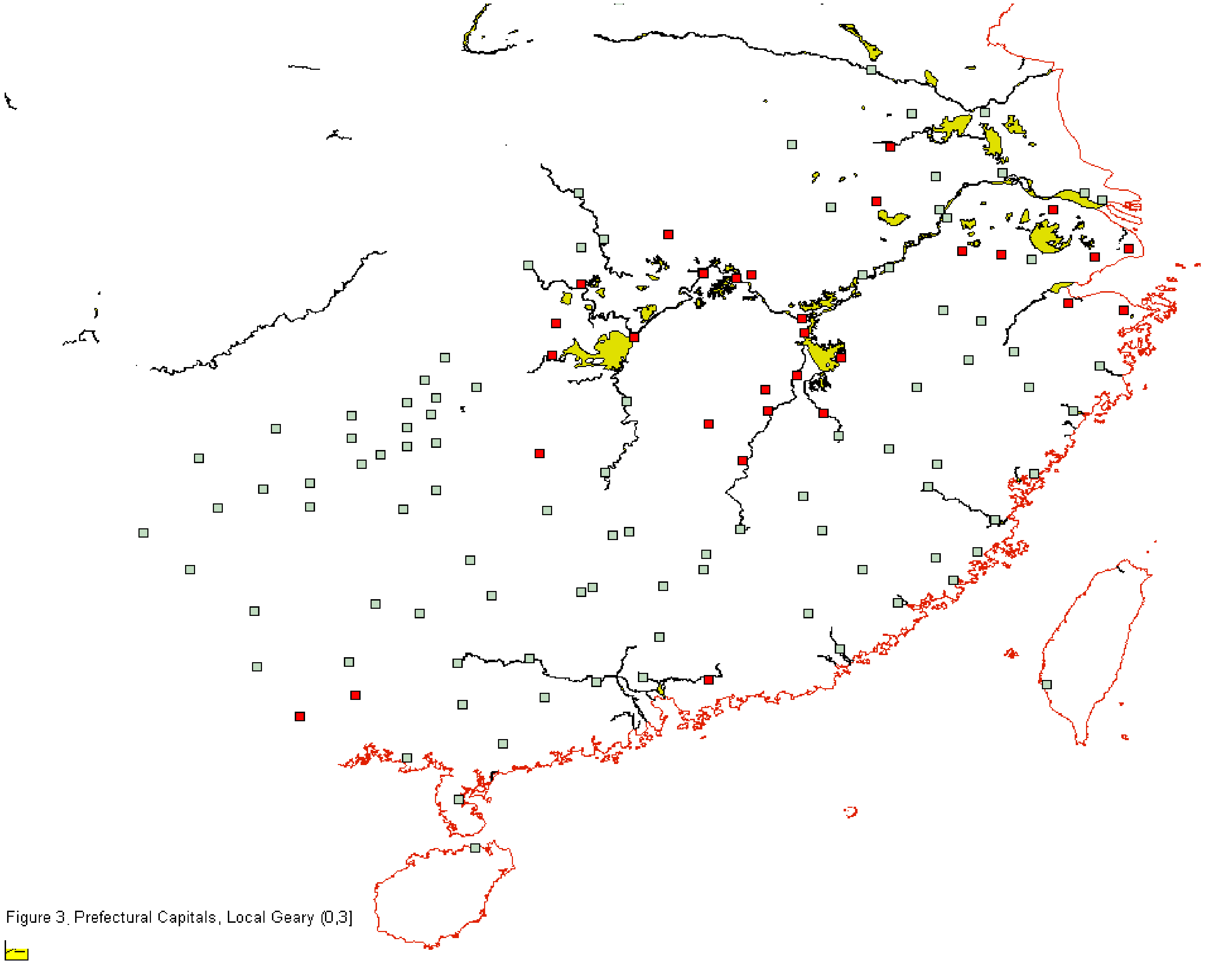


Figure 3. Prefectural Capitals, Local Geary (0,3]

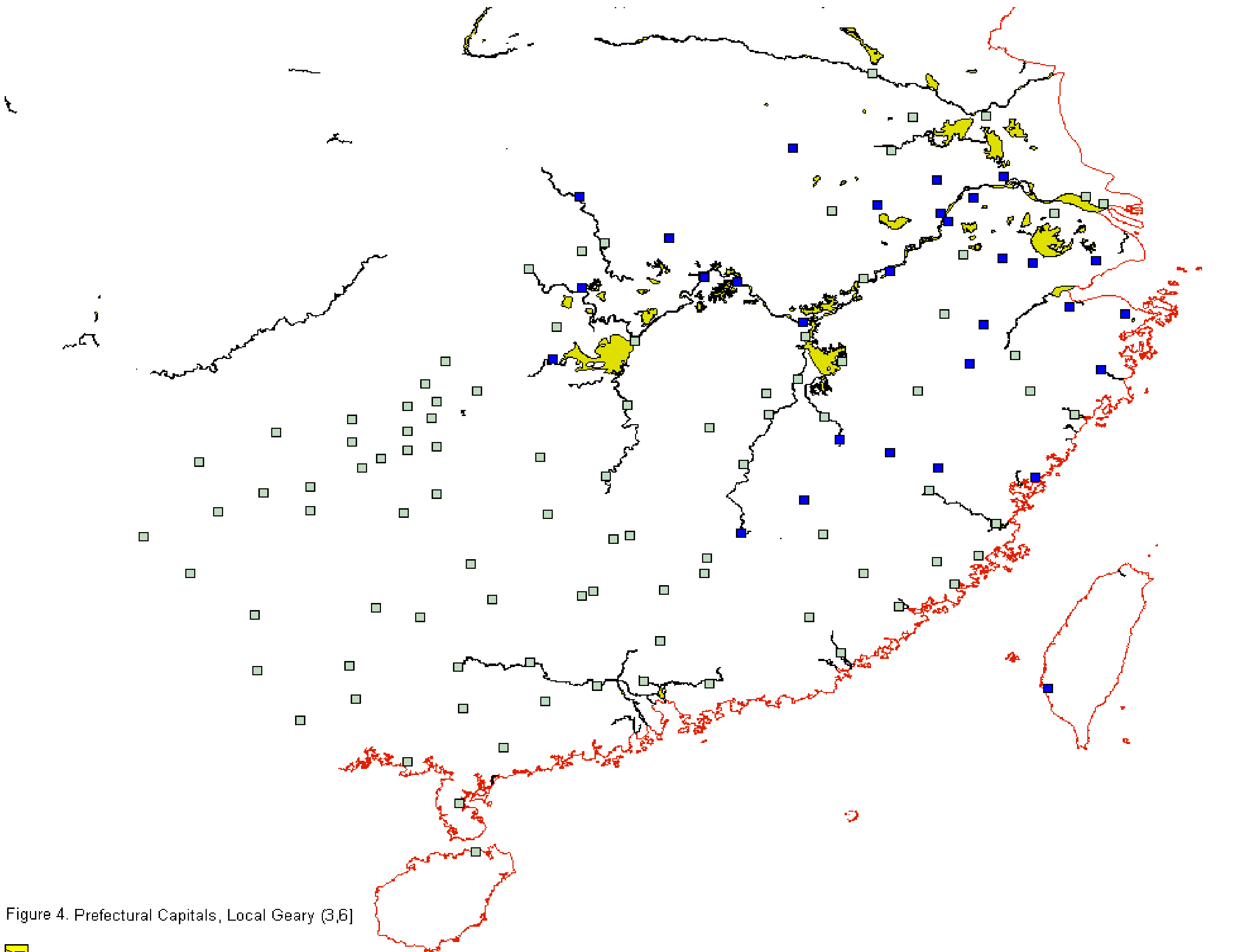


Figure 4. Prefectural Capitals, Local Geary (3,6)

>=