8-2023

The Economic Regions of Chinese Civilization: A GIS-Based Analysis of Grain Markets in China, 1736-1842

Karl Ryavec
University of California, kryavec@ucmerced.edu

Mark Henderson
University of California

Rocco Bowman
University of California

Follow this and additional works at: https://scholarsarchive.byu.edu/CCR

Part of the Comparative Literature Commons, History Commons, International and Area Studies Commons, Political Science Commons, and the Sociology Commons

Recommended Citation
Available at: https://scholarsarchive.byu.edu/ccr/vol89/iss89/15

This Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Comparative Civilizations Review by an authorized editor of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
The Economic Regions of Chinese Civilization: A GIS-Based Analysis of Grain Markets in China, 1736-1842

Cover Page Footnote
This study presents a new spatial analytic method for mapping regional marketing systems of an entire civilization by focusing on China's High Qing period based on monthly grain price time series for 235 prefectures from 1736-1842. Network analyses of prefectural grain price time series and their cross-correlations detected eleven or twelve primary macro-markets. These empirically derived regions generally align with the Late Qing urban systems nested within physiographic macroregions as posited by Skinner (1977), but with some notable differences, especially in North China. This delineation of the traditional marketing systems of agrarian China applicable to a 107-year time span provides a baseline for measuring the subsequent effects of mechanized transportation on Chinese Civilization's further integration into the modern world-system.

This article is available in Comparative Civilizations Review: https://scholarsarchive.byu.edu/ccr/vol89/iss89/15
The Economic Regions of Chinese Civilization: A GIS-Based Analysis of Grain Markets in China, 1736-1842

Karl Ryavec
Anthropology and Heritage Studies Department, University of California, Merced, USA
ORC-ID 0000-0002-5691-2135
Corresponding author: kryavec@ucmerced.edu

Mark Henderson
Anthropology and Heritage Studies Department, University of California, Merced, USA
ORC-ID 0000-0003-2952-0942

Rocco Bowman
Anthropology and Heritage Studies Department, University of California, Merced, USA
ORC-ID 0000-0002-3732-7886

Abstract

This study presents a new spatial analytic method for mapping regional marketing systems of an entire civilization by focusing on China's High Qing period based on monthly grain price time series for 235 prefectures from 1736-1842. Network analyses of prefectural grain price time series and their cross-correlations detected eleven or twelve primary macro-markets. These empirically derived regions generally align with the Late Qing urban systems nested within physiographic macroregions as posited by Skinner (1977), but with some notable differences, especially in North China. This delineation of the traditional marketing systems of agrarian China applicable to a 107-year time span provides a baseline for measuring the subsequent effects of mechanized transportation on Chinese Civilization’s further integration into the modern world-system.

NB: For your reference, a series of illustrative maps and charts are printed at the end of this article.

Spatial Analysis of Regional Grain Markets in China, 1736-1842

The potential for sophisticated spatial analyses of historical phenomena has long been constrained by the availability of digital data and tools, but both have become much more accessible in recent years. An infrastructure of digital humanities data producers, repositories, and statistical methods is emerging that allows researchers to revisit old questions and see historical patterns in a new light. In this study, we link digitized records from the Qing Dynasty Grain Price Database, as published online by Academia Sinica, with urban center locations in the China Historical Geographic Information System (CHGIS).
We compare the results of newer clustering methods to reexamine one of the most influential hypotheses in the study of Chinese civilization and history, the “macroregion” framework, which has attracted renewed attention in recent comparative studies of market integration.

Qing China, an empire of continental scale that endured over two and a half centuries (1644-1911), sustained some of the highest and densest human population levels of any preindustrial civilization with a complex web of governmental, economic, and societal institutions. The geographical structures of these institutions and their effects on helping or hindering economic development has been a matter of debate among China scholars. G. William Skinner (1977) postulated that, in late imperial China, hierarchical market networks were nested within physiographic macroregions, distinct from the administrative hierarchy of empire and provinces. Skinner’s spatial framework remains hugely influential and has been adapted in more recent comparative studies of historical market integration (e.g., J. Li 2014; Bernhofen et al. 2015; 2018), but critics have questioned the empirical support for that approach.

In this paper, we employ an empirical approach in mapping regional markets using long-term time series data of grain prices. We find evidence of regional market behavior for roughly the middle third of the Qing dynasty, prior to the foreign incursions and development of mechanized transport in the later years of the empire. Our analysis generally supports and refines Skinner's spatial framework for this period but, moreover, provides a new evidence-based method for mapping China's historical spatial structure even as markets (and regions) are reshaped by societal and technological changes.

**Fragmented Markets under a Unified Civilization**

How a populous and prosperous preindustrial civilization succumbed to internal fractures and outside pressures while the Industrial Revolution propelled Western powers to global dominance has been one of the major puzzles of modern world history (Pomeranz 2000). As economist Loren Brandt (1997, 282–83) summarized the “conventional wisdom” of a previous generation of scholars, late imperial China struggled to feed a growing population, with economic innovation constrained by tradition that was only broken by the Communist revolution of 1949; that view of stagnation and increasing inequality has given way to an understanding of spatially differentiated development processes and innovations that “frequently drew on — and were not impeded by — the strengths of China's traditional economy.”
A Civilization of Macroregions?

Those strengths, argued anthropologist G. William Skinner, included the resilience of a national economy that “formed not a single integrated urban system but several regional systems, each only tenuously connected with its neighbors” (Skinner 1977, 211). Over a territory as vast and varied as China, periodic natural disasters or economic depressions in one region could be balanced by prosperous conditions in others, which in Skinner's view were likely to be linked to climatic as well as social cycles.

Skinner argued against “the usual focus on the empire as a whole or on its component political divisions,” contending that the “macrocycles” of Chinese history are best understood in terms of “macroregional economies — not of provinces and not of the Chinese empire per se” (Skinner 1985, 275).

Indeed, most academic studies of China, imperial or contemporary, are framed at the national or provincial level; as historian Paul Cohen (1984, 166) put it, most scholars as well as bureaucrats, “for reasons of both habit and convenience, find it hard to resist dividing China up into provinces.”

These political divisions, like states of the United States or nations of the world, come to feel “natural” in their ubiquity (Ludden 2003, 1058). That convenience could only be improved by reducing the number of units; thus, in the Qing dynasty, eight governors-general served to oversee and coordinate the administration of more manageable groupings of provinces (Barnett 1963; Qu 1962; Watt 1972). Skinner too had previously developed a system of eight provincial groupings for indexing field studies, clustering provinces that largely shared river basins and dominant cropping patterns, as well as linguistic and cultural affinities (Henderson 2004, 52; Skinner 1973, I:xxxii–xxxvii).

But putting provincial boundaries aside, Skinner introduced the concept of “physiographic macroregions” in an essay in his edited volume, The City in Late Imperial China, writing that “In tracing out the overlapping hinterlands of [China's major] cities … I came to the realization that they … coincided with minor exceptions to a physiographic unit.” (Skinner 1977, 211). His analysis of China circa 1893 identified eight such macroregions covering agrarian China (plus Manchuria, or Northeast China), highlighting the core zones of each of these macroregions and delineating boundaries among them based on “transport efficiency and trade flows as well as with physiography per se” (Skinner 1977, 212).
Scholars have subsequently pointed out that the physiographic macroregions and their core zones show “a large measure of agreement” with the marketing districts of the Qing salt monopoly and with Republican-era key economic development areas (Sun 1922; Whitney 1970, 34, 53–55; Cartier 2002, 127), lending support to the premise that “these regions were [at least] partially self-contained areas of social and economic development” (Blunden and Elvin 1998, 25). Historians as well as anthropologists largely embraced the concept of macroregions, hailing Skinner's framework as providing “a new conceptual vocabulary that enables us to look at old problems in new ways and to see connectedness where previously we had been blind to it” (Cohen 1984, 165; see also Crissman 2010). Historical studies that had focused on the social fault lines between central and local, urban and rural, or gentry and peasant could be recast in the language of urban hierarchies, cores, and peripheries.

While Skinner's spatial approach to Chinese civilization and history was informed by specific historical circumstances, he was clearly influenced by the language of systems theory. Skinner cited the central place theory of Walter Christaller in earlier works, then referenced regional systems models dating back to Johann Heinrich von Thünen and invoked Torsten Hägerstrand's spatial diffusion theory as he continued to elaborate the model he ultimately named Hierarchical Regional Space (Christaller 1966; Hagerstrand 1965; Skinner 1964; Skinner, Henderson, and Yuan 2000; Thünen 1966). At the widest scale, the notion of cores and peripheries is a pillar of world-systems theory, which focuses on the relative integration of regions into a global economy (Timberlake 1985; Wallerstein 1974).

Theorist Immanuel Wallerstein characterized entire nation-states as “core,” “semiperipheral,” or “peripheral” based on their functions in global networks of capital; in this view, peripheral states become trapped in an exploitative cycle of underdevelopment as their resources are exploited for the benefit of the core. Subsequent authors hypothesized that certain “world cities” were the critical nodes in these networks, suggesting a bridge to Skinner's mode of analysis of core cities and peripheral hinterlands within states (Friedmann 1986; Friedmann and Wolff 1982; Ryavec 2021; Simon 1995).

However, in contrast to the largely qualitative, Marxist approach of world-systems theory, Skinner is more closely aligned with the regional systems theorists who employ mostly quantitative land use, central place, and time geography approaches to map and measure trade and other social interactions in traditional and developing societies (Fujita, Krugman, and Mori 1999; Marshall 1989; Smith 1976). Regional systems theory highlights historical legacies of uneven development, due in part to the limitations of pre-mechanized transportation technologies, which have resulted in systematic geographical differences in culture and demography, from life expectancy and fertility rates to gender differences in educational access and occupations (Skinner 2003; Skinner, Henderson, and Yuan 2000).
Skinner's formulation of the Hierarchical Regional Space model for China (he also proposed versions for nineteenth century Japan and France) attempts to trace economic activity through multiple nested levels, from metropolitan systems contained within macroregions down to local periodic markets of the type he studied for his doctoral fieldwork in 1949 (Skinner 2017; Skinner, Henderson, and Yuan 2000).

The advantage of this approach for world-systems theorists and historians lies in its ability to guide the construction of detailed Historical Geographic Information System (HGIS) models of societies, settlement systems, and economies so that a more detailed spatial history of changes wrought by increasing modes of capitalism in the Modern Period may be mapped and compared with earlier patterns. Subsequent research has probed the limits of Skinner's regional systems approach to Chinese history and, especially, the degree to which the macroregions he delineated coincide with observable social and economic behaviors (Henderson 2018). “While the macroregion model served to shift attention from the imperial to the local level and to regional history,” argued geographer Carolyn Cartier, “it does not aid in understanding human processes and their variation that underlie urbanization and regional formation, especially social and cultural practices, long-distance trade, and associated economic activities” (Cartier 2002, 81).

A key question for economists has been whether macroregions constituted discrete markets for goods and services or were linked to wider networks. Several studies have examined commodity prices in different urban markets, or other evidence of trade networks extending beyond the macroregional scale. For economists Barbara Sands and Ramon Myers (Sands and Myers 1986, 724), “examples of interregional trade . . . raise strong doubts about the historical independence of each macroregion” (p. 724). They maintain that markets, rather than macroregions, should be the basis for socioeconomic analysis: to understand complex, changing economic systems, the “concept of markets satisfies this need whereas macroregions do not” (Sands and Myers 1990, 346).

**Responding with Empirical Data**

Even researchers who appreciated the elegance of Skinner’s formulation acknowledged that his theory was difficult to test: as sociologist William Lavely put it in a response to Sands and Myers, “more data must be found and better tests devised” (Lavely 1989, 112; see also Little and Esherick 1989). Indeed, in the decades since the publication of *The City in Late Imperial China*, access to Chinese archives from the national to the local levels has expanded dramatically, helping fuel the rise of digital humanities platforms for displaying and querying digitized data sources.
Not long after Sands and Myers (1990) called for markets, not macroregions, to be the organizing principle for historical studies of the Chinese economy, contributors to the edited volume *Chinese History in Economic Perspective* showed what that could look like. Writing that “historians of China may be discouraged from pursuing economic topics because of the apparent lack of data” the volume's editors concluded that, broadly defined, there are “more data than meet the eye” (Rawski and Li 1992, 16). Several of the contributors gravitated toward analyses of grain prices, not only because detailed price data was becoming available from Qing archives, but because of the critical role of grain markets in the economy.

As economist Yeh-chien Wang (1992, 35) wrote, “a clear knowledge of the trends of grain prices will provide not only a key to understanding the state of economy and society but also a basis for further research in real wages, the standard of living, and many other areas.” This would have come as no surprise to Qing bureaucrats, who took a keen interest in recording grain prices as an indicator of social stability and to support the operation of grain reserves intended as a hedge against famine; by the start of the Qianlong period (1736), monthly reporting had become a regular practice (Will and Wong 1991; Yeh-chien Wang 1992; Roehner and Shiue 2000).

Historian Peter C. Perdue demonstrated what could be done with a time series of grain price data in his study of Gansu Province (and one neighboring prefecture), mapping the correlations in the grain price time series for 14 prefectures and finding that a majority of them were highly correlated ($r^2 > 0.8$), especially along major transport routes. This allowed Perdue to demonstrate a price gradient across the province, accounting for the transport cost of grain and the cyclical nature of cropping cycles. Though data on grain trade volumes *per se* was scant or unreliable, the correlated movement of prices is consistent with a market model: “correlation coefficients indicate …a web of market relationships…and accurately define the boundaries of macroregions” (Perdue 1992, 123).

With the benefit of more comprehensive, digitized data and maps, our task is to expand on Perdue's approach to map the correlations among reported grain prices for the prefectures in Qing China and using those results to delineate functional marketing regions. We will then be in a position to see how well those macro-markets align with physiographic macroregions. (Another line of inquiry compares the degree of market integration in China with Europe, typically using a priori regional definitions for markets, including Skinner’s; see e.g., Shiue and Keller 2007; Bernhofen et al. 2015; Keller, Shiue, and Wang 2021.)

The caveat, however, to all data-driven “bottom-up” studies is the variability of outcomes provided by different methods of detecting and generating relationships.
For example, Pearson product-moment correlation coefficients are not the only approach to measuring the strength of a relationship between price trends and not necessarily the best method given the constraints of that test which typically compares two different variables. We discuss our approach below.

**Data, Methods and Analysis**

**Data sources and preparation**

In this study we made use of a digitized version of the *shupi yuzhi* (Vermillion Endorsements and Palace Memorials) archive listing the market prices for various grains by lunar month starting in 1736, the first full year of the Qianlong Emperor's reign (D. Wang 1987; Yejian Wang 2014). These records were compiled from reports sent to the imperial capital from each prefecture (*fu*, an administrative region below the provincial level), and contain the highest and lowest prices, in taels of silver, of the grains among the counties (*xian*) in the prefecture.

Prior studies have examined these data for accuracy and consistency; while there are known cases of falsified data (Perdue 1992 points to a Gansu official who wrote fraudulent reports to abscond with grain relief funds) the importance to imperial officials of maintaining accurate data in a common currency makes these an unusually robust data source for this time period (Roehner and Shiue 2000; Shiue 2013; Bernhofen et al. 2015). The Qing Dynasty Grain Price Database published by Academia Sinica has made these invaluable data available to researchers in digital form (Yejian Wang 2014). Retrieving all available records for each province from the database, we found 260 *fu* with grain price data from the beginning of 1736 through the end of 1842 (see Figure 1).

Nonetheless, the dataset is far from complete: while half of the *fu* represented contained at least some monthly records for 100 or more of the 107 years, on average they reported complete 12-month runs of data for only 53 of those years. (Some of the “missing” years can be attributed to changes in the cities' status as *fu* during the mid-Qing.) The average year saw reports of data from 222 of the 260 *fu*. All but 12 years saw reports from over 200 *fu*; only the earliest years (1736-37) and a few years (1822-24) near the start of the Daoguang Emperor's reign saw reports from fewer than 100 *fu* included in the dataset (see Figure 2).

We matched the *fu* in this data source to the prefectural capitals of 1820 represented in the China Historical Geographic Information System (CHGIS) or, in a few cases, to cities at different administrative levels in 1820 that had served as *fu* seats at other times during the Qing dynasty (Bol et al. 2012).
Geographic matching entailed confirming the consistency of place names within each province in both CHGIS and the Qing Dynasty Grain Price Database (Yejian Wang 2014), correcting minor inconsistencies, and assigning a unique identifier from the ChinaW spatial data file (Skinner, Yue, and Henderson 2007) to each fu. Thus matched, grain price data could be mapped and the spatial relationships among fū could be returned for statistical analysis.

Grain price data availability proved to be unevenly distributed not only temporally but also spatially (Figure 3). We selected the prices for wheat in all prefectures except five southern provinces, Fujian, Guangdong, Guangxi, Guizhou, and Yunnan, where only rice price data are available. Cases could be made for analyzing sorghum or millet instead of wheat, or rice due to the extensive secondary trade in “tribute grain” brought to the imperial capital, but the price records for these are less complete in most provinces. In any case, as Li (1992) showed, various grain prices tended to rise and fall in tandem, and we argue that, as a relative luxury, wheat is more likely to be the subject of interregional trade.

We then constructed a monthly time series of mean grain prices (averaging the high and low prices reported for each fū) for each of the 260 fū with price data from the beginning of 1736 through the end of 1842. We selected 1842 as the end date for this analysis, as mentioned above, to focus on regional trading patterns prior to the widespread adoption of steamships and, later, railroad and telegraph lines that followed from foreign incursions under the so-called unequal treaties following the first Opium War (1839–42) and subsequent Treaty of Nanking (1842).

We do not mean to imply that trading patterns were static prior to that date — clearly, challenges like the shifting courses of the Yellow River, the varying states of repair of infrastructure like the Grand Canal, and the functioning of the Ming and Qing imperial courier networks have changed the costs and speed of transporting goods and market information throughout history (Mostern 2021; Skinner, Henderson, and Yue 2006; Hao, Li, and Nye 2020). Nor, for that matter, would markets remain static throughout the entire 107-year span of this study (see Bernhofen et al. 2018 on changes in regional market integration). Nonetheless, the conditions that prevailed during this time period can provide a baseline against which subsequent effects of mechanized transportation and foreign trade can be measured. Of the 260 fū with valid price data, we restricted this further to only time series with at least 20 percent or more data, resulting in a final collection of 235 viable series.

**Spatial Relationships**

In primarily terrestrial empires like Qing China, macro-markets are likely spatially dependent and contiguous. Traditional overland transportation restricted how much and how far grain could travel.
Therefore, any approach to delineating market regions requires spatial adjacency to be explicitly modeled to delimit which price time series could be plausibly related. Luc Anselin (2003) acknowledged that, within the field of econometrics, standard econometric techniques often fail in the presence of spatial autocorrelation, commonly found in geographic data sets, thus requiring explicit inclusion of spatial relationships within a model. Without spatial constraints, potentially spurious regressions could correlate distant locations based on statistical regularities and yield erroneous regions.

To display the correlations on maps, we constructed Thiessen polygons around the 260 fu points as mapped by CHGIS. Thiessen polygons are space-filling polygons with edges situated midway between neighboring points. We used Thiessen polygons rather than the prefectoral boundaries of 1820 mapped by CHGIS to avoid gaps or inconsistencies with our 1736-1842 dataset. Additionally, Thiessen polygons allow a regional structure to emerge from disaggregated data points and the complex web of their interrelations without interference from pre-existing boundaries.

From the Thiessen polygons, we determined which pairs of fu were neighbors via a spatial weight matrix. We chose to apply a queen's case edge contiguity matrix where a fu is only considered connected to its immediate neighbors that share an edge (Figure 4). The weights were simply reported as a “1” if two polygons shared an edge or a “0” if they did not. Multiplying the spatial weights matrix by the correlation matrix (discussed below) yielded our final set of connections and thus our network model.

Edge contiguity is preferable here as we are investigating the relationship between polygons or an “area of influence” of each fu point rather than the non-hegemonic points themselves. A queen's case contiguity matrix yields a normal distribution of link frequency with a mean of 5.04 connections with few polygons having few or many neighbors. Isolates detected in the northeast and Taiwan (then a prefecture of Fujian province) were manually linked for the final analysis.

While other adjacency matrices such as K nearest neighbors or inverse distance might also be viable, a more complex theory of how prefectures were connected would be required to account for connections that might “leapfrog” over other prefectures given traditional transportation technology. River and coastal transport could possibly connect very distant prefectures though, even then, not without stops along the way.

**Time Series Correlation Analysis**

We generated a correlation matrix between each prefecture pair using the cross-correlation method to elucidate the strength of their market relationships. While prior studies of grain price relationships (e.g., Perdue 1992) have calculated Pearson product-moment correlations among prefectural average prices, that method may find spurious correlations between time series or, at the very least, inflate the resulting coefficients.
This is particularly concerning when the same variable is tested (see Yule 1926; Granger and Newbold 1974). By contrast, cross-correlations integrate the temporal dimension inherent in time series data and are more apt to find similarities and differences in price dynamics, especially lagged dynamics, rather than simply comparing distributions (see Shen and Zheng 2009 comparing contemporary Chinese stock market indices).

Standard Pearson correlations have three main disadvantages. First, auto-correlation and time-dependencies of times series are explicitly handled in time series analysis and cross-correlation tests. In contrast, the standard linear correlation tests do not adjust for internal dependencies since independence is assumed. These dependencies can inflate coefficients (0.9 and above) when the signals only share a fraction of parallel fluctuations. Second, standard Pearson tests can often mistake trends in absolute price with relationships. Two time series can exhibit almost identical dynamics, but changes in absolute prices over time (a trend) can yield a strong negative coefficient. In contrast, cross-correlations on detrended data reveal that these time series are highly correlated. Third, even when using standard Pearson coefficients as edge weights in a network graph, we received a much lower modularity score (strength of cluster connections on a scale of −1 to 1) than when using cross-correlation coefficients. Especially without the application of a spatial weights matrix to constrain connections, the community detection algorithm reported a 0.38 modularity score for standard Pearson correlations while the cross-correlation coefficient network received a 0.63 modularity score. Thus, even without modeling the spatial relationships, clusters emerged from the cross-correlation edges in which the prefectures within them were more correlated to each other than other clusters. This indicates that spurious regressions and inaccurate correlations are more common when using standard Pearson tests.

Therefore, we elected to calculate cross-correlations, which are specifically designed to find correlations in time series such as stock prices or signals at different time lags. Before calculating the correlations, we detrended the time series via differencing $t_x - t_{x-1}$ as advocated by Dean and Dunsmuir (2016). We took a first difference for each time series to make them stationary then found the cross-correlation coefficients for lags −1, 0, and 1, though we primarily looked at lag 0 as we are not investigating time lags in grain price changes here. Most lagged correlations of strength primarily exist between neighboring $fu$ and are very scarce. However, these asynchronous correlations warrant a closer look in future research to determine the effect of time-lagged transportation.

Ultimately, we calculated the cross-correlations between the time series of every pair of neighboring $fu$, excluding pairs that were not spatial neighbors by multiplying the spatial weights matrix to all possible connections without duplicate or self-connecting pairs (A-B=B-A) ($n = 27469$) and excluding adjacent, negative correlations ($n = 7050$).
All cross-correlations with p-values greater than 0.05 were removed as well, yielding a total of 498 spatially valid, positive connections among the 235 fu.

**Price Networks**

Visualizing and connecting these relationships was accomplished by constructing an undirected network graph. The network graph and subsequent network analysis was a necessary intermediate step before creating a legible map of macro-markets given reproducible induction is impossible when presented with so many heterogenous connections. Nodes in the graph represent the fu while edges represent the cross-correlation coefficient at lag 0 between them and thus the “weight” of each edge and the strength of each relationship.

Figure 5 shows the undirected network graph of price correlations among all 235 selected fu. The largest connecting lines indicate the strongest correlations between adjacent fu (nodes) while progressively thinner lines indicate medium to weak connections; five classes are produced using the “natural Jenks” method, which is intended to convey the maximum information about the distribution of data across a limited number of classes by maximizing the variance between classes and minimizing the variance between classes (Jenks and Caspall 1971).

At first glance, we can see clusters of strong positive correlations within provinces. For example, the twelve fu of Fujian province show strong correlations among its members with a mean correlation coefficient of 0.46, which is above the global mean. Although some Fujian fu have connections to neighboring provinces, these are much weaker than intra-cluster correlations, with a mean of 0.14 and with the strongest of those being 0.36.

However, not all strong price correlations are limited to provincial boundaries. The largest mass of strong correlations (the red and orange cluster at center-right) straddles Anhui, Jiangsu, and Zhejiang provinces in the Lower Yangzi region, which refutes the hypothesis that fu are most strongly correlated with provincial neighbors. In fact, Skinner argued one reason the imperial Chinese field administration intentionally carved-up rich core regions such as the Lower Yangzi among different provincial units was to prevent collusion among rich merchants and powerful gentry in challenging imperial power (Skinner 1977, p.342-343).

**Community Detection and Finding Clusters**

With nodes connected via edge weights, we proceeded to apply a community detection algorithm to determine which groups of spatially adjacent nodes were closely related and thus belonging to a cluster or, by extension, a macro-market.
A common criticism of regions identified according to one specific clustering method is that other methods may show significantly different regional patterns (Yang, Algesheimer, and Tessone 2016). To test the sensitivity of our approach to changes in algorithm, we compared the outputs of two clustering algorithms — Louvain and Spinglass — that have become widely used in the past decade or so.

The Louvain algorithm, named for Louvain University, was proposed by Blondel et al. (2008) as an efficient method for maximizing the modularity of communities (clusters or macro-markets in this case) in a large network. In other words, the goal of this and any other community detection algorithm is to partition nodes into dense communities where their members are highly connected within their clusters and weakly connected to other clusters. The Louvain algorithm does this by first placing all nodes into their own arbitrary communities and then iteratively places each node into other communities and tests if there is an increase in modularity. If there is an increase, then that node changes communities and if it leads to no increase, then the node stays in its initial community.

Community detection and the resulting modularity scores are common in network research, but there is substantial evidence that modularity is not a completely “objective” measure and that algorithms that seek to find a global maximum module arrangement could potentially find “bad communities” (Kehagias and Pitsoulis 2013). In addition, missing data reduce the number of potential connections and thus potentially produce a different network structure. However, we found that the differences in the algorithms used, and even the use of K-nearest neighbor spatial adjacency matrix to increase the transitivity (the probability of a node to have connections to neighbors in a network graph), made only small differences to the cluster landscape. The same eleven clusters, or slight variations of them, emerged from many trials (Figure 6a). What is obscured in a cluster map produced by the Louvain algorithm, but evident in the network graph, is a certain amount of fuzziness around the borders of each cluster. If an edge's value changes as the result of a perturbation in the network, clusters may change the result.

The Spinglass algorithm (Reichardt and Bornholdt 2006) purposely introduces perturbations — stochastic noise — to the edge values to find, among many possible states, an optimal community structure with high modularity. The algorithm takes its name and logic from a spin glass which, in contrast to a crystal or ferromagnet, is very heterogenous which reflects the structures of many real-world complex systems. The algorithm simulates an annealing process of “heating” up or making many, large changes to the edge values, then “cooling” or settling them until “stronger” modularity results. Because the algorithm must simulate random noise and thus each run is likely to be at least slightly different, we ran this algorithm 100 times and took the most common number of clusters, which turned out to be 12 (Figure 6b).
Results suggest that there are relative “weak points” in the larger clusters and thus, given enough noise, they may split into two systems, perhaps, of the same larger macro-community. This implies that there is a hierarchical structure of communities. Whether the strong connections around, say, Xi’an or Kaifeng are identified as their own market clusters or are subsumed into a larger cluster may change with perturbations to the complex network. However, even given stochastic changes to the edge connections, the same general pattern emerges, and in multiple runs the same clusters are split or subsumed. Radically different arrangements are, at best, very unlikely in this robust model.

**Markets and Macroregions of a Civilization**

The eleven clusters — or macro-markets — that emerge in our analysis bear some striking similarities to Skinner's physiographic macroregions. Figure 7 superimposes these two sets of boundaries for systematic comparison. We will address each region in turn with reference to the results of the Louvain and Spinglass to describe the spatial extent of market clusters and the cross-correlation analyses to characterize the relative level of market integration within and across regional boundaries.

**North China: Beijing, Kaifeng and Taiyuan**

We find the greatest discrepancies from Skinner’s physiographic regionalization in North China. Rather than one integrated system, we find three distinct market systems which roughly correspond to the hinterlands of (1) Beijing; (2) Kaifeng, encompassing parts of Henan and Shandong provinces; and (3) Taiyuan, in Shanxi province. The North China cluster around Beijing has strongly interconnecting nodes within its regional market with moderate cross-market connections to the Kaifeng cluster, but no direct relationship was found with the Taiyuan cluster (see Figure 5).

The cluster around Kaifeng appears to be moderately connected to the Beijing cluster and to the Lower Yangzi, but weakly connected to the Middle Yangzi and, again, unconnected to the Taiyuan cluster in Shanxi. Shanxi is almost an island, only weakly connected to the Middle Yangzi and Northwest China.

The spatial pattern of the strongest level of grain price correlation coefficients \((r \geq 0.8)\) extended from Beijing south along the route of the Grand Canal. It seems likely that this pattern stemmed from the trade in grain along this transportation corridor connecting the Lower Yangzi with Shandong and the imperial capital. Originally built to transport tribute tax grain from the Lower Yangzi to Beijing, the Grand Canal facilitated the movement of many goods while reducing transport costs between north and central China, supporting the growth of long-distance trade. In fact, nearly a third of China’s customs duties in 1753 derived from this canal trade (Cheung 2008, 6).
As imperial officials and soldiers in this region received some of their salary in tribute tax grain, a thriving secondary market in grain would have affected prices in the immediate vicinity. But the sharp discontinuity with grain prices in Henan suggests that market did not extend up the Yellow River to Kaifeng or over the mountains to Taiyuan, whether due to higher transportation costs or to policies restricting trade.

**Northwest China, the Middle Yangzi, and the Upper Yangzi**

The grain market in Northwest China, as seen from strong positive correlations around Xi'an, centered on the valley of the Wei River, a tributary of the Yellow River and the cradle of Chinese civilization. Distant markets across the Gobi in Ningxia and the oasis cities of the Gansu corridor to the west were moderately integrated into one regional market, consistent with Perdue's detailed study of Gansu province. The grain price time series of two prefectures in the Ordos Plateau are more highly correlated with Beijing than the rest of Northwest China, despite being separated from North China by the poorly integrated markets of Shanxi; this suggests either that prices in both areas were affected by similar weather patterns, or possibly by the existence of a frontier market in grain with largely pastoral Inner Mongolian areas not included in the Qing price data nor part of Chinese civilization traditionally.

Just southeast of Xi'an, the prefecture of Shang Zhou shows a moderate correlation with Xi'an, but even stronger correlations with two neighboring prefectures in the upper reaches of the Han River (a tributary of the Yangzi). This suggests some interregional trade across the divide between the Yangzi and Yellow River basins. But we find much less evidence for trade across the physiographic divide separating Northwest China and the Middle Yangzi macroregions from the Upper Yangzi (encompassing the Sichuan Basin around Chengdu). Correlations between neighboring prefectures across that divide are nonsignificant, weak, or negative, apart from the link along the Yangzi itself, bifurcated by the Three Gorges. While the Louvain algorithm keeps the Sichuan Basin prefectures together in a cluster around Chengdu, the Spinglass algorithm splits off a second cluster including Chongqing on the Yangzi River.

Within the Middle Yangzi macroregion, Changsha (see Map 5) and Nanchang emerge as hubs of integrated regional markets less strongly correlated with Hankou (now part of Wuhan) and other prefectures north of the Yangzi. The Louvain and Spinglass algorithms group the Hankou and Changsha networks together, separate from the Nanchang cluster. (Skinner alternately identified the Gan River basin, including Nanchang, as a separate macroregion or as a subsystem of the Middle Yangzi macroregion in later works: see Skinner 1985; Skinner, Henderson, and Berman 2013.)
The Lower Yangzi and Southeast Coast

In the Lower Yangzi, the most advanced economic region of the Qing dynasty, we see a network of strongly correlated neighboring prefectures, suggesting a well-integrated grain market centered on Suzhou. Those strongly correlated pairs extend to the northern boundaries of Anhui and Jiangsu provinces—further north than Skinner's physiographic macroregional boundary, which reflected the course of the Yellow River as it ran south of the Shandong peninsula until 1855. (This accords with Pomeranz 1993; as well as Skinner’s twentieth century analysis: see Skinner, Henderson, and Berman 2013.) We also find stronger pairwise correlations linking southern Zhejiang province with the Lower Yangzi (Skinner drew that area into the Southeast Coast); Chuzhou is equally strongly correlated with Fuzhou (on the Southeast Coast) and Nanjing (on the Yangzi), but with stronger northward correlations, our analysis places it in the Lower Yangzi.

For its part, Fuzhou is the hub of a strongly correlated network extending across Fujian province, including across the strait to Taiwan. The moderate to strong correlations with Fuzhou's time series seen from Guangzhou, northern Jiangsu province, and up the Yangzi to Wuhan suggests an active waterborne grain market extended along the coast and major waterways. But the mountains separating Fujian from Jiangxi provinces posed a barrier to that trade, with weak or negative correlations crossing that physiographic divide.

Lingnan and Southwest China

Almost all of the major prefectures of Skinner's Lingnan microregion — more or less conterminous with Guangdong and Guangxi provinces — showed moderate to strong correlations with the grain price time series for Guangzhou. An exception is Chaozhou, on the border with Fujian, which Skinner drew into the Southeast Coast macroregion; its correlation with Guangzhou is weak but its correlation with Fuzhou is nonsignificant, so our analysis places it in the Lingnan macro-market.

The neighboring prefectures of Lianshan Ting and Liao Zhou (present-day Lianzhou) make for another exception, displaying a nonsignificant correlation with Guangzhou: Lianshan shows strong correlations heading north into the Middle Yangzi, while Liao Zhou's stronger ties connect eastward toward the Southeast Coast.

Skinner grouped most of Guizhou and Yunnan provinces into what he called the Yungui macroregion, which others have variously described as a “nascent,” “remnant,” or “frontier” region (Blunden and Elvin 1998; Wright 2000).
We label this region “Southwest” China, though our dataset does not include sufficient grain prices from Yunnan to calculate correlations in this period, but the time series for prefectures in Guizhou, despite limited temporal coverage, display a regional market of strongly intercorrelated markets around Guiyang and sharp discontinuities across most of the physiographic divides surrounding the province. Consistent with economic historian Tim Wright's study of southwestern China in the early twentieth century, we find that Guizhou was perhaps the region least integrated into a national market for grain (Wright 2000).

The westernmost prefecture of Pu'an is strongly correlated with Guiyang but also shows a strong correlation down the Xi River towards Guangzhou, suggesting that it was a portal for interregional trade. Interestingly, the Spinglass algorithm (Figure 6b) places the Guizhou prefectures in a cluster with Guangxi, while the Louvain analysis (Figure 6a) paired Guangxi with Guangdong, as Skinner had it.

Along the provincial border with Sichuan, Skinner had included Renhuai prefecture in the Upper Yangzi macroregion, but in this analysis it shows a strong positive correlation with Guiyang and negative or nonsignificant correlations to the north. Songtao prefecture, along the Henan border, is the only Guizhou prefecture with a strong correlation with its neighbor in Sichuan, but it has stronger correlations with its immediate Guizhou neighbors. The band of western Hunan prefectures with insufficient data complicate the analysis, but skipping over those immediate neighbors, the prefectures on the eastern edge of Guizhou have very weak correlations with the next prefectures to the east, suggesting that the regional boundary in this area is not simply an artifact of imperfect data.

**Implications and Prospects**

While we have found some strong evidence of interregional trade — identifying prefectures with strong grain price correlations across physiographic boundaries — we depart from critics like Sands and Myers (1986) who aimed to refute Skinner’s macroregional model altogether. Rather, the prevalence of sharp discontinuities in grain prices along physiographic boundaries, with weak, nonsignificant, or even negative correlations, generally supports Skinner's theoretical framework. In addition, cross-correlations show that one-month lags do not exhibit strong relationships and only very few cross regional boundaries.

In short, there is little evidence that paying greater attention to long-distance trade would result in a dramatically different regionalization. Correlations along the Grand Canal, major rivers, and coastline indicate that interregional trade routes allowed market signals — and grain supplies, per se — to mediate prices in the most populous regions and helped account for some of the long-term social stability in the imperial period, as Skinner (1985) anticipated.
We differ from Skinner in finding that provincial boundaries are more salient in defining several regional markets. Shanxi and Guizhou provinces have weak external correlations and strong internal correlations, even across physiographic barriers identified by Skinner. So too we find the Southeast Coast macro-market to be conterminous with Fujian province, excluding the neighboring prefectures from Zhejiang and Guangdong that Skinner considered as part of the Southeast Coast macroregion.

It may be that provincial policies governing grain markets—or practices affecting how prices were reported at the prefectural level in this dataset—explain these spatial patterns. Alternatively, in cases where the delineation of macroregional boundaries was more of a judgement call—as in the Southeast Coast, which consists of multiple coastal river basins—Skinner may have avoided following provincial boundaries to emphasize his point about the difference between market and administrative geographies, but in our dataset the two turn out to be more closely aligned.

This analysis considers correlations among the 1736-1842 time series of prefectural grain prices and finds evidence for regional macro-markets—patterns in space—but does not attempt to examine diachronic change—patterns in time. This reflects a necessity of the mode of analysis: we need to assemble a fairly lengthy time series to detect significant correlations.

But, of course, the Chinese economy at the dawn of the Qianlong emperor’s reign differed in many respects from that of the era of the Opium Wars. Economist Daniel Bernhofen and colleagues suggest that regional markets in China were trending toward a lower degree of national integration by 1820 (Bernhofen et al. 2015). Economists Yanfeng Gu and James Kai-sing Kung argue that market integration followed a “U” shape: markets were more integrated in 1736, then population growth in grain-producing areas strained integration by consuming potential exports, and only after the Taiping Rebellion did market integration rebound (Gu and Kung 2021).

Further work will be needed to determine how the regional markets we identified here were in the process of evolving and how they changed with the advent of treaty ports and mechanized transportation later in the nineteenth century. Statistically, each decade is likely to have, even slightly, different correlation coefficients, connections, and thus regional structures.

The macro-markets that we have identified suggest that Chinese civilization during the Qing period indeed consisted of multiple regional markets, many of which were nested within physiographic macroregions and/or within provincial boundaries. Some were “tenuously connected with their neighbors” (Skinner 1977, 211), some more so, some hardly at all.
While further work will be needed to identify temporal trends, this analysis supports the conception of spatially differentiated markets undergoing uneven growth processes (Brandt 1997, 282–83), with regional, somewhat hierarchical markets shaped in part by interactions with the administrative state. Fortunately for this kind of analysis, that state kept rather good price records even if it did not survive the endogenous and exogenous shocks that were yet to come.

**Statements and Declarations**

The authors declare no competing financial or non-financial interests. Research for this paper did not involve humans or animals.

**References**


Figure 1. Map of prefectures with Qing grain price data.

- • Prefectures with archived grain prices
- --- Qing province/dependency boundary
- || Provinces with archived prices for rice, all other Qing provinces archived prices for wheat
Figure 2. Bar chart of total number of reporting prefectures per year. Vertical lines show the beginning of the Daoguang Emperor's reign (1820) and the beginning of the Taiping Rebellion (1850-64), potential causes of missing data.
Figure 3. Thiessen polygons of prefectures (fu) shaded by relative monthly grain price data availability.
Figure 4. Map of edge contiguity matrix applied to distinguish spatial neighbors using the “queen’s case” method.
Figure 5. Network mapping of adjacent nodes. Line size and shade indicates cross-correlation coefficient strength. Data categorized using the natural Jenks method.
Figure 6. Maps of a) 11 clusters produced by the Louvain algorithm and b) 12 clusters produced by typical run of the Spinglass algorithm.

(a)
Figure 7. Grain macro-markets of 1736-1842 and Skinner's physiographic macroregions.