

NBER WORKING PAPER SERIES

CAPITAL MARKETS AND GRAIN PRICES:  
ASSESSING THE STORAGE APPROACH

Wolfgang Keller  
Carol H. Shiue  
Xin Wang

Working Paper 24388  
<http://www.nber.org/papers/w24388>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2018, Revised January 2019

We thank Stephen Broadberry, Howard Bodenhorn, Jan Luiten van Zanden, and participants at the CEPR Economic History Symposium 2018 for comments. Support by the National Science Foundation (SES grant #1124426) is gratefully acknowledged (Keller and Shiue). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Wolfgang Keller, Carol H. Shiue, and Xin Wang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Capital Markets and Grain Prices: Assessing the Storage Approach  
Wolfgang Keller, Carol H. Shiue, and Xin Wang  
NBER Working Paper No. 24388  
March 2018, Revised January 2019  
JEL No. G12,G21,N2,N61

**ABSTRACT**

This paper evaluates a well-known approach from the economic history literature that uses grain prices to shed light on interest rates. Although this method has been applied in influential work starting with McCloskey and Nash (1984) and has potentially wide applicability in situations where interest rates are not available, this paper provides the first analysis of how well the storage cost approach captures actual capital market performance on a number of different dimensions. Using matched data on bank interest rates and grain prices for early 19th century U.S. regions, we find that the storage cost approach is useful for quantifying the performance of capital markets. The storage cost approach captures well regional differences in market performance, also in comparison to interest rate figures derived from the prices of bills of exchange. The paper also assesses the storage cost approach's robustness to measurement error, incomplete information, outliers, and other factors.

Wolfgang Keller  
Department of Economics  
University of Colorado, Boulder  
Boulder, CO 80309-0256  
and NBER  
Wolfgang.Keller@colorado.edu

Xin Wang  
Institute of New Structural Economics  
Peking University  
Beijing 100871  
China  
xin.wang@nsd.pku.edu.cn

Carol H. Shiue  
Department of Economics  
University of Colorado at Boulder  
Boulder, CO 80309  
and NBER  
carol.shiue@colorado.edu

# 1. Introduction

Capital markets play an essential economic role in channeling current income to projects that pay returns in the future. Interest rates are important indicators of both the scarcity of capital and the riskiness of capital market transactions. And yet, in many historical contexts it is difficult to quantify the development of capital markets in an economy because there are no interest rate records for a sufficiently large number of transactions for which comparable averages can be computed.<sup>5</sup> We do know, however, that even before the appearance of bank data, farmers often traded grain for cash, a basic inter-temporal transaction that allows us to use grain price movements over time to shed light on interest rates. This argument forms the basis of McCloskey and Nash's (1984) influential analysis of English Medieval interest rates.<sup>6</sup> In the early 19<sup>th</sup> century, data for both grain prices and bank interest start to become available in the United States. This paper uses these data to assess the validity of the storage cost approach by asking how close grain prices come to revealing information about actual capital markets at this time.

The approach of employing grain prices to analyze capital markets relies on the idea that in equilibrium, holding grain in storage to sell at a later point will be no more or less profitable than selling grain immediately, which in turn implies that the rate of grain price appreciation is close to the prevailing rate of interest (plus other storage costs). We therefore refer to this as the storage cost approach. Although the storage cost approach is theoretically well-founded both in terms of asset pricing (Working 1933, 1949, Kaldor 1939, Samuelson 1957) and in the analysis of commodity storage (Williams and Wright 1991), to date it has not yet been established how accurately the storage cost approach describes various aspects of capital market development.

---

<sup>5</sup> Typically, interest rate quotes for pre-modern economies cannot be used for systematic comparisons because they omit information on borrower identity, security, and other determinants, and there tend to be too few that are strictly comparable; e.g., Pomeranz (1993), p.32. See also the comparison of bank rates and bills of exchange data for direct evidence on this, Tables A.1 and A.2 in the appendix.

<sup>6</sup> For example, on Chinese farmers trading grain back and forth see the memorial from Tang Pin for the case of 18<sup>th</sup> century China, *Da Qing li chao shilu*, Gaozong reign, 286: 24b-25a (4154-55); Pomeranz (1993), p.32. The link between the intertemporal market in agriculture and other parts of the economy is also confirmed in the description of Chen's (2010) description of the Xu family in Fujian (Chen 2010, p. 433, based on Lin and Liu 2006). Also see Zhang (1996), Pan (1996) on rural borrowing and merchant credit. Outside of China, we know that intertemporal trade in agriculture was prevalent at English town markets and fairs that had been in operation already over the 16<sup>th</sup> and 17<sup>th</sup> centuries (Everitt 1967). There, travelling merchants and salesmen purchased in advance grains and other goods, connecting the village peasant to capital markets. Additional anecdotal evidence on the connection between grain prices and capital markets in England is given in Brunt and Cannon (2009, pp. 34-35).

This paper fills this gap by studying regional capital market performance with matched grain prices (Jacks 2005, 2006, and Shiue and Keller 2007) and interest rates (Bodenhorn 1992, Bodenhorn and Rockoff 1992) for U.S. regions in the early 19<sup>th</sup> century.

We find that although it can be difficult to estimate interest rates specific to a particular time and place, the storage cost approach produces a reasonable estimate of broad interest rates. Furthermore, the storage cost approach captures well differences in the capital development of regions, both in terms of their interest rate levels and in terms of the integration of capital markets. The paper also finds that the storage cost approach is quite robust to the limitations that are often present with historical data.

To be sure, the storage cost approach does not yield results that are as accurate as studying capital market performance with comparable interest rate data. Furthermore, because our analysis requires interest rates to which the storage cost-based rates can be compared, by construction we evaluate the storage cost approach at a time when comparable interest rates are present. Our analysis is thus based on the hypothesis that if grain prices provide valid information on capital market performance right after comparable interest rates have become available, continuity in the development of agricultural markets means that grain prices are also informative just before comparable interest rates are available.

This paper makes two contributions. First, by evaluating a method to assess capital market performance that does not rely on the availability of comparable interest rates, this paper helps to raise the credibility of the method for research on earlier periods when grain prices are available but not bank data—perhaps the Middle Ages or earlier—a significant gain, since high-quality regional bank interest rate data becomes typically available only in the late 19<sup>th</sup> century (e.g., Mitchener and Ohnuki 2009 on Japan in the late 19<sup>th</sup> century).<sup>8</sup> Because of this promise, the storage cost approach has attracted much interest among researchers (including McCloskey and Nash 1984, Taub 1987, Pomeranz 1993, Brunt and Cannon 1999, 2009, Clark 2001, and Shiue 2002); and yet, an empirical validation of the approach has been lacking.

Second, this paper sheds new light on the extent to which individual behavior is driven by incentives in line with economic optimization in historical contexts. If the storage cost approach is

---

<sup>8</sup> An alternative is to use proxies other than interest rates to study capital markets. Outside of London, e.g., Buchinsky and Polak (1993) consider the number of property transactions to study capital markets in England.

misspecified, be it because farmers do not store grain as an asset, or because high frictions in the capital market prevent arbitrage, or even because the farmer in the historical context does not act as *homo oeconomicus* for other reasons, the storage cost approach will lead to very different results as those directly based on interest rates.<sup>9</sup> Along these lines, our paper also aims to quantify the size of the barriers that stood in the way of the frictionless economy and economically rational decision-maker, as specified in an optimal storage model.

## 2. Intertemporal arbitrage and the costs of storage

This section formalizes our approach by establishing the relationship between grain prices and interest rates as one element of storage costs. Consider a farmer in region  $i$  who must decide between selling a unit of grain in period  $t$  for the current market price  $P_{it}$ , or storing the same unit and selling it at  $t + 1$ , for forward price  $F_{it,t+1}^k$ . Selling in period  $t$  would give the farmer revenue that could be used to buy consumption goods, for example. In equilibrium, the following no arbitrage condition must hold for any forward contract  $k$ :

$$(1) \quad F_{it,t+1}^k = P_{it}(1 + \mu_{it} + \varphi_{it}^k + s_{it} - b_{it} + \omega_{it})$$

where  $\mu_{it}$  is the risk-free interest rate in region  $i$  and period  $t$ ,  $\varphi_{it}^k$  is transaction-specific risk,  $s_{it}$  denotes physical storage cost, and  $b_{it}$  denotes the convenience yield. The term  $\omega_{it}$  is a wedge that captures potential barriers between the intertemporal agricultural and other parts of the region's capital market. Furthermore, in this simple framework we abstract from inter-regional trade. Allowing for grain markets to be linked across regions would lead to additional no-arbitrage equations, as discussed, for example, in Shiue (2002). The next step is to take the average over all transactions  $k$  and substitute the future spot price,  $P_{it+1}$ , for the average of the forward prices  $F_{it,t+1}^k$ , which are unobserved. Given these assumptions, we see that the simple intertemporal no-arbitrage condition (1) implies that

---

<sup>9</sup> Komlos and Landes (1991, p.43), for example, criticize McCloskey and Nash's (1984) application of the storage cost approach as anachronistic and forgetting the "social, cultural, intellectual, and institutional realities of the past."

$$(2) \quad \frac{P_{it+1}}{P_{it}} = (1 + \mu_{it} + \varphi_{it} + \epsilon_{it}),$$

that is, the price gradient  $P_{it+1}/P_{it}$  is an increasing function of the risk-inclusive interest rate ( $\mu_{it} + \varphi_{it}$ ) plus other factors ( $\epsilon_{it} = s_{it} - b_{it} + \omega_{it}$ ).

This intertemporal no-arbitrage relationship is at the center of models of optimal storage. The following presents results from a simple competitive storage model along the lines of Williams and Wright (1991). To further simplify, we assume that agents have perfect foresight and the world is deterministic.<sup>11</sup> Figure 1a depicts the equilibrium sequence of prices and storage levels in each period, given other parameters such as physical storage costs, storage capacity, the cost of injection (harvest), and withdrawal, as well as the implied return holding inventories. Notice that both prices and storage levels follow a cyclical pattern. Prices are at their low point once the harvest has come in, and they rise during the period between harvests. Prices must rise during this time because holding grain means not to have the cash the grain can sell for, and the grain price increase has to be in line with the return to postponing consumption (or, capital).<sup>13</sup>

Figure 1b compares the price pattern of two economies, one with a higher interest rate and one with a lower interest rate. The price gradient in the high-interest rate economy is steeper than in the low-interest rate economy. This confirms equation (2) and shows that optimal storage implies, all else equal, the steepness of the price gradient is increasing in the economy's interest rate. The intuition is that when the interest rate is higher, the value of grain between two harvests must rise faster because the opportunity cost of tying up resources is higher.

---

<sup>11</sup> The framework can be extended to include expectation formation and stochastic shocks without altering key relationships.

<sup>13</sup> Storage levels hit zero when prices reach their maximum, indicating that storage takes place to reduce price fluctuations.

**Figure 1a: Storage model with low interest rate**



**Note:** Prices are on the left axis, storage levels on the right.

**Figure 1b: Storage model with high interest rate**



Taking the model to data, one would not expect that the relationship between grain prices and interest rates is always as shown in these figures. First, the patterns shown in Figures 1a and 1b are specific to a particular set of model parameters (governing physical storage cost, storage capacity, etc.), and the exact shape of the cyclical patterns can change for different model

parameters. Second, there are additional factors affecting grain prices, such as the transaction specific risk or capital market imperfections shown in equation (1)). Ideally, in the empirical analysis those variables are explicitly included, however, because many of these factors tend to be unobserved in historical settings, we start out in our baseline analysis by making a number of assumptions about them. In particular,  $\omega_{it}$ ,  $\rho_{it}$ ,  $\varphi_{it}^k$ ,  $s_{it}$  and  $b_{it}$  in equation (1) are allowed to vary between regions and over time, but we restrict these influences to be uncorrelated, i.e. they are white noise. Further below, we allow these influences to vary in a number of ways to assess the robustness of the storage cost approach.

Third, we recognize that grain prices are affected by factors that are not captured by the model underlying Figures 1a and 1b and the no-arbitrage equation (1). One factor is weather shocks, and there could also be systematic variation in the relationship between interest rates and grain prices due to a variety of reasons, for example for political reasons or because of endogenous default. The analysis below will address these potential factors by conducting a number of important extensions.

### **3. Data**

Central to our assessment is to see how well the results from grain price variation together with the storage model match up with results based on directly available interest rates. For the baseline analysis we employ bank interest rates for six U.S. regions during the years 1815 to 1855. This analysis is complemented by alternatively utilizing interest rates for bills of exchange in the U.S. during the years 1836 to 1855. Results based on these interest rates are compared with results from the storage cost approach.

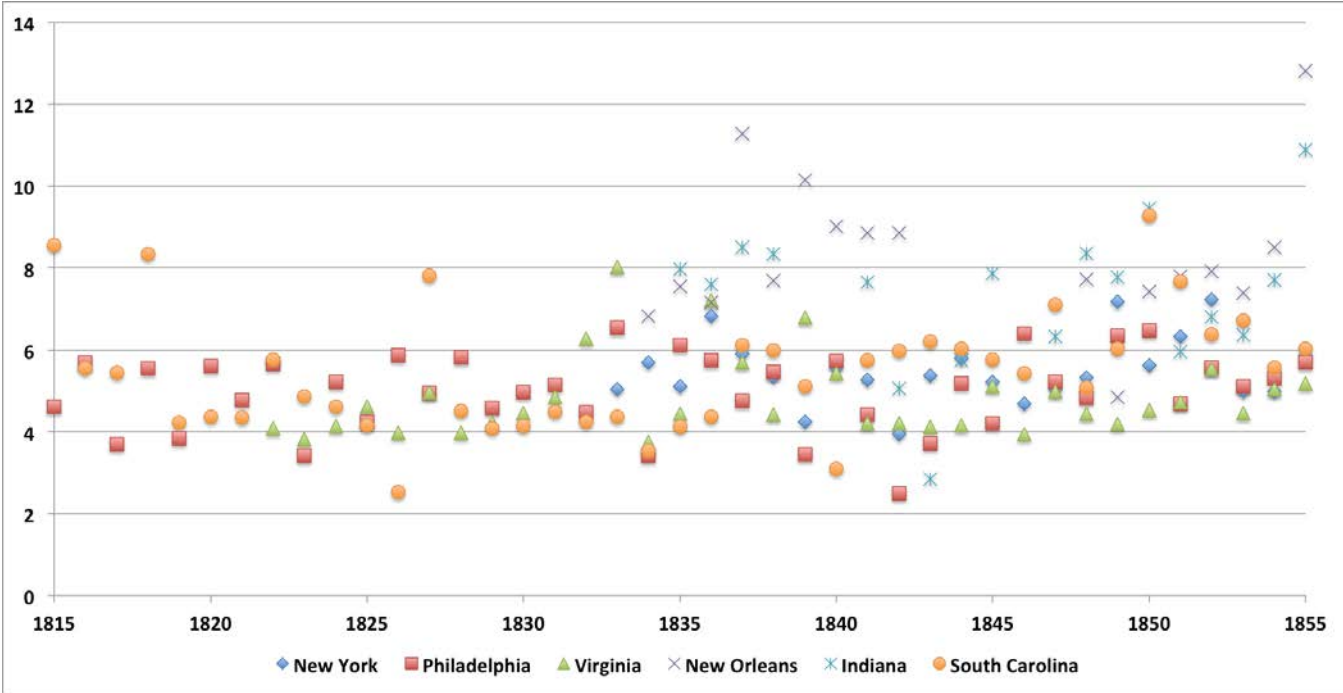
#### **3.1 U.S. Early Regional Capital Markets Data**

For our regional bank interest rate data, we rely on the pioneering work by Bodenhorn (2000) and Bodenhorn and Rockoff (1992). For the earlier part of the 19<sup>th</sup> century, these authors have estimated annual bank interest rates for a number of U.S. regions: Philadelphia, New York City, Indiana, South Carolina, Virginia, and New Orleans. It is apparent that some of these regions are cities and others are states, which means that there is a mix of regions in terms of size in the



sample. This will typically be the case in actual applications. In the following, we will typically refer to a series by the name of the corresponding city for which we have grain price information.<sup>14</sup> Figure 2A shows the bank interest rates that we employ in the benchmark analysis; the values are tabulated in Table A.1 in the Appendix.

**Figure 2A : U.S. Bank interest rates, 1815 - 1855**



**Notes:** The source of the data is Bodenhorn and Rockoff (1992), Table 5.2.

Overall the interest rate on average in our sample has been equal to 5.7 percent, with a standard deviation of 1.68 percent (see Table A.1). The average across all years ranges on the low side from 4.82% and 5.00% (for Alexandria and Philadelphia, respectively), and on the high side from 7.35% and 8.33% (for Indianapolis and New Orleans, respectively). There is also substantial year-to-year variation; for example, the interest rate in New York City moved from 5.32% in year 1848 to 7.17% and then 5.62% in the two following years. The data availability varies, ranging from a minimum of 21 to a maximum of 41 annual observations.<sup>17</sup>

<sup>14</sup> Indianapolis in the state of Indiana, Alexandria in the state of Virginia, and Charleston in the state of South Carolina.  
<sup>17</sup> One might be concerned that the coverage in terms of years varies across cities, but in fact the correlation of the average rates for all years and for common years is high (99.5%).

We complement this data from early U.S. banks with interest rates calculated from U.S. bills of exchange (Bodenhorn 1992). As explained by Bodenhorn (1992), holders of these bills often did not want to wait (typically 30 or 60 days) for funds, so intermediaries often purchased the bills at a discount that represented the implicit interest charge, the costs of collection, and a commission charge. These interest rates derived from bills of exchange data is available for New York City, Philadelphia, Charleston, and New Orleans, that is, four of the same regions for which there are also bank interest rates. Across the four regions, bills of exchange interest rates average between 7.1% for New York City to 10.1% in New Orleans, see Table A.2 in the Appendix.

Note that the bank interest rates are not based on actual transactions at these banks but rather, they are based on the bank's balance sheets and dividend data. Furthermore, due to missing data, Bodenhorn and Rockoff (1992) have to make a number of simplifying assumptions in their estimation of interest rates.<sup>19</sup> For this reason, the bank rates are likely to contain measurement error. If the measurement error were not systematic (classical) it would tend to lower the correlation of bank rates and storage cost rates. However, the mismeasurement in the bank rates could also be systematically related to the grain prices, for example in periods with high inflation. In the benchmark analysis we abstract from bank rate mismeasurement. To the extent that the potential biases stemming from bank rate estimation are similar for all regions, mismeasurement will not affect the part of our assessment that is based on comparing the *pattern* of bank and grain rate correlations.<sup>20</sup>

Similarly, the calculation of the bills of exchange interest rates depends on a number of simplifying assumptions, such as that the cost of collection and the commission charge do not vary with the waiting period for the funds. This means that these interest rates as potentially measured with error as well. The implications of mismeasurement, both classical and systematic, are examined in sections 5.2 and 5.3.

---

<sup>19</sup> For example, assumptions include holding taxes constant and no interest payments are withdrawn..

<sup>20</sup> Another caveat is that at this relatively early stage the US banking system was far from perfectly competitive, there were wildcat banks, and the period was characterized by the occasional crisis, such as the Panic of 1837. Furthermore, our banks did not necessarily account for the majority of all investments that were being made. In addition to Bodenhorn (2000) and Bodenhorn and Rockoff (1992), see Hammond (1957) and Bodenhorn (n.d.) for more details and additional references on US banking during this period.

### 3.2 Grain Price Data

We have obtained observations on monthly grain prices for six U.S. markets during the sample period: Philadelphia, New York City, Alexandria, New Orleans, Indianapolis and Charleston. The grain is wheat except for Charleston for which rice prices are employed.<sup>21</sup> Recall that in principle, the approach should work with any storable commodity.<sup>22</sup> Due to lack of detailed information, we assume that all non-interest factors influencing storage decisions were the same for wheat and rice. All of the series are considered market prices for grain. Wheat prices come from Jacks (2005, 2006) while the Charleston rice prices are from Shiue and Keller (2007). Additional detail on the characteristics of these price series is given in these papers.

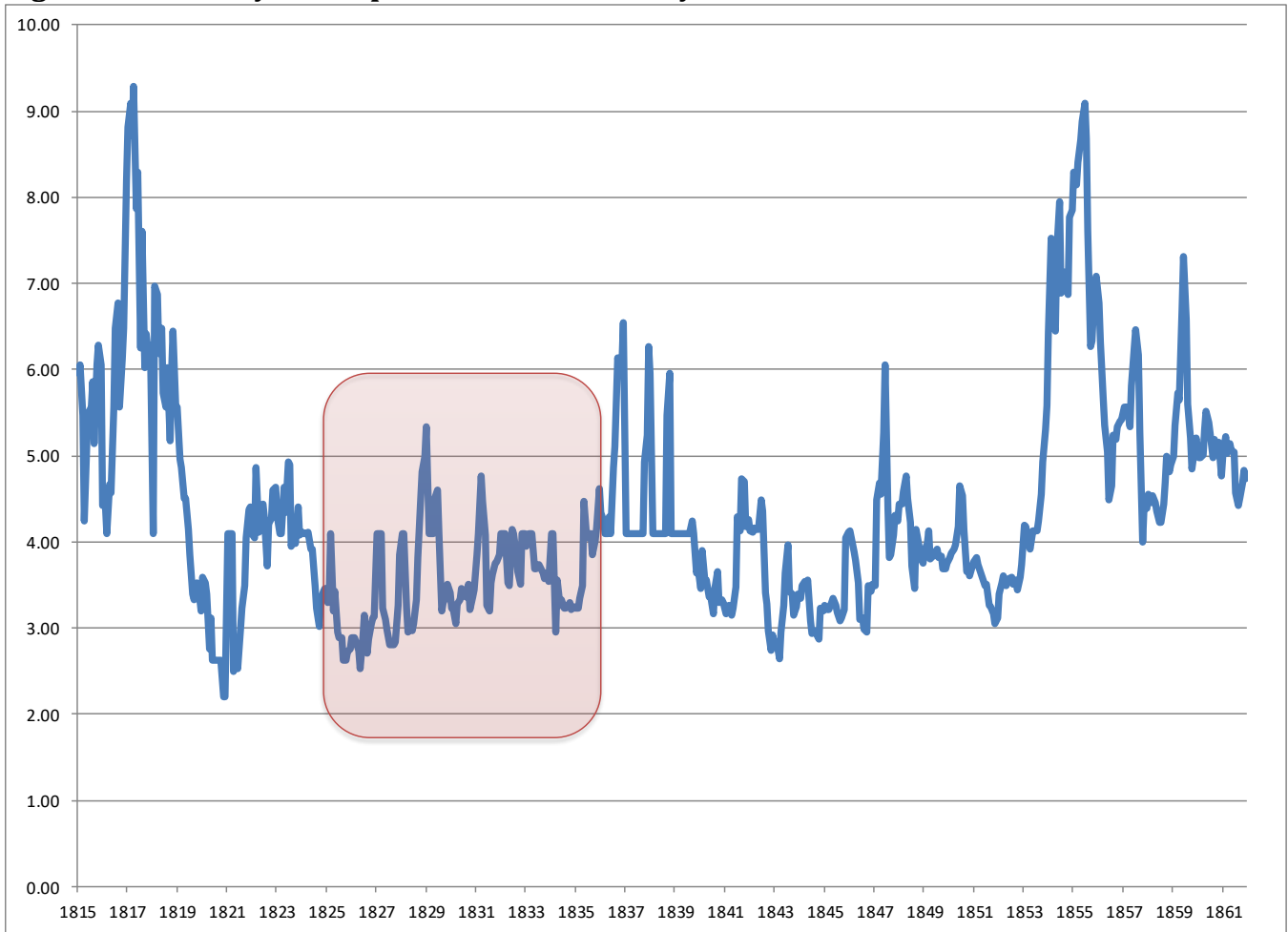
Grain prices reflect more than the movements implied by optimal storage shown in Figures 1a and 1b. Therefore it is useful to see whether there is any evidence for a cyclical pattern in the raw data. Figure 2B shows monthly prices for New York City for the years 1815 to 1861. Upon closer inspection there seem to be periods in which prices move cyclically up and down, but there are also secular trends over several years as well as a considerable amount of noise. To see the pattern more clearly, we average the monthly prices for the decade highlighted in Figure 2 (years 1825-34). The result of this is shown in Figure 3. A cyclical pattern emerges, not unlike the price dynamics implied by the storage cost approach (see Figures 1a and 1b). This provides some initial evidence that the storage cost approach might provide information on interest rate levels.

---

<sup>21</sup> Dropping Charleston would increase the homogeneity of the analysis, at the cost of reducing the sample size. We show in Table 2 how the results change as we drop the rice series from the sample.

<sup>22</sup> Among the authors who have considered more than one commodity before us are McCloskey and Nash (1984) and Taub (1987).

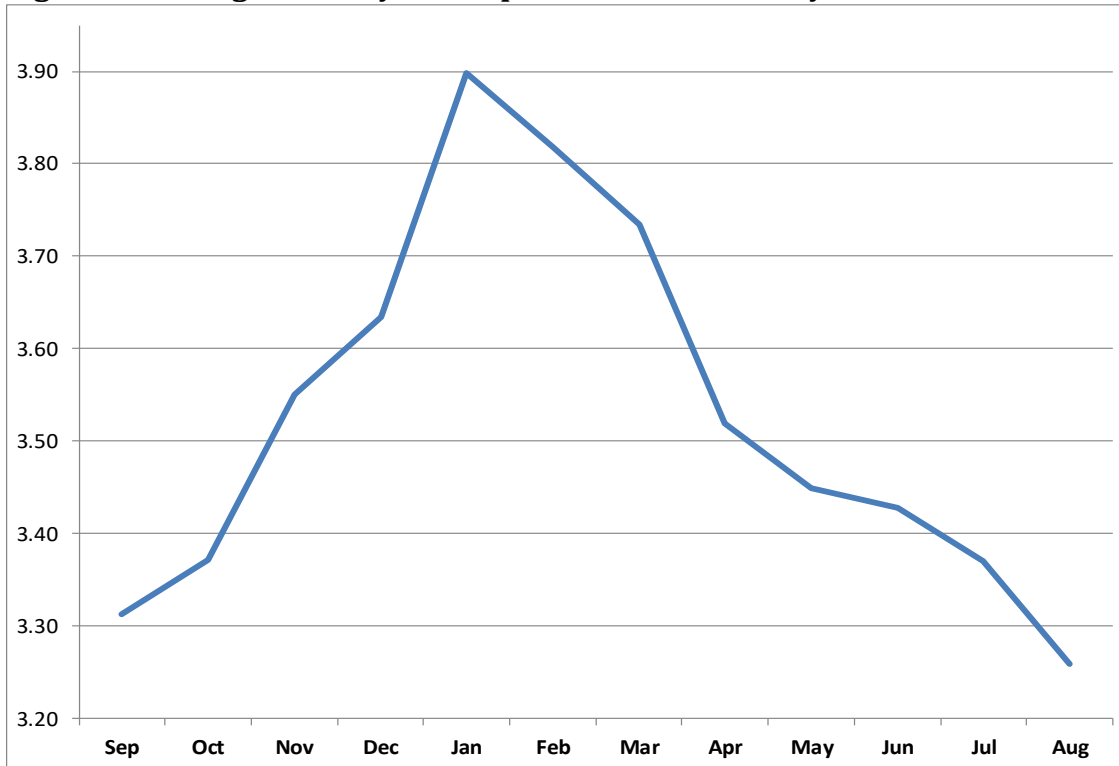
**Figure 2B. Monthly wheat prices in New York City, 1815 to 1861**



**Notes:** The price is in US dollar per 100kg of wheat.

Figure 3 indicates that the appreciation of grain prices is quite consistent from August through December. Therefore, in our benchmark analysis we compute the grain-price based interest rate in region  $i$  and year  $t$  as the average of all one-month log price changes from August to December. Alternative approaches are considered in section 4.3.2 (Table 6).

**Figure 3- Average monthly wheat price in New York City, 1825 to 1834**



**Notes:** The price is in US dollar per 100kg of wheat., averaged over 10 years.

## 4. Empirical Results

### 4.1 Criteria for Assessing Capital Market Performance

Our goal is to compare the performance of capital markets as implied by available interest rates with what is implied by the storage cost-rates. To begin, we first consider which criteria of capital market performance should be adopted. While we could assess the storage cost approach by a single criterion, it is difficult to know which of the criteria is best suited for present purposes, and there is little existing research to guide our choices. We therefore consider several criteria by which to assess the storage cost approach.

Interest rate levels provide information on both capital scarcity and transactional risk, and time-series variation in interest rates sheds light on how this varies from year to year. Additionally, comparing interest rate levels across regions reveals whether the storage cost approach correctly identifies differences in capital scarcity and risk. Furthermore, because

regional interest rate averages are affected by the composition of transactions in a particular location—they vary, e.g., with industry, maturity, borrower, and lender—we also study the degree to which regional markets co-move, that is, the integration of capital markets.

Given the goal of comparing capital market performance based on bank rates with the performance implied by the storage-cost rates, a particular criterion will compare some aspect (or, technically, a moment) of the distribution of bank rates with the same moment of the distribution of storage cost-rates. For example, one criterion would be to compare the region-and-year specific storage cost rates with the corresponding bank rates.

Estimating interest rates from grain price changes that are specific to a particular year and region sets a fairly high bar for the storage cost approach. By contrast, the influential paper by McCloskey and Nash (1984) employed the storage cost approach to estimate the interest rate level in Medieval England—which encompasses a relatively large region and many years of data. The criteria for the assessment of the storage cost approach go therefore beyond the region-and-year specific interest rates. The following describes the criteria employed in our assessment.

Let the available bank interest rate in region  $i$  and year  $t$  be denoted by  $r_{it}$ , with  $i = 1, \dots, 6$ , and  $t = 1815, \dots, 1855$  (as noted above, in an extension also bills of exchange rates will be employed). The corresponding storage cost-rates are denoted by  $\rho_{it}$ . Further, let the average bank rate across all years be denoted by  $\bar{r}_i$ , and, correspondingly,  $\bar{\rho}_i$  is the average storage-cost rate of region  $i$  across all years. Further, let the overall bank rate average and the overall storage rate average be denoted by  $\bar{r}$  and  $\bar{\rho}$ , respectively, with  $\bar{r} = \sum_i \bar{r}_i$  and  $\bar{\rho} = \sum_i \bar{\rho}_i$ . The first criterion we adopt is to compare  $\bar{\rho}$  with  $\bar{r}$ . This provides an estimate of whether storage cost rates match the overall capital scarcity and transactional risk in these areas during the sample period.

The second criterion examines how strongly bank rates and storage-cost rates correlate with each other from year to year. To do this, we stack observations of all six regions and report the t-statistic of the slope coefficient from running an OLS regression of storage-cost rates on bank rates from all observations. The higher the t-statistic, the stronger is the relationship reflected between storage-cost rates and year-to-year changes in bank interest rates, and correspondingly, the stronger is the evidence for the storage cost approach.

Our third criterion examines the pattern of interest rates across regions. Recall that bank interest rates in Philadelphia were on average lower than in New Orleans (Table A.1, bottom), and

an important question is whether this is also the case for storage-cost rates. More generally, our third criterion is the strength of the correlation between average bank and storage-cost rates, denoted by  $Corr(\bar{r}_i, \bar{\rho}_i)$ . Instead of time series variation—as in the previous t-statistic criterion—this measure captures broad differences in interest rate levels across regions.

The integration of markets is frequently studied in analyses of market performance, and integration is also a frequently employed measure of capital market development. A widely-used measure examines the extent to which interest rates co-vary across regions, and the strength of the response to shocks in other markets. The higher the co-variance and the stronger the reaction to shocks elsewhere, the more strongly are capital markets integrated. A very simple integration measure is the bilateral correlation of interest rates across two regions  $i$  and  $j$ . Table 1 shows these bilateral correlations based on Bodenhorn and Rockoff's (1992) bank interest rates. The bilateral correlations range between 0.68 and -0.30. The average of the bilateral correlations of bank interest rates is around 0.13. According to these figures, the integration of capital markets between Philadelphia and New York, as well as Philadelphia and Indianapolis is quite high, while the integration of New Orleans and Philadelphia's capital markets is substantially lower.

**Table 1: Bilateral correlations between regional U.S. bank interest rates**

	Philadelphia	New York City	Alexandria	Indianapolis	Charleston
New York City	<b>0.65</b>				
Alexandria	0.24	0.51			
Indianapolis	<b>0.68</b>	0.29	0.18		
Charleston	-0.00	0.07	-0.27	-0.04	
New Orleans	-0.30	-0.30	0.26	0.27	0.02

**Notes:** Authors' calculations based on Bodenhorn and Rockoff (1992). Shown is bilateral correlation between two log series for the period 1835-55 (n=21). PHI is Philadelphia, NYC is New York City, ALEX is Alexandria, IND is Indianapolis, CHA is Charleston, and NO is New Orleans. Bold indicates a significant OLS coefficient at a 5% level.

For the storage cost approach to capture this difference in capital market integration, it would have to be the case that the correlation of storage-cost rates between Philadelphia and New

Orleans is substantially lower than the correlation of storage-cost rates between Philadelphia and New York City. More generally, analogous to Table 1 we compute the fifteen bilateral correlations between storage cost-rates.

Finally, our fourth criterion considers the correlation between the bilateral bank rate correlations with the bilateral storage-cost correlations; we refer to this criterion as the correlation of bilateral correlations. It is bounded between minus one and plus one. Positive values that are relatively close to zero indicate that the extent of regional integration assessed by either using bank rates or storage cost rates is similar. This would mean that the storage cost approach captures differences in the degree of capital market integration.<sup>24</sup>

## 4.2 Main Findings

We begin with the benchmark case, where the storage cost rates are computed as the average of the first-differences of log grain prices from August to December. Results are given in Table 2. The storage cost rates yield an overall average of 7.25%, see column (1).<sup>26</sup> Storage cost rates are on average 1.55 percentage points higher than bank rates. Compared to the difference in the means, the degree to which bank interest rates and storage-cost rates vary is larger. One reason for that may be shocks and stochastic trends affecting the grain prices. Bank rates, in contrast, are computed from bank balance sheet information (not individual transactions) and appear to be relatively stable.

The next criterion sheds light on the extent to which the storage cost approach captures short-run time series variation in interest rates. The t-statistic for the regression of storage cost-rates on bank rates (and a constant) is 1.86. While this means that the correlation is weakly significant at standard levels, it also suggests that other temporary influences make it difficult for the storage cost rates to closely track the year-to-year variation in interest rates.<sup>27</sup>

We now turn to comparing bank rates and storage cost rates in terms of criteria that are based on regional capital market differences. The first of these compares the regional averages of

---

<sup>24</sup> An extension of this approach to market integration is to examine evidence for cointegration between two series using the autoregressive, distributed lag (ARDL) error-correction framework introduced by Pesaran and co-authors (Pesaran and Shin 1999, Pesaran, Shin, and Smith 2001). While it turns out that it appears to be the case that the ARDL approach is too demanding in our setting, a description and some ARDL results can be found in Appendix III.

<sup>26</sup> We compute all storage-cost rates as 12 times the average monthly rate.

<sup>27</sup> Running this regression with fixed effects for each region yields with a t-statistic of 1.83 to similar results.



the six storage cost rates and the six bank rates,  $\bar{\rho}_i$  and  $\bar{r}_i$ , respectively. As shown in column (3), the correlation between these two sets of regional averages is 0.79. This correlation suggests that capital market performance in terms of interest rate differences across regions over forty years or so is captured relatively well by the storage cost approach.

Finally, we turn to the market integration criterion of assessing capital market performance. Recall that this is the extent to which differences in bilateral correlations of bank interest rates is similar to differences of bilateral correlations of storage cost rates. On the right side of Table 2 in row II, we see that the correlation of bilateral correlations is positive, at 0.64. Figure 4 gives a scatter plot of the relationship between bilateral correlations of bank rates and bilateral correlations from the storage cost rates.

The results so far indicate that while the storage cost approach appears to track some aspects more than others it does seem to yield valuable information on capital market performance. Of course, it is important to quantify how much information the storage cost approach provides. To gauge this, we have applied the four criteria for the evaluation of the storage cost approach—results see row (II), Table 2—analogously to the bills of exchange interest rates that are available for this sample. Results are shown in row (III) of Table 2.

First, the overall average of bills of exchange rates for all years and all regions is 8.94%, compared to the average bank interest rate of 5.70%, a difference of 3.24 percentage points. It is not uncommon for rates for different financial instruments to vary, if only because the level of risk underlying the transactions might vary. At the same time, note that the difference between the typical bills of exchange interest rate is more than twice the difference of the typical storage cost rate, both compared to the typical bank rate. Put differently, the typical storage cost rate is roughly equal to the average rate of *all* financial instruments for these regions, bank rates and bills of exchange rates taken together.

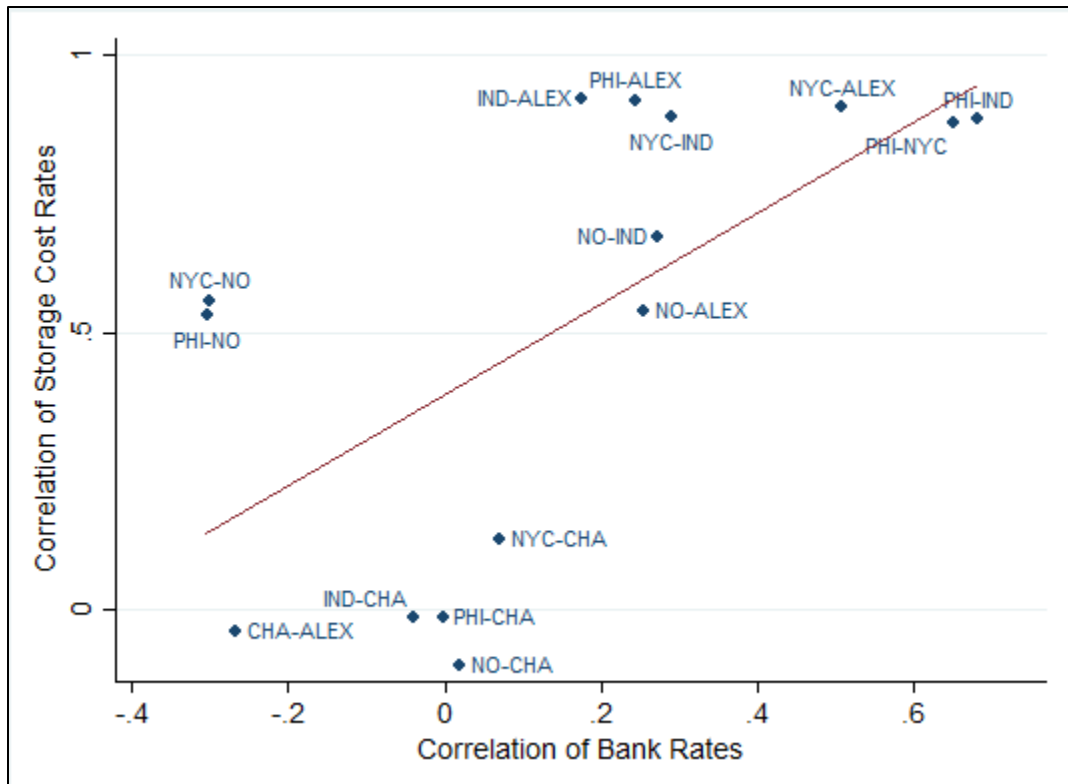
**Table 2: The Storage Cost Approach and Capital Market Performance**

	(1) Interest Rate Average [s.d.]	(2) T-statistic of time series regression	(3) Correlation Of Average Rates Across Regions	(4) Correlation of Bilateral Correlations
(I) Bank Rates	5.70 [1.68]			
(II) Storage Cost Rates Benchmark	7.25 [46.12]	1.86	0.79	0.64
(III) Bills of Exchange Rates	8.94 [3.26]	1.04	0.77	0.21
(IV) Storage Cost Rates Years 1835-1855	9.77 [48.46]	1.64	0.77	0.64
(V) Storage Cost Rates Years 1835-55, Wheat	12.19 [48.83]	1.75	0.80	0.69

**Notes:** Storage Cost Rates are computed as the average of the first-differences of log grain prices from August to December. The Storage Cost Rates Benchmark (II) employs all years (1815-1855) and all regions (PHI, NYC, ALEX, IND, NO, and CHA; n=181); (III) Bills of Exchange Rates employs the years 1836-1855 and the cities NYC, PHI, CHA, and NO; n=74; (IV) Years 1835-55 employs bank and storage cost figures for the years 1835-55 (n=109); (V) Years 1835-55, Wheat uses bank and storage cost figures for years 1835-55 for the five wheat series (PHI, NYC, ALEX, IND, and NO; n = 88). PHI is Philadelphia, NYC is New York City, ALEX is Alexandria, IND is Indianapolis, CHA is Charleston, and NO is New Orleans. Colum (1) reports average and standard deviation of bank, bills of exchange, and storage cost rates across all regions and years; column (2) reports t-statistic of slope coefficient from a stacked OLS regression of storage cost and bills of exchange rates, respectively, on log bank rates; column (3) reports the correlation of average bank rates with average storage cost rates and bills of exchange rates, respectively, across regions; and column (4) shows the correlation of bilateral interest rate correlations of bank rates with the bilateral interest rate correlations from the storage cost rates and bills of exchange rates, respectively. See text for further details.

Second, we turn to the time series variation in the interest rates. A regression of the bills of exchange rates on bank rates yields a positive coefficient with a t-statistic of 1.04, see column (2). This indicates that year-to-year changes in bills of exchange rates are not very similar to fluctuations in bank rates in the same years. Our third criterion is the correlation of regional average rates, see column (3). Now, capital market performance based on bank rates versus bills of exchange rates is, with a correlation of 0.77, quite similar. Basically, rates are relatively high in New Orleans and relatively low in New York City, both according to bank data and according to bills of exchange rates.

**Figure 4: Capital Market Integration: Bilateral Correlations of Bank and Storage Cost Rates**



Our fourth and final criterion considers the degree of market integration by comparing bilateral correlations based on bank rates with bilateral correlations based on bills of exchange. The correlation of bilateral correlations we find, 0.21, is positive but lower than the 0.64 we obtained for storage cost rates.

Taking the results in rows (II) and (III) of Table 2 at face value, there is more similarity in our capital market results for storage cost rates than for bills of exchange rates; only in the case of the differences in average interest rates across regions, column (3), does the storage cost approach not do better matching the moments of the bank rates than what one obtains employing bills of exchange rates. Of course, given that banks and bills of exchange serve different purposes and are subject to different influences it is not surprising that the corresponding interest rates are not identical, and the relatively high correlation in column (3) does suggest that there is information on capital scarcity and risk differences that is both in the bank and the bills of exchange. However, the extent of overlap appears to be limited, at least compared to what we derive from changes in grain prices.

Could the stronger results for storage cost rates compared to bills of exchange rates be due to size and differences in samples? Two additional sets of results reported in Table 2 shed some light on this. First, we focus the attention on the years 1835 to 1855 for which data availability of bank rates is close to complete, and which is virtually identical to the sample period of the bills of exchange (1836 - 55). The average of the storage cost rate for 1835 - 55 is 9.8%, see row (IV), which is higher than the average bank rate for this period (6.1%; not shown). The t-statistic of the regression of storage cost rates on bank rates is somewhat lower while the correlation of average interest rates across regions and the correlation of bilateral correlations both remain very similar to before. If we reduce the sample size further by dropping Charleston (South Carolina), for which the storage cost approach employs rice, not wheat prices, this raises the overall average while the t-statistic of the OLS regression remains similar to before, as does the correlation of regional averages and the correlation of bilateral correlations, see column (V). Overall, the storage-cost approach levels are affected by sample changes (column (1)), however, our results on market integration (column (4)) are quite robust and indicate that storage-cost rates capture capital market differences as implied by bank rates better than bills of exchange rates even when we restrict the sample for the analysis of bank interest rates.

Another possibility would be to discard the bank rates in favor of the bills of exchange rates for capital market analysis, arguing that the latter are superior to the former. One disadvantage of doing so would be that the sample for our assessment shrinks by one-third (from six to four regions), weakening inferences. Abstracting from this, we find that using bills of exchange rates for our assessment changes the findings somewhat; for example, the correlation of bilateral correlations (column (4)) of storage cost rates with bills of exchange rates is 0.91, compared to 0.64 with the bank interest rates. This is possible because the common element in bank and bills rates is limited, as we have seen (row (III) of Table 2). More generally, it is impossible to say whether the early U.S. regional bills of exchange rates are preferred to the regional bank rates, or vice versa. Overall, this underlines that, often, the number of interest rates available for early capital markets may be too limited to allow a powerful quantitative analysis, and storage-cost rates derived from grain prices may be a useful complement to other estimates.

To summarize, this analysis of the benchmark as well as the two extensions suggests that the storage cost approach provides useful information to assess capital market performance. While it can be challenging to estimate region-by-year specific interest rates, capital market performance differences across regions are captured fairly accurately, both in terms of average interest rates and in terms of capital market integration. In the next section we analyze the robustness of these results in a number of important dimensions.

### **4.3 Robustness of the Storage Cost Approach**

Our robustness checks can roughly be divided into two categories: those that deal with issues concerning the bank interest rates, such as endogeneity and measurement error, and those that address potential problems with storage cost rates, such as the timing of the harvest or the treatment of outliers.

#### **4.3.1 Bank Interest Rates: Measurement Error and Endogeneity**

In our benchmark analysis, we make use of bank interest rate data provided by Bodenhorn and Rockoff (1992). Since these authors did not have all the relevant information, they had to make a number of simplifying assumptions. For our empirical analysis, this means that the bank rates are measured with error. Ideally, we would address their assumptions one by one, however, we do not have access to better information than Bodenhorn-Rockoff. Consequently, we must take a different approach.

Note that to the extent that measurement error is unsystematic (“classical”), this would bias us *against* finding a strong correlation between bank rates and storage cost rates, and similarly, against a strong correlation of measures derived from bank rates and storage cost rates, respectively. While we cannot un-do the measurement error that is in the bank rates we can examine to what extent the storage cost approach yields results in line with the bank rates as we *increase* the amount of measurement error in the bank rates. Further below we address the possibility that there are systematic biases in the bank rates, which can be seen as a form of non-classical measurement error.

**Table 3: Measurement Error in Bank Rates**

	(1) Bank Interest Rate Average	(2) T-statistic of time series regression	(3) Correlation of Average Rates Across Regions	(4) Correlation of Bilateral Correlations
<i>(I) Bank Rates</i>	5.70	1.86	0.79	0.64
<i>(II) Error (0,0.5)</i>	5.72	1.44	0.78	0.51
<i>(III) Error (0,1)</i>	5.71	1.70	0.78	0.31

**Notes:** Column (1) reports the average of bank rates across all years and regions. Columns (2), (3), and (4) report statistics comparing (results based on) bank rates with storage cost rates. Row *(II)* adds a normally distributed error term with mean 0 and standard deviation 0.5 to the bank rates, and reports the average values of 100 simulations. Row *(III)* is analogous except with a standard deviation of 1.

Table 3 shows the results. First, we see that classical measurement error does essentially not affect the correlation of average bank and storage cost rates, see column (3). This makes sense because since the measurement error is classical, as long as the sample size is large enough (and we report average results from 100 simulations) the averages of the bank rates are not affected. The time series regression with measurement error yields somewhat weaker results, see column (2), and the fit for our correlation of bilateral correlation criterion worsens, as one would expect from the least squares result that classical measurement error biases the coefficient towards zero.

Non-classical measurement error may arise in a number of ways. An important possibility has to do with the potential endogeneity of bank interest rates because some of the regional banks were state banks that might be more prone to political influence than other banks. To evaluate this possibility, we exclude data from state-owned banks in Indiana and South Carolina. In addition, we simulate stronger endogeneity by increasing the estimated bank rates in years with relatively high grain prices or high changes in grain prices to see how this affects the match between bank interest rates and storage cost rates.

While we have taken the bank interest rates as exogenous so far, there are reasons to believe that in fact they are endogenous in our analysis. One potential source of endogeneity might be the influence of state-owned banks on the calculated bank rates. In particular, the data from Bodenhorn and Rockoff (1992) includes interest rates from the state Bank of Indiana (established in 1833) and the Bank of the State of South Carolina (established in 1812). Since it is possible that those banks adjusted interest rates in response to changes in grain prices for political reasons,

including to keep rates low when grain prices spiked up- this could lead to an endogeneity bias. To examine the influence of this on the results, in Table 4 we repeat the benchmark analysis after excluding either Charleston (row II), or Indianapolis (row III), or both from the analysis (row IV).

In general, the results are quite robust to those changes, and there is no obvious pattern in how the results change. The overall average storage cost rate can be higher *or* lower when state banks are excluded, see column (1), and the same is true for the t-statistic in the time series regression and the correlation of the bilateral interest rate correlations (columns (2) and (4)). In contrast, the correlation of storage cost averages with bank rate averages is higher once the state banks are excluded from the analysis (column (3)). Overall, this indicates that any endogeneity in the analysis that might be arising through state bank behavior is not qualitatively affecting our results.

**Table 4: State Banks and Endogeneity**

	(1) Storage Cost Rate Average	(2) T-statistic of time series regression	(3) Correlation Of Average Rates Across Regions	(4) Correlation of Bilateral Correlations
<i>(I) Benchmark</i>	7.25	1.86	0.79	0.64
<i>(II) Excluding Charleston</i>	9.82	1.97	0.87	0.69
<i>(III) Excluding Indianapolis</i>	6.96	1.82	0.84	0.35
<i>(IV) Excluding Charleston &amp; Indianapolis</i>	9.71	2.09	0.99	0.50

**Notes:** The average bank rate across all regions and years is 5.70.

Another potential endogeneity concern might be due to omitted variables that affect both bank rates and grain price changes at the same time. For example, a bad crop year might lead to a flat grain price gradient while at the same time depressing effective bank interest rates, due to a rise in loan default rates. As a second example, general inflation within the harvest year could lead to overstated growth in storage costs as well as larger changes in crop prices. Both of these

omitted variables might increase the correlation between bank rates and grain prices, but for a reason other than that the storage approach performs well.

To assess how important such issues might be for our comparison of bank and storage cost rates, we simulate this source of endogeneity by using data in which the correlation between bank rates and grain prices has been systematically increased. Table 5 shows the results. In particular, for each market we increase bank rates by 10 percent for years in which the average grain price or the change in grain prices was relatively high (rows *II* and *III*). Increasing bank rates by 10 percent at times of high grain price changes increases the t-statistic in the time series regression (column 2, row *III*), which may not be too surprising given picking up common year-to-year changes of storage cost and bank interest rates is what this measure is designed for. At the same time, even though we have deliberately increased these co-movements, there is little change in our correlation of correlation measure (column (4)), and the correlation of average bank and average storage cost rates actually falls slightly (column (3)). Based on these results it does not appear that endogeneity due to omit variables is an important driver of our findings.

In a related analysis, we split the sample into years with above-average inflation and below-average inflation (as measured by grain price changes), and we repeat the benchmark analysis (row *IV*). As should be expected, the average bank rate is higher in the high-inflation sample (5.82 vs 5.59), confirming the main premise of the storage cost approach. The correlation of average storage cost with interest rates is relatively high for both high- and low-inflations samples (column (3)), while the time series correlation as well as the correlation of bilateral correlations is higher for the low-inflations sample (columns (2) and (4), respectively).<sup>30</sup> Generally, these results do not suggest that inflation is important in bringing about the relation between storage cost and bank interest rates that we find. Analogously to the high- versus low inflation analysis, row *V* reports results for high versus low grain price levels.<sup>31</sup> The lower time series length affects some of our results (especially columns (2) and (4)), however there is little indication that either relatively high or relatively low prices are very important for our results.

---

<sup>30</sup> The generally lower values in row *IV*, column (4) also suggest that market integration analysis with the storage cost approach performs better when the time series is longer; this confirms to some extent our findings with the ARDL cointegration approach, see the Appendix.

<sup>31</sup> Furthermore, periods of high grain prices tend to be periods of high convenience yields ( $b_{it}$  in equation (1), because if grain prices are high inventories tend to be low so that the benefit of holding grain is relatively high. Thus, rows *II* and *V* shed also light on the influence of time-varying convenience yields.



Overall, our analyses so far have shown that the possible endogeneity of bank rates due to a number of reasons does not seem to be affecting our results in a major way.

**Table 5: Endogeneity through Omitted Variables**

	(1) Bank Rate Average	(2) T-statistic of time series regression	(3) Correlation Of Average Rates Across Regions	(4) Correlation of Bilateral Correlations
<i>(I) Benchmark</i>	5.70	1.86	0.79	0.64
<i>(II) ↑ 10% if price is high</i>	5.91	1.82	0.76	0.66
<i>(III) ↑ 10% if price change is high</i>	5.91	3.15	0.77	0.67
<i>(III) High inflation sample</i>	5.82	0.41	0.83	0.17
<i>Low inflation sample</i>	5.59	2.41	0.89	0.27
<i>(IV) High grain price sample</i>	5.93	1.24	0.68	0.38
<i>Low grain price sample</i>	5.50	1.37	0.77	0.31

**Notes:** In all rows, storage cost rates are computed as the average of first-differences of log grain prices in August to December. Row *II* increases the bank rate by 10% whenever the yearly average grain price is above the median grain price. Row *III* increases the bank rate by 10% whenever the yearly change in grain prices is above the median grain price change. Row *IV*, *High inflation sample* only uses observations from years in which the average grain price change was above the median, and vice versa below the median grain price change for *Low inflation sample*. Row *V*, *High grain price sample* only uses observations from years in which the average grain price was above the median grain price, and vice versa for the *Low grain price sample*.

#### 4.3.2 Storage Cost Rates: Storage Months, Outliers, and Measurement Error

The storage cost rates have been calculated as the average of changes in grain prices from August to December, for a given region and year. It is likely, however, that price gradients are affected by factors that are unobserved to us, for example year-to-year variation in the timing of the harvest or changes in the cost of storage through weather shocks. In terms of our theoretical framework of section 2, such shocks would induce time-variation in  $s_{it}$  around the regional average physical storage cost. Table 6 shows the results. We begin by computing the price gradient by taking the median instead of the average of the monthly price changes. In this case, the storage cost rates do not replicate the behavior of the bank rates as well (compare rows *II* and *III* of Table 6), although the extent of this is smaller when criteria are employed that consider

differences in regional capital market performance (columns (3) and (4)). In many empirical applications there may be uncertainty with regards to the harvest time in a given region and year. We therefore consider adding another month to the period from which the price gradient is computed. As Table 6 shows, this raises somewhat the average rate, while the time series t-statistic falls (columns 1 and 2, Row IV, respectively). In contrast, adding another storage month does not worsen the fit in the case of our market integration criterion (column 4).

While we have seen that, in general, the variation in storage cost rate estimates is greater than for bank rates (Table 2, column 1), it is important to see whether the correlations we find are mostly due to a number of extreme observations or whether they reflect a broader pattern. Winsorizing our benchmark storage cost rates at the 1<sup>st</sup> and 99<sup>th</sup> percentile, we see that in terms of most criteria the storage cost approach performs somewhat better than before (compare rows V and II, respectively). This indicates that extreme values do not drive our results. Discarding additional information does not necessarily improve the performance of the storage cost approach, as shown in row VI.

Another question is whether one should focus the analysis on positive price changes, given that interest rates are typically greater than zero. Our results indicate that this does not seem to be a good idea (row VII). While it is clear that dropping negative price changes will increase the overall interest rate average (column 1), it also lowers the t-statistic of the time series regression as well as the correlation between bank rate and storage cost rate averages (columns 2 and 3, respectively). Furthermore, when only using positive price changes the storage cost approach does not capture differences in regional market integration anymore (see column 4). Thus, when applying the storage cost approach it is important to preserve the symmetry of the analysis, using the full distribution of storage cost rate estimates.

The last row of Table 6 shows results for applying a particular threshold for the price gradient calculation. Specifically, for computing the region-by-year specific price gradient we only include months for which typically there is a one-month price change of 0.4% or more. In empirical applications, it might be difficult at time to distinguish low interest rates from noise, and applying a threshold can be beneficial in these cases. Given that in a particular month the price change is above the threshold, we use all data to include both high and low values in the calculation of the average price gradient. We see that while applying the threshold lowers the t-

statistic in the time series regression, differences in regional market performance continue to be captured as they were before (see row VIII, columns 3 and 4).

**Table 6: Alternative Storage Cost Rate Estimates**

	(1) Storage Cost Rate Average	(2) T-statistic of time series regression	(3) Correlation of Average Rates Across Regions	(4) Correlation of Bilateral Correlations
<i>(I) Benchmark</i>	7.25	1.86	0.79	0.64
<i>(II) Median</i>	10.12	0.98	0.77	0.54
<i>(III) Storage Months</i>	7.67	1.29	0.62	0.65
<i>(IV) Winsorize 1/99</i>	7.40	1.96	0.86	0.62
<i>(V) Winsorize 5/95</i>	7.43	1.46	0.89	0.53
<i>(VI) Positive</i>	60.86	0.71	0.53	0.06
<i>(VII) Exceeds 4.8%</i>	25.02	0.65	0.77	0.62

**Notes:** Average bank rates are 5.70 across all regions and years. *Row II* computes price gradient as median instead of the average off August to December one-period log price differences; *Row III* changes the period from which price gradients are computed from August to December to August to January; in *Row IV* price changes below 1<sup>st</sup> percentile are replaced by 1<sup>st</sup> percentile, price changes above 99<sup>th</sup> percentile are replaced by 99<sup>th</sup> percentile; in *Row V* price changes below 5<sup>th</sup> percentile are replaced by 5<sup>th</sup> percentile, price changes above 95<sup>th</sup> percentile are replaced by 95<sup>th</sup> percentile; *Row VI* drops non-positive price changes in gradient calculation; in *Row VII* to compute the price gradient we use only months that on average have monthly price changes of or above 0.4%.

Finally, in addition to the robustness analysis of Tables 3 to 6, we have verified—by applying a number of influential time-series filtering techniques—that stochastic shocks and cycles do not unduly affect our results; interested readers may refer to these results in the working paper version of this paper, Keller, Shiue, and Wang (2018).

## 5. Conclusions

This paper has employed regional bank interest rates and matching grain prices for the early 19<sup>th</sup> century in the United States to investigate how well the storage cost approach captures the actual level of capital market development using a number of different criteria. The analysis has shown that the storage cost approach is useful for quantifying the performance of capital markets. While the estimation of region- and year-specific interest rates can be challenging, the approach reflects differences in capital market performance quite well.

This may not be too surprising, after all. While there are important differences in grain price determinants—including storage technology, data collection, and institutions—and explicitly modeling all of these is often impossible due to lack of data, it is also often the case that many determinants are common to larger regional areas and change only slowly over time. It is then the case that spurious influences can often be eliminated by a comparison across regions, and as a consequence the storage cost approach to capital markets works well when taking a comparative approach. We conclude that the storage cost approach is a useful tool in contexts when reliable capital market information is limited.

Future research may examine broader implications. Grain prices are frequently available for historical economies, and because agricultural markets tend to account for a large fraction of economic activity in the past, purely idiosyncratic factors are unlikely to bias the analysis. In principle, however, the method is applicable to any storable commodity. Finally, the storage cost approach may have applications also to contemporaneous economies in less developed countries.

## References

- Bodenhorn, H. (n.d.), "Antebellum Banking in the United States", <https://eh.net/encyclopedia/antebellum-banking-in-the-united-states-2/> , accessed September 5, 2018.
- Bodenhorn, H. (1992), "Capital Mobility and Financial Integration in Antebellum America", *The Journal of Economic History*, Vol. 52, No. 3: 585-610.
- Bodenhorn, H. (2000), *A History of Banking in Antebellum America: Financial Markets and Economic Development in an Era of Nation-Building*, Cambridge: Cambridge University Press.
- Bodenhorn, H., and H. Rockoff (1992), "Regional Interest Rates in Antebellum America." In *Strategic Factors in American Economic History: A Volume to Honor Robert W. Fogel*, 159-187. Edited by Claudia Goldin and Hugh Rokoff. Chicago: University of Chicago Press, 1992.
- Brunt, L., and E. Cannon (2009), "How does finance generate growth? Evidence from the first Industrial Revolution", working paper, NHH Bergen and University of Bristol, February.
- Brunt, L, and E. Cannon (1999), "A Grain of Truth in Medieval Interest Rates? Re-examining the McCloskey-Nash Hypothesis", Oxford working paper, February.
- Buchinsky, M., and B. Polak (1993), "The Emergence of a National Capital Market in England, 1710 – 1880", *Journal of Economic History* Vol. 53, No. 1: 1-24.
- Clark, G. (2001), "Markets and Economic Growth: The Grain Market of Medieval England", working paper, UC Davis.
- Christiano, L. J., and T. J. Fitzgerald (2003), "The band pass filter", *International Economic Review* 44: 435-465.
- Davis, L. (1965), "The Investment Market, 1870 – 1914: The Evolution of a National Market", *Journal of Economic History*, Vol. 25, No. 3: 355-399.
- Good, D. F. (1977), "Financial Integration in Late-Nineteenth-Century Austria", *Journal of Economic History* 37(4): 890-910.
- Gurley, J. G. and E. S. Shaw (1955), "Financial Aspects of Economic Development." *American Economic Review* 45(4): 515-38.
- Hammond, B. (1957), *Banks and Politics in America from the Revolution to the Civil War*. Princeton: Princeton University Press.
- Jacks, D. (2006), "What Drove Nineteenth Century Commodity Market Integration?" *Explorations in Economic History* 43(3): 383-412.

Jacks, D. (2005), "Intra- and International Commodity Market Integration in the Atlantic Economy, 1800-1913", *Explorations in Economic History* 42(3): 381-413.

Komlos, J., and R. Landes (1991), "Anachronistic economics: grain storage in medieval England", *Economic History Review* XLIV, 1: 36 – 45.

Kaldor, N. (1939), "Speculation and Economic Theory", *Review of Economic Studies* Vol. 7, pp.1-27.

Keller, W., C. H. Shiue, and X. Wang (2018), "Capital Markets and Grain Prices: Assessing the Storage Approach", NBER Working Paper # 24388, March.

McCloskey, D., and J. Nash (1984), "Corn at Interest: The Extent and Cost of Grain Storage in Medieval England", *American Economic Review* 74: 174-187.

Mitchener, K. J., and M. Ohnuki (2009), "Institutions, Competition, and Capital Market Integration in Japan", *Journal of Economic History* 69(1): 138-171.

Narayan, P.K. (2005), "The Saving and Investment Nexus for China: Evidence from Cointegration Tests", *Applied Economics*, 37 (17), 1979-1990.

Pesaran, M.H. and Y. Shin (1999): An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. In: Strom, S. (Ed.): *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*. Cambridge, UK: Cambridge University Press.

Pesaran, M.H., Shin, Y. and R.J. Smith (2001), "Bounds Testing Approaches to the Analysis of Level Relationships", *Journal of Applied Econometrics*, 16 (3), 289-326.

Pomeranz, K. (1993), *The Making of a Hinterland: State, Society, and Economy in Inland North China, 1853-1937*, University of California Press.

Rousseau, P. L. (1999), "Finance, investment, and growth in Meiji-era Japan", *Japan and the World Economy* 11: 185-198.

Rousseau, P. L. , and P. Wachtel (1998), "Financial Intermediation and Economic Performance: Historical Evidence from Five Industrialized Countries", *Journal of Money, Credit, and Banking*, Vol. 30(4): 657-678.

Samuelson, P. A. (1957), "Intertemporal Price Equilibrium: A Prologue to the Theory of Speculation", *Weltwirtschaftliches Archiv* 79: 181-219.

Shiue, C. H., and W. Keller (2007), "Markets in China and Europe on the Eve of the Industrial Revolution", *American Economic Review* 97.

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

Sylla, R. (1969), "Federal Policy, Banking Market Structure, and Capital Mobilization in the United States, 1863-1913", *Journal of Economic History* 29(4): 657-686.

Taub, B. (1987), "A Model of Medieval Grain Prices: Comment", *American Economic Review* Vol. 77, 5: 1048 – 1053.

Williams, J. C. and B. D. Wright (1991), *Storage and Commodity Markets*, Cambridge University Press.

Working, H. (1949), "Theory of Price of Storage", *American Economic Review* 39(6): 1254-62.

Working, H. (1933), "Price Relations between July and September Wheat Futures at Chicago since 1885", *Wheat Studies of the Food Research Institute* 9(6): 187-274.

# Appendix

## I. Regional U.S. bank interest rates

**Table A.1: Regional U.S. bank interest rates, 1815-1855**

Year	New York City	Philadelphia	New Orleans	Indiana	South Carolina	Virginia
1815		4.62			8.55	
1816		5.70			5.55	
1817		3.69			5.45	
1818		5.55			8.35	
1819		3.84			4.23	
1820		5.60			4.36	
1821		4.78			4.34	
1822		5.65			5.77	4.08
1823		3.42			4.86	3.81
1824		5.21			4.62	4.14
1825		4.24			4.15	4.61
1826		5.86			2.53	3.97
1827		4.95			7.81	4.97
1828		5.82			4.50	3.97
1829		4.58			4.09	4.23
1830		4.97			4.14	4.45
1831		5.15			4.49	4.84
1832		4.48			4.24	6.28
1833	5.03	6.54			4.37	8.02
1834	5.69	3.41	6.82		3.54	3.75
1835	5.11	6.12	7.54	7.97	4.12	4.43
1836	6.82	5.74	7.16	7.60	4.37	7.22
1837	5.91	4.75	11.28	8.50	6.11	5.70
1838	5.33	5.47	7.68	8.35	6.00	4.41
1839	4.24	3.44	10.15		5.11	6.78
1840	5.57	5.73	9.01		3.10	5.43
1841	5.27	4.41	8.86	7.65	5.75	4.21



<b>1842</b>	3.95	2.50	8.85	5.05	5.97	4.20
<b>1843</b>	5.37	3.72		2.85	6.2	4.12
<b>1844</b>	5.80	5.18		5.74	6.03	4.15
<b>1845</b>	5.21	4.20		7.86	5.76	5.10
<b>1846</b>	4.69	6.39			5.42	3.95
<b>1847</b>	5.04	5.21		6.32	7.11	4.99
<b>1848</b>	5.32	4.83	7.73	8.36	5.07	4.43
<b>1849</b>	7.17	6.35	4.84	7.77	6.03	4.19
<b>1850</b>	5.62	6.47	7.42	9.45	9.28	4.53
<b>1851</b>	6.32	4.69	7.79	5.95	7.67	4.72
<b>1852</b>	7.23	5.56	7.91	6.81	6.38	5.53
<b>1853</b>	4.99	5.10	7.38	6.37	6.71	4.46
<b>1854</b>	4.98	5.31	8.50	7.70	5.57	5.04
<b>1855</b>	5.87	5.70	12.81	10.89	6.03	5.18
<b>Mean</b>	5.50	5.00	8.34	7.29	5.46	4.82
<b>Std. Dev.</b>	0.82	0.94	1.81	1.78	1.47	0.99
<b>P-value for test of equal mean</b>	< 0.001	< 0.001	n/a	0.30	< 0.001	< 0.001

**Notes:** Source is Bodenhorn and Rokoff (1992). P-value for test of equal mean is compared to New Orleans.

## II. Regional U.S. bills of exchange rates

**Table A.2: U.S. Bills of Exchange Interest Rates, 1836 - 1859**

	New York City	Philadelphia	Charleston	New Orleans
1836			6.367	11.000
1837		12.956	8.871	18.920
1838			18.583	15.400
1839	13.575	13.709	13.408	17.200
1840		7.342	12.925	12.842
1841	7.175	9.517	11.642	9.117
1842	8.013	10.350	9.092	13.642
1843	3.250	4.592	7.375	8.092
1844	4.658	5.092	6.333	6.625
1845	5.883	7.017	5.650	6.317
1846	6.800	8.533	7.575	7.233
1847	7.167	8.392	6.250	10.983
1848	9.825	11.242	11.658	12.950
1849	6.425	8.058	7.842	7.767
1850	5.667	6.975	6.000	7.417
1851	7.017	9.175	5.525	8.950
1852	5.075	6.600	6.325	7.083
1853	7.525	9.225	6.142	10.008
1854	9.392	13.000	9.708	11.933
1855	6.808	9.292	8.000	7.467
Average	7.141	8.948	8.764	10.547

**Notes:** Authors calculation as the simple average of monthly rates reported in Bodenhorn (1992).

### III. The ARDL Cointegration Approach

An extension of the approach to market integration in the text is to examine evidence for cointegration between two series using the autoregressive, distributed lag (ARDL) error-correction framework introduced by Pesaran and co-authors (Pesaran and Shin 1999, Pesaran, Shin, and Smith 2001). One attractive feature of the ARDL cointegration framework is that it can be applied to variables regardless of their underlying stationary properties, that is, they could be either integrated of order zero ( $I(0)$ ; stationary) or integrated of order one ( $I(1)$ ; non-stationary). In contrast, other co-integration approaches require all variables to be integrated of order one. This limits their applicability in many settings because unit root tests for determining the order of integration of at times series often produce mixed results, with some variables stationary while others are non-stationary.

As an example of the ARDL approach the following sketches the analysis in an ARDL(1,1) framework, where both the dependent and the independent variable (the interest rate in one region, and the interest rate in the other region, respectively) enter with one lag. The specific form of the ARDL framework actually applied for a given pair depends on the optimal number of lags which we choose using measures of best fit (information criteria).<sup>34</sup> The regression equation for the case of an ARDL (1,1) process, can be written as

$$(3) \quad y_t = c + a_1 y_{t-1} + b_0 x_t + b_1 x_{t-1} + e_t$$

The long-run equilibrium relationship is obtained when  $y_{t-1} = y_t, \forall t$ , and  $x_{t-1} = x_t, \forall t$ . The long-run coefficient is equal to

$$(4) \quad y_t = \frac{b_0 + b_1}{1 - a_1} x_t, \forall t.$$

A reparameterization that substitutes  $y_t$  with  $y_{t-1} + \Delta y_t$  and  $x_t$  with  $x_{t-1} + \Delta x_t$  yields the error-correction model (ECM) representation:

$$(5) \quad \Delta y_t = c - (1 - a_1) \times \left[ y_{t-1} - \frac{b_0 + b_1}{1 - a_1} x_t \right] - b_1 \Delta x_t + e_t,$$

---

<sup>34</sup> The specific form of the deterministic component is chosen in this way as well.

where the short-run adjustment coefficient equals  $(a_1 - 1)$ . Key to testing for cointegration in this framework is the ARDL bounds test. It is so called because one compares the F-statistic of a joint cointegration test with not one but two critical values, a lower one for the case that all variables are stationary and a higher for the case that all variables are non-stationary. If the F test statistic is either below the lower critical value or above the higher critical value, the cointegration test yields an unambiguous result: no cointegration in the former and cointegration in the latter case. In our application of this framework, the fact that the test might not give an unambiguous answer plays only a minor role.

We have employed the ARDL cointegration approach in this context, beginning with all bilateral pairs of the bank rates. Employing the ARDL approach has a number of limitations in our context. First, as detailed in Table A.3, the ARDL approach is relatively indiscriminate, as we find that only 20% of the bank region pairs are not cointegrated. Furthermore, the result of pervasive cointegration of bank rates is surprising in the light of the average of the bilateral bank rate correlations, which is with 0.15 not far from zero (see Table 1 above).

We also find that across different pairs the strength of the evidence for cointegration is increasing with the data's time series length.<sup>35</sup> Part of the explanation for these findings surely is the relatively short time series length in many of our pairs. Accounting for the optimally chosen number of lags, we have 18 observations or less in 60% of the bilateral pairs, at the same time when statistics for the ARDL small-sample case apply in the case of 30 to 80 observations (Narayan 2005). When we focus on the three bilateral pairs for which there are more than thirty bank rate observations (Philadelphia-Charleston, Philadelphia-Alexandria, and Alexandria-Charleston), the rank correlation between the cointegration F statistics for the bank rates and the storage cost rates is 0.71; however, employing the ARDL criterion on the basis of such a small sample (three bilateral pairs) cannot support strong inferences. All in all, there is too little data in the time series dimension for employing the ARDL approach, and the analysis will focus on our other criteria for this assessment.

---

<sup>35</sup> For every 10 more observations, on average the F-statistic of the ARDL bounds test increases by about 4, when the critical value is between 4 and 5 in our case; see Table A.3 and Narayan (2005).

**Table A.3 ARDL Cointegration Test Results**

	Long Run	T-stat (Long Run)	Short Run Adjustment	T-stat (Short Run)	ARDL Bounds F-test	# of Lags (y)	# of Lags (x)	N
NYC/NO	-0.777	[-7.65]	-1.843	[-8.68]	37.84	2	3	18
PHI/CHA	0.021	[0.17]	-1.183	[-7.19]	27.20	1	0	41
PHI/ALEX	0.100	[0.57]	-1.145	[-6.11]	19.78	1	0	34
PHI/NYC	0.801	[2.23]	-0.999	[-5.12]	15.52	1	0	23
CHA/PHI	0.523	[1.26]	-0.712	[-4.77]	12.99	1	1	41
ALEX/CHA	-0.015	[-0.10]	-0.915	[-4.74]	12.40	1	0	34
NYC/ALEX	-0.380	[-1.25]	-0.982	[-4.78]	12.03	1	3	23
CHA/NO	-1.184	[-4.67]	-1.51	[-4.61]	10.84	2	3	18
ALEX/NO	0.308	[1.74]	-1.015	[-4.48]	10.78	1	0	18
CHA/NYC	0.825	[1.90]	-0.8	[-4.32]	10.41	1	1	23
ALEX/NYC	-0.061	[-0.19]	-0.89	[-4.22]	9.83	1	0	23
NYC/PHI	0.396	[2.69]	-1.249	[-4.33]	9.36	2	2	23
CHA/ALEX	-0.541	[-1.24]	-0.69	[-4.17]	8.90	1	1	34
NO/NYC	-0.603	[-1.93]	-1.38	[-3.52]	7.07	2	3	18
NO/ALEX	0.822	[1.57]	-0.726	[-2.93]	6.71	1	0	18
PHI/NO	-0.475	[-1.25]	-0.86	[-3.50]	6.61	1	0	18
NYC/CHA	0.297	[1.54]	-1.017	[-3.56]	6.45	2	3	23
ALEX/PHI	0.710	[1.39]	-0.698	[-2.34]	6.34	3	1	34
IND/CHA	0.915	[1.21]	-0.638	[-3.12]	5.89	1	3	18
IND/NYC	1.058	[1.39]	-0.827	[3.39]	5.78	2	3	18
NYC/IND	0.079	[0.51]	-0.894	[-3.20]	5.46	1	0	18
IND/ALEX	0.809	[1.40]	-0.829	[-3.00]	4.88	2	0	18
NO/PHI	-0.265	[-1.19]	-0.838	[-3.08]	4.87	1	0	18
CHA/IND	-0.057	[-0.23]	-0.743	[-2.96]	4.38	1	0	18
IND/NO	-0.327	[0.585]	-0.875	[-2.88]	4.15	2	1	17
NO/CHA	-0.171	[-0.64]	-0.884	[-2.55]	4.07	1	0	18
ALEX/IND	0.253	[1.52]	-0.754	[-2.16]	2.96	3	0	18
IND/PHI	1.619	[2.21]	-0.556	[-2.14]	2.81	1	4	18
PHI/IND	-5.04	[-0.19]	-0.113	[-0.22]	2.57	3	3	18
NO/IND	0.264	[0.68]	-0.546	[-1.68]	2.46	1	0	17

**Notes:** Missing bank interest rates (see Table A.1) are linearly interpolated for the purposes of this analysis. The optimal lag structure has been determined using the Akaike Information Criterion (AIC), with the maximum number of lags (dependent and independent variables) equal to 4. The 10% critical values for the Bounds test are 4.29 (I(0)) and 5.08 (I(1)) for n = 30, 4.23 (I(0)) and 5.05 (I(1)) for n=35, and 5.24 (I(0)) and 5 (I(1)) for n = 40 (Narayan 2005, p.1988). Critical values for n < 30 are not available.