HIAS-E-74

Capital Markets and Grain Prices: Assessing the Storage Cost Approach

Wolfgang Keller[(a),(b),(c),(d)], Carol H. Shiue[(a),(b),(c),(d)] and Xin Wang[(e)]

(a) University of Colorado
(b) CEPR
(c) NBER
(d) CESifo
(e) Peking University

September, 2018

Hitotsubashi Institute for Advanced Study, Hitotsubashi University
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
tel:+81 42 580 8604 http://hias.ad.hit-u.ac.jp/

Capital Markets and Grain Prices:

Assessing the Storage Cost Approach[1]

Wolfgang Keller[2]
University of Colorado, CEPR, NBER, and CESifo

Carol H. Shiue[3]
University of Colorado, CEPR, NBER, and CESifo

Xin Wang[4]
Peking University

Abstract

This paper evaluates an approach to shed light on capital markets using grain prices, since stored grain incurs interest costs as part of the storage costs. Though this storage cost approach has been applied in McCloskey and Nash (1984) and has potentially wide applicability in situations where interest rate data is not available, this paper provides the first analysis of how well the storage cost approach captures actual capital market developments. 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. While the estimation of region- and year-specific interest rates can be challenging, the approach grain price approach accurately reflects differences in capital market development. Furthermore, the approach is robust to employing time series filtering techniques as well as dealing with unavailable information on harvest times, outliers, and a range of other factors.

[1] 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).

[2] Email: kellerw@colorado.edu
[3] Email: carol.shiue@gmail.com
[4] Email: xin.wang@nsd.pku.edu.cn

1. Introduction

Capital markets play a major role in the economy because they channel income that is not currently spent to projects that pay off their returns in the future. Interest rates are important indicators of both the scarcity of capital and the riskiness of transactions in the capital market. 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 that would allow computing comparable averages.[5] Before that already, farmers traded grain for cash back and forth inter-temporally implying that grain prices might shed light on interest rates, the basis of McCloskey and Nash's (1984) analysis of English Medieval interest rates.[6] This paper assesses the validity of the approach of employing grain prices to examine capital market development—storage cost approach for short-- by asking how close it comes to actual capital development in the early 19th century United States.

The storage cost approach relies on the notion that in equilibrium holding grain in storage to sell it at a later point will be no more or less profitable than selling grain immediately for money, so that the rate of grain price appreciation is a good approximation of the interest rate (plus other storage costs). While in principle applicable to any storable commodity, market prices for grain are relatively often available for historical economies, and furthermore, agriculture is a large fraction of such economies, with a large number of market participants, making it unlikely that idiosyncratic factors introduce biases into the analysis. 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 behavior (e.g., Williams and Wright

[5] 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 are typically too few that are strictly comparable; e.g., Pomeranz (1993), p.32.

[6] For example, on Chinese farmers trading grain back and forth see the memorial from Tang Pin for the case of 18th 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 16th and 17th 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).

1991), to date it has not been established how accurately the storage cost approach describes various aspects of capital market development. This paper fills this gap.

By employing regional data for an economy in which comparable interest data is still limited and capital markets still developing, the early 19th century US, this paper assesses the storage cost approach in a setting where it has bite.[7] We find that although it can be difficult to estimate interest rates specific to a particular time and place, provided one uses a sizable amount of data the storage cost approach produces a reasonable estimate of broad interest rate levels. Furthermore, the storage cost approach captures well differences in the capital development of regions, be it in terms of their interest rate levels or their integration of capital markets. The paper also finds that the storage cost approach is quite robust to the limitations that are often present when using historical data, and identifies a number of factors where the availability of data is most beneficial.

This paper makes two contributions. First, by evaluating a method to assess capital market development prior to the availability of comparable interest rates this paper potentially pushes back the quantitative study of capital markets to a time for which there are systematic records of prices on storable commodities at a high frequency, perhaps the Middle Ages or earlier. In contrast, high-quality regional bank interest rate data becomes typically available only in the late 19th century (e.g. Mitchener and Ohnuki 2007, 2009 on Japan in the late 19th century).[8] Because of its promise the storage cost approach has attracted much interest in the literature (McCloskey and Nash 1984, Taub 1987, Pomeranz 1993, Brunt and Cannon 1999, 2009, Clark 2001, and Shiue 2002), however, until now an empirical validation of the approach has been lacking. Furthermore, the storage cost approach has to potential to be applied in today's less developed countries for which comparable interest rates are absent.

Second, the paper sheds new light on the extent to which in historical contexts individual behavior is driven by incentives in line with economic optimization. Put simply, if in the historical context the storage cost approach is misspecified, be it because farmers do not store grain as an

---

[7] For example, the number of banks quadrupled in the US from 1820 to 1855 (Bodenhorn (n.d.)). Once the banking system has fully matured, interest rates comparable across time and space are typically available.

[8] An alternative is to use proxies other than interest rates to study capital market development. Outside of London, e.g., Buchinsky and Polak (1993) employ the quantity of property transactions to study the emergence of a national capital market in England.

asset, because high frictions in the capital market prevent arbitrage, or because the farmer in the historical context for other reasons does not act as *homo oeconomicus,* the storage cost approach will fail to yield results comparable to those based on bank interest rates.[9] Along these lines our paper quantifies the size of the barriers that stand in the way of frictionless economically rational behavior from the viewpoint of prices in an optimal storage model.[10]

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

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

where μ_{it} is the risk-free interest rate in region *i* and period *t*, φ_{it}^{k} is transaction-specific risk, s_{it} denotes physical storage cost, and b_{it} denotes the convenience yield. The term ω_{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

[9] Komlos and Landes (1991, p.43), for example, criticize McCloskey and Nash's (1984) application of the grain price approach as anachronistic and forgetting the "social, cultural, intellectual, and institutional realities of the past."
[10] The analysis below will also consider the possibility of false positives, that the storage cost approach yields results similar to those using bank interest rates for spurious reasons.

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

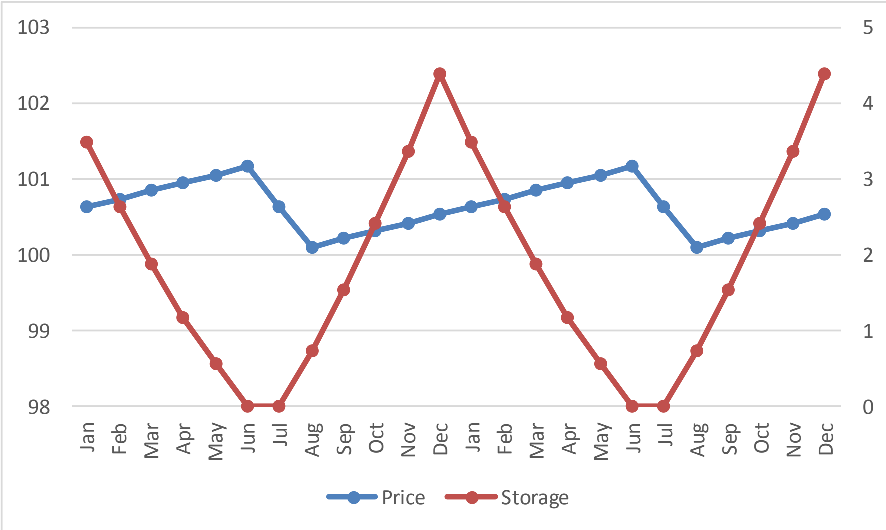
that is, the price gradient P_{it+1}/P_{it} is an increasing function of the risk-inclusive interest rate $(\mu_{it} + \varphi_{it})$ plus other factors $(\epsilon_{it} = s_{it} - b_{it} + \omega_{it})$.

This intertemporal no-arbitrage relationship is at the center of models of optimal storage. The following presents simulations of 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.[11] 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.[12] 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 is worth, and the grain price increase has to be in line with the return to postponing consumption (or, capital).[13]

Figure 1b compares the price pattern of two economies, one with a higher and one with a lower interest rate. We see that 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 behavior implies that all else equal the steepness of the price gradient is increasing in the economy's interest rate. In the benchmark analysis below, we will capture this price gradient by the average price change from August to December, for a given region and year. 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.

[11] The framework can be extended to include expectation formation and stochastic shocks without altering key relationships.
[12] The exact shape of the cyclical price pattern is determined by the parameters of the model.
[13] 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 exactly as shown in Figure 1b. First, this is due to the influence of factors such as changes in storage cost, transaction specific risk, or capital market imperfections, as it can be seen in equation (1). In the empirical analysis, ideally those variables should be explicitly modeled to bring the framework closer to the real world setting. However, because many of these factors essentially unobservable in the historical period that we consider, to be able to empirically assess the storage cost approach we make certain assumptions about them. In particular, for our benchmark analysis we allow $\omega_{it}, \rho_{it}, \varphi_{it}^{k}, s_{it}$ and b_{it} to vary between regions and over time, but we assume that those variations are uncorrelated, i.e. they are white noise. Following the benchmark analysis we will conduct additional analyses in which these influences are allowed to vary systematically.

Second, we recognize that grain prices are affected by factors outside of our storage cost model, such as weather or demand shocks for example. There could also be systematic variation in the relationship between interest rates and grain prices due to political reasons, or endogenous default. The analysis below will address these potential factors by conducting a number of important extensions.

3. Data

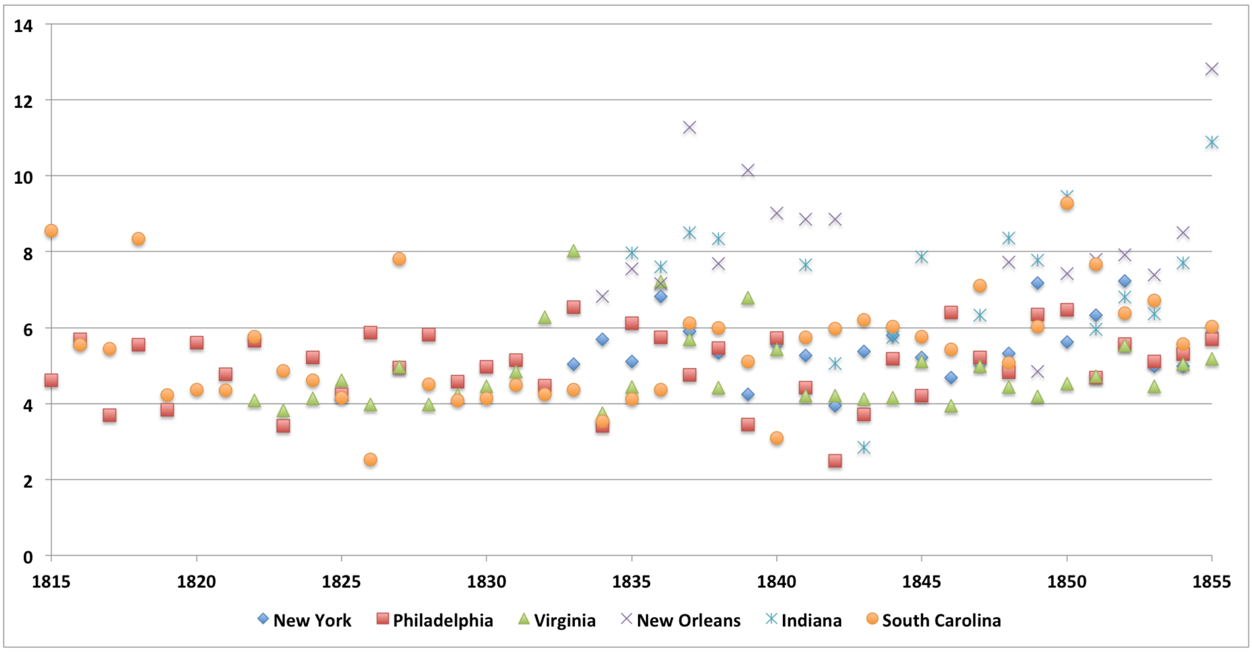
Central to our assessment is to see how well the information from monthly grain price data together with the storage model matches up with information on capital markets based on more directly measured or estimated interest rates. Bank interest rates provide a relatively good measure of prevailing rates against which we assess the interest inferred from the storage cost approach. The latter is most valuable in relatively early times, when comparable directly observed interest rate data is not yet broadly available. We therefore pick for our assessment the earliest possible setting for which we are able to collect both high-frequency grain prices and interest rates for the same regions, which turns out to be for parts of the U.S. during the years 1815 to 1855.

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 Rokoff (1992). These authors have estimated annual interest rates for a number of U.S. cities and states during the earlier part of the 19th century. The series are for the following 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 U.S. 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.[14] Figure 2 shows the bank data that we employ, with the actual values tabulated in Table A.1 in the Appendix.

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.[15]

Figure 2: Bank interest rates, 1815 - 1855



Notes: The source of the data is Bodenhorn and Rokoff (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

[14] Indianapolis in the state of Indiana, Alexandria in the state of Virginia, and Charleston in the state of South Carolina.
[15] In addition to Bodenhorn (2000) and Bodenhorn and Rokoff (1992), see also Hammond (1957) and Bodenhorn (n.d.) for more details and additional references on US banking during this period.

from 7.35% and 8.33% (for Indianapolis and New Orleans, respectively).[16] 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.

Note that these bank interest rates are not based on actual transactions at these banks but they are rather based on the bank's balance sheets and dividend data. Furthermore, due to missing data, Bodenhorn and Rokoff (1992) have to make a number of simplifying assumptions in their estimation of interest rates, such as holding taxes constant and assuming that no interest payments are withdrawn. 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. The implications of mismeasurement in the bank rates, both classical and systematic, are examined in sections 5.2 and 5.3.

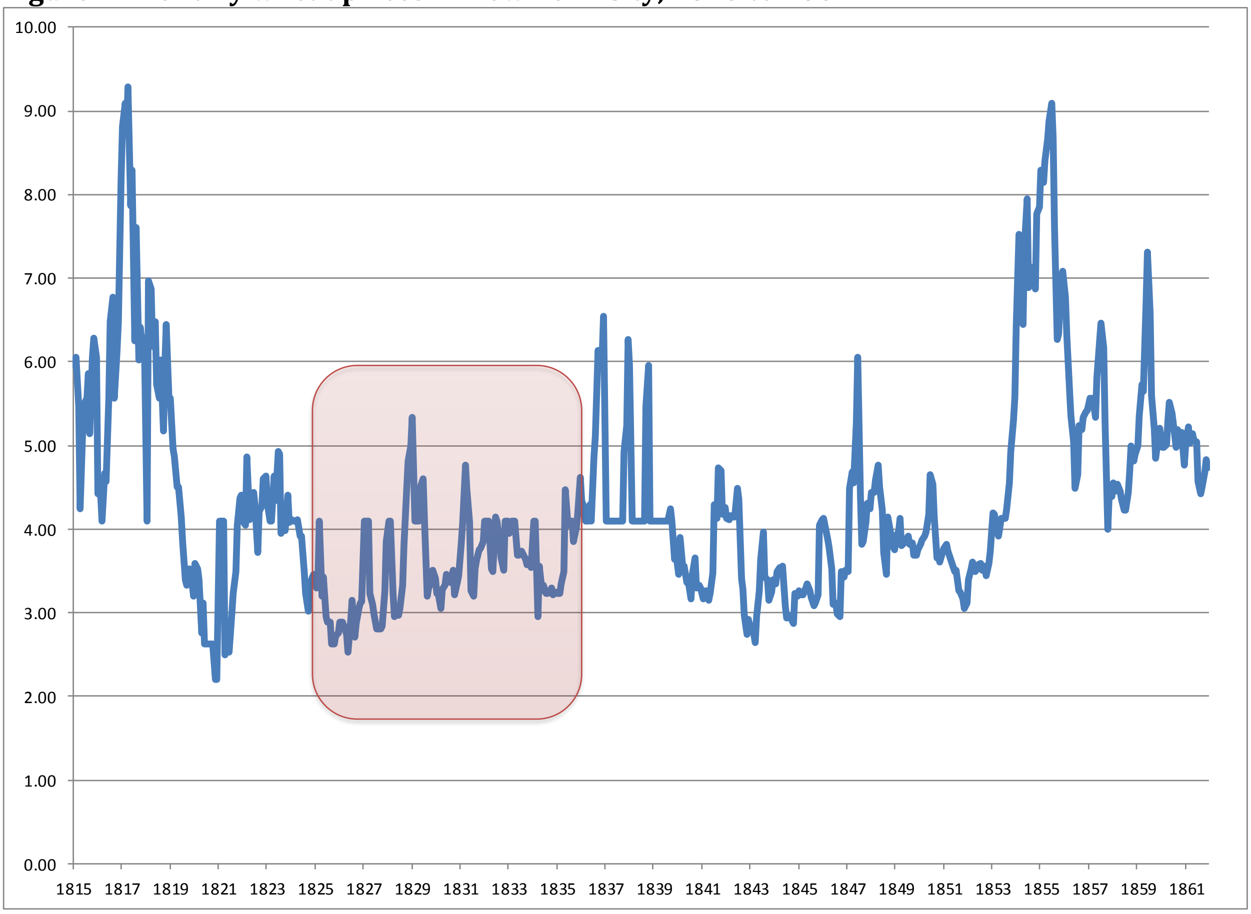
3.2 Grain Price Data

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.[17] Recall that in principle the approach should work with any storable commodity.[18] 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 prices are from Shiue and Keller (2007). Additional detail on the characteristics of these price series is given in these papers.

[16] 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%).
[17] 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.
[18] Several authors before us have considered more than one commodity (McCloskey and Nash 1984, Taub 1987).

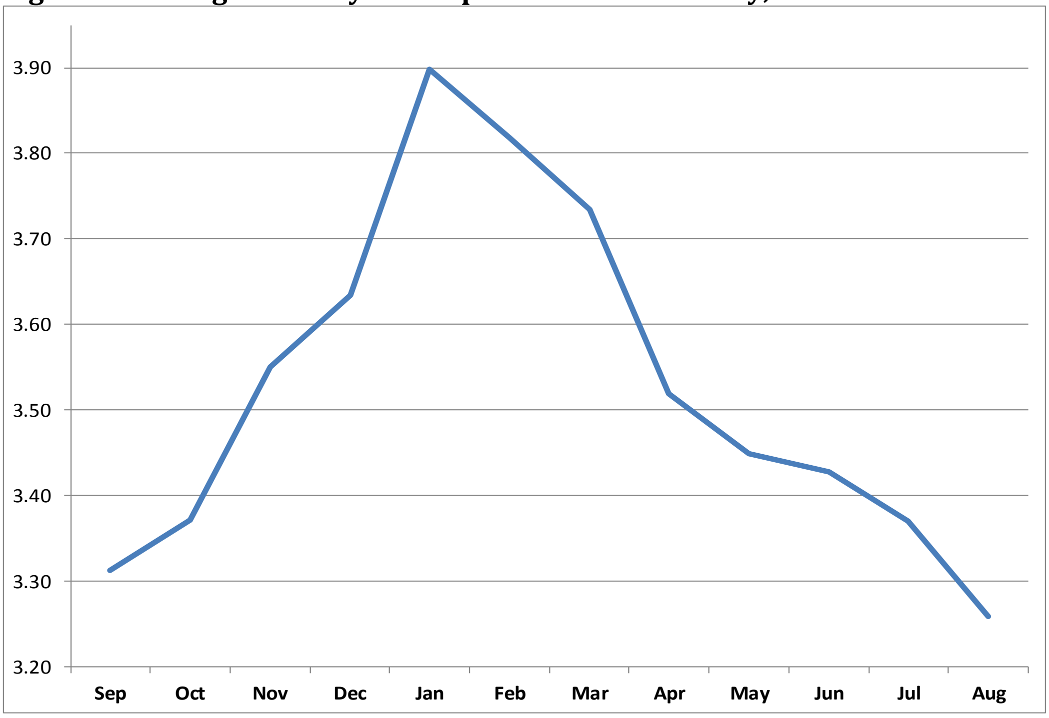
Grain prices reflect no doubt more than the movements implied by storage shown in Figures 1a and 1b. Nevertheless, it is interesting to see whether there is any evidence for a cyclical pattern in the raw data. Figure 2 shows monthly prices for New York City for the years 1815 to 1861. Upon closer inspection there seem to 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.

Figure 2- Monthly wheat prices in New York City, 1815 to 1861



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

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.

Thus, in our benchmark analysis we compute our interest rate in region *i* and year *t* based on the storage cost approach as the average of all grain price changes from August to December of year *t* in region *i*. Alternative approaches are considered in section 4.3.2 (Table 6).

4. Empirical Results

4.1 Criteria for Assessing Capital Market Performance

Given our goal of comparing the performance of capital markets as implied by the bank rates versus as implied by the storage cost-rates, the first question is which criteria of capital market performance should be adopted? 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—varying, e.g., with industry, maturity, borrower, and lender—much emphasis is placed on the degree to which regional markets co-move with each other, that is, the integration of capital markets.

While we could assess the storage cost approach by the best of these criteria, the difficulty with this is that it is unknown which of these approaches is best suited for our purposes. In part this is due to the fact that the present paper is the first to empirically assess the storage cost approach. Another consideration is that data availability in historical contexts tends to be more limited than in other contexts, which affects, for example, the length of the time series, T. For these reasons, we think it to be prudent to at least initially consider several criteria.

Recall that the main goal is to compare capital market development as implied by the bank rates with that implied by the storage-cost rates. Thus any criteria will involve comparing some aspect (or, technically, some moment) of the distribution of bank rates with the same moment of the distribution of storage cost-rates. Importantly, an implication of that is that while the moment in itself matters, it matters less than the comparison of the same moment between bank rate and storage-cost rate distributions. To take a purely hypothetical example, we could assess the validity of the storage-cost approach simply by whether it predicts well the interest rate in Philadelphia in the year 1834. It turns out that according to the bank rates estimated by Bodenhorn and Rokoff (1992), in that year the interest rate in Philadelphia was 3.41% (see Table A.1). One may actually believe that this figure may be somewhat off. In 1833, the year before, the Philadelphia bank rate was 6.54% according to Table A.1, and in the year after, 1835, the Philadelphia bank rate was 6.12%, and one might argue that such large year-to-year swings are implausible. And yet, these bank rates are among the best evidence that we have for this era. Arguably, none of the figures in Table A.1 are obviously wrong. Thus, key to our assessment would be to see how similar to 3.41% is the storage-cost rate for Philadelphia in 1834.

Naturally, estimating interest rates that are specific to a particular year and region is a challenge. For some perspective, McCloskey and Nash (1984) employed the approach to get an idea of the general interest rate level in Medieval England (a relatively large region and many years). Thus we adopt criteria that are both less stringent but at the same time more general than the interest rate in a particular year and region (such as Philadelphia in 1834). Broadly speaking, our criteria can be divided into two sets. The first is simply based on the *overall* characteristics of all regions, while the second exploits *differences* between some versus other regions.

Formally, let the bank interest rate in region *i* and year *t* be denoted by r_{it}, with *i* = 1,…,6, and *t* = 1815,…,1855. The corresponding storage cost-rates are denoted by ρ_{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. Our first criterion giving an overall measure is the average interest rate across all regions and years. These interest rate means give an overall sense of capital scarcity and transactional risk in these areas during the sample period. For reference, the overall bank interest rate average is equal to 5.7% in our sample.

The second criterion examines how strongly bank rates and storage-cost rates correlate from year to year. We employ OLS regressions to estimate this correlation. The presence of stochastic shocks and other influences will prevent that the storage cost model will fit exactly (coefficient equal to one). 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 do storage-cost rates reflect year-to-year changes in bank interest rates.

Our next criterion examines the pattern of interest rates across regions. For example, 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 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 fluctuations this measure captures broad differences in interest rate levels across regions.[19]

The integration of capital markets is also a frequently employed measure of capital market development. One influential measure is the extent to which interest rates co-vary across regions, and how strongly they respond 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 Rokoff's (1992) bank interest rates. The bilateral correlations range between 0.68 and -0.30. The average of the bilateral

[19] For example, if bank rate averages would vary across regions and storage cost-rates would always underestimate these means by 10% in one year and overestimate these means by 20% in the next, plus some noise, the time series correlation between bank rates and storage cost rates would be low but the cross-sectional correlation between bank rate and storage-cost averages may well be quite high.

correlations of bank interest rates is around 0.13, and typically the correlation is not significant at standard levels. 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	**0.65**				
Alexandria	0.24	0.51			
Indianapolis	**0.68**	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: 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: OLS coefficient is significant 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 all fifteen bilateral correlations between storage cost-rates. Our fourth criterion to assess the storage cost approach is then how strong the correlation between the bilateral correlations implied by bank rates and implied by storage-cost rates is; 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.

A powerful extension of the bilateral correlation approach 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 regions, 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 information criteria.[20] The regression equation for the case of an ARDL (1,1) process, can be written as

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

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

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

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

[20] The specific form of the deterministic component is also chosen using information criteria.

produces 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.

4.2 Main Results

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).[21] Storage cost rates are on average 1.55 percentage points higher than bank rates. There is a substantially greater difference in the degree to which bank and grain-based interest rates vary, as seen from the standard deviations in brackets. 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), which appears to be relatively stable from year to year.

With the next criterion we shed light on the extent to which the storage cost approach captures the time series variation in regional bank 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, it also suggests that other temporary influences make it difficult for the storage cost rates to closely track the year-to-year variation in bank interest rates.[22]

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 storage cost and bank rates. As shown in column (3), the correlation between these two sets of region averages is with 0.79 quite high. On the right side of Table 2 in row II, we see that storage cost rates imply differences in bilateral interest correlations that are positively correlated, with a value of 0.64, with the bilateral interest rate correlations based on bank rates. Figure 4 gives a scatter plot of the relationship between correlations implied by the bank rates and correlations implied by the storage cost rates.

[21] We compute the grain-based rates as 12 times the average monthly rate.

[22] Running this regression with fixed effects for each region yields with a t-statistic of 1.83 to similar results.

Table 2: Storage Cost Approach and Capital Markets – Main Findings

	(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) Storage Cost Rates Years 1835-1855	9.77 [48.46]	1.64	0.77	0.64
(IV) 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 Benchmark (II) employs all years (1815-1855) and all regions (PHI, NYC, ALEX, IND, NO, and CHA; n=181); Years 1835-55 (III) employs data only for the years 1835-55 (n=109); Years 1835-55, Wheat (IV) uses data only 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 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 rates on log 17bank rates; column (3) reports the correlation of average bank rates with average storage cost rates across regions; and column (4) shows the correlation of bilateral interest rate correlations implied by bank rates with the bilateral interest rate correlations implied by storage cost rates. See text for further details.

The two right-most columns of Table 2 (and row II) suggest that the storage cost approach captures the main differences in the development of regional capital markets.

We have also 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 Appendix 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.[23] 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.

Table 2 also reports two other sets of results in rows (III) and (IV). First, during the later part of the sample period there is generally more data available. For example, during the years 1835 to 1855 there is data on about 87% of all possible bank interest rates for these six regions. To examine the influence of this change in data availability we focus the attention on the years 1835 to 1855. Another advantage is that the analysis gives more uniform weight across regions. The average interest rate is now 9.8%, see row (III), which is higher than the average bank rate for 1835 to 1855 (6.1%; not shown). Further, the t-statistic of the regression of storage cost rates on bank rates is somewhat lower. In contrast, the correlation of average interest rates across regions and the correlation of bilateral correlations are both very similar to before. Overall, changes in the availability of data and sample composition have a limited impact on the results.

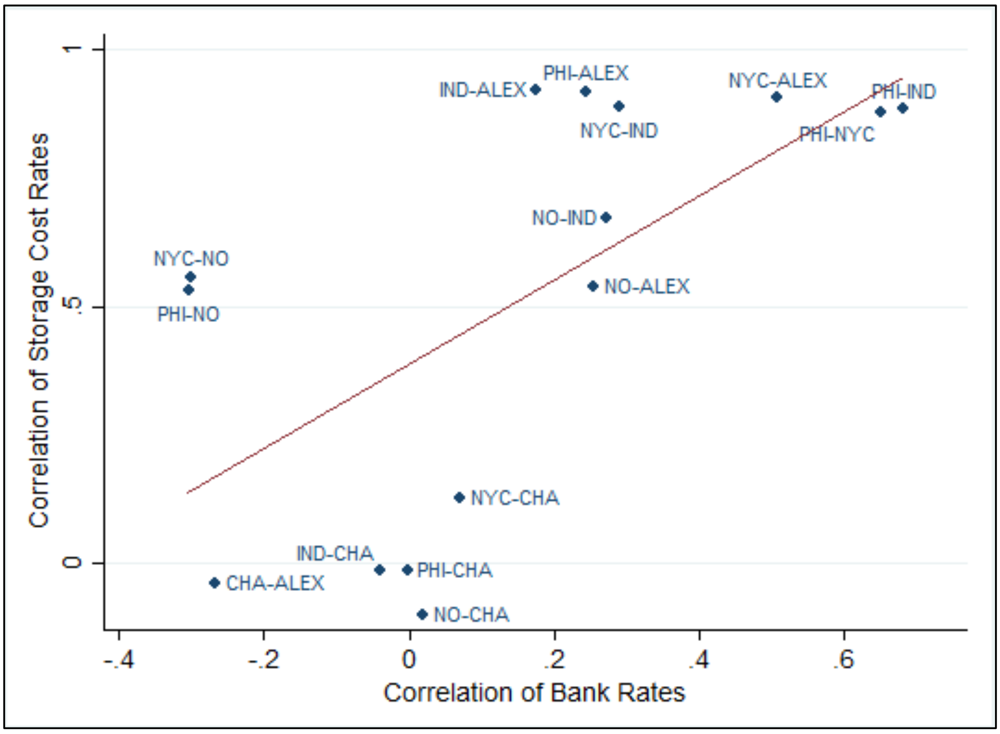
We reduce the sample size further by dropping Charleston (South Carolina), for which the storage cost approach employs rice and not wheat prices. As shown in row (IV) this raises the average rate while the t-statistic of the OLS regression of storage cost on bank rates is similar to before. The correlation of the regional averages is now 0.80, and also the correlation between bilateral correlations based on bank rates and storage cost rates is slightly higher than before. This suggests that while employing price data for the same grain is helpful our results for criteria

[23] 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 Appendix A.3 and Narayan (2005).

that capture differences in regional capital market development (columns (3) and (4)) are not much affected.

To summarize, this analysis of the benchmark as well as the two extensions suggests that the storage cost approach works quite well. While it can be challenging to estimate region-by-year specific interest rates, capital development 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.

Figure 4: Capital Market Integration with Bank vs. Storage Cost Rates



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 Rock off (1992). Since these authors have not all the relevant information, Bodenhorn and Rokoff (1992) have to make a number of simplifying assumptions. For our empirical analysis this means that the bank rates are measured with error. Ideally we would un-do their assumptions one by one, however, apart from the fact that we do not have better information than Bodenhorn-Rokoff at our disposal, the primary focus of this paper is to assess the storage cost approach, not to improve on Bodenhorn and Rokoff's (1992) estimation of bank interest rates. Consequently, we thus 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 classical 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) Bank Rates	5.70	1.86	0.79	0.64
(II) Error (0,0.5)	5.72	1.44	0.78	0.51
(III) Error (0,1)	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 several reasons why they may in fact be 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 Rokoff (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) Benchmark	7.25	1.86	0.79	0.64
(II) Excluding Charleston	9.82	1.97	0.87	0.69
(III) Excluding Indianapolis	6.96	1.82	0.84	0.35
(IV) Excluding Charleston & Indianapolis	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 using data in which the correlation between bank rates and grain prices has been systematically increased; see Table 5 for 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 results.

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 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).[24] 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.[25] 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. 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.

[24] 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.

[25] 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.

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) Benchmark	5.70	1.86	0.79	0.64
(II) ↑ 10% if price is high	5.91	1.82	0.76	0.66
(III) ↑ 10% if price change is high	5.91	3.15	0.77	0.67
(III) High inflation sample	5.82	0.41	0.83	0.17
Low inflation sample	5.59	2.41	0.89	0.27
(IV) High grain price sample	5.93	1.24	0.68	0.38
Low grain price sample	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), generally 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 1st and 99th 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) Benchmark	7.25	1.86	0.79	0.64
(II) Median	10.12	0.98	0.77	0.54
(III) Storage Months	7.67	1.29	0.62	0.65
(IV) Winsorize 1/99	7.40	1.96	0.86	0.62
(V) Winsorize 5/95	7.43	1.46	0.89	0.53
(VI) Positive	60.86	0.71	0.53	0.06
(VII) Exceeds 4.8%	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 1st percentile are replaced by 1st percentile, price changes above 99th percentile are replaced by 99th percentile; in *Row V* price changes below 5th percentile are replaced by 5th percentile, price changes above 95th percentile are replaced by 95th 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%.

4.3.3 Time Series Filtering and Storage Cost Rates

Usually employed in business cycle analysis, time series filtering techniques are designed to separate the trend from cyclical components in time series behavior. See Hamilton (1994) and Wei (2006) for a discussion of time series filtering techniques. In the context of the storage cost approach they might be helpful to suppress stochastic shocks and trends other than the cyclical harvest pattern of Figure 1.[26] We have employed a number of them to the logged price data before calculating storage cost rates as the average of log price differences in August to December for a given region and year, as before. Table 7 shows results for the Christiano-Fitzgerald (2003) and the Butterworth (1930) filters; additional details are given in Appendix A.2.

[26] For example, the Butterworth (1930) filter reweights frequencies so as to bring out the desired cyclical properties in the data both at the low-frequency and the high-frequency end, making it a “bandpass” filter. The assumptions underlying any of these time series filters make them more or less well suited to approximate the cyclical price pattern of Figure 1 depending on parameters of the storage model.

Table 7: The Storage Cost Approach and Time Series Filtering

	(1) Number of Parameter Settings	(2) Storage Cost Rate Average	(3) T-statistic of time series regression	(4) Correlation of Average Rates Across Regions	(5) Correlation of Bilateral Correlations
(I) Benchmark	1	7.25	1.86	0.79	0.64
(II) Christiano-Fitzgerald (Central 90% range)	100	5.74 (2.23-8.34)	1.88 (1.26-2.53)	0.23 (-0.67-0.80)	0.60 (0.56-0.67)
(III) Butterworth (Central 90% range)	96	5.08 (2.79-6.72)	1.77 (1.08-2.55)	0.48 (0.22-0.70)	0.65 (0.57-0.70)

Notes: Average of the bank rates is 5.70. Column (1) shows for each time series filter the number of different parameter settings we have employed. Full results are given in Appendix A.2.

As seen in row II, the Christiano-Fitzgerald filtered data produces good results in a number of dimensions, including the overall average (column 2) and the correlation of bilateral correlations (column 4). In parentheses, the table shows the lower 5th and upper 95th percentile of our results across different parameter settings, to provide information on how widely the results vary. The main exception to the good performance of the Christiano-Fitzgerald filter is that storage rates based on it often do not pick up the correlation of bank rates across regions (column 4). This is important because optimal parameter settings of the filter are typically unknown. In this respect the Butterworth filter is more robust because even when the parameter settings of the filter are off it tends to produce at least acceptable results according to all criteria (the lowest correlation of average bank and storage rates is 0.22, see column 4 of row III).

Overall, these results indicate that our findings are generally robust to filtering the grain price data using well-known time series techniques. At the same time, we note that the benchmark storage cost approach, without any filtering, appears to work quite well.[27] This is encouraging

[27] The full results in Appendix A.2 show that most other filtering techniques perform no better than the two filters considered so far.

because it shows that the results are not predominantly driven by common shocks that are left in the unfiltered data series, which further increases our confidence in the robustness of our results.

5. Conclusions

This paper has employed regional bank rates and matching grain prices for the early 19th 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 development 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 other reliable capital market information is not available.

References

Baxter, M., and R. G. King (1999), "Measuring business cycles: Approximate band-pass filters for economic time series", *Review of Economics and Statistics* 81: 575-593.

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. (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. Rokoff (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.

Butterworth, S. (1930), "On the Theory of Filter Amplifiers", *Experimental wireless and the wireless engineer*, October 1930, pp. 536 – 541.

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.

Hamilton, J. D. (1994), *Time Series Analysis*. Princeton: Princeton University Press.

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.

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.

Mitchener, K. J., and M. Ohnuki (2007), "Capital Market Integration in Japan", *Monetary and Economic Studies* 25(2): 129-153.

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.

Wei, W. W. S. (2006), *Time Series Analysis: Univariate and Multivariate Methods*. 2nd ed. Boston: Pearson.

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.

Online Appendix

For Purposes of the Referees Only

Appendix

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

1842	3.95	2.50	8.85	5.05	5.97	4.20
1843	5.37	3.72		2.85	6.2	4.12
1844	5.80	5.18		5.74	6.03	4.15
1845	5.21	4.20		7.86	5.76	5.10
1846	4.69	6.39			5.42	3.95
1847	5.04	5.21		6.32	7.11	4.99
1848	5.32	4.83	7.73	8.36	5.07	4.43
1849	7.17	6.35	4.84	7.77	6.03	4.19
1850	5.62	6.47	7.42	9.45	9.28	4.53
1851	6.32	4.69	7.79	5.95	7.67	4.72
1852	7.23	5.56	7.91	6.81	6.38	5.53
1853	4.99	5.10	7.38	6.37	6.71	4.46
1854	4.98	5.31	8.50	7.70	5.57	5.04
1855	5.87	5.70	12.81	10.89	6.03	5.18
Mean	5.50	5.00	8.34	7.29	5.46	4.82
Std.Dev.	0.82	0.94	1.81	1.78	1.47	0.99
P-value for test of equal mean	< 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. For the co-integration analysis, missing values for New Orleans and Indiana during the years 1835 to 1855 are linearly interpolated.

A.2 Extended Results for Filtered Grain Price Data

The following presents results for each individual parameter setting for the Christiano-Fitzgerald, Baxter-King, Butterworth, and Moving-Average filtering techniques. Detailed results for the Hodrik-Prescott, Non-linear, and Exponential filters are available upon request.

We begin by eliminating stochastic shocks and trends that may overlap the see-saw pattern of Figure 1 by applying a number of standard time series filtering techniques, namely the following: (i) Christiano-Fitzgerald (2003), (ii) Butterworth (1930), (iii) Baxter-King (1999), and (iv) Moving average filters. In each case, we feed the log monthly grain prices through one of the filters before calculating the within-harvest year price gradient, as before. As an extension we will also show results based on (a) Hodrick-Prescott, (b) Exponential, and (c) Nonlinear time series filtering techniques. The filters (i) to (iv) have the advantage that the degree of smoothing depends on one (or a small number) of parameters that substantially change the time series properties of the series, which allows for an expanded robustness analysis. This is not the case for the filters (a) to (c).

Table A.2.1 shows the average results for each filtering technique, together with results based on the benchmark grain rates as well as the bank rates.

Table A.2.1: The Grain Price Approach and Time Series Filtering

	(1) Number of Parameter Settings	(2) Storage Cost Rate Average	(3) T-statistic of time series regression	(4) Correlation of Average Rates Across Regions	(5) Correlation of Bilateral Correlations
Storage Cost Rate Benchmark		7.25	1.86	0.79	0.64
Christiano-Fitzgerald	100	5.74	1.88	0.23	0.60
Butterworth	96	5.08	1.77	0.48	0.65
Baxter-King	144	5.56	1.53	0.61	0.57
Moving-Average	72	0.94	0.58	0.24	0.44
Hodrik-Prescott	4	6.27	1.61	0.70	0.65
Exponential	18	-0.36	0.78	-0.36	0.41
Nonlinear	5	5.03	1.33	0.65	0.58

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

Column 1 in Table A.2.1 shows, for each filter, the number of methods (corresponding to parameter settings) for which the reported means are based (for Bank rates and Benchmark, N = 1). From these results we conclude, first, that across the board, filtering typically lowers the grain-price based interest rate estimate; the means in column (2) are all below those for both the bank rates and the benchmark method (unfiltered first-difference).

Second, exponential and moving-average smoothed data does not do well capturing year-to-year changes in interest rates, as indicated in column (3), in contrast to some of the filters which have a mean t-statistics of closer to 2. We also see that cross-regional interest rate differences are often not captured very well using exponential and moving-average smoothers (column 4), however it should be noted that Table A.2.1 reports averages; below we will see that filtering techniques can 'work' when the appropriate degree of smoothing is applied.

Column (5) confirms that bilateral correlations of grain-price based interest rates are higher than among bank rates on average, likely the result of common high-frequency shocks that affect grain-based rates but not bank rates. Compared to the mean level of bilateral correlation, the pattern of bilateral correlations is picked up better using some filtering techniques, with mean correlations as high as 0.65 for the Hodrick-Prescott and Butterworth filters (see column 5), except for moving average and exponential filters.

Of course, the optimal setting of smoothing parameters for a filtering technique is unknown in a specific empirical application. Given that, Table A.2.2 shows the full range of outcomes, using all parameter values that we have considered, for a subset of filtering techniques, those with a N (the number of different smoothing parameter combinations) higher than 50. In particular, it is useful to examine the worst-case scenario, when the researcher errs in setting the smoothing parameters. We see that both the Butterworth and the Christiano-Fitzgerald filters never go much below a correlation of 0.6 in accounting for the pattern of bilateral interest rate correlations (see table A.2.2, column (5)), and of the two filters, the Butterworth filter performs considerably better in terms of accounting for cross-regional variation in mean interest rates (see column 3)).

Table A.2.2: Time series filtering and robustness to parameter settings

	(1) Storage Cost Rate Average	(2) T-statistic of time series regression	(3) Correlation of Average Rates Across Regions	(4) Correlation of Bilateral Correlations
Storage Cost Rate Benchmark	7.25	1.86	0.79	0.64
Christiano- Fitzgerald	2.23- 8.34	1.26- 2.53	-0.67- 0.80	0.56- 0.67
Butterworth	2.79- 6.72	1.08- 2.55	0.22- 0.70	0.57- 0.70
Baxter- King	3.06- 7.56	0.78- 2.88	0.28- 0.80	0.25- 0.67
Moving- Average	-1.21- 4.09	-0.65- 2.19	-0.57- 0.82	0.30- 0.59

Notes: The average bank rate across all regions and years is 5.70. For each filtering technique, reported is the range of the central 90% of the results for different smoothing parameter combinations. The different combinations of smoothing parameters are: N = 100 for Christiano-Fitzgerald, N = 96 for Butterworth, N = 144 for Baxter-King, and N = 72 for Moving Average filter.

A.2.3 Christiano-Fitzgerald (2003) filter
Parameter choices are as follows:
[2] Max: filters out stochastic cycles at periods larger than #
[3] Min: filters out stochastic cycles at periods smaller than #
[4] Order: Number of observations in each direction that contribute to each filtered value
Columns [5] to [8] as in Tables 2 to 5

No	Max	Min	Order	Overall Mean	T-stat	Corr of Means	Pattern of Bilat Correlation
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
1	12	2	1	1.155	2.612	0.886	0.540
2	12	2	2	3.320	2.247	-0.702	0.543
3	12	2	3	3.377	2.313	-0.668	0.544
4	12	2	4	1.152	2.610	0.885	0.540
5	12	3	1	0.717	1.933	0.797	0.600
6	12	3	2	4.835	1.227	-0.836	0.609
7	12	3	3	4.847	1.256	-0.825	0.609
8	12	3	4	0.714	1.931	0.796	0.600
9	12	4	1	2.347	2.246	0.746	0.577
10	12	4	2	4.641	1.757	-0.475	0.577
11	12	4	3	4.590	1.800	-0.407	0.577
12	12	4	4	2.344	2.243	0.745	0.577
13	12	5	1	3.238	2.474	0.802	0.558
14	12	5	2	7.319	1.965	-0.344	0.560
15	12	5	3	7.374	2.037	-0.270	0.560
16	12	5	4	3.236	2.472	0.801	0.558
17	12	6	1	2.230	2.600	0.796	0.574
18	12	6	2	7.117	1.973	-0.511	0.574
19	12	6	3	7.084	2.022	-0.446	0.574
20	12	6	4	2.228	2.598	0.795	0.575
21	15	2	1	6.256	2.536	0.661	0.594
22	15	2	2	3.303	2.476	-0.611	0.585
23	15	2	3	3.312	2.460	-0.615	0.587
24	15	2	4	6.253	2.534	0.660	0.594
25	15	3	1	5.818	1.894	0.507	0.653
26	15	3	2	4.818	1.513	-0.719	0.650
27	15	3	3	4.783	1.457	-0.742	0.652

28	15	3	4	5.815	1.892	0.506	0.653
29	15	4	1	7.448	2.105	0.474	0.612
30	15	4	2	4.624	1.984	0.144	0.608
31	15	4	3	4.525	1.935	0.128	0.609
32	15	4	4	7.445	2.103	0.473	0.612
33	15	5	1	8.339	2.332	0.547	0.603
34	15	5	2	7.302	2.203	0.235	0.597
35	15	5	3	7.309	2.181	0.231	0.597
36	15	5	4	8.337	2.330	0.546	0.603
37	15	6	1	7.331	2.450	0.497	0.621
38	15	6	2	7.100	2.234	0.106	0.603
39	15	6	3	7.020	2.185	0.090	0.603
40	15	6	4	7.328	2.448	0.496	0.621
41	18	2	1	6.500	2.199	0.651	0.621
42	18	2	2	3.538	2.263	-0.498	0.603
43	18	2	3	3.534	2.278	-0.493	0.606
44	18	2	4	6.496	2.197	0.649	0.621
45	18	3	1	6.062	1.598	0.493	0.670
46	18	3	2	5.052	1.387	-0.607	0.656
47	18	3	3	5.004	1.369	-0.630	0.659
48	18	3	4	6.058	1.596	0.492	0.670
49	18	4	1	7.691	1.735	0.460	0.596
50	18	4	2	4.858	1.770	0.292	0.585
51	18	4	3	4.747	1.762	0.283	0.588
52	18	4	4	7.688	1.733	0.459	0.596
53	18	5	1	8.583	1.935	0.535	0.603

54	18	5	2	7.537	1.988	0.371	0.585
55	18	5	3	7.531	2.003	0.373	0.588
56	18	5	4	8.580	1.933	0.534	0.603
57	18	6	1	7.574	2.006	0.484	0.622
58	18	6	2	7.335	2.004	0.282	0.592
59	18	6	3	7.241	1.998	0.272	0.595
60	18	6	4	7.571	2.004	0.483	0.622
61	21	2	1	6.395	2.054	0.674	0.628
62	21	2	2	3.301	2.062	-0.399	0.639
63	21	2	3	3.313	2.078	-0.401	0.639
64	21	2	4	6.391	2.053	0.673	0.628
65	21	3	1	5.957	1.480	0.518	0.671
66	21	3	2	4.816	1.238	-0.543	0.680
67	21	3	3	4.783	1.222	-0.579	0.680
68	21	3	4	5.953	1.478	0.517	0.671
69	21	4	1	7.586	1.549	0.483	0.574
70	21	4	2	4.622	1.529	0.381	0.587
71	21	4	3	4.526	1.524	0.368	0.590
72	21	4	4	7.583	1.548	0.482	0.574
73	21	5	1	8.478	1.724	0.557	0.583
74	21	5	2	7.301	1.733	0.453	0.588
75	21	5	3	7.310	1.750	0.452	0.590
76	21	5	4	8.475	1.723	0.556	0.584
77	21	6	1	7.469	1.772	0.509	0.595

78	21	6	2	7.098	1.735	0.381	0.591
79	21	6	3	7.020	1.732	0.367	0.593
80	21	6	4	7.466	1.771	0.508	0.595
81	24	2	1	6.166	1.876	0.684	0.629
82	24	2	2	3.055	1.981	-0.358	0.618
83	24	2	3	3.054	1.966	-0.349	0.617
84	24	2	4	6.162	1.875	0.683	0.629
85	24	3	1	5.728	1.293	0.525	0.672
86	24	3	2	4.569	1.174	-0.525	0.658
87	24	3	3	4.525	1.132	-0.554	0.657
88	24	3	4	5.724	1.291	0.524	0.672
89	24	4	1	7.358	1.355	0.489	0.577
90	24	4	2	4.375	1.439	0.414	0.566
91	24	4	3	4.268	1.405	0.406	0.565
92	24	4	4	7.354	1.353	0.488	0.578
93	24	5	1	8.249	1.523	0.564	0.588
94	24	5	2	7.054	1.633	0.481	0.566
95	24	5	3	7.051	1.618	0.484	0.566
96	24	5	4	8.245	1.521	0.563	0.588
97	24	6	1	7.241	1.564	0.516	0.601
98	24	6	2	6.851	1.631	0.414	0.568
99	24	6	3	6.762	1.598	0.405	0.567
100	24	6	4	7.237	1.563	0.514	0.601

A.2.4 Baxter-King (1999) filter
Parameter choices are as follows:
[2] Max: filters out stochastic cycles at periods larger than #
[3] Min: filters out stochastic cycles at periods smaller than #
[4] Order: Number of observations in each direction that contribute to each filtered value
[5] Stationarity assumption: 1 = nonstationary, 2 = stationary
Columns [6] to [9] as in Tables 2 to 5

No	Max	Min	Order	Version	Overall Mean	T-stat	Corr of Means	Pattern of Bilat Corr
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
1	12	3	6	1	3.336	1.746	0.673	0.720
2	12	3	6	2	3.120	1.681	0.637	0.709
3	12	3	12	1	3.659	2.340	0.827	0.609
4	12	3	12	2	3.728	2.410	0.834	0.605
5	12	3	18	1	3.408	2.503	0.788	0.691
6	12	3	18	2	3.342	2.447	0.775	0.689
7	12	3	24	1	4.431	2.276	0.790	0.596
8	12	3	24	2	4.450	2.299	0.794	0.595
9	12	3	30	1	4.489	2.139	0.798	0.498
10	12	3	30	2	4.463	2.110	0.795	0.498
11	12	3	36	1	4.582	2.120	0.786	0.480
12	12	3	36	2	4.601	2.135	0.787	0.479
13	12	6	6	1	5.326	2.180	0.575	0.644
14	12	6	6	2	5.050	2.119	0.535	0.640
15	12	6	12	1	3.746	2.522	0.774	0.614
16	12	6	12	2	3.789	2.577	0.785	0.611
17	12	6	18	1	4.353	2.894	0.679	0.661
18	12	6	18	2	4.269	2.808	0.658	0.660
19	12	6	24	1	4.545	2.885	0.663	0.650
20	12	6	24	2	4.557	2.902	0.670	0.649
21	12	6	30	1	5.184	2.583	0.795	0.608

22	12	6	30	2	5.152	2.536	0.787	0.608
23	12	6	36	1	5.109	2.498	0.825	0.585
24	12	6	36	2	5.121	2.509	0.827	0.585
25	12	9	6	1	1.937	1.648	0.322	0.650
26	12	9	6	2	1.803	1.683	0.281	0.676
27	12	9	12	1	2.927	2.968	0.496	0.637
28	12	9	12	2	3.054	3.174	0.558	0.624
29	12	9	18	1	3.446	3.163	0.618	0.623
30	12	9	18	2	3.346	3.011	0.584	0.615
31	12	9	24	1	3.185	2.844	0.599	0.596
32	12	9	24	2	3.211	2.895	0.620	0.596
33	12	9	30	1	2.914	2.762	0.748	0.598
34	12	9	30	2	2.904	2.745	0.744	0.598
35	12	9	36	1	2.981	2.685	0.779	0.567
36	12	9	36	2	2.987	2.691	0.780	0.567
37	24	3	6	1	5.095	1.689	0.490	0.713
38	24	3	6	2	5.143	1.685	0.497	0.713
39	24	3	12	1	6.683	1.789	0.616	0.700
40	24	3	12	2	6.516	1.673	0.573	0.705
41	24	3	18	1	6.616	1.662	0.654	0.716
42	24	3	18	2	6.601	1.654	0.651	0.715
43	24	3	24	1	6.086	1.516	0.733	0.699
44	24	3	24	2	6.127	1.549	0.745	0.699
45	24	3	30	1	6.308	1.324	0.787	0.515
46	24	3	30	2	6.304	1.321	0.786	0.515
47	24	3	36	1	7.411	1.251	0.803	0.487
48	24	3	36	2	7.362	1.224	0.798	0.487
49	24	6	6	1	7.086	1.886	0.439	0.655
50	24	6	6	2	7.073	1.888	0.437	0.655

51	24	6	12	1	6.770	1.703	0.513	0.654
52	24	6	12	2	6.577	1.558	0.462	0.657
53	24	6	18	1	7.560	1.657	0.572	0.650
54	24	6	18	2	7.528	1.640	0.566	0.651
55	24	6	24	1	6.201	1.671	0.600	0.652
56	24	6	24	2	6.235	1.699	0.614	0.652
57	24	6	30	1	7.003	1.387	0.713	0.540
58	24	6	30	2	6.992	1.378	0.711	0.539
59	24	6	36	1	7.938	1.278	0.730	0.518
60	24	6	36	2	7.882	1.245	0.721	0.518
61	24	9	6	1	3.696	1.453	0.237	0.640
62	24	9	6	2	3.827	1.399	0.256	0.619
63	24	9	12	1	5.952	1.451	0.338	0.529
64	24	9	12	2	5.842	1.366	0.309	0.527
65	24	9	18	1	6.653	1.418	0.516	0.396
66	24	9	18	2	6.605	1.390	0.505	0.391
67	24	9	24	1	4.840	1.393	0.555	0.479
68	24	9	24	2	4.888	1.438	0.577	0.490
69	24	9	30	1	4.732	1.262	0.653	0.324
70	24	9	30	2	4.744	1.272	0.657	0.327
71	24	9	36	1	5.810	1.129	0.661	0.336
72	24	9	36	2	5.747	1.089	0.646	0.336
73	36	3	6	1	5.313	1.670	0.468	0.712
74	36	3	6	2	5.627	1.618	0.506	0.705
75	36	3	12	1	6.336	1.291	0.559	0.716
76	36	3	12	2	6.262	1.245	0.537	0.718
77	36	3	18	1	6.414	1.253	0.686	0.717
78	36	3	18	2	6.247	1.167	0.653	0.711
79	36	3	24	1	6.019	1.081	0.764	0.687

80	36	3	24	2	5.977	1.047	0.753	0.686
81	36	3	30	1	6.114	1.048	0.763	0.596
82	36	3	30	2	6.138	1.065	0.768	0.597
83	36	3	36	1	7.062	0.955	0.781	0.570
84	36	3	36	2	7.126	0.988	0.789	0.570
85	36	6	6	1	7.303	1.850	0.422	0.655
86	36	6	6	2	7.556	1.780	0.447	0.639
87	36	6	12	1	6.423	1.160	0.453	0.631
88	36	6	12	2	6.323	1.094	0.425	0.631
89	36	6	18	1	7.359	1.210	0.601	0.628
90	36	6	18	2	7.174	1.113	0.565	0.624
91	36	6	24	1	6.134	1.203	0.642	0.617
92	36	6	24	2	6.085	1.163	0.624	0.616
93	36	6	30	1	6.809	1.094	0.676	0.563
94	36	6	30	2	6.826	1.107	0.680	0.564
95	36	6	36	1	7.589	0.961	0.695	0.545
96	36	6	36	2	7.646	0.993	0.705	0.545
97	36	9	6	1	3.914	1.428	0.227	0.638
98	36	9	6	2	4.310	1.251	0.279	0.576
99	36	9	12	1	5.605	0.882	0.273	0.468
100	36	9	12	2	5.588	0.871	0.269	0.468
101	36	9	18	1	6.451	0.915	0.547	0.273
102	36	9	18	2	6.251	0.801	0.500	0.239
103	36	9	24	1	4.773	0.859	0.604	0.256
104	36	9	24	2	4.739	0.828	0.590	0.248
105	36	9	30	1	4.538	0.899	0.603	0.277
106	36	9	30	2	4.578	0.929	0.615	0.286
107	36	9	36	1	5.461	0.753	0.608	0.273
108	36	9	36	2	5.512	0.782	0.622	0.274

109	48	3	6	1	5.368	1.665	0.463	0.712
110	48	3	6	2	5.843	1.567	0.520	0.698
111	48	3	12	1	6.161	1.132	0.539	0.719
112	48	3	12	2	6.236	1.176	0.561	0.717
113	48	3	18	1	6.391	1.129	0.694	0.722
114	48	3	18	2	6.276	1.072	0.672	0.718
115	48	3	24	1	6.012	1.006	0.778	0.712
116	48	3	24	2	5.927	0.940	0.755	0.711
117	48	3	30	1	5.965	1.010	0.779	0.657
118	48	3	30	2	5.895	0.961	0.766	0.657
119	48	3	36	1	7.062	0.905	0.786	0.622
120	48	3	36	2	7.048	0.897	0.784	0.622
121	48	6	6	1	7.359	1.841	0.418	0.655
122	48	6	6	2	7.772	1.710	0.458	0.627
123	48	6	12	1	6.248	0.994	0.432	0.622
124	48	6	12	2	6.297	1.024	0.445	0.622
125	48	6	18	1	7.336	1.081	0.609	0.625
126	48	6	18	2	7.203	1.013	0.584	0.622
127	48	6	24	1	6.127	1.120	0.658	0.621
128	48	6	24	2	6.034	1.047	0.625	0.619
129	48	6	30	1	6.661	1.050	0.692	0.592
130	48	6	30	2	6.583	0.995	0.675	0.588
131	48	6	36	1	7.590	0.904	0.703	0.551
132	48	6	36	2	7.569	0.893	0.699	0.551
133	48	9	6	1	3.969	1.422	0.224	0.638
134	48	9	6	2	4.526	1.167	0.297	0.553
135	48	9	12	1	5.430	0.716	0.250	0.451
136	48	9	12	2	5.562	0.800	0.286	0.457
137	48	9	18	1	6.428	0.780	0.555	0.300

138	48	9	18	2	6.280	0.700	0.522	0.281
139	48	9	24	1	4.766	0.778	0.623	0.299
140	48	9	24	2	4.688	0.711	0.591	0.287
141	48	9	30	1	4.390	0.854	0.625	0.331
142	48	9	30	2	4.336	0.813	0.608	0.323
143	48	9	36	1	5.461	0.700	0.624	0.272
144	48	9	36	2	5.434	0.685	0.617	0.272

A.2.5 Butterworth (1930) filter

The Butterworth filter is a bandpass filter in which the smoothing is performed in two steps. Max1 and Order1 govern the filtering in step 1, while Max2 and Order2 control the filtering in step 2.
[2] Max1: filters out stochastic cycles at periods larger than #
[3] Order1: Number of observations in each direction that contribute to each filtered value
[4] Max2: filters out stochastic cycles at periods larger than #
[5] Order2: Number of observations in each direction that contribute to each filtered value
Columns [6] to [9] as in Tables 2 to 5

No	Max1	Order1	Max2	Order2	Overall Mean	T-stat	Corr of Means	Pattern of Bilat Corr
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
1	12	2	3	2	3.068	1.743	0.596	0.709
2	12	2	3	4	2.878	1.645	0.577	0.713
3	12	2	3	6	2.872	1.621	0.569	0.711
4	12	2	3	8	2.799	1.581	0.564	0.708
5	12	2	6	2	3.598	2.176	0.517	0.658
6	12	2	6	4	3.944	2.227	0.536	0.659
7	12	2	6	6	3.947	2.206	0.540	0.659
8	12	2	6	8	3.935	2.179	0.545	0.658
9	12	4	3	2	3.065	1.938	0.641	0.687
10	12	4	3	4	2.876	1.841	0.625	0.690
11	12	4	3	6	2.870	1.816	0.618	0.688
12	12	4	3	8	2.797	1.777	0.613	0.685
13	12	4	6	2	3.529	2.396	0.556	0.643
14	12	4	6	4	3.897	2.437	0.581	0.643
15	12	4	6	6	3.900	2.408	0.585	0.642
16	12	4	6	8	3.889	2.380	0.590	0.641
17	12	6	3	2	3.001	2.103	0.702	0.671
18	12	6	3	4	2.812	2.006	0.691	0.675
19	12	6	3	6	2.807	1.980	0.685	0.672
20	12	6	3	8	2.733	1.941	0.681	0.670
21	12	6	6	2	3.456	2.566	0.624	0.624
22	12	6	6	4	3.830	2.614	0.651	0.620
23	12	6	6	6	3.834	2.584	0.655	0.619
24	12	6	6	8	3.823	2.555	0.659	0.618
25	12	8	3	2	2.851	2.088	0.737	0.654

26	12	8	3	4	2.656	1.989	0.726	0.658
27	12	8	3	6	2.651	1.964	0.721	0.655
28	12	8	3	8	2.577	1.927	0.717	0.652
29	12	8	6	2	3.336	2.555	0.688	0.602
30	12	8	6	4	3.672	2.579	0.697	0.602
31	12	8	6	6	3.678	2.551	0.697	0.602
32	12	8	6	8	3.667	2.521	0.700	0.601
33	18	2	3	2	5.190	1.816	0.564	0.698
34	18	2	3	4	5.001	1.755	0.552	0.704
35	18	2	3	6	4.996	1.741	0.546	0.704
36	18	2	3	8	4.922	1.715	0.543	0.703
37	18	2	6	2	5.603	1.994	0.507	0.640
38	18	2	6	4	6.025	2.067	0.518	0.647
39	18	2	6	6	6.032	2.065	0.521	0.650
40	18	2	6	8	6.021	2.051	0.524	0.650
41	18	4	3	2	5.551	1.590	0.352	0.672
42	18	4	3	4	5.362	1.537	0.338	0.678
43	18	4	3	6	5.356	1.525	0.332	0.678
44	18	4	3	8	5.283	1.501	0.330	0.677
45	18	4	6	2	5.946	1.731	0.309	0.622
46	18	4	6	4	6.375	1.787	0.322	0.633
47	18	4	6	6	6.383	1.788	0.325	0.636
48	18	4	6	8	6.372	1.778	0.330	0.637
49	18	6	3	2	5.899	1.777	0.442	0.674
50	18	6	3	4	5.712	1.723	0.430	0.680

51	18	6	3	6	5.706	1.710	0.425	0.680
52	18	6	3	8	5.632	1.686	0.422	0.679
53	18	6	6	2	6.248	1.918	0.388	0.634
54	18	6	6	4	6.725	1.989	0.406	0.642
55	18	6	6	6	6.732	1.987	0.409	0.645
56	18	6	6	8	6.722	1.977	0.413	0.645
57	18	8	3	2	6.218	1.937	0.538	0.676
58	18	8	3	4	6.034	1.884	0.529	0.681
59	18	8	3	6	6.029	1.870	0.524	0.681
60	18	8	3	8	5.955	1.846	0.521	0.680
61	18	8	6	2	6.516	2.058	0.471	0.639
62	18	8	6	4	7.050	2.157	0.497	0.645
63	18	8	6	6	7.055	2.155	0.501	0.646
64	18	8	6	8	7.045	2.143	0.504	0.646
65	24	2	3	2	5.857	1.646	0.590	0.691
66	24	2	3	4	5.668	1.602	0.580	0.699
67	24	2	3	6	5.663	1.594	0.575	0.701
68	24	2	3	8	5.589	1.573	0.572	0.700
69	24	2	6	2	6.238	1.712	0.540	0.603
70	24	2	6	4	6.684	1.802	0.547	0.615
71	24	2	6	6	6.692	1.808	0.549	0.621
72	24	2	6	8	6.682	1.799	0.553	0.623

73	24	4	3	2	5.741	1.338	0.398	0.685
74	24	4	3	4	5.552	1.298	0.384	0.693
75	24	4	3	6	5.546	1.291	0.379	0.694
76	24	4	3	8	5.472	1.272	0.377	0.694
77	24	4	6	2	6.128	1.381	0.354	0.614
78	24	4	6	4	6.564	1.461	0.365	0.629
79	24	4	6	6	6.572	1.469	0.368	0.634
80	24	4	6	8	6.562	1.461	0.372	0.635
81	24	6	3	2	5.445	1.017	0.240	0.628
82	24	6	3	4	5.255	0.982	0.225	0.638
83	24	6	3	6	5.250	0.976	0.220	0.641
84	24	6	3	8	5.176	0.958	0.219	0.639
85	24	6	6	2	5.854	1.039	0.208	0.511
86	24	6	6	4	6.268	1.108	0.220	0.529
87	24	6	6	6	6.276	1.117	0.224	0.537
88	24	6	6	8	6.266	1.112	0.228	0.540
89	24	8	3	2	5.669	1.150	0.323	0.651
90	24	8	3	4	5.481	1.115	0.309	0.659
91	24	8	3	6	5.475	1.109	0.304	0.661
92	24	8	3	8	5.402	1.091	0.302	0.659
93	24	8	6	2	6.043	1.160	0.278	0.555
94	24	8	6	4	6.493	1.247	0.296	0.577
95	24	8	6	6	6.501	1.256	0.299	0.584
96	24	8	6	8	6.491	1.250	0.304	0.586

A.2.6 Moving-Average filter
Parameter settings are as follows.
[2] Lag: Number of past terms used to filter
[3] Forward: Number of future terms used to filter
[4] Current: Whether current period is used or not; 0 = exclusion, 1 = inclusion
Columns [5] to [8] as in Tables 2 to 5

No	Lag	Forward	Current	Overall Mean	T-stat	Corr of Means	Pattern of Bilat Corr
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
1	2	2	0	5.936	1.067	0.560	0.368
2	2	2	1	6.181	1.281	0.625	0.484
3	2	4	0	4.440	0.734	0.795	0.283
4	2	4	1	4.812	0.975	0.789	0.436
5	2	6	0	3.398	-0.105	0.818	0.168
6	2	6	1	3.795	0.215	0.827	0.330
7	2	8	0	0.837	-0.757	0.690	0.484
8	2	8	1	1.385	-0.391	0.813	0.560
9	2	10	0	-0.594	-0.560	0.037	0.654
10	2	10	1	-0.026	-0.235	0.477	0.686
11	2	12	0	-0.065	-0.124	0.065	0.613
12	2	12	1	0.391	0.142	0.442	0.619
13	4	2	0	1.069	0.412	-0.634	0.371
14	4	2	1	1.939	0.689	-0.457	0.440
15	4	4	0	1.135	0.250	0.360	0.309
16	4	4	1	1.792	0.514	0.653	0.403
17	4	6	0	0.948	-0.397	0.525	0.255
18	4	6	1	1.495	-0.092	0.846	0.344

19	4	8	0	-0.787	-0.981	-0.452	0.445
20	4	8	1	-0.199	-0.638	-0.348	0.482
21	4	10	0	-1.782	-0.819	-0.592	0.548
22	4	10	1	-1.211	-0.509	-0.563	0.566
23	4	12	0	-1.168	-0.408	-0.597	0.492
24	4	12	1	-0.701	-0.151	-0.553	0.489
25	6	2	0	0.571	0.658	-0.575	0.487
26	6	2	1	1.303	0.887	-0.488	0.483
27	6	4	0	0.718	0.497	0.068	0.422
28	6	4	1	1.291	0.714	0.475	0.437
29	6	6	0	0.628	-0.099	0.153	0.382
30	6	6	1	1.113	0.153	0.373	0.392
31	6	8	0	-0.830	-0.662	-0.319	0.494
32	6	8	1	-0.319	-0.374	-0.237	0.487
33	6	10	0	-1.709	-0.553	-0.509	0.537
34	6	10	1	-1.211	-0.282	-0.482	0.527

35	6	12	0	-1.177	-0.203	-0.545	0.491
36	6	12	1	-0.760	0.030	-0.515	0.452
37	8	2	0	0.525	1.191	-0.188	0.514
38	8	2	1	1.110	1.360	0.032	0.523
39	8	4	0	0.594	0.961	0.389	0.467
40	8	4	1	1.079	1.129	0.578	0.485
41	8	6	0	0.504	0.354	0.339	0.422
42	8	6	1	0.927	0.563	0.474	0.438
43	8	8	0	-0.783	-0.221	-0.075	0.514
44	8	8	1	-0.339	0.032	0.033	0.520
45	8	10	0	-1.586	-0.150	-0.355	0.539
46	8	10	1	-1.149	0.098	-0.302	0.538
47	8	12	0	-1.138	0.151	-0.411	0.471
48	8	12	1	-0.763	0.366	-0.356	0.448
49	10	2	0	2.130	2.095	0.603	0.476
50	10	2	1	2.486	2.171	0.729	0.500

51	10	4	0	1.907	1.748	0.718	0.418
52	10	4	1	2.232	1.842	0.774	0.460
53	10	6	0	1.638	1.108	0.634	0.331
54	10	6	1	1.940	1.255	0.697	0.381
55	10	8	0	0.344	0.561	0.386	0.389
56	10	8	1	0.679	0.760	0.500	0.441
57	10	10	0	-0.507	0.612	0.084	0.395
58	10	10	1	-0.164	0.816	0.239	0.425
59	10	12	0	-0.215	0.828	0.115	0.341
60	10	12	1	0.087	1.008	0.352	0.351
61	12	2	0	3.326	2.637	0.780	0.423
62	12	2	1	3.545	2.654	0.817	0.452
63	12	4	0	2.941	2.213	0.783	0.363
64	12	4	1	3.163	2.262	0.808	0.412
65	12	6	0	2.569	1.544	0.726	0.264
66	12	6	1	2.788	1.653	0.761	0.324
67	12	8	0	1.300	1.021	0.610	0.343
68	12	8	1	1.556	1.181	0.683	0.394
69	12	10	0	0.432	1.065	0.508	0.329
70	12	10	1	0.703	1.235	0.659	0.350
71	12	12	0	0.612	1.234	0.728	0.316
72	12	12	1	0.856	1.389	0.877	0.311

A.3 ARDL Cointegration Test Results

	Long Run	T-stat (Long Run)	Short Run Adj	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: Results are for linearly interpolated bank interest rates (see Table A1). 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.