STORAGE, SLOW TRANSPORT, AND THE LAW OF ONE PRICE: THEORY WITH EVIDENCE FROM NINETEENTH-CENTURY U.S. CORN MARKETS

Andrew Coleman*

Abstract—This paper argues that localized price spikes should be a regular feature of competitive commodity markets. It develops a rational expectations model of physical arbitrage in which trade takes time, and shows that inventory management plays a crucial role in the way regional prices are determined. In equilibrium, arbitrageurs choose export quantities to ensure inventories in the importing center regularly fall to 0. They earn enough profits from high prices on these occasions to offset small losses at other times. An analysis of detailed data from Chicago and New York corn markets provides empirical support for the model.

I. Introduction

In the late nineteenth century, complex marketing infrastructures were developed in Chicago and New York to facilitate the export of grain from the Great Plains to Europe. Each year, millions of bushels of corn were sent to Chicago for sale to shipping agents, who transported the grain to New York, where it was resold, transferred to elevators, and then shipped to European markets. Both cities had elaborate facilities for receiving, storing, and forwarding grain, and financial exchanges that offered an array of liquid spot and futures markets. In such a setting it would be imagined that prices always obeyed the law of one price, that is, that the price of corn in New York was always equal to the price of corn in Chicago plus the cost of shipping. However, this was not the case. Although the law of one price held most of the time, on numerous occasions—approximately 10% of the weeks in the period under analysis—the New York spot price spiked upwards to a level considerably higher than the Chicago price plus the transport cost. Although infrequent, the fact that these spikes occurred throughout the period in one of the world's most organized markets suggests that they were not incidental deviations from the law of one price but an essential feature of the arbitrage process.

This paper develops a model of commodity prices that explains why localized price spikes occur, where a localized price spike is defined as an occasion when the price in one location temporarily exceeds the price in another by more than the costs of shipping the goods. While such price spikes appear to contradict the law of one price, this paper develops a model of commodity price arbitrage in which these price spikes must occur if arbitrage is to take place. Indeed, if markets are competitive, arbitrage activity will not occur unless there are localized price spikes at times when supplies are unusually tight, because arbitrageurs will not cover their costs.

The key feature of the model is that shipping takes time. This means arbitrageurs cannot immediately obtain additional supplies in a center where local demand is unusually high or local supply is unusually low, so inventory management has a vital role in the arbitrage process. Indeed, prices in different cities will depend critically on inventory levels if competitive arbitrageurs simultaneously manage their shipments and inventories to maximize profits. If two cities both have large inventories, their spot prices will never exceed the difference between the transport cost and the storage cost. But if an importing city runs out of inventories, its spot price will rise upwards and exceed the spot price in the other center by more than the difference between the transport cost and the storage cost. The higher prices will last until supplies are replenished. In addition, shipments will be managed to ensure that inventory levels regularly fall to 0 in the importing city. Consequently, there will be occasions that an importing city will have inventory levels that regularly fall to 0, and on these occasions the spatial price difference will exceed the difference between the transport cost and the storage cost.

The intuition of the model is straightforward. Rational, risk-neutral arbitrageurs ship goods between centers if the expected future price in the importing center is at least as large as the price in the exporting center plus the cost of transport. If there are sufficient inventories in the importing center when the goods are sent, arbitrageurs in the importing center run down their inventories until the spot price equals the expected future price minus storage costs, and thus the spot price difference between centers equals the transport cost minus the storage cost. If inventories in the importing center are zero, however, the spot price cannot be arbitraged down and the price difference between the centers exceeds the difference between the transport cost and the storage cost until the shipment arrives. The spatial price difference should exceed the transport cost minus the storage cost on occasion because exporters limit their shipments to ensure that random variation in local supply or demand regularly causes inventories to fall to 0 in the importing center. Unless they do this, the exporters make negative profits because the price in the importing center does not cover the full cost of exporting if the exports arrive when supplies are plentiful.

Localized price spikes were a regular feature of the New York and Chicago corn markets in the late nineteenth

Received for publication August 9, 2004. Revision accepted for publication November 7, 2007.

* Motu Economics.

An early version of the paper was circulated under the title "Storage, Arbitrage, and the Law of One Price," and I am grateful to Professors Angus Deaton, Kenneth Rogoff, and especially Chris Paxson for comments and suggestions. I would also like to thank Chris House, Stephen Salant, Linda Tesar, and Jeffrey Williams for comments on the full draft. The paper has benefited from comments by participants at seminars at Boston University, UCLA, the Kennedy School, the University of Michigan, and Princeton University. I would like to thank Ed Knotek and the librarians at the University of Michigan for help collecting the data. The European Science Foundation provided financial support to attend the July 2004 conference "Market Performance and the Welfare Gains of Market Integration in History."

The Review of Economics and Statistics, May 2009, 91(2): 332–350

century. I examine these markets because they have all the data needed to detect the price relationships implied by the model, including spot and future prices, storage quantities, shipping quantities, and transport costs. Such data are not often available, but exist in this setting because both cities had active futures markets for corn. The price relationships are strikingly consistent with the model. The paper shows the New York price for future delivery was always very close to the Chicago spot price plus the cost of the lake and canal transportation, but that on numerous occasions the New York spot price exceeded the Chicago price plus the transport cost. These high prices were temporary and mostly occurred when corn inventories were extremely low in New York. The data therefore show that regular deviations from the law of one price occurred in one of the world's most organized markets, deviations that are consistent with the theoretical model and not explained by models of price arbitrage that ignore inventories.

This paper is organized as follows. In section II, the model of price arbitrage is specified and the solution technique is outlined. Two analytic results concerning the relationships between prices, inventories, and trade flows are derived in the first half of section III, while the results of some numerical simulations are presented in the second half. In section IV the main features of the Chicago–New York corn trade are described and it is established that grain was predominantly shipped by the slowest and cheapest mode of transport. The relationship between New York and Chicago prices, the cost of shipping, and inventories is examined in section V, and a discussion of the results is offered in section VI.

II. A Structural Model of Storage and Commodity Arbitrage

Models of price arbitrage can be traced to Cournot (1838). The basic approach, exemplified by Samuelson (1952), has been to link a set of equations representing supply and demand curves in different locations with a set of no-arbitrage conditions that preclude excess profits from the instantaneous shipment of goods. A set of prices P_t^i are found that equate aggregate demand with aggregate supply while ensuring profitable arbitrage opportunities do not exist:

$$P_t^i - P_t^j \leq K^T; \quad (P_t^i - P_t^j - K^T) \cdot T_t^j = 0, \tag{1}$$

where P_t^i is the price in center i at time t, T_t^j is the quantity of exports from j to i at time t, and K^T is the transport cost.

Models of rational storage have a similar structure (see Williams, 1936; Williams & Wright, 1991; or Deaton & Laroque, 1992). Risk-neutral arbitrageurs are assumed to predict future prices and purchase and hold inventories until the expected price increase just offsets the cost of storage; conversely, if the expected appreciation is less than the cost of storage, inventories will be 0. There are three possible storage costs. First, there can be an elevator charge K^S per unit to store goods each period. Secondly, the commodity depreciates at rate δ so if S_t is stored in period t, $(1 - \delta)\, S_t$ will be available in period $t + 1$. Thirdly, there is an interest cost r foregone when storage is undertaken. These relationships are represented by the following equations:

$$\left(\frac{1-\delta}{1+r}\right) E_t[P_{t+1}^i] - P_t^i \leq K^S; \qquad \left[\left(\frac{1-\delta}{1+r}\right) E_t[P_{t+1}^i] - P_t^i - K^S\right] \cdot S_t^i = 0, \tag{2}$$

where E_t is the expectations operator conditioning on information known at t, and S_t^i is the quantity of inventories held at time t.

Williams and Wright (1991) solved a model combining both spatial arbitrage and storage. They investigated how prices in two locations would be determined if rational, forward-looking, and risk-neutral arbitrageurs could store goods or transport them instantaneously from one location to the other. They used numerical techniques to find the set of optimal storage and trade functions that generate a stationary rational expectations equilibrium. The model developed in this paper copies their approach but relaxes the assumption that trade is instantaneous. This is a crucial modification, for if transport is instantaneous and there are no transport capacity constraints, the price in one center will never exceed the price in the other by more than the cost of transportation.

The model is also related to those developed in the logistics management literature. This literature has examined the optimal ways for a company with sales in several locations to minimize total procurement, transport, and storage costs (see Baumol & Vinod, 1970, for an original statement, or Tyworth, 1991, McGinnis, 1989, or de Jong, 2000, for a review). The literature shows that careful inventory management enables a firm to substitute low cost but slow transport systems for faster but higher priced systems. Indeed, a cost-minimizing firm will use fast transportation only when inventories unexpectedly fall to such low levels that it cannot use slow transportation to replenish them before they run out. The model in this paper extends this literature, as prices are determined endogenously in markets, rather than being set exogenously by the firm. In this model, therefore, unexpected shortages lead to price spikes rather than rationing.

A. *The Model*

There are two centers, A and B, each with a separate inverse demand function for a commodity:

$$P_t^i = D_i^{-1}(Q_j^i): \quad D_i^{-1}(0) < \infty, \quad \lim_{Q\to\infty} D_i^{-1}(Q) = 0, \tag{3}$$

where Q_t^i is the amount purchased for final use at time t and $i = A, B$. The quantity of output produced each period, X_t^i, is assumed to be price inelastic but stochastic, because it has a long gestation period. There are two ways that output is usually modeled in the literature: either it is serially autocorrelated or it varies seasonally. The empirical section of this paper studies a market in which output is seasonal. Nonetheless, since the analytical structure of the model does not depend on how output is modeled, in this paper I follow Deaton and Laroque (1996) and assume that output in each center follows an independent first-order autoregressive process around a constant mean:

$$(X_t^i - \bar{X}^i) = \rho^i(X_{t-1}^i - \bar{X}^i) + e_t^i, \quad i = A, B, \tag{4}$$

where e_t^i is a white-noise process and $|\rho^i| < 1$.[1] By adopting this assumption, the Deaton and Laroque results can be viewed as the limiting case of this model that occurs when transport is instantaneous and transport costs are 0.

The length of a period is the time that it takes to ship goods from one center to another. All production, consumption, storage, and trade activity takes place at the beginning of the period.[2] A good shipped at the beginning of period t arrives at its destination at the beginning of period $t + 1$. It is assumed that unlimited quantities of the good can be stored in either center, that goods produced at different times are indistinguishable and have the same price, and that the storage cost K^S is less than the transport cost K^T.

Let product availability, M_t^i, be the total quantity of stored and imported goods in a center at the beginning of the period,

$$M_t^i = (1 - \delta)(S_{t-1}^i + T_{t-1}^j), \tag{5}$$

where S_{t-1}^i is the nonnegative quantity stored in center i and T_{t-1}^j is the nonnegative quantity exported from center j. The quantities stored and exported are such that $0 \leq S_t^i + T_t^i \leq X_t^i + M_t^i$. The vector of state variables y_t comprises the quantity of product availability and the amount of output in each center, $y_t = [M_t^A, M_t^B, X_t^A, X_t^B]$.

It is assumed that risk-neutral, profit-maximizing, and rational speculators in both cities undertake a mixture of trade and storage to take advantage of expected price differences. The speculators have rationally determined expectations about future prices that incorporate all information about output, storage, and trade in both centers. The values of prices, inventories, and trade quantities at any point in the state space are determined by the four no-arbitrage conditions that must hold to prevent excess profits: at each point y_t,

$$\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^A|y_t] - P^A(y_t) \leq K^S;$$
$$\left[\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^A|y_t] - P^A(y_t) - K^S\right] \cdot S^A(y_t) = 0, \tag{6a}$$

$$\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^B|y_t] - P^B(y_t) \leq K^S;$$
$$\left[\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^B|y_t] - P^B(y_t) - K^S\right] \cdot S^B(y_t) = 0, \tag{6b}$$

$$\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^B|y_t] - P^A(y_t) \leq K^T;$$
$$\left[\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^B|y_t] - P^A(y_t) - K^T\right] \cdot T^A(y_t) = 0, \tag{6c}$$

$$\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^A|y_t] - P^B(y_t) \leq K^T;$$
$$\left[\left(\frac{1-\delta}{1+r}\right) E[P_{t+1}^A|y_t] - P^B\right) - K^T] \cdot T^B(y_t) = 0, \tag{6d}$$

where

$$P^i(y_t) = D_i^{-1}(X_t^i + M_t^i - S^i(y_t) - T^i(y_t)),$$
$$M_{t+1}^i(y_t) = (1 - \delta)(S^i(y_t) + T^j(y_t)), \text{ and}$$
$$E[P_{t+1}^i|y_t] = \iint_X D_i^{-1}(X_{t+1}^i + M_{t+1}^i(y_t) - S^i(y_{t+1}) - T^i(y_{t+1}))f(X_{t+1}^i, X_{t+1}^j) \times dX_{t+1}^i dX_{t+1}^j. \tag{7}$$

The first two of these inequalities are the conditions for profitable storage in either center, while the second two are the conditions for profitable trade between centers. Each of the four inequalities holds with equality if the control variables (storage or trade) are nonzero.

The model solution, which is found numerically, comprises two parts. The first part is the set of optimal storage and trade functions [$S^A(.)$, $S^B(.)$, $T^A(.)$, $T^B(.)$] that satisfy the four inequalities (6a)–(6d). Each function depends on the vector of four state variables. The second part of the solution is the distribution of the state variables that occurs in equilibrium, which depends on the assumed stochastic

[1] The assumption that mean output is predetermined and price inelastic can be relaxed; see Williams and Wright (1991), where they have a model of commodity prices in one center where future output is responsive to expected future prices. Williams and Wright also analyze a model of storage in which output is stochastic and seasonal, rather than stochastic and serially correlated. Bailey and Chambers (1996) and Deaton and Laroque (1996) argue that if demand is linear any effect on prices stemming from changes in the mean or variance of output can be replicated by a change in the demand function, so the shocks can be considered either demand shocks or supply shocks.

[2] This assumption is made for analytic convenience, as storage and trade decisions could be made at higher frequencies than the time it takes to ship goods. This would substantially complicate the model as all decisions would then depend on extra state variables representing quantities at different stages of transit.

TABLE 1.—ANALYTICAL RESULTS CORRESPONDING TO EQUATIONS (6a–6d)

T_t^A	T_t^B	S_t^A	S_t^B	Prob(.) $X^A = 100$	Prob(.) $X^A = 50$	$P_t^A - P_t^B$	$E[P_{t+1}^A - P_{t+1}^B \mid y_t]$
=0	=0	>0	>0	69.0%	0.9%	$\leq (K^T - K^S)$	$= \frac{1+r}{1-\delta}(P_t^A - P_t^B)$
=0	=0	=0	>0	0.3%	0%	$\geq -(K^T - K^S)$	$\leq \frac{1+r}{1-\delta}(K^T - K^S)$
=0	=0	>0	=0	0.3%	0%	$\leq (K^T - K^S)$	$\geq -\frac{1+r}{1-\delta}(K^T - K^S)$
=0	=0	=0	=0	0.4%	0%	Uncertain	Uncertain
=0	>0	>0	>0	10.3%	80.4%	$= (K^T - K^S)$	$= \frac{1+r}{1-\delta}(K^T - K^S)$
=0	>0	=0	>0	3.1%	1.6%	$\geq (K^T - K^S)$	$= \frac{1+r}{1-\delta}(K^T - K^S)$
=0	>0	>0	=0	0.4%	13.3%	$= (K^T - K^S)$	$\geq \frac{1+r}{1-\delta}(K^T - K^S)$
=0	>0	=0	=0	1.2%	3.8%	$\geq (K^T - K^S)$	$\geq \frac{1+r}{1-\delta}(K^T - K^S)$
>0	=0	>0	>0	10.3%	0%	$= -(K^T - K^S)$	$= -\frac{1+r}{1-\delta}(K^T - K^S)$
>0	=0	=0	>0	0.4%	0%	$= -(K^T - K^S)$	$\leq -\frac{1+r}{1-\delta}(K^T - K^S)$
>0	=0	>0	=0	3.1%	0%	$\leq -(K^T - K^S)$	$= -\frac{1+r}{1-\delta}(K^T - K^S)$
>0	=0	=0	=0	1.2%	0%	$\leq -(K^T - K^S)$	$\leq -\frac{1+r}{1-\delta}(K^T - K^S)$

The table gives the value of the price differential $P_t^A - P_t^B$ and the expected future price differential $E[P_{t+1}^A - P_{t+1}^B|y_t]$ where they can be determined exactly by equations (6a–6d), the arbitrage conditions describing storage and trade. The probabilities of each set of conditions occurring refer to the cases when A produces 50% of output ($X^A = 100$) and when A produces 25% of output ($X^B = 50$), using the standard parameters.

process determining output and the optimal storage and trade functions. The solution fulfills two conditions: first, that storage and trading decisions are profit-maximizing conditional on expectations of future prices; and, secondly, that price expectations are consistent with the storage and trading decisions and expectations of future output quantities.

B. *The Model Solution Technique*

A numerical solution to the model is calculated over a discrete four-dimensional grid corresponding to the four state variables. The solution technique has four steps. First, a discrete joint probability distribution over the grid values for the stochastic variables X^A and X^B is chosen, and the double integral formula in equation (7) is replaced by the equivalent summation formula. The joint probability density for X is chosen to mimic an autocorrelated process with normal innovations, and is represented by an $m^2 \times m^2$ Markov transition matrix Π specifying the probability of going from one point (X_{i1}^A, X_{j1}^B) to a second point (X_{i2}^A, X_{j2}^B).

The second step is to solve the optimal storage problem for the limiting "combined center" case when transport costs are 0 and trade is instantaneous, using techniques similar to those used by Deaton and Laroque (1995, 1996).[3] Linear demand functions for each center are specified:

$$D_i^{-1}(Q_t^i) = \begin{cases} \alpha^i & \text{if } Q_t^i = 0 \\ \alpha^i - \beta^i Q_t^i & \text{if } 0 < Q_t^i \leq \alpha^i/\beta^i. \\ 0 & \text{if } Q_t^i > \alpha^i/\beta^i \end{cases} \quad (8)$$

The third step is an algorithm that calculates the optimal amounts of storage and trade in the two centers. The algorithm, which is briefly described in appendix A, constructs a series of successive approximations to the optimal storage and trade functions, and is repeated until the difference between successive values of the control values is small.

The fourth step, once the optimal storage and trade functions are found, is the calculation of the invariant probability distribution of the model solution. The invariant distribution of the model is the unconditional probability of being at a particular grid point, which is used to calculate various statistics about the price distributions in each center. The method used to calculate the invariant distribution is also described in appendix A.

III. Properties of the Model

A. *Analytical Results*

The key results of the paper are derived analytically by considering various combinations of the complementary conditions associated with equations (6a)–(6d). The twelve possible combinations of storage and trade quantities, along with the value of the price differential $P_t^A - P_t^B$ and the expected price differential $E[P_{t+1}^A - P_{t+1}^B \mid y_t]$, are shown in table 1.[4] The optimal storage and trade functions fulfill the set of conditions (6a)–(6d) irrespective of the assumed stochastic process or parameters, but the distribution of the state variables in equilibrium changes as these assumptions change. Table 1 also shows the probability of each of the twelve combinations of complementary conditions occurring when center A produces 50% or 25% of total output.[5]

[3] The centers are combined to have a single demand function and a single production function, and the optimal storage function is calculated using equation (6a) applied to the combined centers.

[4] There are only twelve combinations because the centers will not export simultaneously as $K^S < K^T$.

[5] The probabilities are calculated using the parameters described in footnote 7, with $(X^A, X^B) = (100,100)$ if the centers are the same or (50, 150) if center B mainly produces. The probabilities are representative of

There are two key analytic results. First, if inventories in both centers are positive, the prices in each center will differ by no more than $K^T - K^S$, and prices will be expected to increase over time. Secondly, exports will only occur at time t if there is some probability that inventories in the importing center will be 0 when the goods arrive at time $t + 1$.

To establish the first result, consider a point y_t at which inventories are positive in each center (rows 1, 5, or 9 of table 1). Because inventories are positive, the price in each center is expected to increase and the price difference is expected to increase; more precisely, by equations (6a) and (6b)

$$E[P^i_{t+1}|y_t] = \left(\frac{1+r}{1-\delta}\right)(P^i(y_t) + K^S), \quad i = A, B, \tag{9a}$$

and consequently

$$E[P^A_{t+1} - P^B_{t+1}|y_t] = \left(\frac{1+r}{1-\delta}\right)(P^A(y_t) - P^B(y_t)). \tag{9b}$$

Furthermore, equations (6b) combined with (6c) and (6a) combined with (6d) imply the spatial price difference lies within the range

$$-(K^T - K^S) \leq P^A_t - P^B_t \leq (K^T - K^S) \tag{10}$$

with the lower inequality holding when exports from A to B are just profitable and the upper inequality holding when exports from B to A are just profitable. If this were not the case, profits could be made by reducing inventories in the low-priced center and exporting to the high-priced center.

To establish the second result, consider a point y_t at which arbitrageurs in center B export to center A (rows 5–8 of table 1). (A parallel set of results hold when center A exports to center B.) There are four possible combinations of inventories. If both center A and center B have positive inventories, equations (6a), (6b), and (6d) hold with equality and imply

$$P^A(y_t) - P^B(y_t) = K^T - K^S, \tag{11a}$$

$$E[P^A_{t+1} - P^B_{t+1}|y_t] = \frac{1+r}{1-\delta}(K^T - K^S). \tag{11b}$$

If center B but not A has positive inventories, equations (6b) and (6d) hold and imply

$$P^A(y_t) - P^B(y_t) \geq K^T - K^S, \tag{12a}$$

$$E[P^A_{t+1} - P^B_{t+1}|y_t] = \frac{1+r}{1-\delta}(K^T - K^S). \tag{12b}$$

If center A but not B has positive inventories, equations (6a) and (6b) hold with equality and imply

$$P^A(y_t) - P^B(y_t) = K^T - K^S, \tag{13a}$$

$$E[P^A_{t+1} - P^B_{t+1}|y_t] \geq \frac{1+r}{1-\delta}(K^T - K^S). \tag{13b}$$

When neither center has inventories, equation (6d) holds and

$$P^A(y_t) - P^B(y_t) \geq K^T - K^S, \tag{14a}$$

$$E[P^A_{t+1} - P^B_{t+1}|y_t] \geq \frac{1+r}{1-\delta}(K^T - K^S). \tag{14b}$$

Equations (11b), (12b), (13b), and (14b) all indicate that exports to A only occur if at time t the expected future price difference at $t + 1$ is at least as large as $\frac{1+r}{1-\delta}(K^T - K^S)$. However, it is evident from considering equations (11a), (12a), (13a), and (14a) and the remaining combinations listed in table 1 that the price differential $P^A_t - P^B_t$ only exceeds $K^T - K^S$ if the storage quantity in center A is 0. It follows that goods will only be exported from center B at time t if it is expected that on some occasions they will arrive when center A has zero inventories and prices are unusually high. To ensure that there are some of these occasions, shippers limit their exports so that on some occasions inventories in the importing market fall to 0 and prices temporarily spike upwards. They would not cover their interest and depreciation costs on average if this did not happen.

Because commodity prices are a convex function of commodity availability when inventory management is rational, the distribution of returns to the shipper is skewed.[6] The shipper exports a quantity of goods that ensures the average price difference in the subsequent period will be at least $\frac{1+r}{1-\delta}(K^T - K^S)$. If local output is higher than expected in the destination center when the goods arrive, the surplus can be stored and prices will only be a little lower than expected; but if output is lower than expected, there will be a shortage and prices will be sharply higher than expected. For the expected zero profit condition to hold, the quantity shipped must be sufficiently large that small numbers of sharp price increases are offset by large numbers of small price decreases. If output shocks are symmetric, arbitrageurs will choose a quantity that ensures the median trade is unprofitable, but which will generate large profits on the few occasions that the exports arrive when supplies are unusually low. In turn, the importing center will normally have positive inventories, because on most occasions the imported shipment will be larger than needed for immediate consumption.

those pertaining to a wide variety of parameters, for a given level of regional specialization.

[6] See Deaton and Laroque (1992) for the convexity result that holds in the one-center storage model.

TABLE 2.—PRICE, STORAGE, AND TRADE STATISTICS CORRESPONDING TO THE MODEL
($\bar{X}^A = 100$, $\bar{X}^B = 100$, CHANGING TRANSPORT COSTS)

Statistic	$K^T = 1$	$K^T = 2.5$	$K^T = 5$	$K^T = 10$	$K^T = 20$	No Trade
Prices						
Mean (P^A)	100.07	100.07	100.08	100.10	100.11	100.12
S. Dev. (P^A)	6.2	6.2	6.2	6.7	7.6	8.1
Mean ($P^A - P^B$)	0	0	0	0	0	0
S. Dev. ($P^A - P^B$)	1.8	2.8	4.3	6.9	10.1	11.3
S. Dev. ($\Delta P^A - \Delta P^B$)	2.0	2.4	2.9	3.3	4.3	4.3
% ($\|P^A - P^B\| > K^T - K^S$)	9%	8%	9%	9%	6%	0%
Storage						
Mean (S^A)	64	70	80	94	111	116
S. Dev. (S^A)	81	86	92	102	113	119
% ($S^A = 0$)	8%	7%	7%	7%	11%	12%
% (S^A, $S^B = 0$)	5%	4%	3%	2%	1%	1%
Trade						
Mean (T^A)	3	2	2	1	0.2	0
S. Dev. (T^A)	6	6	5	4	2	0
% ($T^A = 0$)	77%	81%	85%	88%	97%	100%
% (T^A, $T^B = 0$)	54%	61%	70%	77%	94%	100%

P^A: the price in center A. S^A: storage in center A. T^A: trade from center A to center B.
% ($|P^A - P^B| > K^T - K^S$): the fraction of time the price difference exceeds the difference between the trade cost and the storage cost. Note $K^S = 0$ in these simulations.
% ($S^A[T^A] = 0$): the fraction of time storage [exports] = 0.
The "no trade" column refers to the case when transport costs are so high that trade does not occur.

B. *Numerical Results*

In the rest of this section, results from numerical simulations are presented to illustrate how trade flows, storage quantities, and prices depend on transport costs. The results depend on the extent to which output is regionally specialized. When the centers are identical, they both tend to have large inventories, and goods are exported as frequently one direction as the other; but when one center produces a disproportionate fraction of output, it tends to hold most of the inventories and export most of the time. The solutions are sufficiently distinctive that results corresponding to the two cases are presented.

For the simulations, the model parameters are chosen so that the centers have identical demand functions, the price elasticity in each center at average consumption is −1, the storage cost K^S is 0, the period is one week, and the annual interest and depreciation rates are 5%.[7] These parameters are chosen not because they are representative of any particular market—they are not—but for their ease of interpretation as average consumption and prices are both approximately 100, and the price elasticity at average consumption is −1.

C. *The Two-Center Model with Equal Centers*

Table 2 shows how trade flows, storage quantities, and prices depend on transport costs when the centers are identical—that is, when each center produces 50% of total output. Three features of the results stand out. First, prices are primarily smoothed through the adjustment of inventories, not through the transport of goods. Because it is less expensive to adjust inventories than to ship goods, inventories are adjusted to keep the spatial price difference less than the cost of transporting goods whenever possible. In the simulations, the frequency that each center exports declines sharply as transport costs increase, from 23% when transport costs are 1% of average prices to 3% when transport costs are 20% of average prices. Even when transport costs are minimal, the mean value of exports is very small, in all cases less than 25% of the standard deviation of weekly output.[8] In contrast, inventories are held in each center approximately 90% of the time and the mean value of inventories in each center is at least four times the standard deviation of weekly output. Even though trade is relatively infrequent, the spatial price difference never exceeds $K^T - K^S$ more than 10% of the time because of the way that inventories are managed.

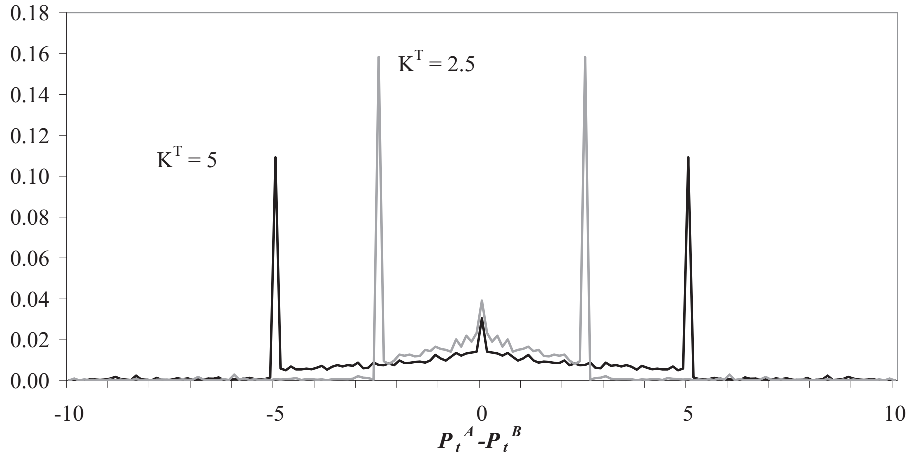
Secondly, mean inventory levels and mean availability decline as transport costs decline. Inventories are reduced by a substantial amount as transport costs decline, even though mean trade volumes remain tiny. The reduction occurs because it becomes cheaper to import from the other center in times of shortage rather than to accumulate large inventories in anticipation of a shortage. The reduction in inventories is offset by the higher average consumption that stems from the reduction in average prices.

Thirdly, the variance of the price difference is increasing in transport costs. The relationship between transport costs and the distribution of $P_t^A - P_t^B$ is best understood by examining figure 1, which shows the complete distribution

[7] In the baseline case, the following model parameters are used: the demand function $D_i^{-1}(Q) = \alpha - \beta Q = 200 - Q$; the production conditional variance $\sigma^2 = 36$; the production autocorrelation $\rho = 0.9$; the weekly interest rate $r = 0.001$; the weekly depreciation rate $\delta = 0.001$; $K^S = 0$; and $K^T = 5$.

[8] In the simulations, the standard deviation of shocks to output is 6, but because output is serially correlated, the standard deviation of output is 13.8.

FIGURE 1.—PROBABILITY DENSITY OF THE SPATIAL PRICE DIFFERENCE
$X^A = 100$, $X^B = 100$, $K^T = 2.5$ OR 5, $K^S = 0$



function of $P_t^A - P_t^B$ when $K^T = 2.5$ and $K^T = 5$. The distribution is symmetric with a local maximum at $(P_t^A - P_t^B) = 0$.[9] The main features of the distribution are the spikes at $\pm (K^T - K^S)$ corresponding to occasions when one center exports and both have inventories, and the density outside these spikes corresponding to occasions when the importing center has zero inventories and a localized price spike occurs. As transport costs increase, the probability mass at $\pm (K^T - K^S)$ decreases, although the probability that the spatial price difference lies outside these bands changes little, at least until transport costs become very large.

The supply and demand parameters of the model can also be manipulated to demonstrate the effect of adjusting the transport speed, holding transport costs constant. For example, the effect of doubling the transport speed can be estimated by halving the time period, halving mean output, and doubling the slope of the demand curve.[10] The simulations (not reported) show that mean inventory levels, trade volumes, the probability that a center has zero inventories, and the variance of $P_t^A - P_t^B$ all decrease as transport speed increases.

D. *The Two-Center Model with a Dominant Exporter*

There are two cases to note when one center is a dominant exporter: first, when the production asymmetry between the centers is small, so that both centers regularly export; and secondly, when production is sufficiently specialized that trade almost always flows from one center to other. These cases are modeled by altering the fraction of total output produced in center A from 50% to 15%, and the results are presented in tables 3 and 4.

When the production asymmetry is small, the solution is similar to the symmetric case and goods are regularly transported in both directions. The mean spatial price difference is less than the transport cost because inexpensive storage technology is used more frequently than expensive transport technology to arbitrage prices. Average inventory levels remain large in both centers. As output in center A decreases, the probability mass at $P_t^A - P_t^B = K^T - K^S$ becomes larger than the probability mass at $P_t^A - P_t^B = -(K^T - K^S)$ and the density in the region $|P_t^A - P_t^B| \leq K^T - K^S$ is increasing in the spatial price difference.

As output becomes more regionally specialized, one center exports to the other most of the time and reverse direction exports almost never occur (see table 3). When A's average production declines from 50% to 45% of the total, the fraction of the time that A exports declines from 15% to 1%, while the fraction of the time it imports increases from 15% to 53%. When A produces 40% of average production, it never exports and imports 86% of the time, the exception being periods of exceptional harvests. Mean trade volumes from the exporting center are close to the level that would prevail if there were no uncertainty. As center A's average production declines further, it imports nearly every period, and average imports increase one for one with the reduction in local production.

Table 4 shows how storage, trade, and prices depend on transport costs when center A produces 25% of total output and imports nearly every period. Three features of this table stand out. First, inventories in the importing center are positive, although small, most of the time. Each period arbitrageurs in center B export a quantity of goods that ensures that inventories in A only fall to 0 if it has a particularly bad harvest in the subsequent period; consequently, when harvests are normal or large, there are suffi-

[9] The accuracy of the price distributions is limited because of the discrete grid. The sharp spike at $P_t^A - P_t^B = 0$ in figure 1 reflects the occasions when output and storage are the same in each center (an event with zero probability in the continuous case) and the spatial price difference is exactly 0.

[10] It is assumed that there is only one shipment per period no matter the length of the transport period.

Table 3.—Price, Storage, and Trade Statistics Corresponding to the Model ($K^T = 5$, $K^S = 0$, mean output changes across columns)

Statistic $\bar{X}^A =$ $\bar{X}^B =$	100 100	90 110	80 120	70 130	50 150	30 170
Prices						
Mean (P^A)	100.08	102.20	102.69	102.75	102.77	102.77
S. Dev. (P^A)	6	6	6	6	6	6
Mean(P^B)	100.08	97.97	97.48	97.43	97.43	97.45
S. Dev.(P^B)	6	6	6	6	6	6
Mean($P^A - P^B$)	0	4.2	5.2	5.3	5.3	5.3
S. Dev.($P^A - P^B$)	4.3	2.9	2.1	2.3	2.4	2.0
S. Dev($\Delta P^A - \Delta P^B$)	2.9	2.7	2.8	3.2	3.3	2.7
% ($\lvert P^A - P^B \rvert > K^T - K^S$)	9%	7%	6%	5%	5%	6%
Storage						
Mean(S^A)	80	33	20	19	20	19
S. Dev.(S^A)	92	41	16	13	16	14
% ($S^A = 0$)	7%	6%	6%	6%	5%	6%
Mean(S^B)	80	123	134	134	131	124
S. Dev.(S^B)	92	132	146	147	142	138
% ($S^B = 0$)	7%	10%	16%	17%	17%	18%
% (S^A, $S^B = 0$)	3%	3%	4%	5%	4%	4%
Trade						
Mean(T^A)	1.6	0.1	0	0	0	0
S. Dev.(T^A)	4.9	0.9	0	0	0	0
% ($T^A = 0$)	85%	99%	100%	100%	100%	100%
Mean(T^B)	2	8	17	27	47	67
S. Dev.(T^B)	5	10	12	14	15	13
% ($T^B = 0$)	85%	47%	14%	3%	1%	0%
% (T^A, $T^B = 0$)	70%	46%	14%	3%	1%	0%

P^A: the price in center A. S^A: storage in center A. T^A: trade from center A to center B.
% ($\lvert P^A - P^B \rvert > K^T - K^S$): the fraction of time the price difference exceeds the difference between the trade cost and the storage cost. Note $K^T = 5$ and $K^S = 0$ in these simulations.
% ($S^A[T^A] = 0$): the fraction of time storage [exports] = 0.

Table 4.—Price, Storage, and Trade Statistics Corresponding to the Model ($\bar{X}^A = 50$, $\bar{X}^B = 150$, changing transport costs)

Statistic	$K^T = 1$	$K^T = 2.5$	$K^T = 5$	$K^T = 10$	$K^T = 20$
Mean(P^A)	100.7	101.5	102.8	105.4	110.5
S. Dev. (P^A)	6.6	5.6	6.5	6.5	7.2
Mean(P^B)	99.5	98.7	97.4	94.9	89.8
S. Dev.(P^B)	6	6	6	5	5
Mean($P^A - P^B$)	1.2	2.8	5.3	10.5	20.7
S. Dev.($P^A - P^B$)	1.4	1.8	2.4	3.4	4.7
S. Dev($\Delta P^A - \Delta P^B$)	1.7	2.4	3.3	4.8	6.7
% ($\lvert P^A - P^B \rvert > K^T - K^S$)	5.1%	5.0%	5.3%	6.3%	5.3%
Mean(S^A)	19	19	20	22	23
S. Dev.(S^A)	14	15	16	20	18
% ($S^A = 0$)	5%	5%	5%	6%	5%
Mean(S^B)	99	110	131	183	305
S. Dev.(S^B)	116	126	142	171	184
% ($S^B = 0$)	20%	19%	17%	13%	4%
% (S^A, $S^B = 0$)	5%	4%	4%	3%	1%
Mean(T^A)	0	0	0	0	0
S. Dev.(T^A)	0	0	0	0	0
% ($T^A = 0$)	100%	100%	100%	100%	100%
Mean(T^B)	49	49	47	45	40
S. Dev.(T^B)	14	14	15	16	19
% ($T^B = 0$)	0.1%	0.5%	1%	1.5%	3.6%
% (T^A, $T^B = 0$)	0.1%	0.5%	1%	1.5%	3.6%

P^A: the price in center A. S^A: storage in center A. T^A: trade from center A to center B.
% ($\lvert P^A - P^B \rvert > K^T - K^S$): the fraction of time the price difference exceeds the difference between the trade cost and the storage cost. Note $K^T = 5$ and $K^S = 0$ in these simulations.
% ($S^A[T^A] = 0$): the fraction of time storage [exports] = 0.

cient quantities of output available that a small surplus is held over to the next period. The level of inventories in the importing center varies little with transport costs.[11] Rather, it is primarily determined by the standard deviation of output in the importing center. In the simulations, inventories in the importing center fall to 0 about 5% of the time, and thus the spatial price difference exceeds $K^T - K^S$ about 5% of the time.

Secondly, the mean spatial price difference exceeds $K^T - K^S$ by a small amount that is increasing in the transport cost.

[11] In addition, the mean level of inventories in the importing center varies little with average production levels once the importing center imports most of the time.

FIGURE 2.—PROBABILITY DENSITY OF THE SPATIAL PRICE DIFFERENCE
$X^A = 50$, $X^B = 150$, $K^T = 2.5$ OR 5, $K^S = 0$

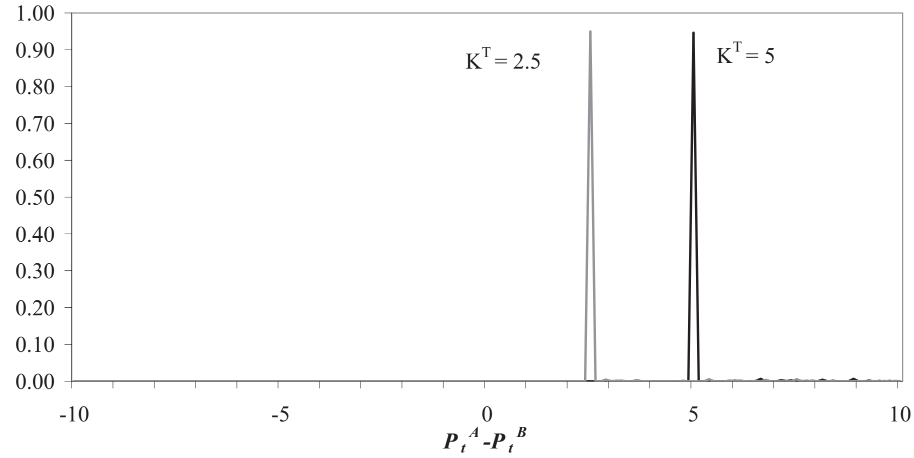


Figure 2 shows the distribution of the spatial price difference. $P_t^A - P_t^B$ equals $K^T - K^S$ most of the time, and exceeds it on the occasions that center A runs out of inventories. The variance of $P_t^A - P_t^B$ also increases with transport costs, but the relationship is much weaker than when the centers are equal size. The variance increases slightly with transport costs because average availability in the importing center decreases slightly with transport costs; as a result, prices increase by a greater amount when the importing center runs out of inventories, as available supplies are smaller.

Thirdly, inventories in the exporting center decrease as transport costs fall. The decline is substantial. As transport costs decrease, prices in the importing center decrease and prices in the exporting sector increase. The increase in average prices in the exporting center makes it more expensive to store goods, and thus the quantity of inventories held decreases. The decrease in average inventories is offset by higher average consumption. Since average trade volumes increase only modestly as transport costs decline, total availability also declines.

E. *Transport Costs, Prices, and Output Specialization*

The relationship between transport costs and the spatial distribution of prices in this model is of some interest as several papers have tried to estimate transport costs from various moments of the joint price distribution. For instance, Engel and Rogers (1996) analyzed prices in different North American cities and showed the standard deviation of $\Delta P_{t+1}^A - \Delta P_{t+1}^B$ was an increasing function of the distance between cities. They used this relationship to estimate the extent to which the U.S.-Canada border disrupted trade.

However, this model suggests that it is difficult to make strong inferences about transport costs from price data alone as the relationship between the standard deviations of $P_t^A - P_t^B$ and of $\Delta P_{t+1}^A - \Delta P_{t+1}^B$ depend on the extent to which output is regionally specialized. For example, tables 2 and 4 indicate that the standard deviation of $P_t^A - P_t^B$ is much higher when the two centers are equal than when one center produces most of the output, for any level of transport costs. This is because prices usually vary within the whole range of the band $|P_t^A - P_t^B| \leq (K^T - K^S)$ when average output is the same in each center, but the spatial price difference is normally equal to $K^T - K^S$ when one center produces most of the output.

Figure 3 shows how the standard deviation of $\Delta P_{t+1}^A - \Delta P_{t+1}^B$ depends on transport costs when A produces either 50% or 25% of total output. It shows the standard deviation of $\Delta P_{t+1}^A - \Delta P_{t+1}^B$ increases faster as a function of transport costs when output is regionally specialized than when it is equally divided between the centers. Consequently, one needs to be careful making the inference that transport costs between a pair of cities are high because the standard deviation of $\Delta P_{t+1}^A - \Delta P_{t+1}^B$ is high. It could simply be the case that output for that good has a large degree of regional specialization.

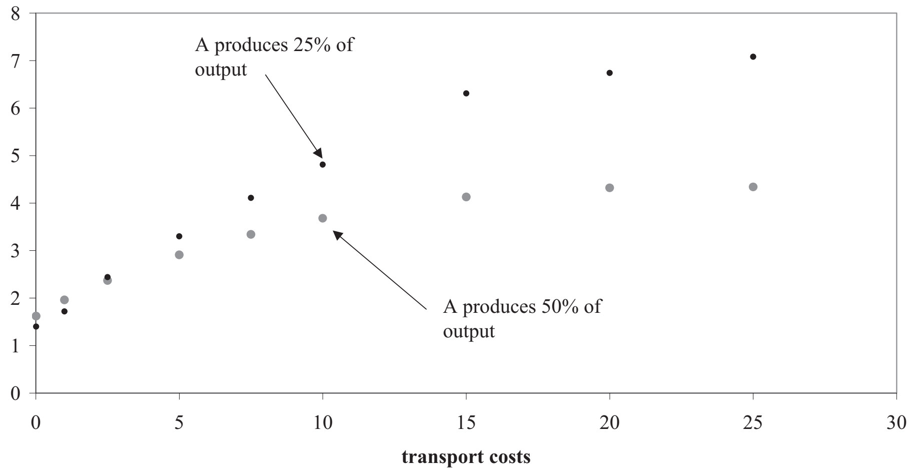
IV. The Late-Nineteenth-Century U.S. Corn Trade

During the late nineteenth century, corn was shipped from all over the Great Plains to Europe. Chicago was the preeminent inland shipping port, receiving and shipping 67 million bushels per year (on average) between 1878 and 1890. Most of this grain was sent to East Coast ports prior to export. New York was the most important port, exporting as much as Boston, Baltimore, and Philadelphia combined.[12]

In the next section, weekly data are used to test two of the arbitrage relationships implied by the set of equations (6a–6d). In particular, the data are used to examine whether New York prices were equal to Chicago prices plus transport

[12] Most corn was consumed where it was grown, and only a small fraction was shipped any distance. Illinois produced some 225 million bushels of corn annually, over 10% of the U.S. crop. In contrast, total production in New York, New Jersey, and Pennsylvania was 75 million bushels per year.

Figure 3.—Standard Deviation of ($\Delta P_T^A - \Delta P_T^B$) as a Function of Transport Costs



costs on weeks during which trade occurred. As a precursor to this examination, in this section it is shown that trade occurred continuously during the open water season, but hardly at all during the winter months, and that the appropriate transport cost was the cost of shipping grain over the Great Lakes and the Erie Canal. Indeed, I do not analyze the winter month prices because in the absence of trade there is no reason why the New York price would equal the Chicago price plus the transport cost.[13]

Corn was forwarded from Chicago to New York in one of three ways. The slowest method was to ship corn to Buffalo via the Great Lakes, and then to send it to New York via the Erie Canal. This method took three weeks, but was unavailable between November and late April when the lakes were frozen. A faster method, taking ten days, was to ship corn to Buffalo via the Great Lakes and then to send it to New York by rail. The fastest and most expensive method was to send corn to New York by rail, a trip that took three or four days. Between 1881 and 1891, when average annual costs were reasonably stable, the average cost of shipping a bushel of corn from Chicago to New York was 7.7 cents by lake and canal, 10.3 cents by lake and rail, and 14.6 cents by rail.[14] The average price of a bushel of corn in Chicago during this time was 45 cents.

Between 1878 and 1890, the vast majority of the corn sold in Chicago and then shipped to New York was shipped by lake and canal. This fact has to be demonstrated indirectly, by examining Chicago exports and New York imports separately, as bilateral trade data from Chicago to New York do not exist. First, it is shown that most corn leaving Chicago went by ship. On average, 67 million bushels of corn were shipped from Chicago each year between 1878 and 1890. Of these, 45 million bushels of corn per year were shipped from Chicago by lake, and 22 million were transported by rail. Of the latter, however, 15 million bushels were "through-shipments"—shipments that started west of Chicago, that were routed through Chicago, but that were never sold or handled in Chicago.[15] These through-shipments were included in the annual and weekly shipping statistics, creating a misleading picture of the true extent to which corn sold in Chicago was transported by rail. When the through-shipments are excluded from the Chicago shipments, the fraction of corn sold in Chicago that was shipped by lake increases from 67% to 86%. Lake transport was even more dominant during the open water season. Of the 7 million bushels of corn sold in Chicago and sent by rail, 4.3 million bushels were sent between December and March. Consequently, 95% of corn that was transported in the open water season was shipped by lake.[16] Figure 4 shows the seasonal pattern of the shipments.

Secondly, it is shown that most Chicago corn was shipped through Buffalo, and that most New York corn came from Buffalo. Detailed shipping statistics for New York are available for the years 1878–1881 and 1888–1890. During these years, New York received 40 million bushels of corn (on

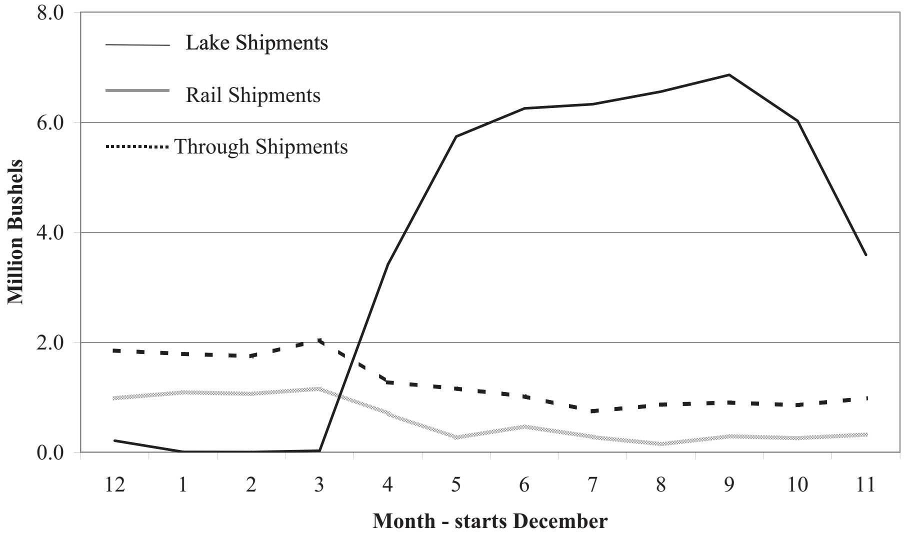
[13] In the winter New York was supplied from other places as the New York price was normally lower than the Chicago price plus the cost of rail transport. Many of these places were west of Chicago as in some places the all rail route to New York was cheaper than the rail route to Chicago plus the cost of lake and canal transport from Chicago to New York when Chicago handling fees were added (U.S. Bureau of Statistics, Treasury Department, 1879, pp. 103–106; and Tunell, 1897). Chicago shippers preferred to store goods in Chicago through the winter and ship during spring rather than to ship during the winter.

[14] Chicago Board of Trade (1892, p. 122).

[15] The through-shipments are calculated as the total of the monthly through-shipments on the Chicago and Northwestern; Illinois Central; Chicago, Burlington, and Quincy; Chicago, Rock Island, and Pacific; and Chicago and Alton railroads that are reported in the Chicago Board of Trade annual reports each year.

[16] Unless most rail shipments were through-shipments there are two puzzles. First, rail and lake shipments took place simultaneously even though rail rates were considerably higher than lake rates. There are some reasons why rail shipment was preferred—it was faster, and had less risk of additional heat damage if the grain were already damaged—but ordinarily these did not justify the extra cost. Secondly, rail shipments during winter occurred when the price gap between New York and Chicago was lower than the rail cost.

FIGURE 4.—SEASONAL PATTERN OF CHICAGO CORN SHIPMENTS, 1878–1890



average), of which 21 million arrived by canal boat, 16 million arrived on trains coming from Buffalo, and 3 million came from elsewhere. In the same period, 39 million bushels of corn (on average) were shipped by lake from Chicago to Buffalo, and another 10 million bushels were sent by rail.[17] It is established in appendix C that most of the corn sent from Chicago to Buffalo by rail was forwarded to New York by rail, meaning that at most 6 million bushels of the grain arriving in New York by rail from Buffalo could have arrived in Buffalo by ship from Chicago. Hence, 21 million or 78% of the 27 million bushels of corn that arrived in Buffalo by water were forwarded to New York by canal. The lake and canal route was more dominant than this fraction suggests, however, because the lake and canal season was three weeks shorter than the lake and rail season. Since a typical season lasted thirty weeks, the lake and canal route accounted for 87% of the grain arriving in New York during the period that the lake and canal route was open.[18]

High-frequency shipping data suggest that corn was sent from Chicago on most of the weeks that the lakes were open. The lowest monthly shipment by lake from Chicago between May and November in any year was 2.2 million bushels, and there were only three weeks between May and mid-November during the entire period when the weekly shipment by either lake or rail was less than 0.5 million bushels. The weekly import statistics for New York, which are not available between 1882 and 1888, suggest some of this grain arrived in New York each week. Between the middle of May and the end of November, corn receipts in New York exceeded 0.25 million bushels on all but one of the weeks in 1879, 1880, 1881, 1889, and 1890. In 1891, when overall corn shipments were smaller, there were nine weeks when weekly imports were less than 250,000 bushels, but in all of these cases imports exceeded 150,000 bushels.

A. *Transport Prices between Chicago and New York*

There was a marked seasonal pattern in shipping costs (see figure 5). Lake and canal and lake and rail prices were typically high at the opening of the season, but they declined during early summer before rising steadily between July and the end of the shipping season. Rail rates varied seasonally between winter and summer, particularly before 1886 when railroads competed aggressively with each other and with shipping lines for the grain business. The competition was sufficiently fierce that large quantities of grain were shipped by rail as through-shipments from locations west of Chicago, although not from Chicago itself (Tunell, 1897; United States Treasury, 1898). It is likely that the Chicago Board of Trade data on rail prices overstates the true rail prices, as much of the business was transacted at lower, unrecorded prices.[19]

B. *New York and Chicago Inventories*

Grain arriving in New York was transferred to an elevator or a lighter. Grain was often stored temporarily, but the storage capacity was rarely fully utilized, even in winter.[20]

[17] The data are from the annual reports of the Chicago Board of Trade. Only 70% of corn sent by ship from Chicago went to Buffalo. The rail shipments are the sum of the shipments sent on the Michigan Central and the Lake Shore and Michigan railroads. The averages are calculated for the years 1878–1881 and 1888–1890 since New York data for 1882–1887 are unavailable.

[18] That is, 21 million out of the 24 million bushels that arrived in New York during the lake and canal season.

[19] See the discussion by Nimmo in his reports on the internal commerce of the United States (United States Bureau of Statistics, 1877, 1881, 1884). Porter (1983, 1985) discusses the pre-1886 price wars.

[20] In 1888 there were 24 million bushels storage capacity in New York and Brooklyn, and a further 3 million in New Jersey. However, peak grain storage in New York and Brooklyn between 1887 and 1889 was just over 16 million bushels, of which 11 million bushels were wheat and 4 million bushels were corn.

FIGURE 5.—AVERAGE TRANSPORT COSTS, 1880–1891, BY WEEK

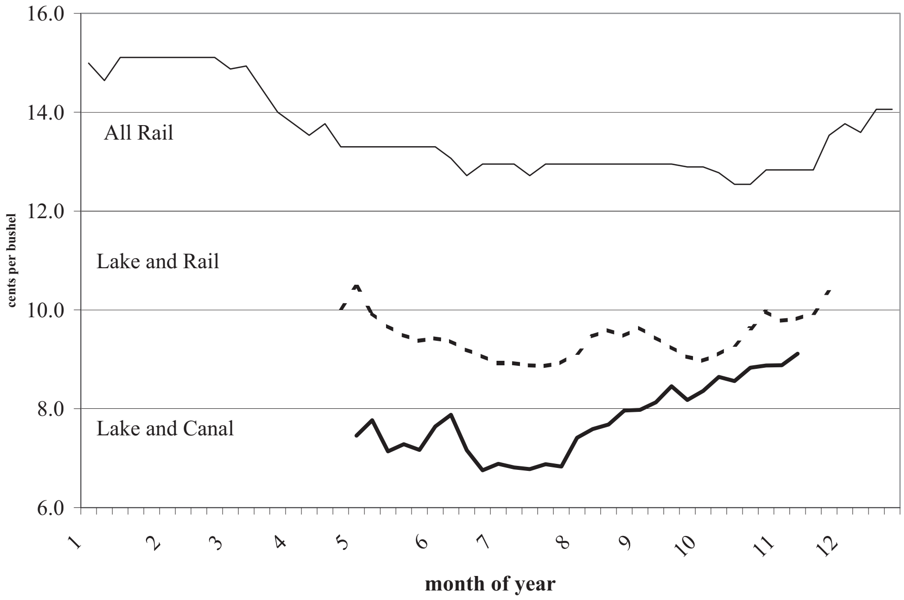
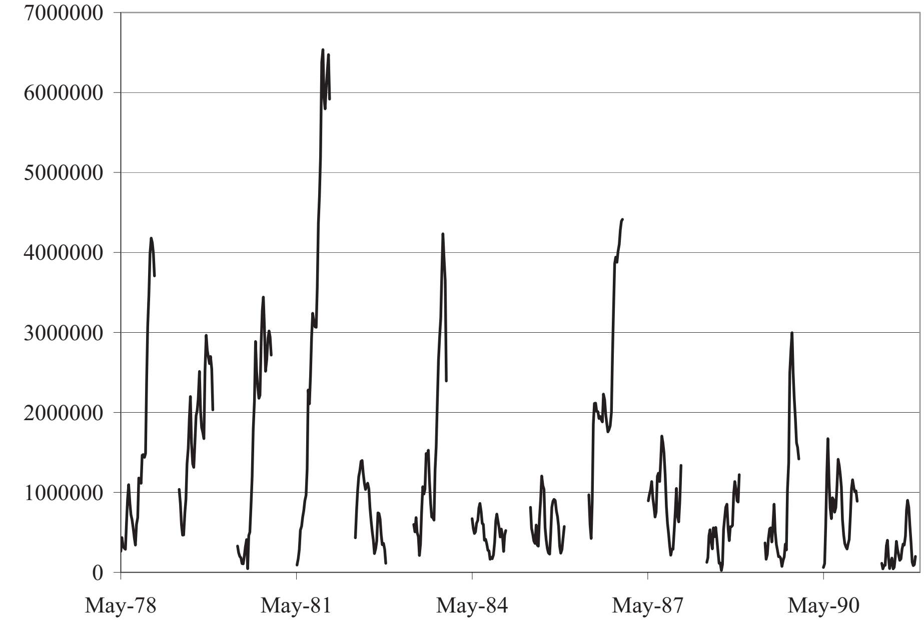


FIGURE 6.—"SUMMER" STORAGE (MAY–NOVEMBER), NEW YORK, 1878–1891



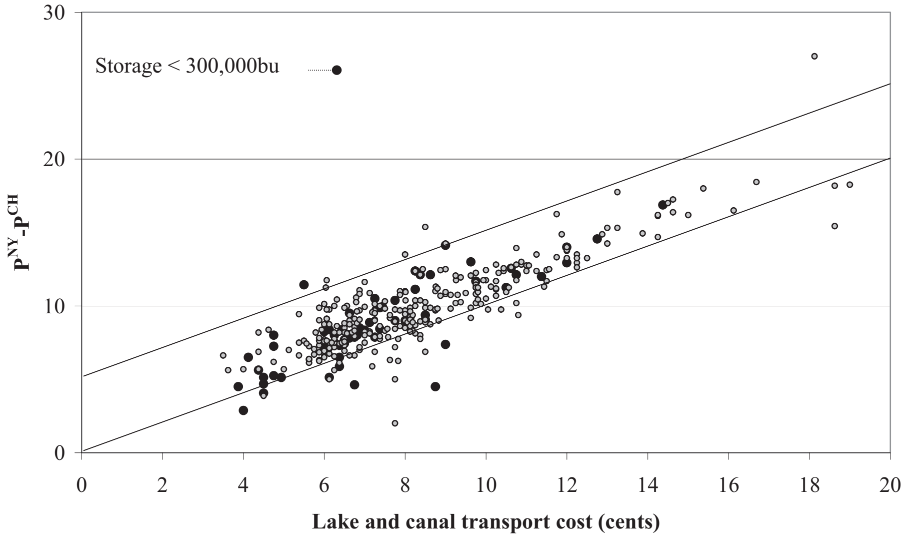
Inventory levels never fell to 0, because the elevators always contained some grain in transit as they were used to transfer grain from arriving canal boats and rail cars to departing ocean vessels.[21] For this reason, in the empirical work that follows I distinguish inventory levels by whether they were less than or greater than 300,000 bushels, this level being an estimate of the amount of grain that would be in elevators even if none were held for speculative purposes.

[21] See Harley (2004) for a discussion about the New York–Europe grain trade. Scheduled liners often used grain as ballast, and on occasion would offer to purchase grain at high prices (or to ship grain at very low rates) to obtain it immediately before departure.

Corn inventories typically fell to their low points each year in the middle of May, prior to the opening of the waterways, and in August. Corn inventories between 1878 and 1891 are shown in figure 6.

Corn inventories in Chicago had a marked seasonal pattern. Receipts in Chicago occurred throughout the year, but were higher than shipments between December and March, when the lakes were closed, and in August and September. Inventories peaked in March, and averaged 4 million bushels on shore, with a further 2.2 million bushels on ships. Inventories reached seasonal lows in late July and November, when they averaged 1.4 million bushels.

FIGURE 7.—NY FUTURE AND CHICAGO SPOT PRICE DIFFERENCE VERSUS TRANSPORT COST, 1878–1891
(NEW YORK FUTURE IS "DELIVERY THIS MONTH" IF DATE IS BEFORE THE TENTH; OTHERWISE "DELIVERY NEXT MONTH.")



Note: The diagonal lines indicate $P_t^{NY} - P_t^{CH} = K^T$ and $P_t^{NY} - P_t^{CH} = K^T + 5$.

Storage charges varied little during the period (Goldstein, 1928; Ulen, 1982). In 1888, it cost 0.625 cents per bushel to deposit grain in an elevator, including the cost of ten days' storage; thereafter, storage cost 0.25 cents per bushel per ten days. There were additional charges for trimming from canal boats and to ocean ships. Charges in Chicago were similar. Other costs of holding inventories included insurance and the opportunity cost of buying grain. Working (1929) estimated that these costs were approximately 1.4 cents a bushel per month in 1913.

C. Market Prices

The standard spot and futures contracts in both New York and Chicago were settled by the delivery of grain to a warehouse or elevator. The main contracts were for immediate delivery (the spot contract) or for delivery at any time within the current month, the next month, two months' time, or in May of a particular year (the futures contracts). The seller had the option as to the date in a particular month the grain was delivered, so spot prices normally exceeded or were equal to the zero-mouth future price. This paper uses Wednesday closing prices for all of the analysis. (See appendix B for the precise definition and source of the data.)

V. Spatial Corn Price Arbitrage and Transport Costs, 1878–1891

In this section I use the New York and Chicago data to examine two implications of the model developed in section II. First, the relationship between the Chicago spot price, the transport cost, and the New York future price is analyzed. Since exports left Chicago practically every week, the price expected to prevail in New York three weeks after the exports were sent—here proxied by the price for future delivery in New York—should have equaled the Chicago spot price plus the transport cost, irrespective of inventory levels in New York. Secondly, the difference between the Chicago and New York spot prices and the transport cost is examined. According to the model, the New York price should have equaled the Chicago spot price plus the transport cost minus the storage cost on any day corn was shipped if there were positive inventories in New York, but it should have exceeded this level if inventories in New York were 0.

A. Transport Costs, the Chicago Spot Price, and the New York Future Price

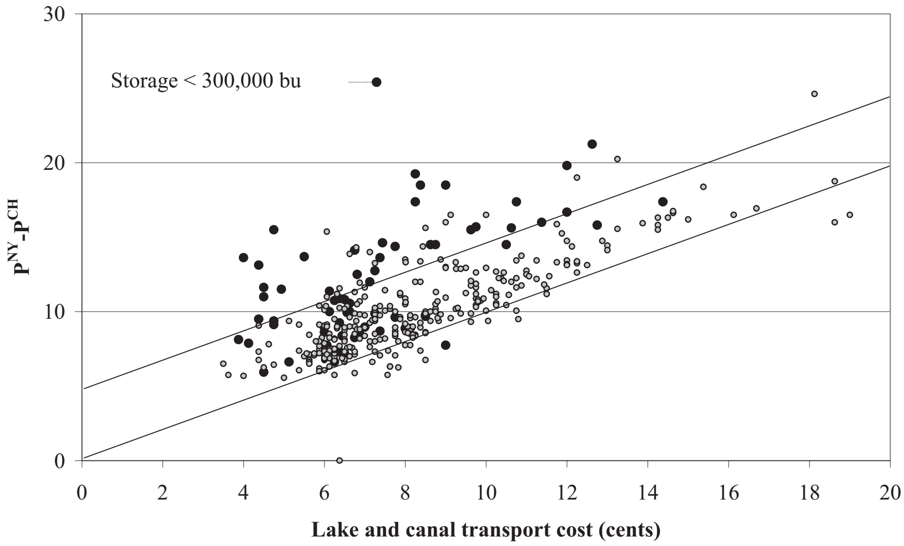
According to equation (6c), when corn was shipped from Chicago at time t the spot price in New York three weeks later was expected to equal the spot price in Chicago at time t plus the cost of transport. Although the expected spot price in New York at time $t + 3$ is not known, the price for future delivery in New York can be used as a proxy. Since the seller had the option of delivering any time during the month, and the trip took three weeks, the price for delivery "this month" was used if the date of the month was before the eleventh, and the price for delivery in the subsequent month was used if the date occurred on or after the eleventh.[22] Consequently, in the regression

$$(F_{t,t+3}^{NY} - P_t^{CH}) = \alpha + \beta \; Transport\ Cost_t + u_t,$$

the coefficient β should equal 1 and the error should be uncorrelated with inventory levels. The data are plotted in figure 7 for each week that lake and canal transport cost data

[22] Ten of the futures prices were not available, and these observations were omitted.

FIGURE 8.—NY SPOT AND CHICAGO SPOT PRICE DIFFERENCE VERSUS TRANSPORT COST, 1878–1891



Note: The diagonal lines indicate $P_t^{NY} - P_t^{CH} = K^T$ and $P_t^{NY} - P_t^{CH} = K^T + 5$.

are available, 1878–1891. Port handling charges in New York are not included in the transport cost, so the coefficient α is likely to be greater than 0.[23] The graph distinguishes observations for which New York inventories were less than 300,000 bushels from those for which inventories exceeded 300,000 bushels. The observation for September 23, 1884, is omitted as there was a corner in the Chicago market and the Chicago price was 16 cents higher than the New York price.

The best fitting regression line including an indicator variable for weeks with low storage is

$$(F_{t,t+3}^{NY} - P_t^{CH}) = \underset{(0.23)}{1.11} + \underset{(0.044)}{0.95}\ Transport\ Cost_t + \underset{(0.22)}{0.22}\ 1(Storage_t < 300{,}000) + u_t$$

$$u_t = 0.41u_{t-1} + e_t \qquad R^2 = 0.74\ N = 358$$

where I ($Storage < 300{,}000$) has a value of 1 if inventories were less than 300,000 and 0 otherwise. The regression was estimated using feasible generalized least squares to take into account first-order autocorrelation in the error process.[24] The slope of the coefficient on the transport cost variable is very close to, and insignificantly different from, 1. In addition, the coefficient on the low-inventory variable is both tiny and insignificantly different from 0, indicating that there was no systematic tendency for the future spot price gap to be high when inventories were low. These estimates therefore provide clear evidence that the law of one price as described by equation (6c) held in these markets.

B. Transport Costs, the Chicago Spot Price, and the New York Spot Price

If the model is accurate, $P_t^{NY} - P_t^{CH}$ should have equaled $K^T - K^S$ when New York inventories were positive, but exceeded $K^T - K^S$ when New York inventories were 0. To examine this implication of the model, the spot price difference $P_t^{NY} - P_t^{CH}$ is plotted against the cost of lake and canal transport in figure 8. Again, observations when New York inventories were less than 300,000 bushels are distinguished from those when inventories exceeded 300,000 bushels.

Two features of the graph stand out. First, as with figure 5, most of the observations lie above the 45-degree line, indicating that the spot price difference normally exceeded the transport cost. Secondly, there were a large number of occasions when the spot price difference exceeded the transport cost by more than 5 cents, and these were much more likely to occur when inventories were less than 300,000 bushels than when they exceeded 300,000 bushels. This is documented in table 5, which analyzes the distribution of the variable $P_t^{NY} - P_t^{CH} - K^T$ for six different inventory levels. On 62% of the 21 occasions when inventories were less than 150,000 bushels, and 31% of the 45 occasions when inventories were between 150,000 and 300,000 bushels, the New York price exceeded the Chicago price plus transport cost by 5 cents or more. In contrast, the New York price exceeded the Chicago price plus transport costs by more than 5 cents on only 4% of the 224 occasions when inventories exceeded 600,000 bushels.

[23] It cost 0.15 cents per bushel for canal boat trimming, 0.625 cents for receiving, and discharging, and 0.25 cents for screening and blowing (New York Produce Exchange Annual Report, 1892). Official grain inspection, if required, cost $3 per canal boat. It cost 0.25 cents per bushel to store grain for ten days.

[24] The errors are possibly autocorrelated because the data are sampled weekly, while shipping took three weeks.

TABLE 5.—DISTRIBUTION OF SPOT PRICE DIFFERENCE ADJUSTED FOR TRANSPORT COSTS ($P_T^{NY} - P_T^{CH} - K^T$) BY STORAGE LEVEL

Storage level bu	< 150,000	< 300,000	< 600,000	< 1,000,000	< 2,000,000	2,000,000 +
N	21	45	79	62	87	75
% obs. < 0	0%	2%	2.5%	3%	10%	15%
% obs. 0 ≤ x < 5	38%	67%	85%	90%	87%	83%
% obs. ≥ 5	62%	31%	12.5%	6.5%	2%	3%
Mean	6.01	3.84	2.70	1.73	1.45	1.10
Std.	2.93	2.57	1.89	2.82	1.61	1.81
WMW test 1		−2.77*	−4.44*	−5.23*	−5.95*	−5.96*
WMW test 2		−2.77*	−2.43*	−2.52*	−1.94	−1.03

The table shows the fraction of observations in each group that are less than 0, between 0 and 5 cents, and more than 5 cents.

The first Wilcoxon-Mann-Whitney test, WMW test 1, tests whether the distribution is the same as the distribution of the group for which storage is less than 150,000 bushels. The second Wilcoxon-Mann-Whitney test, WMW test 2, tests whether the distribution is the same as the distribution of the group immediately to the left.

The WMW test is asymptotically distributed as N(0, 1) and a * indicates significance at the 5% critical level, that is, that the two cumulative distributions lie above each other.

TABLE 6.—DISTRIBUTION OF SPOT-FUTURE PRICE DIFFERENCE ADJUSTED FOR TRANSPORT COSTS ($F_{t,t+3}^{NY} - P_t^{CH} - K^T$) BY STORAGE LEVEL

Storage level bu	< 150,000	< 300,000	< 600,000	< 1,000,000	< 2,000,000	2,000,000 +
N	19	44	77	60	84	74
% obs. < 0	21%	7%	10%	8%	7%	11%
% obs. 0 ≤ x < 5	74%	91%	86%	92%	92%	85%
% obs. ≥ 5	5%	2%	4%	0%	1%	4%
Mean	1.04	1.60	1.60	1.29	1.45	1.55
Std.	2.12	1.33	1.76	3.03	1.12	2.09
WMW test 1		1.00	1.07	0.92	0.73	0.90
WMW test 2		1.00	0.07	−0.31	−0.49	0.56

See table 5. The table shows the fraction of observations in each group that are less than 0, between 0 and 5 cents, and more than 5 cents.

Wilcoxon-Mann-Whitney statistics were calculated to test the hypotheses that the cumulative distribution functions of $P_t^{NY} - P_t^{CH} - K^T$ were the same for each of the six inventory groups, against the alternative that one distribution lay above the other. For each inventory group, the hypothesis that the distribution of $P_t^{NY} - P_t^{CH} - K^T$ was the same as that of the lowest inventory group is rejected. In addition, the hypotheses that the cumulative distribution functions of the first four inventory groups were the same were rejected, although the distributions were similar for all groups with more than a million bushels. Consequently, it is possible to conclude that the large spatial price differences that occurred when inventories in New York were low were not due to simple random variation.

In table 6, the distribution of $F_{t,t+3}^{NY} - P_t^{CH} - K^T$ for the six different inventory categories is calculated. In contrast to the results for the New York spot price, the distributions show that the New York future price did not spike upwards when inventories were low; indeed, it is not possible to reject the hypothesis that the cumulative distributions of $F_{t,t+3}^{NY} - P_t^{CH} - K^T$ were the same for any of the six inventory categories.

The different behavior of the New York spot and future prices can be further demonstrated by estimating the best-fitting regression line corresponding to figure 8. The line, estimated using feasible generalized least squares to take into account first-order autocorrelation, is[25]

$$(P_t^{NY} - P_t^{CH}) = \underset{(0.29)}{1.75} + \underset{(0.06)}{0.85}\, Transport\ Cost_t + \underset{(0.29)}{2.05}\, 1(Storage_t < 300{,}000) + u_t.$$

$$u_t = 0.45u_{t-1} + e_t \qquad R^2 = 0.69\ N = 368$$

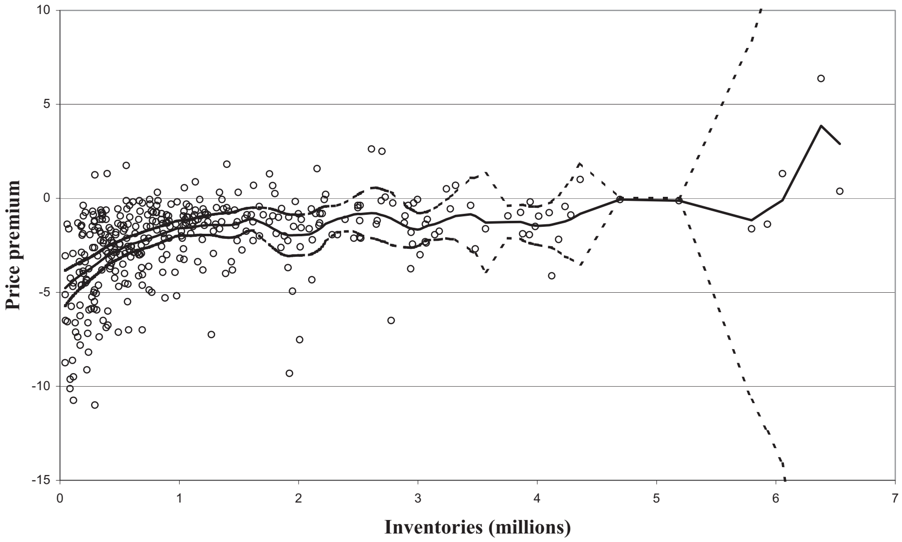
The positive, statistically significant, and economically large coefficient on the low-inventory dummy variable is in sharp contrast to the small and statistically insignificant coefficient estimated when the future price was used. The difference between the two coefficients further confirms that the New York spot price, but not the future price, spiked upwards relative to the Chicago price and the transport cost when inventories were low.

The relationships evident in figure 8 and table 5 can be summarized by constructing a "spatial arbitrage-storage" curve that plots the difference between the New York spot price and the Chicago spot price plus transport costs as a function of inventory levels (see figure 9).[26] The graph resembles a traditional supply of storage curve, a graph of the difference between the spot price and the future price of a commodity versus the quantity of storage. In a supply of storage curve the spot price is typically higher than the

[25] The observation for September 23, 1884, was omitted.

[26] The scatter plot of points is accompanied by a nonparametric kernel regression showing the average relationship between the future premium and the storage quantity. The kernel regression is estimated with an Epanechnikov kernel with bandwith 300,000 bushels.

FIGURE 9.—SPATIAL ARBITRAGE: STORAGE CURVE
($P^{CH} + K^{T} - P^{NY}$) VERSUS INVENTORIES



Note: The solid line is a nonparametric kernel regression (with 95% confidence intervals) using an Epanechnikov kernel with 300,000 bandwith.

future price when storage volumes are low, but when storage quantities are high the future price exceeds the spot price (Working, 1949; Brennan, 1958). The theoretical model suggests that the spatial arbitrage-storage curve and the supply of storage curve should be closely related because when a center has zero storage the spot price should exceed both the local future price and the spot price in the other center plus the transport cost.

VI. Conclusions

This paper has attempted to enhance economists' understanding of how spatial arbitrage occurs by examining how the interaction of storage and trade affects prices in different locations. Its theoretical contribution has been to link the economics literature analyzing spatial price arbitrage with the logistics management literature analyzing how the speed of transport determines inventory holdings. It has made the link by relaxing the standard assumption in models of spatial arbitrage that transport is instantaneous; in doing so, it has emphasized the manner in which inventories are used to smooth price fluctuations. This role had long been recognized in the logistics management literature, but only at the level of the firm, rather than across competitive markets. Its empirical contribution has been to assemble a set of price, transport cost, and storage data that is detailed enough to detect how logistics issues have affected commodity prices in one specific market. It has shown that the spatial price difference frequently exceeded the cost of shipping goods, in a heavily traded commodity market in which there were large investments in logistics infrastructure and thick financial markets. In doing so, it has established that prices exceeded the transport cost when there were low supplies in the importing city, causing a temporary price spike followed by a price decrease when new supplies arrived from the exporting center. The result can perhaps be best summarized by plotting a spatial arbitrage-storage curve that shows how the spatial price difference depends on inventory levels in the importing center.

The theoretical model suggests that the process of physical arbitrage is significantly more complex than has previously been recognized. The central feature of the model is that the quantity of goods arbitrageurs ship each period will not be the amount necessary to ensure that the spatial price difference is always equal to the cost of transport. Rather, in order to make normal profits on average, arbitrageurs will ship an amount such that there will be insufficient supplies and high prices in the destination center on a regular but infrequent basis. The profits they make on these occasions offset the small losses they make on the more frequent occasions that the exports arrive when local supplies are adequate. Put more starkly, traders make normal profits on average only by obtaining high prices and extraordinary profits on occasions that supplies in the importing market are low. If this phenomenon is generally true in practice, it provides an argument against governments attempting to stabilize prices in times of shortage.

The paper has several other implications for time series analyses of regional price dynamics. First, the data requirements to analyze episodes of physical arbitrage are large. In addition to the price data usually used in studies of arbitrage, it is necessary to know whether trade actually occurs between a pair of cities, the extent to which transport prices vary, and the level of inventories in each market. More generally, if cheap transport is slow, studies of market integration need to focus more on logistics issues—the way arbitrageurs use transport systems and storage to integrate markets—rather than just transport issues.

Secondly, models of market integration that assume that the spatial price difference is bounded by the transport cost need modification. For example, Spiller and Wood (1988) developed a popular methodology to estimate spatial transactions costs from price data that is explicitly predicated on the assumption that the price difference is always equal to the cost of transport, but that the cost of transport is variable. If that methodology were used on this data set, it would overestimate the mean and variability of transport costs because it wrongly assumes that the large price differences that occurred when one center had low storage were due to temporarily high transport costs.

Thirdly, it is not possible to use price data to make strong inferences about transactions costs without a lot of additional information about production patterns and trade flows. The model suggests that a large number of factors including transport costs, transport speed, the cost of storage, and the extent to which output is regionally specialized determine the overall distribution of prices in two locations. The simulations suggest that a small change in regional specialization can have a similar effect on the distribution of prices as a large change in transport costs. If this is the case, studies linking moments of the price distribution in different cities to the distance between cities need careful interpretation, as the result may indicate the pattern of regional specialization across space rather than the pattern of transactions costs.

REFERENCES

Bailey, Roy E., and Marcus J. Chambers, "A Theory of Commodity Price Fluctuations," *Journal of Political Economy* 104:5 (1996), 924–957.

Baumol, W. J., and H. D. Vinod, "An Inventory Theoretic Model of Freight Transport Demand," *Management Science* 16:7 (1970), 413–421.

Brennan, Michael J., "The Supply of Storage," *American Economic Review* 48:1 (1958), 50–72.

Chicago Board of Trade, *Annual Report of the Trade and Commerce of Chicago* (Chicago: JMW Jones Stationery and Printing Co, 1876–1892).

Coleman, Andrew M. G., "Storage, Arbitrage, and the Law of One Price," Princeton University, thesis (1998).

Commercial and Financial Chronicle (New York: William B. Dana Co., weekly: 1882–1887, vols. 35–43).

Cournot, Augustin, *Researches into the Mathematical Principles of the Theory of Wealth* [translated by Nathaniel T. Bacon, 1897] (New York: Augustus M. Kelley, reprints of economic classics, 1838, 1960).

Deaton, Angus, and Guy Laroque, "On the Behavior of Commodity Prices," *Review of Economic Studies* 59 (1992), 1–23.

——— "Estimating a Nonlinear Rational Expectations Commodity Price Model with Unobservable State Variables," *Journal of Applied Econometrics* 10 (1995), S9–S40.

——— "Competitive Storage and Commodity Price Dynamics," *Journal of Political Economy* 109:5 (1996), 896–923.

de Jong, Gerard, "Value of Freight Travel-Time Savings" (pp. 553–564), in D. A. Hensher and Kenneth J. Button (Eds.), *Handbook of Transport Modeling* (Oxford: Pergamon, 2000).

Dumas, Bernard, "Pricing Physical Assets Internationally," NBER working paper no. 2569 (1988).

Engel, Charles, and John H. Rogers, "How Wide Is the Border?" *American Economic Review* 86:5 (1996), 1112–1125.

Goldstein, Benjamin, *Marketing: A Farmer's Problem* (New York: MacMillian, 1928).

Hansen, Bruce E., "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis," *Econometrica* 64:2 (1996), 413–430.

Harley, Knick, "The North Atlantic Wheat Trade and Its Institutional Consequences," Oxford University mimeograph (2004).

Lee, Guy A., "The Historical Significance of the Chicago Grain Elevator System," *Agricultural History* 11:1 (1937), 16–32.

McGinnis, Michael A., "A Comparative Evaluation of Freight Transportation Models," *Transportation Journal* 29:2 (1989), 36–46.

New York Produce Exchange, *Annual Report* (New York: Jones Printing Company, 1876–1892).

New York Times (1870–1892).

Porter, Robert H., "A Study of Cartel Stability: The Joint Executive Committee, 1880–1886, *Bell Journal of Economics* 14:2 (1983), 301–314.

——— "On the Incidence and Duration of Price Wars," *Journal of Industrial Economics* 33:4 (1985), 415–426.

Samuelson, Paul A., "Spatial Price Equilibrium and Linear Programming," *American Economic Review* (1952), 283–303.

Spiller, Pablo T., and Robert O. Wood, "The Estimation of Transactions Costs in Arbitrage Models," *Journal of Econometrics* 39 (1988), 309–326.

Tunell, George, "The Diversion of the Flour and Grain Traffic from the Great Lakes to the Railroads," *Journal of Political Economy* 5 (1897), 339–375.

Tyworth, John, "The Inventory Theoretic Approach in Transportation Selection Models: A Critical Review," *Logistics and Transportation Review* 27:4 (1991), 299–318.

Ulen, Thomas S., "The Regulation of Grain Warehousing and Its Economic Effects: The Competitive Position of Chicago in the 1870s and 1880s," *Agricultural History* 56:1 (1982), 194–210.

United States Bureau of Statistics, *First Annual Report on the Internal Commerce of the United States, by Joseph Nimmo Jr., Part 2 of the Annual Report of the Chief of the Bureau of Statistics on the Commerce and Navigation of the United States for the Fiscal Year Ending June 30, 1876* (Washington, DC: U.S. Government Printing Office, 1877).

United States Bureau of Statistics, Treasury Department, *Report on the Internal Commerce of the United States, December 1879, by Joseph Nimmo* (Washington, DC: U.S. Government Printing Office, 1879).

United States Bureau of Statistics, Treasury Department, *Report on the Internal Commerce of the United States, 1881, by Joseph Nimmo* (Washington, DC: U.S. Government Printing Office, 1881).

United States Bureau of Statistics, Treasury Department, *Report on the Internal Commerce of the United States, 1881–1882, by Joseph Nimmo* (Washington, DC: U.S. Government Printing Office, 1884).

United States 52nd Congress, 2nd Session, *Senate Report 1394: Wholesale Prices, Wages, and Transportation, Report by Mr. Aldrich from the Committee on Finance, March 3, 1893, Part 1* (Washington; DC: U.S. Government Printing Office, 1893).

United States Treasury, *Letter from the Secretary of the Treasury Transmitting a Report Made to the Bureau of Statistics by Mr. George G. Tunell of Chicago on Lake Commerce* (Washington, DC: U.S. Government Printing Office, 1898).

Williams, Jeffrey, *The Economic Function of Futures Markets* (Cambridge: Cambridge University Press, 1986).

Williams, Jeffrey C., and Brian D. Wright, *Storage and Commodity Markets* (Cambridge: Cambridge University Press, 1991).

Williams, John Burr, "Speculation and the Carryover," *Quarterly Journal of Economics* 50:3 (1936), 436–455.

Working, Holbrook, "The Post Harvest Depression of Wheat Prices," *Wheat Studies of the Food Research Institute* 6:1 (1929), 1–40.

——— "The Theory of Price of Storage," *American Economic Review* 39:6 (1949), 1254–1262.

APPENDIX A

Solving the Rational Expectation Model

This appendix contains more details about the solution technique used to solve the model in section II. A solution is found by constructing an algorithm that calculates a series of successive approximations to the optimal storage and trade functions, $S_k^i(y_t)$ and $T_k^i(y_t)$, where k refers to the kth approximation. The iteration process is as follows. First, given the kth value of these functions, the price functions for each center are calculated at all grid points, using the inverse demand function $P_k^i(y_t) = D_i^{-1}(D_k^i(y_t))$, where

$$D_k^i(y_t) = X_t^i + M_t^i - S_k^i(y_t) - T_k^i(y_t).$$

Secondly, given the kth storage and trade rules $S_k^i(y_t)$ and $T_k^i(y_t)$, a schedule of the expected future price conditional on being at M_{t+1}^A and M_{t+1}^B in the subsequent period is calculated. At each point $(M_{t+1}^A, M_{t+1}^B, X_t^A, X_t^B)$ the expected future price is calculated by multiplying the price functions by the transition matrix probabilities π:

$$\begin{aligned} &E_X[P_{k,t+1}^i | M_{t+1}^A, M_{t+1}^B, X_t^A, X_t^B] \\ &= \sum_{X_{i,t+1}^A, X_{j,t+1}^B} D_i^{-1}(D_k^i(M_{t+1}^A, M_{t+1}^B, X_{i,t+1}^A, X_{j,t+1}^B))\pi \qquad \text{(A1)} \\ &\quad \times (X_{i,t+1}^A, X_{j,t+1}^B | X_t^A, X_t^B). \end{aligned}$$

Thirdly, the values of the four complementary conditions associated with each of the equations (6a–6d) are calculated at each point y_t. The value of the expected future price $E[P_{k,t+1}^i|y_t]$ is equal to $E_X[P_{k,t+1}^i|((1-\delta)(S_k^A(y_t) + T_k^B(y_t)),\ (1-\delta)(S_k^B(y_t) + T_k^A(y_t)), X_t^A, X_t^B)]$ and is calculated using linear interpolation of equation (15). Linear interpolation is used because it ensures that if a linear inequality restriction holds at contiguous grid points, it will also hold in between the grid points. If the value of the complementary condition is inconsistent with the value of the associated control variable—for example, if in inequality (6a), $S_k^A > 0$ but $\frac{1-\delta}{1+r} E[P_{t+1}^A \mid y_t] - P_t^A(y_t) - K^S \neq 0$, or $S_k^A = 0$ but $\frac{1-\delta}{1+r} E[P_{t+1}^A|y_t] - P_t^A(y_t) - K^S > 0$—then the value of the control variable is recalculated at the grid point; otherwise, it remains unchanged. The new $k+1$th values of the control variables associated with the inconsistent inequalities are simultaneously calculated at the grid point using an optimizing routine such as the Newton Rhapson or the secant method. The process is repeated at each grid point until S_{k+1}^i and T_{k+1}^i are calculated over the whole domain.

The whole algorithm is repeated until the difference between successive values of the control values is small, typically a price difference of less than 0.01 compared with an average price of 100. It proved necessary to sacrifice accuracy at some points of the domain of y_t simply to construct a four-dimensional grid small enough to be modeled by a computer; in the end, the grid structure consisted of some 20,000 to 40,000 points. The algorithm was solved using Gauss software. Typically convergence took three hundred iterations to achieve acceptable accuracy.

The starting value for the algorithm was calculated by splitting the combined center solution into two, that is, by estimating storage functions S^A and S^B such that $S^A(M^A, M^B, X^A, X^A) + S^B(M^A, M^B, X^A, X^A) = S(M^A, M^B, X^A, X^A)$, where $S(\cdot)$ is the solution for storage in the combined center problem. The initial values for the trade variables were 0.

The Invariant Distribution

The invariant distribution of the model is the unconditional probability of being at a particular grid point. Deaton and Laroque (1995) suggest a method for finding the invariant distribution as follows. Suppose Y is a $n \times 1$ vector of all possible grid points corresponding to the four state variables. Let the function ω: $(Y \times Y) \rightarrow \mathbf{R}$ be the conditional transition probabilities of moving from one point one period to another in the next period:

$$\omega(y_1, y_2) = \text{Prob}(Y_{t+1} \in B(y_2)|Y_t = y_1),$$

where $B(y_2)$ is a region around y_2 defined so that the regions form a partition of the space and only include one grid point. Let Ω be the $n \times n$ matrix of these probabilities, with $\Omega_{ij} = \omega(y_i, y_j)$. Ω has at least one unit eigenvalue, since all the rows of Ω sum to 1 and all elements of Ω are strictly less than 1 by construction. The eigenvector corresponding to this eigenvalue will be the invariant probability distribution of Y.

Deaton and Laroque suggest that the eigenvalue and its corresponding eigenvector can be found by inverting the $n \times n$ matrix Ω. This is not practical in this case as n typically exceeds 20,000. The alternative method for finding the invariant distribution is simply to multiply an arbitrary initial distribution on Y by the transition matrix Ω' until some successive values of the product meet some convergence criteria. (In practice, the loop was stopped at the $j+5$th iteration if the average difference between the n values of the vectors Y_{j+5} and Y_j was less than $10^{-5}/n$.) This method proves to be fast, for even though Ω has several hundred million elements, the vast majority of these are zeros and it is straightforward to devise an algorithm that only uses the several hundred thousand positive elements in the iterative procedure.

APPENDIX B

Data Sources

Six kinds of data have been assembled for this project: the spot price of corn in New York and Chicago; the future price of corn in New York and Chicago; transport costs between Chicago and New York; transport volumes between Chicago and New York; storage prices in Chicago and New York; and storage volumes in Chicago and New York.

1. Spot Price of Corn

Prices were collected for number 2 yellow corn. Number 2 corn was the primary future grade and comprised a large fraction of the spot market. Grades were defined as follows.

New York: "YELLOW CORN shall be sound, dry, plump and well cleaned; an occasional white or red grain shall not deprive it of this grade. No. 1 CORN shall be mixed corn of choice quality, sound, dry and reasonably clean. No. 2 CORN shall be mixed corn, sound, dry and reasonably clean" (New York Produce Exchange, 1882, p. 207).

Chicago: "No. 1 YELLOW CORN shall be yellow, sound, dry, plump and well cleaned. No. 2 CORN shall be dry, reasonably clean, but not plump enough for No. 1" (Chicago Board of Trade, 1882, pp. 79–80).

Spot prices for both cities were collected in the Thursday edition of the *New York Times,* 1878–1891. The prices were for the preceding Wednesday. If the Wednesday were a public holiday, the Tuesday price was collected. If the markets were closed on both Wednesday and Tuesday, the data were skipped for that week.

Daily spot prices for New York are also available in some years in the annual report of the New York Produce Exchange. However, since the *New York Times* had to be used to collect the Chicago spot price and the New York future price, as well as the New York spot price in years where it was not reported in the annual report, the weekly *New York Times* data were used.

2. Future Price of Corn

Prices were collected for number 2 yellow corn. The Chicago future price was collected from the annual report of the Chicago Board of Trade. The quotes are for seller delivery: the seller could choose any day to deliver within the said month. Wednesday quotes were collected.

The New York Wednesday future prices were collected from the Thursday edition of the *New York Times.* The seller also had the option as to the delivery date.

3. Corn Trade and Storage Data

Storage and trade data for Chicago was sourced from the Chicago Board of Trade annual reports. The New York data came from a variety of sources. Where possible, it came from the New York Produce Exchange annual reports, but these documents had little data between 1882 and 1887. Storage data for these years came from the weekly *Commercial and Financial Chronicle.* Export data for several years came from the Chicago Board of Trade annual reports. Storage cost data come from the Chicago Board of Trade and New York Produce Exchange annual reports, and from Goldstein (1928).

4. Transport Data

The transport cost data were published by the Chicago Board of Trade and New York Produce Exchange annual reports. They are similar but not identical to the data published in the Aldrich Report (United States 52nd Congress, 2nd Session, 1893).

APPENDIX C

Chicago-Buffalo–New York Trade Flows

It can be demonstrated that grain sent by rail from Chicago to Buffalo mainly went to New York by comparing the monthly shipments of corn along the railroads connecting Chicago and Buffalo with the monthly shipments along the railroads connecting Buffalo and New York during the years that such data exist, 1877 to 1881.[27] New York receipts along these lines were 50% higher on average than rail shipments from Chicago to Buffalo, because the Hudson and Erie lines were used to rail some of the corn shipped to Buffalo on the Great Lakes. Nonetheless, there was an almost one-for-one correspondence between the variation in rail flows from Chicago to Buffalo and flows from Buffalo to New York. Formally, a regression of New York rail receipts from Buffalo with Chicago rail shipments to Buffalo has a slope estimate of 1.20, with a 95% confidence interval of 0.93 to 1.46.[28]

$$(NY\ rail\ receipts)_t = \underset{(130{,}000)}{-125{,}000} + \underset{(0.131)}{1.20}\ (Chicago\ rail\ shipments)_t + \underset{(0.020)}{0.087}\ (lake\ shipments)_t + e_t.$$

(130,000) $R^2 = 0.67$; 60 observations.

Similar data can be used to show that most of this corn was part of a through-shipment. The data suggest a near one-for-one correspondence between the variation in rail flows from Chicago to Buffalo and the through-shipments through Chicago. A regression of Chicago rail shipments to Buffalo against Chicago through-shipments has a slope estimate of 0.85, with a 95% confidence interval of 0.72 to 0.99.[29]

$$(Chicago\ rail\ shipments)_t = \underset{(86{,}000)}{205{,}000} + \underset{(0.067)}{0.85}\ (Chicago\ through\text{-}shipments)_t + e_t.$$

$R^2 = 0.65$; 60 observations.

These two regressions suggest the corn sent by rail to Buffalo was forwarded to New York and that the corn originating in Chicago was part of a through-shipment. A third regression linking New York rail receipts to Chicago through-shipments confirms this link:[30]

$$(NY\ rail\ receipts)_t = \underset{(180{,}000)}{215{,}000} + \underset{(0.131)}{1.127}\ (through\text{-}shipments)_t + \underset{(0.024)}{0.038}\ (lake\ shipments)_t + e_t.$$

$R^2 = 0.57$; 60 observations.

[27] Data for these years were published in the annual reports of the Chicago Board of Trade and the New York Produce Exchange, but I have been unable to find it for other years. The Chicago data are the monthly shipments along the Michigan Central and the Lake Shore and Michigan Railroads. The New York data are the monthly shipments along the Erie and the New York Central and Hudson railroads. For information on through-shipments, see footnote 15.

[28] OLS regression with standard errors was calculated using the Newey-West method using five lags. The regression was also estimated using a feasible generalized least squares estimate that corrected for first-order serial correlation and assumed the variance of errors was a linear function of Chicago shipments. In this case the slope estimate was 1.16 with standard error of 0.13.

[29] OLS regression with standard errors was calculated using the Newey-West method using five lags. The feasible generalized least squares estimate (see footnote 28) was 0.95 with standard error of 0.10.

[30] OLS regression with standard errors was calculated using the Newey-West method using five lags. The feasible generalized least squares estimate (see footnote 28) was 1.10 with a standard error of 0.15.