Spatial Price Equilibrium with Convex Marginal Costs of Transportation: Applications to the Brent-WTI Spread

Max S. Bennett*

March 20, 2012

Abstract

Standard models of spatial price equilibrium assume that transportation costs are constant and therefore conclude that a spatial price spread must be bound by the constant marginal cost of transportation. However, the empirics of many spatial price spreads, such as the Brent-WTI spread, demonstrate clear violations of this condition. This paper explains this phenomenon by instead assuming that the marginal cost of transportation is convex. Furthermore, this paper rationalizes the sensitivity of spatial price spreads to localized demand and supply schedules as well as to expected future costs of transportation, an endogeneity that standard models are unable to explain.

—I would like to thank Bruce Petersen for his incredible support, insight, and inspiration throughout the entire process of writing this paper; he has been a wonderful mentor. I would also like to thank Sebastian Galiani for his valuable feedback and advice. And I want to express the utmost gratitude to Dorothy Petersen for providing me with the opportunity to engage my fascination with economics by writing a thesis, and for helping me navigate my undergraduate academic career.
1 Introduction

Understanding the determinants of commodity prices is central to the study of economics; it allows us to examine and develop policy, construct micro-foundations for macroeconomic theory, and more accurately predict future prices. However, our understanding is incomplete without incorporating the interconnectedness of spatially separated markets. The possibility of transportation makes the price of a commodity in one region endogenous to the demand and supply schedules in other regions. As such, without understanding the relevance of transportation, economics would be wholly unable to study and address many problems in agriculture, energy, and financial markets. In fact, the first explicit description of price determination as the intersection of demand and supply curves was given by A. A. Cournot (1838) partially in attempt to describe price relations between spatially separated markets. The integration of markets is so important because of the value it provides: it allows commodities to be allocated from the cheapest suppliers to the consumers with the highest demand. If spatially separated markets become decoupled from one another, there is value to be earned by reconnecting them. As such, from a policy perspective, understanding the causes of changing spatial price spreads is important because it allows us to more accurately prescribe effective policy to address the issue and reintegrate spatially separated markets.

Models of spatial price equilibrium have been developed in order to incorporate the interconnectedness of spatially separated markets into commodity price determination. A standard assumption in the literature is that the marginal cost of transportation is constant (independent of volume). As such, standard models, as exemplified by Samuelson (1952) and Takayama and Judge (1964), conclude that the price differential between any two regions\(^1\) is bound by the constant marginal cost of transportation. These models thereby conclude that transportation will occur if and only if the spatial price spread is exactly equal to the transportation cost. The key empirical implication of such models is that once transportation occurs, the spatial price spread cannot increase further. However, we somewhat regularly observe dramatic violations of this condition in many commodity markets, such as within the US crude oil market during 2011. To address this empirical problem theorists have attempted to relax assumptions in the standard models. One such example is a paper by Coleman (2009), which shows that if transportation is non-instantaneous, than spatial price spreads can increase without bound and not violate no-arbitrage.

This paper will show that if we instead relax the assumption that the marginal cost schedule of transportation is constant, then even with instantaneous transport spatial price spreads can increase without bound and not violate no-arbitrage. I will go on to show that under such assumptions, spatial price spreads become endogenous to localized demand and supply schedules, an endogeneity that I will provide empirical support for. It is important to note that standard models of which assume that the marginal cost schedule of transportation is constant cannot explain this relationship. I will go on to include the possibility of storage which will illuminate the endogeneity of the current spread to expected future costs of transportation, a relationship that I have not found described in the literature.

I will use the model developed in this paper to examine the spatial price spread between crude oil at Cushing, Oklahoma (known as “West Texas Intermediate” or “WTI”) and that at St. James, Louisiana (known as “Light Louisiana Sweet” or “LLS”). As seen in Figure 1, the two locations are approximately 650 miles apart. This spread (henceforth referred to as the “LLS-WTI spread”) increased dramatically in 2011 (see figure 2b). This received substantial media attention; crude

\(^1\)Otherwise known as the “spatial price spread”.
oil prices and consequently gasoline and other refined products became substantially cheaper in the Midwest than in the rest of the United States and the rest of the world.

In the media this phenomena has generally been referred to as the widening of the “Brent-WTI spread” instead of as the the widening of the LLS-WTI spread. This is largely because the most widely used benchmarks for world crude oil prices are Brent, which is priced in the United Kingdom, and WTI. The widening of the LLS-WTI spread (decoupling of the Midwest from the rest of the US) will obviously also widen the Brent-WTI spread (since Brent moved largely in tandem with LLS). It is more interesting to isolate the closest spatial markets between which spatial price spreads widened; this allows us to isolate a more specific transportation market to analyze.

Interestingly, the hypotheses regarding the causes of the Brent-WTI spread put forth by the media implicitly assume that the marginal cost schedule of transportation is not constant. But the economic literature has not thoroughly investigated this possibility and I have not found any mention of the 2011 Brent-WTI spread in the economic literature. News centers such as the Wall Street Journal and Bloomberg have written numerous articles regarding the causes of the 2011 widening of the Brent-WTI spread. The articles, often quoting analysis performed by financial institutions and energy consultants, largely describe two hypotheses regarding the causes of the spread:

(i) The widening of the Brent-WTI spread was caused by a negative production shock in the Middle East. In particular, Arab Spring and loss of Libyan Oil put upward pressure on Brent prices while transportation constraints between the Cushing and Europe isolated WTI from this effect.

(ii) The widening of the Brent-WTI spread was caused by a positive production shock in the Midwest of the United States, specifically Cushing, Oklahoma. In early 2011, new pipelines were built bringing more Canadian oil into Cushing, and transportation constraints between the Cushing and Europe isolated Brent from this effect.

I will develop a model that will provide a direct way to test the efficacy of each of these hypotheses as well as other potential causes of the widening spread. Additionally, I will attempt to use the model developed in this paper to explain the observed relationship between inventories, transportation, and the Brent-WTI spread (see Figure 3)\(^2\).

\(^2\)It is important to note that the econometric conclusion of this paper is consistent with that of most financial
1.1 Background of Oil Market at Cushing, Oklahoma

The most easily refined crude oil, and thus the most valuable, is light sweet crude. WTI, Brent, and LLS are all light sweet crude oils, and are almost identical in physical composition. As such, any substantial deviation in price between these crude oils can only be a consequence of spatial price equilibrium and not differences in intrinsic value.

Cushing, Oklahoma is the delivery point for the NYMEX oil futures contract and therefore refineries, storage facilities, and pipelines have all developed substantial infrastructure in the periphery of the city. As of 2011, it is estimated that storage capacity at Cushing is around 48 million barrels of crude oil, and as much as 600 thousand barrels flow into Cushing daily.

The transportation infrastructure between Cushing and St. James primarily consists of pipelines. Most pipelines transport liquid at a speed of approximately 3 to 13 miles-per-hour; therefore it should only take between two and nine days to transport crude oil from Cushing to St. James. There are, however, additional modes of transportation between Cushing and St. James: notably, rail and trucking. Transportation from Cushing to St. James is estimated to cost between $7 to $10 per barrel by rail and between $11 and $15 per barrel by trucking, which is much more expensive than the estimated $2 per barrel by pipeline.\(^3\) As noted above, a con-

---

\(^3\)CommodityOnline, quoting analysis by Bank of American Merrill Lynch.
constant marginal cost schedule of transportation has been the standard assumption in the literature, which is clearly at odds with transportation costs of oil from Cushing to the Gulf Coast if pipeline capacity is insufficient and therefore some oil must be shipped by other, more expensive, modes of transportation.

In late 2010 and early 2011, two large new pipelines that directed oil from Canada to Cushing went online, substantially increasing the availability of crude in Cushing. This occurred over the backdrop of steadily increasing production of crude oil in the midwest over the past decade. These two forces dramatically increased the quantity of crude flowing into Cushing. Models of spatial price equilibrium with a constant marginal cost of transportation would suggest that this would have no effect on the LLS-WTI spread, as the additional oil would be immediately directed towards the coast via transportation infrastructure. However, I will argue in this paper that because of rising marginal costs, a positive supply shock at Cushing is a theoretically sound explanation for the widening of the LLS-WTI spread.

2 Building A Model of Spatial Price Equilibrium

2.1 Constant Marginal Costs of Transportation Without Storage

Consider a non-perishable and non-depreciating commodity, called "x", that trades in two locations, point A and point B. Production of x at both A and B during a given period t, denoted $Q_t^A$ and $Q_t^B$ respectively, is assumed to be given as exogenous to arbitrageurs. This is consistent with a commodity whose production occurs in earlier periods, such as oil, or commodities that have perfectly inelastic supply curves. The corresponding price of x at each point is denoted $p_t^A, p_t^B$.

In this model, arbitrageurs are risk-neutral and have the opportunity to transport commodity x from point A to point B at marginal cost $k_{t,AB}^T$, or from point B to point A at marginal cost $k_{t,BA}^T$.

Before building a model with convex marginal costs of transportation with storage, let us first build intuition by describing a model where the marginal cost of transportation between A and B is constant and where there is no option to store commodity x from period t to period $t+1$. These are the assumptions in the standard spatial price equilibrium models described by Samuelson (1952) and Takayama and Judge (1962b).
2.1 Constant Marginal Costs of Transportation Without Storage

It is clear that in order to maintain no-arbitrage, the spatial price spread, given by \( \sigma_t = p_t^B - p_t^A \) must be bound by the marginal costs of transportation. Explicitly the no-arbitrage condition has

\[
-\bar{\kappa}_{t,BA} \leq \sigma_t \leq \bar{\kappa}_{t,AB}
\]

Within these bounds, transportation will only occur from A to B if \( \sigma_t = k_{t,AB}^T \) and transportation will only occur from B to A if \( \sigma_t = -k_{t,BA}^T \). If the spread lies within these bounds then transportation between A and B will yield negative profits, and therefore transportation in either direction will not occur.

No-arbitrage condition (1) is well understood in the literature and is the baseline description of the relationship between the spatial price spread and the marginal cost of transportation. However, it says nothing about the quantity of transportation that will occur. In order to understand this we must understand the direct effect transportation has on the spatial price spread.

We begin by describing the determinants of the absolute price of \( x \) at each point. The price of \( x \) at point \( A \) is given by economic agents' demand for consumption of \( x \) at point \( A \). The demand schedule for \( x \) at \( A \) will be denoted by \( D^A(N_t^A) = p_t^A \), where \( N_t^A \) is the quantity of \( x \) available at \( A \) for consumption during period \( t \). We will maintain the standard assumption for normal goods that the demand schedule is downward sloping. Although the production of \( x \) at \( A \) during period \( t \) is exogenous, the quantity available for consumption is not; arbitrageurs have the option to pull \( x \) out of the market at \( A \) to transport it to \( B \). This gives a relationship described by \( N_t^A = \bar{Q}_t^A - T_t \), where \( T_t \) is the quantity of \( x \) transported from point \( A \) to point \( B \). Arbitrageurs at point \( A \) will optimally transport commodity \( x \) while reacting to their activities' effect on the price of \( x \) at \( A \).

It follows that the price of \( x \) at \( A \) is endogenously defined by the following function:

\[
p_t^A(Q_t^A, T_t) = D^A(\bar{Q}_t^A - T_t)
\]

(2)

The market for \( x \) at \( B \) is different from the market at \( A \) only in that transportation increases the quantity of \( x \) available at \( B \), whereas it decreases the quantity available at \( A \). Specifically, the quantity of \( x \) available at \( B \), denoted \( N_t^B \), is given by the relation \( N_t^B = \bar{Q}_t^B + T_t \). Furthermore, just as with the market at \( A \), we assume the demand function is downward sloping. It then follows that the price of \( x \) at \( B \) is given by

\[
p_t^B(Q_t^B, T_t) = D^B(\bar{Q}_t^B + T_t)
\]

(3)

Therefore the spatial price spread is given by what we will refer to as the "spread function" or "spread curve," denoted \( \sigma_t(T_t, \bar{Q}_t^A, \bar{Q}_t^B) \):

\[
\sigma_t(T_t, \bar{Q}_t^A, \bar{Q}_t^B) = D^B(\bar{Q}_t^B + T_t) - D^A(\bar{Q}_t^A - T_t)
\]

(4)

The spread curve is simply the difference in the equilibrium price of \( x \) in each region, given supply levels and transportation between the regions. This representation of the spread curve illuminates that increases in transportation directly decreases the spatial price spread.

We can now explicitly describe the equilibrium level of transportation between \( A \) and \( B \). If the spatial price spread without transportation would exceed the constant marginal cost of transportation, then arbitrageurs will increase transportation until the spread decreases to \( \sigma_t = \bar{\kappa}_{t,AB}^T \). Additionally, if the spatial price spread without transportation is within the band of marginal costs of transportation between points \( A \) and \( B \), then there will be no transportation because it would provide arbitrageurs with negative profits. Therefore equilibrium is given by the
2.2 Convex Marginal Costs of Transportation Without Storage

following conditions:

\[
\sigma_t(T_t, \ldots)|_{T_t=0} \geq \bar{k}_t,AB \implies T_t^* \in \mathbb{R}^+ \implies D^B(\bar{Q}_t^B + T_t^*) - D^A(\bar{Q}_t^A - T_t^*) = \bar{k}_t,AB \quad (5)
\]
\[
-k_t,BA < \sigma_t(T_t, \ldots)|_{T_t=0} < \bar{k}_t,AB \implies T_t^* = 0 \quad (6)
\]

Clearly, the level of transportation is endogenous to the production levels at both A and B, and furthermore, if the marginal cost of transportation is constant, the equilibrium spread will never exceed \( \bar{k}_t,AB \).

Figure 4: Equilibrium With Constant Marginal Costs of Transportation

![Figure 4](image)

(a) Graph of condition (5). Equilibrium quantity of transportation when \( \sigma_t(T_t, \ldots)|_{T_t=0} > \bar{k}_t,AB \)

(b) Graph of condition (6). Equilibrium quantity of transportation when \( -\bar{k}_t,BA < \sigma_t(T_t, \ldots)|_{T_t=0} < \bar{k}_t,AB \)

Figure 3 graphs the equilibrium as described by conditions (5) and (6). Note that supply shocks in either the Middle East or in the Midwest fail to sufficiently explain the widening of the Brent-WTI spread in the context of the model described here. Given that \( \sigma_t(T_t, \ldots)|_{T_t=0} \geq \bar{k}_t,AB \), changes in production levels can only shift the spread curve out, after which the quantity of transportation will simply adjust to ensure that the spatial price spread stays at \( \bar{k}_t,AB \). Therefore, in this model the only way the spatial price spread can increase beyond the initial \( \bar{k}_t,AB \) is if there is an exogenous increase in the marginal cost of transportation.

2.2 Convex Marginal Costs of Transportation Without Storage

Introducing convex marginal costs of transportation changes the economic story and makes it possible for supply shocks at either point A or point B to influence the equilibrium spatial price spread beyond the initial marginal cost of transportation. I now assume that the marginal cost of transportation is endogenous and convex to the level of transportation with

\[
k_t,AB = k_t,AB(T_t) : \quad \partial k_t,AB(T_t)/\partial T_t > 0, \quad \partial^2 k_t,AB(T_t)/\partial T_t^2 > 0
\]

Admittedly, in the real world it is likely that the marginal cost schedule of transportation is instead piecewise. There would likely be an initial flat region until the cheapest mode of transportation has reached capacity, and then a jump and another flat region at the second cheapest mode of transportation and so on until capacity is reached for all modes of transportation, at which point
2.3 Final Model: Convex Marginal Costs of Transportation With Storage

the marginal cost schedule should be perfectly inelastic. However, in order to approximate the effect of multiple modes of transportation I will simply assume that the schedule is convex. This simplifying approximation makes the mathematics easier to deal with, and does not affect the key results of the paper.

The no-arbitrage condition (1), briefly restated, becomes

\[ -k_{t,BA}(T_{t,BA}) \leq \sigma_t(T_t, \tilde{Q}_t^A, \tilde{Q}_t^B) \leq k_{t,AB}(T_{t,AB}), \]

where we will just refer to \( T_{t,AB} \), the quantity of transportation from \( A \) to \( B \), as \( T_t \). The equilibrium conditions (5) and (6) are no longer based on constant marginal costs of transportation, and instead are described by the following conditional statements

\[
\sigma_t(T_t, \ldots)|_{T_t=0} \geq k_{t,AB}^T(0) \implies T_t \in \mathbb{R}^+ \text{ s.t. } D^B(\tilde{Q}_t^B + T_t) - D^A(\tilde{Q}_t^A - T_t) = k_{t,AB}(T_t) \quad (7)
\]

\[
-k_{t,BA}(0) < \sigma_t(T_t, \ldots)|_{T_t=0} < k_{t,AB}^T(0) \implies T_t = 0 \quad (8)
\]

Where no transportation occurs if and only if the spatial price spread is less than the cost of transporting the first unit of \( x \).

\[\text{Figure 5: Equilibrium With Convex Marginal Costs of Transportation}\]

\[\text{(a) Equilibrium Quantity of Transportation and (b) Effects of an increase in } \tilde{Q}_t^A \text{ or a decrease in } \tilde{Q}_t^B\]

Spatial Price Spread

The key difference in the results of a model with convex marginal costs of transportation is the relationship between the equilibrium level of transportation and the spatial price spread. Specifically, as seen in Figure 4(b), production shocks at both \( A \) and \( B \) and can now directly push the equilibrium spatial price spread above the original marginal cost of transportation.\(^4\)

\[\text{Figure 4: Equilibrium With Constant Marginal Costs of Transportation}\]

\[\text{(a) Equilibrium Quantity of Transportation and (b) Effects of a production shock at } A \text{ and a production shock at } B\]

Spatial Price Spread

2.3 Final Model: Convex Marginal Costs of Transportation With Storage

The inclusion of a storage market for commodity \( x \) significantly complicates the model. The possibility of intertemporal transferring of commodity \( x \) makes the spread curve sensitive to arbitrageurs’ expectations of the future. For the sake of the analysis here, I wish to focus on the intertemporal transferring of \( x \) performed by speculators planning on taking advantage of the

\(^4\)Specifically, a positive production shock at \( A \) and a negative production shock at \( B \) will shift the spread curve out.
spatial price spread in the future. However, it is important to note that there are a plethora of reasons why economic agents would want to store a commodity.

A key result of this analysis will be that the current spatial price spread will decrease as expectations of future transportation costs decrease.\(^5\) Specifically, a storage market allows a speculator to purchase a unit of \(x\) during period \(t\) and store it until period \(t + 1\) at the marginal cost of storage \(k_t^S\). The marginal cost schedule of storage is given by

\[
k_t^S = k_t^S(S_t) : \quad \partial k_t^S(S_t)/\partial S_t > 0, \quad \partial^2 k_t^S(S_t)/\partial S_t^2 > 0
\]

The convexity of the marginal cost schedule of storage is regularly described in the storage literature (see Working, 1948 and Brennan, 1958). The economic theory suggests that infrastructure constraints make the marginal cost of storage increase as the quantity of storage approaches capacity.

The first way the storage market enters our model is by changing the spread function. As speculators increase storage at \(A\), they directly take units of \(x\) out of the market at \(A\), holding all else equal. It follows that the availability of \(x\) at \(A\), denoted \(N_t^A\), now equals \(\bar{Q}_t^A - \Delta S_t - T_t\), where the change in storage from period \(t - 1\) to \(t\) is denoted by \(\Delta S_t = S_t - S_{t-1}\). Therefore, our new spread function is described by \(\sigma_t = D_t^B(\bar{Q}_t^B + T_t) - D_t^A(\bar{Q}_t^A + S_{t-1} - S_t - T_t)\).

Risk-neutral speculators seeking to benefit from expectations about the spread in the future will engage in storage if the net present value of the following risk-free transaction is non-negative: buying \(x\) at \(A\) at price \(p_t^A\), storing it from period \(t\) to \(t + 1\) at marginal cost \(k_t^S(S_t)\), transporting it to point \(B\) at the expected future marginal cost of transportation \(E[k_{t+1}^T]\), and selling the unit of \(x\) at the expected future price of \(x\) at \(B\) during period \(t + 1\), denoted \(E[p_{t+1}^B]\). The net present value of this storage transaction, denoted \(\pi_S\), is given by

\[
\pi_S = \frac{1}{1+r} E[p_{t+1}^B] - \frac{1}{1+r} E[k_{t+1}^T] \equiv k_t^S(S_t) - p_t^A(\bar{Q}_t^A, S_t, T_t)
\]

where \(r\) is the 1-period risk-free interest rate. The no-arbitrage hypothesis suggests that the net present value of this risk-free transaction must be non-positive.\(^6\) Therefore, given small enough values of the first unit of storage \(k_t^S(0)\), arbitrageurs will store positive levels of \(x\) at \(A\) up until they no longer perceive that act of storage to be a positive net present value transaction. This leaves us with the following no-arbitrage condition

\[
\frac{1}{1+r} E[p_{t+1}^B] = \frac{1}{1+r} E[k_{t+1}^T] + k_t^S(S_t) + D_t^A(\bar{Q}_t^A + S_{t-1} - S_t - T_t)
\]

We can now describe general equilibrium in this model with convex marginal costs of transportation and storage. Given sufficient differentials in \(\bar{Q}_t^A\) and \(\bar{Q}_t^B\) such that both transportation and storage are positive, the equilibrium quantities of storage and transportation can be described by the following two conditions:

(i) No-Arbitrage Condition in Transportation Market: \(\sigma_t^* = D_t^B(\bar{Q}_t^B + T_t^*) - D_t^A(\bar{Q}_t^A + S_{t-1} - S_t^*) = k_t^T(T_t^*)\)

(ii) No-Arbitrage Condition in Storage Market: \(\frac{1}{1+r} E[p_{t+1}^B] = \frac{1}{1+r} E[k_{t+1}^T] + k_t^S(S_t^*) + D_t^A(\bar{Q}_t^A + S_{t-1} - S_t^* - T_t)\)

---

\(^5\)One motivation for presenting this result is the announcement that Enbridge made in November 2011 that they plan on adding to the pipeline capacity out of Cushing, which should have decreased expected future costs of transportation.

\(^6\)In other words, the following condition must hold: \(\frac{1}{1+r} E[p_{t+1}^B] \leq \frac{1}{1+r} E[k_{t+1}^T] + k_t^S(S_t) + D_t^A(\bar{Q}_t^A + S_{t-1} - S_t - T_t)\)
2.4 Determinants of The Equilibrium Quantity of Storage

Graphically, the inclusion of a storage market can influence the shape and magnitude of the slope of the spread curve, but will not change the sign of its slope. Most importantly however, the storage market adds two exogenous expectational determinants of the spread curve: expected future costs of transportation and the expected future price of \( x \) at \( B \). This can be seen by substituting (ii) into (i):

\[
\sigma_t^* = \frac{1}{1+r} E[k_t^{T}] + k_t^S(S_t^*) + D^B(\bar{Q}_t^B + T_t^*) - \frac{1}{1+r} E[p_{t+1}^B]
\]  

(9)

I have not found this relationship in the literature: as expectations of future transportation costs decrease, the current spatial price spread falls. The intuition is relatively simple: if you think that future transportation costs will be low, then you will see more of a benefit from storing during period \( t \) and shipping during period \( t+1 \) than transporting during \( t \); therefore you will decrease transportation during period \( t \), which will drive down the cost of transportation during period \( t \).

2.4 Determinants of The Equilibrium Quantity of Storage

The equilibrium quantity of storage at point \( A \) should be sensitive to the exogenous production of \( x \) at both points as well as expectations of future costs of transportation. The sensitivity of storage to the exogenous production at both points can be seen in the re-arranged form of general equilibrium condition (i):

\[
S_t^* = Q_t^A + S_{t-1} - T_t^* - (D^A)^{-1}(D^B(Q_t^B + T_t^*) - k_t^T(T_t^*))
\]  

(10)

Which makes clear the following inequalities:

\[
0 \leq \frac{\partial S_t^*}{\partial Q_t^A} \leq 1 \quad (11)
\]

\[
0 \leq \frac{\partial S_t^*}{\partial Q_t^B} \leq 1 \quad (12)
\]

It follows that although a positive supply shock at \( A \) or a negative supply shock at \( B \) will both similarly increase the price spread, they will have opposite effects on equilibrium quantities of storage. Given (11), a positive supply shock at \( A \) should increase the equilibrium quantity of storage, and given (12), a negative supply shock at \( B \) should decrease the equilibrium quantity of storage.

2.5 Empirical Predictions of Model

(i) Spread Curve: The first prediction of the model is the relationship described by the spread curve: \( \sigma_t = D^B(Q_t^B + T_t) - D^A(Q_t^A + S_{t-1} - S_t - T_t) \). Specifically, we should see a negative relationship between the spread and the variable \( Q_t^B + T_t \) and a positive relationship between the spread and the variable \( Q_t^A + S_{t-1} - S_t - T_t \). The variables \( Q_t^B + T_t \) and \( Q_t^A + S_{t-1} - S_t - T_t \) represent the availability of a commodity at each point, and in some contexts should be exactly equal to consumption at each point. The variables \( Q_t^B \) and \( Q_t^A \), although defined as “production”, can be more generally thought of as supply in a region derived by means other than drawing down storage and transporting from the other considered region.

(ii) Convex Relationship Between Spread and Transportation: The mathematical representation of this is the prediction given by condition (i): \( \sigma_t^* = k_t^T(T_t^*) \). This model predicts
that, in the short-run, where transportation infrastructure is relatively unchanged and thus
the marginal cost schedule of transportation and storage are both not shifting, we should see
a convex relationship between the spatial price spread and the quantity of transportation.

(iii) **Delay Between Increases in Transportation and Widening of Spread:** The convexity of the marginal cost schedule of transportation and storage can explain why storage and transportation spiked before the Brent-WTI spread widened, as seen in Figure 2. Given progressive changes in production differentials, the spread curve would, over time, shift along the flat region of the marginal cost schedule of transportation, increasing transportation substantially but not the price spread. Eventually, the spread curve would reach the convex portion of the marginal cost schedule, and small increases in transportation would correlate with very large increases in the spatial price spread.

(iv) **Direct Relationship Between Current Spread and Expected Future Costs of Transportation:** Given equation (9), we should observe a positive relationship between expected future costs of transportation between a region and the current spatial price spread.

(v) **Distinct Relationships Between Spread, Storage, and Transport Given Different Causes of a Widening Spread:** This model predicts generally distinct combinations of effects on transportation and storage given each possible cause of the Brent-WTI spread. At the end of this paper, we will use these predictions to perform a hypothesis test on the LLS-WTI spread to determine the cause. The predictions are summarized in Table I.

<table>
<thead>
<tr>
<th>Cause of Widening Spread</th>
<th>Interpretation in Model</th>
<th>Effect on Transportation</th>
<th>Effect on Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive production shock in Midwest</td>
<td>Increase in $Q_t^A$</td>
<td>$T_t$ will increase</td>
<td>$S_t$ will increase</td>
</tr>
<tr>
<td>Negative production shock abroad</td>
<td>Decrease in $Q_t^B$</td>
<td>$T_t$ will increase</td>
<td>$S_t$ will decrease</td>
</tr>
<tr>
<td>Increase in marginal cost schedule of transportation</td>
<td>Upwards shift of $k_{t,AB}^T(T_t)$ curve</td>
<td>$T_t$ will decrease</td>
<td>N/A</td>
</tr>
</tbody>
</table>

3 **Empirical Analysis of the LLS-WTI Spread**

3.1 **Data Set**

We will now both test the predictions of the model using the LLS-WTI spread and then use the model to determine the cause of the widening spread. Specifically, we can use a hypothesis test to determine if the widening spread was due to an upward shift in the marginal cost schedule of transportation, an increase in the production of oil at Cushing, or a decrease in the production of oil abroad. As discussed in the introduction, to be accurate with our analysis we will use the LLS-WTI spread as the relevant price spread instead of the Brent-WTI spread in order to isolate an individual transportation market.
3.1 Data Set

If we are to apply our model, we must first specify what we mean by “point A” and “point B” in the context of the LLS-WTI spread. Point A, which in the model was the exporting region, will be defined as PADD2, which is one of the five subregions in the US on which the Energy Information Administration reports detailed production, import, and transportation data. Figure 5 shows the borders of each of the five PADD districts. PADD2 encompasses Cushing, Oklahoma, and therefore transportation out of PADD2 is the best proxy for transportation out of Cushing, the pricing point of WTI, that was readily available. LLS is likely a proxy for oil prices in the rest of the United States, and therefore we will use the summations of data for the other four PADDs (or the rest of the US excluding PADD2) as “point B”. For brevity, we will henceforth call the regions encompassed by all the PADDs other than PADD2 simply the “US”. Undeniably, these regions are larger than I would have liked. I would like to have used just Cushing at “point A” instead of PADD2, but because inflow and outflow transportation data for Cushing is not readily available I will use PADD2 as a proxy for Cushing.

Table II lists summary statistics for the data set I compiled on the two regions considered: the US (excluding PADD2) and PADD2. I compiled monthly spot prices of WTI and LLS from 2004 to 2011 from Bloomberg, and I compiled the rest of the data listed in Table II from the U.S. Energy Information Administration (EIA). I should note that our data on transportation only includes transportation reported to the EIA, and therefore undoubtedly underestimates the actual quantity of transportation from PADD2 to the rest of the US.

The following trends are immediately apparent in Table II: the LLS-WTI spread, transportation, storage in PADD2, and PADD2 production each consistently increased from year-to-year. Alternatively, US production, US imports, and US storage exhibit no clear trend. However, the 2010-2011 time period is of most interest to us, and it should be noted that US imports dropped off sharply in 2011, which potentially supports the hypothesis that there was a negative supply shock abroad. Contrastingly however, PADD2 imports (which accounts for the opening of new pipelines from Canada to Cushing) and production both increased in 2011 as well. This leaves room for the possibility that a simultaneous positive supply shock at ”point A” and a negative supply shock at ”point B” both contributed to the widening LLS-WTI spread. We will econometrically determine the significance of the relationships between these variables in the context of the model developed in this paper.

Table II lists summary statistics for the data set I compiled on the two regions considered: the US (excluding PADD2) and PADD2. I compiled monthly spot prices of WTI and LLS from 2004 to 2011 from Bloomberg, and I compiled the rest of the data listed in Table II from the U.S. Energy Information Administration (EIA). I should note that our data on transportation only includes transportation reported to the EIA, and therefore undoubtedly underestimates the actual quantity of transportation from PADD2 to the rest of the US.

The following trends are immediately apparent in Table II: the LLS-WTI spread, transportation, storage in PADD2, and PADD2 production each consistently increased from year-to-year. Alternatively, US production, US imports, and US storage exhibit no clear trend. However, the 2010-2011 time period is of most interest to us, and it should be noted that US imports dropped off sharply in 2011, which potentially supports the hypothesis that there was a negative supply shock abroad. Contrastingly however, PADD2 imports (which accounts for the opening of new pipelines from Canada to Cushing) and production both increased in 2011 as well. This leaves room for the possibility that a simultaneous positive supply shock at ”point A” and a negative supply shock at ”point B” both contributed to the widening LLS-WTI spread. We will econometrically determine the significance of the relationships between these variables in the context of the model developed in this paper.

---

7PADDs or Petroleum Administration for Defense Districts, were delineated during World War II to facilitate oil allocation. Since then, they have been used to describe intra-country information on the US crude oil industry.

8Although the increase of imports into PADD2 likely understates the dramatic increase in imports specifically into Cushing, since new pipelines were built specifically from Canada to Cushing.
3.2 Empirical Support for the Spread Curve

We will first test the relationship described by the spread curve, \( \sigma_t = D^B (\bar{Q}^B_t + T^*_t) - D^A (\bar{Q}^A_t + S_{t-1} - S^*_t - T^*_t) \). With point \( B \) representing the portion of the US not including PADD2 and point \( A \) representing PADD2, we can econometrically specify the spread function as

\[
\text{Spread} = \beta_0 + \beta_1 (Q_{US} + \text{Imports}_{US} + \text{Transport} - \Delta \text{Storage}_{US}) + \beta_2 (Q_{PADD2} + \text{Imports}_{PADD2} - \text{Transport} - \Delta \text{Storage}_{PADD2})
\]

\( \Delta \text{Storage}_{US} \) and \( \Delta \text{Storage}_{PADD2} \) are the monthly increases in storage within the US and PADD2 respectively. We include \( \text{Storage}_{US} \) because although we assumed away storage at point \( B \) for simplicity in the theoretical section of this paper, the assumption significantly departs from the empirics of the US oil market. The inclusion of storage at \( B \) simply changes the spread function as seen in the specification above. \( \text{Imports}_{US} \) and \( \text{Imports}_{PADD2} \) are included because the US is such a large importer of oil that field production and imports are similar when considering the availability of oil. We expect \( \beta_1 < 0 < \beta_2 \) because the demand curves should be downward sloping. Furthermore, we expect \( |\beta_1| < |\beta_2| \) because PADD2 is a significantly smaller market than the rest of the US, and should therefore have a more elastic demand curve. The regression implicitly assumes that the demand curves for oil in the US and PADD2 have remained constant between 2004 and 2011. Given that this time-series data includes the Great Recession, this assumption can only be justified as an approximation. However, I believe the specification to be a good approximation because it accounts for the effect of supply shocks on price fluctuations (through production and imports) and the effect of expectations of future economic output on price fluctuations (through the inclusion of storage).
3.2 Empirical Support for the Spread Curve

Table 3: Regression Set 1
Dependent variable = "Spread," as defined in Table II

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>(Q_{US} + \text{Imports}_{US} + \text{Transport} )</td>
<td>-0.082***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>(Q_{PADD2} + \text{Imports}<em>{PADD2} - \text{Transport} - \Delta \text{Storage}</em>{PADD2} )</td>
<td>0.616***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>(Q_{US} + \text{Imports}<em>{US} + \text{Transport} - \Delta \text{Storage}</em>{US} )</td>
<td>-0.097***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>(Q_{PADD2} + \text{Imports}<em>{PADD2} - \text{Transport} - \Delta \text{Storage}</em>{PADD2} )</td>
<td>0.616***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.534</td>
</tr>
<tr>
<td></td>
<td>(0.887)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.260</td>
</tr>
<tr>
<td>(\text{AdjR}^2)</td>
<td>0.244</td>
</tr>
<tr>
<td>SE Est.</td>
<td>4.965</td>
</tr>
<tr>
<td>F-Stat</td>
<td>15.981</td>
</tr>
</tbody>
</table>

The symbols *, **, and *** mean statistical significance at the 10%, 5%, and 1% level, respectively.

The results are reported in Table III. We report regression (1), which excludes storage from the independent variables, to demonstrate how including storage in the model increases the explanation power of the specification (\(R^2\) increases from 0.260 to 0.311). In regression (2), the coefficient of \((Q_{US} + \text{Imports}_{US} + \text{Transport} - \Delta \text{Storage}_{US})\) is -0.097 and the coefficient of \((Q_{PADD2} + \text{Imports}_{PADD2} - \text{Transport} - \Delta \text{Storage}_{PADD2})\) is 0.616. This implies that the slope of the demand curve for oil is -0.097 dollars per million monthly barrels in the US (excluding PADD2) and -0.616 dollars per million monthly barrels in PADD2. It would then follow that an increase in transportation from PADD2 by 10 million barrels per month would increase the price of oil in PADD2 by $6 dollars, and decrease the price of oil in the rest of the US by about $1 dollar. The implied price elasticity of demand is -1.96 in the US (excluding PADD2) and -2.02 in PADD2.\(^9\) These are between two and four times larger in magnitude than most econometric estimates of the long-run price elasticity of demand for crude oil in the US.\(^10\) This overestimation is likely caused by the demand curve shifting inward during the Great Recession, which is captured in our data set. In fact, what we would expect to see if there was a recession in the entire United States is that estimations of the demand elasticity in both PADD2 and the US should be overestimated by similar degrees. As such, the remarkable similarity between the elasticities in the US and PADD2, although both likely overestimated, strongly supports the empirical relevance of the spread curve as described in this paper.\(^11\)

\(^9\)Calculated using the average values of the LLS price, WTI price, \((Q_{US} + \text{Imports}_{US} + \text{Transport} - \Delta \text{Storage}_{US})\), and \((Q_{PADD2} + \text{Imports}_{PADD2} - \text{Transport} - \Delta \text{Storage}_{PADD2})\) over the sample data (2004-2011), which were $75.43, $71.61, 396 million barrels per month, and 57.35 million barrels per month respectively.


\(^11\)We make the assumption that the actual demand elasticity shouldn’t vary substantially between the US and PADD2, and therefore simultaneous shifts in the demand curves in both regions should leave the elasticities similar in value, although potentially over or underestimated.
3.3 Empirical Support for the Convex Relationship Between the Spread and Quantity of Transportation

The most obvious econometric specification for equilibrium condition (i), given our assumption that the marginal cost schedule is convex, is

$$\text{Spread} = \beta_0 + \beta_1 \text{Transport} + \beta_2 \text{Transport}^2 + \beta_3 \text{Transport}^3$$

Where Transport is the quantity of transportation out of PADD2 into the rest of the US. I use the cubic function because I want to give the regression room to flatten a large region of the marginal cost schedule before it reaches its convexity. By running this regression I am implicitly assuming that the marginal cost schedule of transportation is not shifting, and rather the spread curve is shifting and equilibrium is operating on the marginal cost schedule. The signs of the coefficients in the above specification, given our assumption of a convex marginal cost schedule of transportation, are expected to maintain $\beta_1 > 0, \beta_2 < 0, \beta_3 > 0$. We can also test a more simplified specification

$$\text{Spread} = \beta_0 + \beta_1 \text{Transport} + \beta_2 \text{Transport}^2$$

where we should expect to see $\beta_1 < 0, \beta_2 > 0$.

Table 4: Regression Set 2

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Transport</td>
<td>-3.072***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>Transport^2</td>
<td>0.466***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
<tr>
<td>Transport^3</td>
<td>0.09***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>6.292***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

| R^2                   | 0.737        | 0.766        |
| Adjusted R^2          | 0.731        | 0.758        |
| SE Est.               | 2.962        | 2.81         |
| F-Stat                | 127.259      | 97.945       |

The symbols *, **, and *** mean statistical significance at the 10%, 5%, and 1% level, respectively.

The results are reported in Table IV. The signs of the coefficients are exactly consistent with the predictions of the model. Although this regression includes no actual data on transportation costs, no-arbitrage suggests that the spread should equal the marginal cost of transportation. Therefore, regressions (1) and (2) should be viewed as indirect estimates of the marginal cost schedule of transporting oil from PADD2 to the rest of the US. As such, the convexity in specifications (1) and (2) supports this paper’s fundamental postulation that the marginal cost schedule of transportation is in fact convex in many real-world markets. Specifically, (1) and (2) both suggest that the widening LLS-WTI spread in 2011 was caused by the spread curve shifting out along the marginal cost schedule.
3.4 The Role of Expectations

The model developed here predicts that a decrease in expected future costs of transportation will decrease the current spatial price spread. But because I have been unable to find a good empirical metric for expected future costs of transportation from Cushing to St. James, I have been unable to develop a robust and direct way to test this prediction. However, the abrupt tightening of the spread on November 16th, 2011, following the announcement that pipeline capacity out of Cushing would be increased in 2012, is at least consistent with the predictions of the model.

**Figure 7:** Daily Percentage Change in LLS-WTI Spread

![Graph showing daily percentage change in LLS-WTI spread with a circled data point on November 16th, 2011.](image)

Figure 6 shows the daily percentage changes in the LLS-WTI spread for each day in 2011. The circled data point is the change in the LLS-WTI spread on the day that Enbridge announced that it would increase pipeline capacity out of Cushing in the future, namely, in 2012. The LLS-WTI spread decreased by 15.8% on the 16th, the 6th largest daily change (in magnitude) for all of 2011. Given the sample data set in 2011, the change observed on the 16th has a t-statistic of -2.39 and a two-sided p-value of 1.7%. This supports the hypothesis that this wasn’t just noise, and instead was a consequence of the mechanism described in this paper that should create a positive correlation between the current spatial price spread and expected future costs of transportation.

3.5 Hypothesis Tests on The Cause of The LLS-WTI Spread

Now that evidence has been presented in support of the model, we can perform hypothesis tests on whether the cause of the LLS-WTI spread was a negative production shock abroad, a positive production shock in the Midwest, or an increase the the marginal cost schedule of transportation. We can test this using the predictions of the model developed in this paper that are summarized in Table (I). We will perform hypothesis tests on the correlations between transportation and the LLS-WTI spread as well as between storage and the LLS-WTI spread. The combination of correlations, if both statistically significant, will provide strong support for a cause of the widening spread.

---

12 Calculated using the standard deviation of daily changes of the spread during 2011 (6.753%) and the mean daily change (0.41%)
Table 5: Hypothesis Tests

<table>
<thead>
<tr>
<th>Hypothesis Test</th>
<th>Sample Correlation</th>
<th>t-statistic</th>
<th>One-sided p-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0 : \rho_{e1, T1} \leq 0$</td>
<td>0.783</td>
<td>12.051</td>
<td>0.000</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>$H_1 : \rho_{e1, T1} &gt; 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0 : \rho_{e1, S1} \leq 0$</td>
<td>0.532</td>
<td>6.183</td>
<td>0.000</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>$H_1 : \rho_{e1, S1} &gt; 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

t-statistics calculated by $t = \frac{\rho \sqrt{df}}{\sqrt{1-\rho^2}}$ where this data sample has $df = n - 2 = 92$

Table V reports the results of the hypothesis tests. I calculated the sample correlation between the spread and transportation (reported in first row) and the sample correlation between the spread and storage (reported in the second row). The first hypothesis test, holds the null hypothesis that the correlation between the spread and transportation was nonpositive between 2004 to 2011. The statistical significance of the observed correlation of 0.783 allows us to reject the null at a 1% level of confidence. We can similarly reject the null hypothesis that the correlation between the spread and storage was nonpositive between 2004 and 2011. We can therefore conclude that both transportation and storage were positively correlated with the LLS-WTI spread between 2004 and 2011. According to the model presented in this paper, this suggests that the LLS-WTI spread was caused by a positive supply shock in the Midwest. However, the results are sensitive to the time period tested, and if expectations of future transportation costs changed substantially within the data set then we could observe these correlations without a positive supply shock in the Midwest.

4 Conclusion

The addition this paper provides to the economic literature is primarily theoretical. The framework of analysis presented in here is a generalized, and in my view much clearer, two region version of that presented by Samuelson (1952). Unlike Samuelson (1952), we explicitly defined a spread function and computed the equilibrium level of transportation as the intersection between the spread curve and the marginal cost schedule of transportation. This perspective made clear the key differences between a model with constant marginal costs of transportation versus one with convex marginal costs of transportation.

The key addition of this paper to the literature is the description of these theoretical differences and the effects that a convex marginal cost schedule of transportation has on spatial price equilibrium between two regions. Specifically, this paper describes the endogeneity of the equilibrium spatial price spread to both the demand curves and production levels in different regions. The applicability of the generalized model presented here is likely not limited to the LLS-WTI spread, and further research should be performed on the applicability of this model to other spatial price spreads that exhibited similar changes.

The second major addition of this paper to the literature is the description of the direct effect expected future costs of transportation has on the current price spread. General equilibrium (ii) provides a channel through which changes in expectations of future transportation costs can influence price spreads. This opens up the literature to explore potential cases of "animal spirits" causing dramatic changes in geographic price differentials. It is reasonable to posit that if some
market participants’ expectations of future costs of transportation were to increase, the subsequent widening of the spread and increase in the cost of transportation today would cause expectations of future costs of transportation to rise even further. This spiraling effect could lead the price spread to increase without bound.

Further theoretical work should be performed to generalize this model to multiple regions to more directly contrast it with the multiple-region models developed by Samuelson (1952) and Takayama and Judge (1964).

My econometric analysis of the LLS-WTI spread could be improved and expanded upon substantially. First, a sufficient metric for expected future costs of transportation should be used to test equilibrium condition (ii) directly. Second, regression set 2 should be more rigorously specified to isolate two distinct locations instead of the massive regions encompassed by PADD2 and the rest of the US. Third, and most important, econometric analysis should be performed on a spatial price spread with data on the marginal cost of transportation. The most convincing empirical support for the model presented in this paper would be through direct data on transportation costs. I was unable to find transportation cost data on the LLS-WTI spread, but if such data were found it would be a powerful test of this model’s empirical legitimacy.

In sum, this paper provides a necessary generalization of the literature that addresses the interconnectedness of spatially separated commodity markets. In doing so this paper has rationalized the empirical phenomena of abruptly widening spatial price spreads, opened the door for further econometric analysis on the effect of expected future costs of transportation on current spatial price spreads, and developed the foundation for developing a model of spatial price equilibrium with multiple regions in the presence of a convex marginal cost schedule of transportation.

References


REFERENCES


