NBER WORKING PAPER SERIES

"THE BIGGER THEY ARE, THE HARDER THEY FALL": HOW PRICE DIFFERENCES ACROSS U.S. CITIES ARE ARBITRAGED

Paul G. J. O'Connell
Shang-Jin Wei

Working Paper 6089

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1997

We are grateful to Jon Faust, John Leahy, Alan Taylor and seminar participants at Harvard University and Georgetown University for helpful comments. All errors are our own. This paper is part of NBER's research programs in International Finance and Macroeconomics and International Trade and Investment. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

"The Bigger They Are, The Harder They Fall":
How Price Differences Across U.S. Cities Are
Arbitraged
Paul G. J. O'Connell and Shang-Jin Wei
NBER Working Paper No. 6089
July 1997
JEL Nos. F31, C32
International Finance and Macroeconomics
and International Trade and Investment

ABSTRACT

Recent empirical work has made headway in exploring the non-linear dynamics of deviations from the law of one price and purchasing power parity that are apt to arise from transaction costs. However, there are two important facets of this work that need improvement. First, the choice of empirical specification is arbitrary. Second, the data used are typically composite price indices which are subject to potentially serious aggregation biases.

This paper examines the evidence for transport-cost-induced nonlinear price behavior within the U.S. We address both of the above shortcomings. First, we use a simple continuous-time model to inform the choice of empirical specification. The model indicates that the behavior of deviations from price parity depends on the relative importance of fixed and variable transport costs. Second, we employ data on disaggregated commodity prices, yielding a "pure" measure of the deviations from price parity. We find strong evidence of nonlinear reversion in these deviations. The nature of this reversion suggests that fixed costs of transportation are integral to an understanding of law-of-one-price deviations.

Paul G. J. O'Connell
Department of Economics
Harvard University
Cambridge, MA 02138
poconnel@kuznets.fas.harvard.edu

Shang-Jin Wei
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
shang-jin_wei@harvard.edu

1 Introduction

The idea that transactions costs are important for understanding deviations from the law of one price (LOP) is not a new one. Obstfeld and Taylor (1997) note that, as far back as 1916, Heckscher argued that transport costs should create some scope for price discrepancies to arise without precipitating goods arbitrage. Notwithstanding its vintage, it is only recently that attention has been devoted to developing the theoretical and empirical implications of this idea.

Williams and Wright (1991), Dumas (1992), Uppal (1993) and Coleman (1995) have shown that, in the presence of proportional transport costs, the incentive to engage in goods arbitrage is tempered.[1] Building on this, recent empirical work has noted that transport costs provide a viable candidate explanation for the chimerical nature of mean-reversion in relative goods prices; if market frictions such as transport costs are present, then the power of standard tests for mean-reversion is diminished. O'Connell (1997a), Obstfeld and Taylor (1997) and Michael, Nobay and Peel (1997) build this insight into their analyses of purchasing power parity.

In this paper we seek to extend the discussion of transport costs and their impact on goods arbitrage along two dimensions. Along the theoretical dimension, we provide a simple continuous-time model that highlights the relative importance of proportional *and* fixed costs of shipping. The main implications of the model are as follows. First, if the only cost of arbitrage is a proportional transport cost, then the process for LOP deviations is confined between reflecting barriers that delimit a "range of no-arbitrage," within which it is not profitable to engage in trade. In these circumstances, when arbitrage does take place, the quantities traded are very small, sufficient to prevent the LOP deviation from going outside the reflecting barrier, but insufficient to drive the deviation back into the interior of the no-arbitrage band. Second, if transport costs are fixed rather than variable, the process for LOP deviations is confined between "resetting" barriers. These too delimit a band of no-arbitrage. However, when arbitrage does take place at the edge of the band, it is sufficient to completely eliminate the LOP deviation. Thus the process for the deviation is reset to zero through trade. Third, if both fixed and variable transport costs are present, an interesting hybrid case emerges. There are two bands for LOP deviations, an inner band within which no

[1] In addition to proportional transport costs, Coleman (1995) assumes that shipping takes time.

arbitrage takes place, and an outer band in which there is some arbitrage activity. Finally, the magnitude of these bands is increasing in the variability of relative goods prices.

The second dimension along which we seek to add value is empirical. There are two shortcomings of the extant empirical work. First, the choice of empirical specification is arbitrary. Second, the data used are typically composite price indices which are subject to potentially serious aggregation biases. We address the former by using our model to inform the choice of empirical specification, and to interpret the results. To tackle the latter, we employ data on disaggregated commodity prices, yielding a "pure" measure of the deviations from price parity. To summarize the results, we find strong evidence to support the existence of market frictions that interfere with goods arbitrage. Many relative price series that appear nonstationary using standard techniques exhibit strong evidence of reversion once the deviation from LOP goes outside an estimated band of no-arbitrage. When we test whether it is the fixed or variable component of market frictions that is most important, the evidence indicates that the fixed component is dominant. This suggest that fixed costs such as those associated with building production or distribution capacity are integral to an understanding of U.S. goods prices.

The paper begins with a brief review of the extant theoretical and empirical work on relative price behavior in the presence of market frictions. Then Section 3 sets out our model and its implications for empirical analysis. Section 4 describes the data and the empirical strategy. Finally, Section 5 contains the results from the battery of empirical tests that we carry out.

2 Extant work on market frictions and relative prices

2.1 Theoretical work

As noted in the introduction, a number of papers have employed proportional transport costs to support deviations from the law of one price. A simple example illustrates the likely effect of such costs. Suppose that there are two countries, A and B, and that the price of a good is 1 in A and P in B. Furthermore assume that there proportional transport costs of the "iceberg" form—if a good is shipped from one location to another, a fraction l melts en route, so that only $(1 - l)$ of the good actually arrives. At what levels of P does it become profitable to ship goods from one country to another?

The profit from shipping one good from B to A is $(1-l)-P$, which is positive for $P<1-l$. The profit from shipping a good in the opposite direction is $(1-l)P-1$, which is positive for $p>1/(1-l)$. Thus the "band of no-arbitrage"—the interval for P within which arbitrage doesn't pay—is $(1-l)<P<1/(1-l)$. The assumption of proportional transactions costs renders this interval symmetric about P *in percentage terms.*[2]

The models of Williams and Wright (1991), Dumas (1992), Uppal (1993) and Sercu, Uppal and Van Hulle (1995) generate a similar band of no arbitrage. Such models carry specific implications for the impact of arbitrage on the real exchange rate. For instance, if P reaches the upper threshold generated by the transport costs, $1/(1-l)$, the amount of trade that takes place is sufficient to ensure that the P does not rise above this level, but not so large as to cause P to fall into the interior of the band. In the jargon of continuous time, the thresholds $(1-l)$ and $1/(1-l)$ are "reflecting barriers" for the real exchange rate process. This means that evidence of arbitrage activity will not be reflected directly in prices, only in quantities. While the barriers render P stationary, it can be difficult to detect this stationarity using standard techniques.[3]

In an important contribution to this topic, Coleman (1995) challenges the assumption of instantaneous trade, and builds a model that instead allows transportation to take time. In his model, arbitrageurs engage in trade whenever the *expected* future price in one country exceeds the spot price in the other country, plus the transport cost. The result is that the real exchange rate can differ from 1 by more than the transport cost. Coleman shows that once P strays outside the band of no arbitrage, it exhibits a tendency to revert to the band. Moreover the strength of this tendency is increasing in the distance from the band. In this setting, arbitrage activity is reflected in both

[2]One interpretation of iceberg transport costs is that they are costs incurred in the destination country. More generally, per-unit costs could comprise l_1 in the export country and l_2 in the import country, while still preserving the symmetry of the band of no-arbitrage.

[3]It is worth noting that these models can also render predictions about the behavior of the real exchange rate *within* the band of no arbitrage. In the models of Dumas (1992) and Uppal (1993), for instance, there is a single good that can be consumed or invested. Each country is endowed with a production technology for the good that is subject to white-noise shocks. Because agents in each country are risk averse, it is optimal to rebalance the (capital) stock of the good in each country after each shock. However, if there are costs of shipping the good, no rebalancing will take place until the marginal benefit of doing so exceeds the marginal cost. The relative shadow value of the good across countries (the real exchange rate) can therefore deviate from 1. Dumas (1992) and Uppal (1993) show that in the equilibrium, the real exchange rate is bounded between $(1-l)$ and $1/(1-l)$. Interestingly, however, the real exchange rate displays a centrifugal tendency within the band, in the sense that the conditional probability of the deviation from parity widening is always larger than the probability of it narrowing. In consequence the ergodic distribution for P is U-shaped: P spends most of the time away from parity, close to the boundary where arbitrage takes place.

prices and quantities, and hence it may be easier to detect using just prices.

In the model set out in Section 3, we consider an alternative modification to the basic iceberg cost setup. Specifically, we add a fixed cost of arbitrage to the proportional transport cost. We show that this supports interesting and realistic dynamics in the process for relative prices.

2.2 Empirical work

Some recent empirical work has sought to modify standard tests of the LOP and purchasing power parity (PPP) to take account of actual or potential market frictions.[4] Parsley and Wei (1996) add a higher-order term to the standard Dickey-Fuller regression in an effort to determine whether mean-reversion is increasing in the size of LOP deviations. Using the same data that this paper employs, they find that convergence is faster for large initial price differences.[5] Michael, Nobay and Peel (1997), employing a substantially similar technique, find that the rate at which deviations from PPP die out is increasing in the size of deviation for interwar CPI data, and for a 200-year data set of UK and French real CPI exchange rates against the dollar. However, when O'Connell (1997a) applies the same technique to post-Bretton Woods real exchange rates, both large and small deviations from PPP appear equally persistent.

An alternative strategy is to fit a threshold autoregression model (described in more detail below) to relative prices. O'Connell (1997a) fits such a model to real effective trade-weighted exchange rates from the 1973–1995 period, and finds no evidence that large deviations from parity die out relatively more quickly. Indeed, if anything, the opposite appears to be true: small deviations tend to die out but large ones are apt to persist indefinitely. This finding may be related aggregation biases in the composite traded goods prices used in the paper. Obstfeld and Taylor (1997) also fit threshold autoregression models to detrended real exchange rates for the U.S. sampled from 1973–1995. They report evidence that large deviations from the linear trend of the real exchange rate revert to parity quite quickly, while small deviations do not.

In the empirical analysis in Section 5, we fit a variety of what may be termed "nonlinear"

[4]Many "standard" tests for relative price stationarity have been carried out of late, especially in panel settings. These include Abuaf and Jorion (1990), Wei and Parsley (1995), Engel, Hendrickson and Rogers (1996), Frankel and Rose (1996), Jorion and Sweeney (1996), Papell (1996), Taylor (1996), O'Connell (1997b) and Papell and Theodoridis (1997a). See Froot and Rogoff (1995) and Rogoff (1996) for surveys of the literature on PPP.

[5]See also Wei and Parsley (1995).

reversion models to detailed commodity price data for the U.S. The hope is that the quality of the data will circumvent some of the potentially serious aggregation biases that arise when using aggregate prices indices. The analysis complements the earlier work of Parsley and Wei (1996), and draws on some of the techniques used in O'Connell (1997a) and Obstfeld and Taylor (1997). Before engaging in the empirical analysis, however, it is useful to set out a theoretical framework for thinking about transport costs and other market frictions, and this is the purpose of the next section.

3 A model of arbitrage with transport costs

In this section, we illustrate the relative importance of fixed and variable transport costs in determining relative prices. The analysis is simplified by not modelling production or intertemporal trade. In principle it would be straightforward to generalize the model to include production, investment and changes in the economy's net international investment position, but the minimal specification is sufficient to highlight the major points. The model draws from the excellent general discussion of optimal control and regulation in Dixit (1993).

3.1 The optimal pattern of trade

We examine the optimal trade strategy in a small open economy that is endowed with two nonstorable commodities, X and Y. The representative agent for this economy has the quasilinear utility function

$$U(C^X, C^Y) = C^X - \frac{1}{\gamma}\exp(-\gamma C^Y). \tag{3.1}$$

Thus the marginal utility of X is fixed at 1. The economy is endowed with a nonstochastic supply of X at each instant. This endowment is abundant, in the sense that there will always be some of X consumed.[6] The endowment of Y is stochastic at each instant, with dynamics given by

$$dY = \sigma dz, \tag{3.2}$$

[6]In this sense, the model is partial equilibrium. The endowment of Y is assumed to be insignificant in size relative to the endowment of X.

where dz is the increment of a standard Wiener process. Consumption C^Y is equal to the endowment plus net imports of the good M from abroad:

$$C^Y = Y + M. \tag{3.3}$$

We assume that X and Y are traded on world markets, and that both are priced at unity. Moreover trade is always balanced.[7] Together these assumptions imply that

$$C^X = X - M. \tag{3.4}$$

Our ultimate interest is in the process for P, the price of good Y in this economy. Since the marginal utility of X is one, $P = \exp(-\gamma C^Y)$.

The transportation technology available to the economy has the following features. First, in order to facilitate imports or exports, trade capacity must be available. We assume that trade capacity is unidirectional—capacity installed for the purposes of exporting cannot be used for imports. The per-unit cost of new capacity is l. Once installed, capacity must be utilized. Capacity may, however, be decommissioned at a cost. To preserve symmetry, we assume that the per-unit decommissioning cost is also l. Second, every time that trade capacity is adjusted, a fixed cost k must be expended.

In the absence of transport costs, the solution to the representative agent's maximization problem is straightforward. At each instant, the marginal utility from consumption of Y must equal 1, so the agent simply sets M equal to $-Y$. In the presence of transport costs, however, C^Y can deviate from 0. To obtain the solution, it is convenient to work with a utility loss function rather than the level of utility itself. The loss function is defined as

$$L(C) = -\frac{1}{\gamma} + C + \frac{1}{\gamma}e^{-\gamma C} \tag{3.5}$$

where $C \equiv C^Y$ (the Y superscript is dropped hereafter to economize on notation). $L(C)$ captures the relative gain in utility from consuming at 0 instead of consuming at C.[8] It is a convex function that attains a minimum 0 at $C = 0$. The reason for working with this loss function is that it is invariant to the scale of the endowment $Y(t)$. Clearly, *total* utility depends on this endowment, but

[7]Implicit here is the assumption that the economies rate of time preference is equal to the world interest rate. If this were not the case, then there would be no well-defined equilibrium, as marginal utility is not diminishing in X.

[8]Consuming at 0 yields $-1/\gamma$ plus Y export revenue, while consuming at C yields $-exp(-\gamma C)/\gamma$ plus $-M$ export revenue. The difference between these is $L(C)$.

given quasilinear preferences and the abundance of X the marginal import decision is independent of this endowment, and so nothing is lost by seeking to minimize $L(C)$ rather than to maximize $U(\cdot)$.

Holding M constant, the process for $C \equiv Y + M$ is simply

$$dC = \sigma dz. \tag{3.6}$$

Define the value function

$$V(C) = \min_{\{M\}} E_C \int_0^\infty e^{-\rho t} L(C) dt, \tag{3.7}$$

where ρ is the discount rate. The problem is to solve for $V(C)$, taking into account the costs associated with trade. The Hamilton-Bellman-Jacobi equation for this problem is $E[dV(C)]/dt + L(C) = \rho V(C)$. Applying Itô's lemma, this can be written as

$$\frac{1}{2}\sigma^2 V_{CC}(C) - \rho V(C) + L(C) = 0. \tag{3.8}$$

This differential equation has a well-known solution (see, for example, Dixit (1993)). The complementary function is

$$Ae^{-\alpha C} + Be^{\alpha C},$$

where $\alpha = \sqrt{2\rho}/\sigma$, and A and B are constants to be determined. For the particular integral, try the linear form

$$V(C) = K_1 e^{-\gamma C} + K_2(C - 1/\gamma).$$

This implies that $V_{CC}(C) = K_1\gamma^2 e^{-\gamma C}$. Substituting this into (3.8) and solving we obtain

$$K_1 = -\frac{1}{\gamma\left(\frac{1}{2}\sigma^2\gamma^2 - \rho\right)}; \qquad \text{and} \qquad K_2 = \frac{1}{\rho}.$$

Thus the full solution is

$$V(C) = Ae^{-\alpha C} + Be^{\alpha C} + \frac{e^{-\gamma C}}{\gamma\left(\rho - \frac{1}{2}\sigma^2\gamma^2\right)} + \frac{1}{\rho}\left(C - \frac{1}{\gamma}\right). \tag{3.9}$$

The solution for $V(C)$ is made up of two parts. The last two terms on the right-hand side comprise the present value of the loss function if net imports are fixed—they are the discounted sum of expected losses $L(C)$. The first two terms represent the change in value that accrues from the ability to control M: the value of the option to import is $AC^{-\alpha}$, while the value of the option to export is BC^{α}. These quantities are negative as they add to utility and hence subtract from the loss function. It can be shown (see Dixit (1993), and the references cited therein) that the optimal trade policy is characterized by four threshold levels of C, $\overline{C}_1 > \overline{C}_2 > \overline{C}_3 > \overline{C}_4$, that have the following features. First, if C rises to $\overline{C}_1$, $(\overline{C}_1 - \overline{C}_2)$ of net export capacity is installed at cost $k + (\overline{C}_1 - \overline{C}_2)l$, and used to export an additional $(\overline{C}_1 - \overline{C}_2)$ of Y to the world market.[9] Second, if C falls to $\overline{C}_4$, $(\overline{C}_3 - \overline{C}_4)$ of net import capacity is installed, and used to import an additional amount $(\overline{C}_3 - \overline{C}_4)$ from the world market. In other words, if C strays too far from it's optimum of 0, the representative agent "resets" it at a value that is closer to the optimum by raising or lowering net imports.

To solve for the these threshold values, we employ some boundary conditions that tie down the constants A and B. At the outer barriers $\overline{C}_1$ and $\overline{C}_4$, it must be true that the increment in value that is brought about by trade equals the cost of trade. Thus

$$V(\overline{C}_1) - k - l(\overline{C}_1 - \overline{C}_2) = V(\overline{C}_2).$$

In addition, it must be true that at both $\overline{C}_1$ and $\overline{C}_2$, the marginal increment to value from further trade is just equal to the marginal cost, or

$$V_C(\overline{C}_1) = V_C(\overline{C}_2) = l.$$

These are the smooth-pasting conditions.[10] Similar conditions obtain at $\overline{C}_3$ and $\overline{C}_4$:

$$V(\overline{C}_4) - k - l(\overline{C}_3 - \overline{C}_4) = V(\overline{C}_3).$$

and

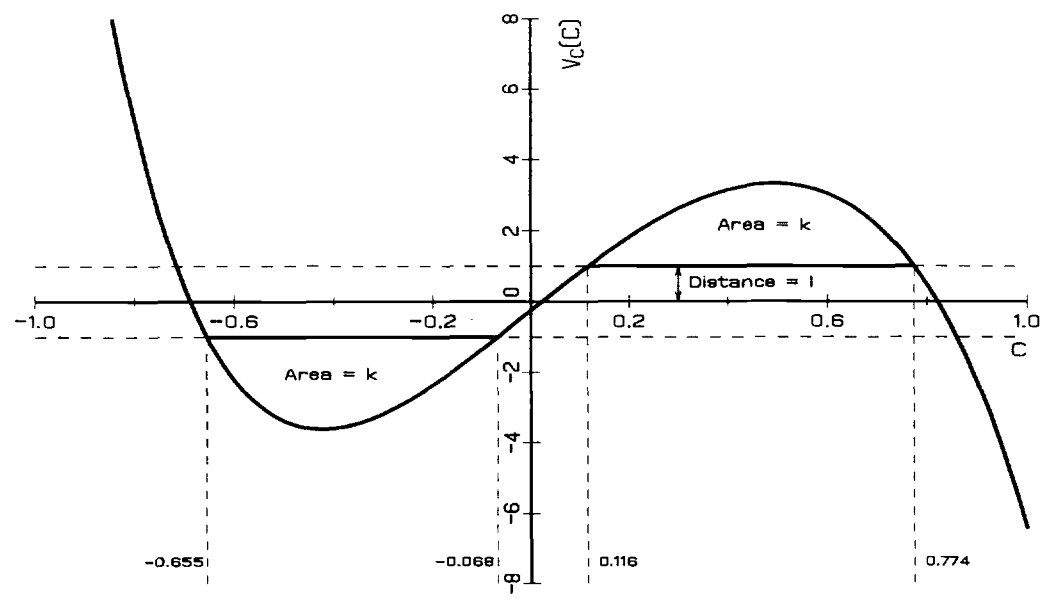
$$V_C(\overline{C}_3) = V_C(\overline{C}_4) = -l.$$

These six boundary conditions can be solved for A, B and the four optimal thresholds. As they are highly nonlinear, it is easiest to obtain the solution numerically. For a benchmark case, let the

[9] If the country is already importing, then some of its import capacity will be decommissioned.

[10] Given that L_C, the marginal utility loss, is monotonically increasing, it may appear surprising that $V_C = l$ at *both* the outer threshold $\overline{C}_1$ and the inner threshold $\overline{C}_2$. The reason is that the marginal value of the ability to control C through trade is monotonically *decreasing* in C. If net import capacity *is* actually adjusted, then the drop in the option value to carry out further adjustments contributes negatively to utility. This is illustrated in Figure 1.

Figure 1: Optimal trade policy in the presence of fixed and variable transport costs



coefficient of absolute risk aversion γ equal 1, the instantaneous variance of the endowment σ equal 0.01, the rate of discount ρ equal 0.05, and the fixed and variable costs of transport be $k = 1$ and $l = 1$. respectively. To solutions for A and B are -0.872 and -0.246, and the four threshold values of C are $\overline{C}_1 = 0.774$, $\overline{C}_2 = 0.116$, $\overline{C}_3 = -0.068$ and $\overline{C}_4 = -0.655$. Constantinides and Richard (1978), Harrison, Sellke and Taylor (1983) and Dixit (1993) all use a simple graph of $V_C(C)$ to convey the nature of this type of solution, which is shown in Figure 1.

For C close to 0, the value of the options to adjust import capacity up or down are relatively small, and $V_C(C)$ mimics $L_C(C)$. As C departs from its optimum level, the trade options become more valuable—the value of the option to export, $B\exp(\gamma C)$, is increasing in C, and the value of the option to import, $A\exp(-\gamma C)$ is decreasing in C. It is variation in the value of these trade options that gives the overall marginal utility curve its characteristic shape. The location of the curve is determined by the value-matching and smooth-pasting conditions. In particular, the smooth-pasting conditions require that the abscissae of the curve's intersections with l are $\overline{C}_1$ and $\overline{C}_2$, and that the abscissae of the curve's intersections with $-l$ are $\overline{C}_3$ and $\overline{C}_4$. The value-matching conditions require that the area below the curve between $\overline{C}_1$ and $\overline{C}_2$ equals $k + l(\overline{C}_1 - \overline{C}_2)$, and that the area above the curve between $\overline{C}_3$ and $\overline{C}_4$ equals $k + l(\overline{C}_3 - \overline{C}_4)$.

3.2 The behavior of relative prices

Having solved for the optimal trade policy, we are now in a position to examine the behavior P, the relative price of good Y in the economy.

A. Costless trade

If all trade is costless ($k = l = 0$), all four trading thresholds collapse to zero. Trade takes place instantaneously whenever the endowment of C differs from 0. P is therefore fixed at unity.

B. Infinite costs of trade

If there are infinite costs of trade, $k \to \infty$ and/or $l \to \infty$. In this case, $\overline{C}_1, \overline{C}_2 \to \infty$, $\overline{C}_3, \overline{C}_4 \to -\infty$, and no trade takes place. The relative price P will equal $U_C(C) \equiv \exp(-\gamma c)$ at all instants. By Itô's lemma,

$$dP = \frac{1}{2}\sigma^2\gamma^2 P dt - \sigma\gamma P dz. \tag{3.10}$$

This implies that $p \equiv \ln(P)$ follows a driftless arithmetic Brownian motion with instantaneous variance $\sigma^2\gamma^2$:

$$dp = \sigma\gamma dz. \tag{3.11}$$

C. Proportional costs of trade

If the only costs of trade are proportional (i.e. $k = 0$), then $\overline{C}_1$ coincides with $\overline{C}_2$, and $\overline{C}_3$ coincides with $\overline{C}_4$. In our benchmark case, the solutions for these thresholds are $\overline{C}_1 = 0.283$, and $\overline{C}_4 = -0.256$. The resulting process for consumption shares many of the features of the solution in the Dumas (1992) model. In particular, $\overline{C}_1$ and $\overline{C}_4$ become *reflecting* barriers for the consumption process. If trade takes place, it will involve infinitesimal quantities at these barriers. The process for U_C and hence P will inherit these properties. Thus P will follow the process (3.10) until such time as C reaches one of the barriers, when sufficient trade will take place to hold P at its barrier level, without driving it back towards parity. In our benchmark case, the reflecting barriers for P are $\exp(-0.283\gamma) = 0.753$ and $\exp(0.256\gamma) = 1.292$.

D. Fixed costs of trade

If the only costs of trade are fixed (i.e. $l = 0$), then the middle two thresholds $\overline{C}_2$ and $\overline{C}_3$ coincide. With the benchmark parameters, the threshold solutions are $\overline{C}_1 = 0.688$, $\overline{C}_2 = \overline{C}_3 = 0.019$, and $\overline{C}_4 = -0.600$. The resulting consumption process differs markedly from the case of proportional trade costs. In particular, whenever C hits either $\overline{C}_1$ or $\overline{C}_4$, a discrete amount of trade will take place that is sufficient to bring consumption back to 0.019.[11] The intuition is that, once the fixed cost has been expended, it would be suboptimal to reset consumption to a point that is away from value-maximizing point close to 0.

The process for U_C and hence P inherit these resetting features. Thus P will follow the process (3.10) until such time as C reaches one of the barriers, at which time P is reset to $\exp(-0.019\gamma) = 0.981$. The resetting barriers for P are 0.549 and 1.990.

E. Fixed and variable costs of trade

Lastly, we consider the behavior of P in the presence of both fixed and proportional costs of transportation. In the benchmark case, when P hits $\exp(-\gamma\overline{C}_1) = 0.461$, arbitrage moves it to $\exp(-\gamma\overline{C}_2) = 0.890$. Correspondingly, when P hits $\exp(-\gamma\overline{C}_4) = 2.168$, arbitrage resets it to $\exp(-\gamma\overline{C}_3) = 1.070$. The interesting aspect of this solution is that it generates two "bands" for the deviations from the LOP. Whenever Y hits the outer barriers, it is reset by trade to the inner barriers. This resetting behavior differs from the infinitesimal arbitrage that characterizes models predicated solely on proportional transport costs.

3.3 Implications for testing the LOP

The model developed above has three important implications for the empirical analysis of the LOP. First, in the presence of transport costs, the stationarity of relative prices may be difficult to detect using conventional tests. Second, by taking advantage of the special structure of price behavior that arises with transport costs, the power to detect stationarity can be increased. And third, the different patterns of price behavior occasioned by fixed and variable costs provide some basis for empirically distinguishing their relative importance.

[11] Notice that the value-maximizing point is not actually 0. This is because of the asymmetry introduced by the Itô or Jensen's inequality term in the value function.

A. Detecting stationarity in the presence of transport costs

Let the true process for the relative price of a good in two locations, p, be

$$\Delta p_t = \begin{cases} \epsilon_t & \text{if } |p_{t-1}| < a \\ b + \epsilon_t & 0 < b < a \text{ otherwise} \end{cases} \tag{3.12}$$

This is the discrete-time analog of the process for p that emerges from our model in the presence of both fixed and flexible costs. The process is globally stationary, but because innovations to the process are i.i.d. for a portion of the time (i.e. whenever $|p_{t-1}| < a$), this stationarity can be difficult to detect using standard techniques. For example, if $a = 4$, $b = 3$ and $\epsilon \sim \text{i.i.d.}N(0,1)$, the power of the Dickey-Fuller test to reject the random walk null with 50 observations on this process is 22 percent at the 5 percent significance level.[12] This can be compared to the power of the Dickey-Fuller test when the true process is $AR(1)$ with the same total variance $V(p_t)$ as the process (3.12). When $a = 4$, $b = 3$ and $\epsilon \sim \text{i.i.d.}N(0,1)$, $V(p_t) = 3.42$.[13] This is matched by an $AR(1)$ process with a disturbance variance of 1 and a root of 0.84.[14] The power of the Dickey-Fuller test to reject the random walk null under this $AR(1)$ alternative is 37 percent at the 5 percent significance level, 15 percent higher than for process (3.12).

It appears, then, that standard tests for stationarity have low power under the alternatives generated by market frictions. This may account for the perennial difficulty of rejecting the unit root null in relative price data.

B. Increased power to detect stationarity

There is, of course, a positive side to the particular structure that transport cost models imply for relative price behavior. It is likely that the power of stationarity tests can be increased by modifying them to take account of the fact that reversion in prices only takes place at certain times. Transport cost models predict that small price discrepancies will not be arbitraged, but that large ones will. This suggests conditioning reversion on the size of the deviation from LOP. As discussed in Section 2, this strategy has been adopted in some very recent work. To illustrate, suppose that only the observations on Δp_t for which $|p_{t-1}|$ exceeds some threshold s are included in

[12] This power calculation is carried out by Monte Carlo simulation, using 1,000 observations on the distribution of the test statistic under the alternative (3.12). The Dickey-Fuller regression run here includes an intercept.

[13] $V(p_t)$ is estimated from 1,000 simulations of the process (3.12).

[14] The total variance of an $AR(1)$ process $p_t = \rho p_{t-1} + \epsilon_t$ is $\sigma_\epsilon^2/(1-\rho^2)$.

the Dickey-Fuller regression. In other words, small deviations from the LOP are excluded from the regression. If the true process for p is (3.12), then these observations contain little or no information on reversion in p; they just add noise to the estimation. It follows that a more precise estimate of reversion is available from the test, which ought to increase power. This in fact turns out to be the case. If, for example, we choose to look only at the upper quartile of LOP deviations, the power to reject nonstationarity under the alternative (3.12) is 31 percent.[15]

C. Distinguishing the importance of fixed versus variable transport costs

The model suggests that fixed costs of transportation tend to lead to large discrete jumps in relative prices back towards parity, while variable costs lead to small continuous changes in price that limit the size of deviations from parity, but do not necessarily reduce them. The hybrid model in which there are both fixed and variable costs predicts that relative prices will exhibit "band-reversion"— large deviations from the LOP will be partially but not completely arbitraged away when they reach a certain size. This different predictions offer a potential way to determine whether fixed or variable costs of arbitrage are dominant. One simply tests whether observed reversion in relative prices is towards parity, or towards some point that is away from parity.

4 Empirical strategy

In this section, we set out our empirical strategy for testing the model's implications

4.1 Data

Our data set is substantially the same as that used in Parsley and Wei (1996). We sample the prices of 48 final goods and services from 24 cities in the United States over the period 1975:1–1992:4. The data is collected from the American Chamber of Commerce Researchers Association publication, *Cost of Living Index* (hereafter, *Index*). Each quarterly issue of *Index* contains comparative average price data for a sample of urban areas, and a cost of living index computed from these data by the Association. In this study we use only the raw price data.

[15]This power calculation is estimated from 5,000 simulations of the test statistic under the null, and 1,000 simulations under the alternative.

The actual data collection is done by the local Chamber of Commerce staff or volunteers for the Chamber, and is voluntary. Explicit instructions and data forms are provided for each data collector by the association. Some prices are obtained by phone and usually the respondents do not know it is for a survey. Once collected, the data is sent to one of nine different regional coordinators for checking. Finally, the data is sent to Houston where it is transferred to computer and subjected to both computer and visual checks for outliers. Publication occurs approximately five and one half months after the original data are collected.

The sample of cities included in each issue of *Index* varies. At the beginning of our sample period there were one hundred sixty six cities and forty four items priced. The number of cities steadily increased to two hundred ninety seven in 1992.4; however each report contains a distinct sample of cities. We choose a sample of 24 cities which appeared in roughly ninety percent of the quarterly surveys.

For this study we select 48 goods and services (hereafter, commodities) with three criteria in mind. First, for each commodity we want wide coverage in terms of availability across cities and over time. Second, we want variation in the degree of tradability of the commodities included in the data set. Finally, we want homogeneity in the definitions of the commodities over time. The definitions of some commodities did change during the sample period, typically as a result of a change in manufacturer packaging. These changes had only small effects on *relative* prices.

For this study, we classify the 48 goods into tradables (23), perishables (15) and services (10). These categories were designed to facilitate discussion of how our results vary with the potential "tradability" of the commodities under consideration. The goods included in each category are described in the Appendix.

4.2 Construction of relative prices

Mindful of the problem of low power that afflicts many empirical analyses of the LOP, we conduct our analysis in a panel setting. We construct two different sets of relative price panels. The first set groups commodities by type. Thus for each of the 48 commodities in the data set, we construct a panel that contains all the price series that are available for that commodity across the country. The second set of panels groups commodities by location. In this case, the panels contain all the price series that are available for a given city. Within these two types of panels, we further subdivide the goods into the categories "traded," "perishables" and "services." Cutting the goods this way allows us to examine how the tradability of a good affects its price behavior.

The absolute prices that are included in each panel must be converted to relative prices in order to test the LOP. This requires choosing a numeraire for each panel. For some of the tests run, the choice of numeraire will prove immaterial, but for others it will make an important difference. Accordingly we choose the numeraire with a eye to its economic meaning. For the panels which group commodities by type, we choose the *average price across all cities* as the numeraire. So for example, each series in the panel of aspirin prices is calculated as

$$q_{\text{Aspirin},jt} = p_{\text{Aspirin},jt} - \frac{1}{M}\sum_{j=1}^{M} p_{\text{Aspirin},jt}, \tag{4.1}$$

where $p_{\text{Aspirin},jt}$ is the price of aspirin in city j at time t, and M is the total number of cities for which aspirin price series are available. One of the $q_{\text{Aspirin},j}$ series is redundant, and we arbitrarily choose this to be the series for Louisville, Kentucky in all of the panels. In addition, we demean each relative price constructed. For the panels which groups commodities by location, we choose prices in New Orleans as the numeraires. This is because the data for New Orleans are relatively complete. So for example each series in the Houston panel is constructed as

$$q_{i,\text{Houston},t} = p_{i,\text{Houston},t} - p_{i,\text{New Orleans},t}. \tag{4.2}$$

Here i indexes each of the commodities. There will be three different panels for Houston, one for traded goods, one for perishables, and one for services. Once again, we demean each relative price series constructed.

In all panels, we delete series that have fewer than 43 quarterly observations. This yields balanced panels, which simplifies the empirical analysis substantially.

4.3 Empirical tests

We carry out three types of tests on each panel of relative prices, a unit-root test, and two threshold autoregression tests. Each is described in turn.

A. Panel unit-root test

The first test is a panel unit root test that takes the form

$$\Delta q_{ijt} = \rho q_{ij,t-1} + \epsilon_{ijt}, t = 1, ..., T, \tag{4.3}$$

where the panel is selected by fixing either i or j constant. The null hypothesis is that q follows a random walk—$\rho = 0$—and the alternative is that $\rho < 0$. To counter the size biases that can arise owing to the cross-sectional dependence that is typically present in panels of relative prices, ρ is estimated by *GLS* (see O'Connell (1997b)). It is assumed that the data-generating process (DGP) for the disturbance terms is a mean-zero $VAR(p)$ process

$$\epsilon_t = \mathbf{\Phi}_1 \epsilon_{t-1} + \mathbf{\Phi}_2 \epsilon_{t-2} + \cdots + \mathbf{\Phi}_p \epsilon_{t-p} + \mathbf{u}_t, \tag{4.4}$$

where $\mathbf{u}_t \sim N(\mathbf{0}, \mathbf{\Sigma})$. Clearly, it would be impossible to estimate all the parameters of an unrestricted *VAR* such as this using only the short spans of price data that are available. For example, even if the $\{\mathbf{\Phi}_i\}_{i=1}^p$ are restricted to be diagonal, identification requires that the lag length p satisfies $Np \leq T - p - 1$.[16] As a result, we restrict the $\{\mathbf{\Phi}_i\}_{i=1}^p$ to be scalar multiples of the identity matrix, but place no restriction on the form of $\mathbf{\Sigma}$. An appealing property of this choice of restrictions is that it renders the test outcome invariant to the choice of numeraire (O'Connell, 1997b).[17]

$\mathbf{\Sigma}$ and the $\{\mathbf{\Phi}_i\}_{i=1}^p$ are estimated by maximum likelihood under the null hypothesis that $\rho = 0$.[18] A separate lag length is chosen for each panel using a data-dependent procedure. The process (4.4) is fitted by maximum likelihood to the first differences of the relative prices for values of p from 0 to 12. Three different criteria are then used to choose the best characterization of the data: the likelihood ratio test (*LLR*), the Akaike Information Criterion (AIC), and the Bayes-Schwartz Information Criterion (BIC). The three statistics often select the same DGP, though when they

[16] For p such that $Np > T - p$, the panel matrix of residuals will be of less than full row rank, precluding estimation of $\mathbf{\Omega}$.

[17] Papell (1997b) has recently suggested that this type of restriction is unwarranted in panel tests of the LOP and PPP. His preferred specification allows for heterogeneous serial correlation in each relative price series. However he does not address the problem of identification in short spans of data. He fits separate univariate $AR(p)$ models to each series, rather than estimating them in a multivariate setting. This is inconsistent with full identification of the multivariate DGP for the disturbance terms when the time span of the data is short (see previous footnote). It would, of course, be possible to allow for much richer forms of serial correlation if parsimony was instead achieved through restrictions on $\mathbf{\Sigma}$. However O'Connell (1997b) has shown that serious size biases can arise if $\mathbf{\Sigma}$ is restricted in such a way as to prevent *GLS* from controlling for the cross-sectional dependence that is normally manifest in relative prices. This is not to dispute Papell's (1997b) assertion that it can be important to allow for heterogeneous serial correlation, but rather to point out that it is often infeasible to do so in a consistent way, given the short spans of data typically available.

[18] Even for panels of moderate size, estimation of these parameters under the alternative would be computationally expensive. The presence of lagged endogenous variables precludes the use of iterative least squares techniques, and therefore the likelihood function would have to be maximized numerically.

differ, the LLR tends to favor less parsimony while the BIC favors more. For each panel, we choose the "middle" estimate of p from these three tests as the best lag length to describe the DGP. These estimates are reported for each panel in Tables 1–6.

Having estimated the serial correlation and cross-sectional dependence properties of the disturbance vector under the null, the test statistic is then formed in the usual fashion. That is to say, the matrices of first differences $\mathbf{Y}$ and lagged levels $\mathbf{X}$ of q are each transformed by the estimated VAR lag polynomial $\tilde{\mathbf{\Phi}}(L)$ to yield $\mathbf{Y}^*$ and $\mathbf{X}^*$, and the $FGLS$ estimate of ρ is calculated as

$$\hat{\rho}_{FGLS} = \mathrm{tr}(\mathbf{Y}^{*\prime}\mathbf{Y}^*\tilde{\mathbf{\Omega}}^{-1})/\mathrm{tr}(\mathbf{X}^{*\prime}\mathbf{X}^*\tilde{\mathbf{\Omega}}^{-1}), \tag{4.5}$$

Critical values for this test statistic are tabulated by parametric bootstrap. This involves drawing bootstrap samples of real exchange rate innovations from the fitted VAR processes (4.4), and simulating the distribution of $\hat{\rho}_{\mathrm{FGLS}}$ estimated from these samples.

B. AR, EQ-TAR and BAND-TAR tests

Having tested for stationarity, we examine whether the behavior of relative prices can be characterized by a band of no-arbitrage for small deviations from the LOP. We do this by fitting threshold autoregressive (TAR) processes to the panels of relative prices. The TAR processes considered have the following general form

$$\Delta q_t = \begin{cases} \rho_1(q_{t-1} - b) + \epsilon_t & \text{if } a < q_{t-1} \\ \rho_0 q_{t-1} + \epsilon_t & \text{if } -a \leq q_{t-1} \leq a \\ \rho_1(q_{t-1} + b) + \epsilon_t & \text{if } q_{t-1} < -a \end{cases} \tag{4.6}$$

where $\rho_0 \leq 0$, $\rho_1 < 0$, $0 < b < a$ and $\epsilon \sim N(0, \sigma_\epsilon^2)$. Thus q_t may revert to 0 whenever $|q_{t-1}| \leq a$. When $|q_{t-1}| > 0$, q_t reverts to the edge of a band around 0 defined by the range $[-b, b]$.

We consider two important special cases of this general form. The first, known as an "equilibrium TAR" (EQ-TAR) specification, constrains b to be 0. This is the model estimated in O'Connell (1997a). If $b = 0$, then deviations from the LOP revert towards parity (zero) whenever they exceed a in absolute value. The second nested model, known as a "band TAR" ($BAND$-TAR) model, constrains b to equal a. This is the form adopted by Obstfeld and Taylor (1997). With $b = a$, reversion takes place towards the threshold a. We do not restrict ρ_0 to be zero, as Obstfeld and Taylor (1997) do in their estimation. Finally, for comparison purposes, we also estimate the model subject to the constraints $\rho_0 = \rho_1$ *and* $b = a = 0$. In this case the TAR reduces to an $AR(1)$. In this guise the specification is similar to the panel unit root test, though as discussed below the treatment of serial and contemporaneous correlation will differ from that in the GLS panel unit root test.

Estimation of these TAR specifications serves two purposes. First, if transport costs create a band of no-arbitrage in relative price series, TAR models provide a more powerful way to detect global stationarity of deviations from the LOP than the standard test run earlier. *This is true even if the true price behavior does not conform to the TAR specification.* For example, it might be that the tendency of q to revert to parity is a continuous increasing function of the distance from parity. A process that gives rise to this behavior is[19]

$$q_t = \gamma_1 q_{t-1} + \gamma_2 |q_{t-1}| q_{t-1} + \epsilon_t. \tag{4.7}$$

Even if this is the true model, the TAR specification will provide a more powerful test of stationarity than the Dickey-Fuller test. The second purpose served by the TAR specifications is that, in view of the model presented earlier, they might in fact offer a good characterization of the true process for q_t. In particular, if fixed costs are an important part of impediments to arbitrage, we might expect to observe EQ-TAR behavior. On the other hand, if variable costs are predominant, then $BAND$-TAR behavior is more likely to be observed. By estimating both types of model, we are able to shed some light on which cost is more important.

The TAR specifications are estimated by maximum likelihood. For a given a, say $\bar{a}$ we split the observations into two subsamples, those for which $|q_{t-1}| \leq \bar{a}$, and those for which $|q_{t-1}| > \bar{a}$. ρ_0 and ρ_1 are estimated by least squares within each subsample. The maximized log-likelihood for the model is then (up to a constant)

$$\ln \mathcal{L}(\rho_0, \sigma_0^2, \rho_1, \sigma_1^2, \bar{a}) = -N_0 \ln(\sigma_0^2) - N_1 \ln(\sigma_1^2), \tag{4.8}$$

where σ_0^2 is the residual variance from the subsample associated with ρ_0, σ_1^2 is the residual variance from the subsample associated with ρ_1, and N_0 and N_1 are the number of observations in each subsample. The value of $\bar{a}$ that maximizes this function can then be found by grid search.

[19]Similar functional forms have been assumed in Wei and Parsley (1995), Parsley and Wei (1996), Michael, Nobay and Peel (1997) and O'Connell (1997a).

One issue that arises is how to control for the serial and contemporaneous correlation that is present in the data. If a *GLS* procedure is used, then it complicates the classification of observations into the two subsamples. For this reason we estimate ρ_0 and ρ_1 in each subsample by *OLS*, and instead seek to control for the serial and contemporaneous correlation in our critical values.

To simplify the presentation, we focus on just two test statistics: (a) $LLR_{EQ} \equiv 2(\ln \mathcal{L}_{\mathcal{EQ}} - \ln \mathcal{L}_{\mathcal{AR}})$, the likelihood ratio for the *EQ-TAR* model relative to the $AR(1)$ model; and (b) $LLR_{BA} \equiv 2(\ln \mathcal{L}_{\mathcal{BA}} - \ln \mathcal{L}_{\mathcal{AR}})$, the likelihood ratio for the *BAND-TAR* model relative to the $AR(1)$ model. These likelihood ratios cannot be used to perform a likelihood ratio test for the respective models in the usual fashion, as the thresholds are not identified under the $AR(1)$ null. Instead, following Obstfeld and Taylor (1997), we derive appropriate critical values for the test statistics by Monte Carlo simulation. Specifically, the distribution of LLR_{EQ} and LLR_{BA} is tabulated under the null that the true process for real exchange rates is $AR(1)$, taking the OLS estimate of the first-order autoregressive coefficient to be the true root. As in the case of the *GLS* panel unit root test, the disturbance innovations used to simulate the $AR(1)$ real exchange rates are derived from the estimated DGP (4.4).

5 Empirical results

The empirical results are presented in six tables. Tables 1 and 2 look at "tradables," Tables 3 and 4 at "nontradables," and Tables 5 and 6 at "services."[20]

5.1 Tradable goods

Column 4 of Table 1 presents the test statistic for the *GLS* panel unit root test as applied to panels of tradable goods grouped by commodity. The null that deviations from the LOP follow a random walk can only be rejected for 7 of 23 commodities at the 10 percent significance level. The same picture emerges from Column 4 of Table 2, which groups commodities by location: there, only 5 of 20 cities appear to have traded goods prices that exhibit reversion to traded goods prices in New Orleans. These findings are surprising: we might have expected that the ostensible ease with which

[20]Note that estimates of ρ_0 are not shown to conserve space.

these goods can be transported would have led to much stronger evidence against the random walk null.

Could it be that this apparent nonstationarity is due to the presence of transport costs? The answer appears to be, Yes. Fully 17 of the 23 commodities in Table 1 reject the AR(1) model in favor of *both* the *EQ-TAR* and the *BAND-TAR* model, typically at better than the 1% level. Only two commodities, beer and toothpaste, provide no support for either of the nonlinear models. In Table 2, 18 of the 20 cities in Table 2 provide support for the *BAND-TAR* specification. From these results we might conclude that small deviations from the LOP for traded goods do not tend to revert to 0, but that large ones do.

A second question worth asking is whether there is more support for the *EQ-TAR* or the *BAND-TAR* specification. Unfortunately, there does not appear to be enough statistical power in the data to distinguish between the models on the basis of a likelihood ratio test. It does appear, however, that the maximized value of the likelihood function is higher for the *EQ-TAR* model than it is for the *BAND-TAR* model. This is prima facie evidence that the *EQ* specification provides a better characterization of the data. [21]

5.2 Perishable goods

Tables 3 and 4 give corresponding results for perishable goods. From Column 4 of Table 3, we see that 8 of the 15 perishable commodities provide evidence against the random walk. From Column 4 of Table 4, the perishable goods in 9 of 20 cities appear to be stationary with respect to New Orleans prices. If anything, these numbers may be higher than our priors might suggest, since the defining characteristic of these goods is that they can be expensive to transport.

With perishable goods, there is resounding evidence of *TAR* behavior. 13 of the 15 goods reject the $AR(1)$ model in favor of *both* the *TAR* alternatives at better than the 1% level. McDonalds™ and Kentucky Fried Chicken™ (labeled "Fr. Chicken") are the exceptions.[22] For 12 of these 13

[21] However, it is also true that the p-values for LLR_{BA} are generally lower than those for LLR_{EQ}. We thank Alan Taylor for pointing this out.

[22] It is interesting that neither McDonalds™ nor Kentucky Fried Chicken™ (labeled "Fr. Chicken") prices reveal *any* evidence of stationarity. Both panels fail to reject the unit-root null with the *GLS* test, and neither shows evidence of *TAR* behavior. This can be contrasted with Cumby's (1996) finding of mean reversion in the relative international prices of McDonalds™ products. It also suggests that the price behavior of nationally branded goods may differ substantially from that of more generic products. We aim to explore this conjecture in future work.

commodites, the maximized likelihood for the *EQ-TAR* model exceeds the maximized likelihood for the *BAND-TAR* model, sometimes by a substantial margin. This lends credence to the *EQ-TAR* model as a description of perishable goods price behavior.

Turning to Table 4, the evidence of *TAR* behavior in perishables grouped by location is quite persuasive. All cities except Salt Lake City UT and Appleton WI evince *TAR* behavior. Once again, the maximized likelihoods favor the *EQ* specification.

5.3 Services

Only for 3 of the 10 services in Table 5 can we reject the unit root null. The corresponding proportion for Table 6, which groups the services by location, is 7 of 20 cities. These numbers are in line with expectations, since it is difficult to arbitrage the prices of these products.[23] Consistent with this prior, the evidence of *TAR* behavior in services is more equivocal than for either tradables or perishables. Only 6 of the 10 services panels in Table 5 support one of the *TAR* models over the $AR(1)$ specification, and only 4 services support both *TAR* specifications over $AR(1)$. 14 of the services panels grouped by location in Table 6 provide evidence of *TAR* effects. Overall, the results leave one with the impression that nonlinear reversion is an important ingredient of the behavior of some, but not all, services. Where services support both types of *TAR* behavior, it is the *EQ-TAR* specification that produces the highest maximized likelihood in nearly all instances. However, the excess margin over LLR_{BA} is on average lower than is true for perishables and tradables.

6 Conclusion

In this paper we have sought to develop the nascent literature on nonlinear commodity price behavior along two dimensions. First, we set out a simple continuous-time framework to gauge the importance of transport costs for relative price behavior. A useful feature of this framework is its ability to distinguish between the impacts of fixed and variable transport costs. In particular, the model shows that in the presence of fixed and proportional transport costs, we can expect to observe "band reversion" rather than "mean reversion" in relative goods prices. That is, when

[23]Services would normally be called nontradables in an international context, for obvious reasons. Within a country like the U.S., we might suspect that factor mobility contributes to price convergence.

deviations from the LOP become large, goods arbitrage will cause these deviations to narrow, but will not eliminate them completely.

Second, we employed a detailed data set on U.S. goods prices to canvas the pattern of reversion exhibited by deviations from the LOP. The data, measured in 24 cities over the period 1975:1–1992:4 afford a "purer" measure of deviations from price parity than is possible with aggregate price indices. Three apparent facts emerged from this analysis. First, using standard tests, the evidence that deviations from the LOP are stationary is weak at best. Second, when we look at only large deviations from the LOP, the evidence of stationarity is very strong. The implication is that small deviations from the LOP tend to be extremely persistent; it is only large deviations that revert towards equilibrium. And third, the weight of evidence is in favor of reversion to the equilibrium of zero—mean-reversion—rather than reversion to a band of no-arbitrage—band reversion. Thus when adjustments in relative prices do take place, they tend to eliminate, rather than reduce, price discrepancies.

There are a number of questions which warrant further attention, three of which are immediate. First, What is the rate of reversion exhibited by LOP deviations at various distances from zero.[24] Second, Is there any relationship between the estimated thresholds and candidate explanatory variables such as distance or price volatility, as suggested by Parsley and Wei (1996) and Taylor and Obstfeld (1997). And third, Is it possible to impose more structure on the data to facilitate direct testing of the model set out in Section 3? We aim to provide answers to these and other questions in future research.

[24]Owing to the serial correlation that is present in our data, the estimates of ρ_1 in Columns 7 and 11 in the tables are *biased* estimates of the rate of reversion, and so alternative estimates must be developed.

References

Abuaf, N. and P. Jorion (1990). "Purchasing power parity in the long run," *Journal of Finance* 45: 157–174.

Benninga, S. and A. Protopapadakis (1988). "The equilibrium pricing of exchange rates and assets when trade takes time," *Journal of International Money and Finance* 7: 129–149.

Coleman, A. M. G. (1995). "Arbitrage, storage, and the 'Law of One Price': new theory for the time series analysis of an old problem," unpublished (September), Princeton University.

Constantinides, G. M. and S. F. Richard (1978). "Existence of optimal simple policies for discounted-cost inventory and cash management in continuous time," *Operations Research* 26: 620–636.

Cumby, R. E. (1996). "Forecasting exchange rates and relative prices with the hamburger standard: Is what you see what you get with McParity," unpublished (May), Georgetown University.

Dixit, A. K. (1993). "The art of smooth pasting," in J. Lesourne and H. Sonnenschein (eds.) *Fundamentals in Pure and Applied Economics*, Volume 55. Chur Switzerland: Harwood Academic Publishers.

Dumas, B. (1992). "Dynamic equilibrium and the real exchange rate in a spatially separated world." *Review of Financial Studies* 5 (2): 153–180.

Engel, C., M. Hendrickson and J. H. Rogers (1996). "Intra-national, intra-continental and intra-planetary PPP," prepared for the TCER-NBER-CEPR conference on "The Exchange Rate and Price Movements: Theory and Evidence," Tokyo, December 20–21, 1996.

Frankel, J. and A. Rose (1996). "A panel project on purchasing power parity: Mean reversion within and between countries," *Journal of International Economics* 40: 209–225.

Froot K. and K. Rogoff (1995). "Perspectives on PPP and long-run real exchange rates," in: G. Grossman and K. Rogoff eds., *Handbook of International Economics.* Amsterdam: North-Holland.

Harrison, J. M., T. M. Sellke and A. J. Taylor (1983). "Impulse control of Brownian motion," *Mathematics of Operations Research* 8: 454–466.

Jorion, P. and R. Sweeney (1996). "Mean reversion in real exchange rates: Evidence and implications for forecasting," *Journal of International Money and Finance*, forthcoming.

Michael, P., A. R. Nobay and D. A. Peel (1996). "Transactions costs and nonlinear adjustment in real exchange rates: An empirical investigation," *Journal of Political Economy*, forthcoming.

Obstfeld, M. and A. M. Taylor (1997). "Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited," prepared for the NBER-TCER-CEPR conference on "Purchasing power parity revisited: The exchange rate and price movements, theory and evidence," Tokyo, December 20–21, 1996.

O'Connell, P. G. J. (1997a). "Market frictions and relative traded goods prices," prepared for the JIMF-LIFE Workshop on "Developments in exchange rate modelling," Maastricht, The Netherlands, April 4–5, 1997.

O'Connell, P. G. J. (1997b). "The overvaluation of purchasing power parity," *Journal of International Economics*, forthcoming.

Papell, D. H. (1996). "Searching for stationarity: Purchasing power parity under the current float," *Journal of International Economics*, forthcoming.

Papell, D. H. and H. Theodoridis (1997a). "Increasing evidence of purchasing power parity over the current float," unpublished (March), University of Houston.

Papell, D. H. and H. Theodoridis (1997b). "The choice of numeraire currency in panel tests of purchasing power parity," unpublished (April), University of Houston.

Parsley, D. and S. Wei (1996). "Convergence to the law of one price without trade barriers or currency fluctuations," *Quarterly Journal of Economics* 108: 1211–1236.

Rogoff, K. (1996). "The purchasing power parity puzzle," *Journal of Economic Literature* 34: 647–668.

Sercu, P, R. Uppal and C. Van Hulle (1995). "The exchange rate in the presence of transactions costs: Implications for tests of purchasing power parity." *Journal of Finance* 50: 1309–1319.

Taylor, A. M. (1996). "International capital mobility in history: Purchasing power parity in the long run," prepared for the NBER conferences on "The Determination of Exchange Rates," Cambridge, MA, May 10–11, 1996.

Uppal, R. (1993). "A general equilibrium model of international portfolio choice," *Journal of Finance* 48: 529–553.

Wei, S. and D. Parsley (1995). "Purchasing power *dis*-parity during the floating rate period: Exchange rate volatility, trade barriers, and other culprits,' NBER Working Paper No. 5032.

Williams, J. C. and B. D. Wright (1991). *Storage and commodity markets.* Cambridge, MA: Cambridge University Press.

Table 1
Results for traded goods grouped by commodity type, 1975–1992:4

	Full-panel $AR(1)$					EQ-TAR model				$BAND$-TAR model			
Good	p	N	T	$t_{\rho_{GLS}}$	$\hat{\rho}_{OLS}$	a	$\hat{\rho}_1$	N_1	LLR_{EQ}	a	$\hat{\rho}_1$	N_1	LLR_{BA}
Aspirin	10	24	43	-5.21	-0.35	0.20	-0.42	79	67.81	0.18	-1.11	97	57.91
				(0.58)	(0.52)				(0.16)				(0.00)
Babyfood	6	23	72	-4.61	-0.27	0.13	-0.36	171	310.38	0.13	-1.01	171	320.59
				(0.89)	(0.05)				(0.00)				(0.00)
Beer	12	24	43	-4.23	-0.40	0.10	-0.44	114	311.17	0.10	-0.79	114	289.36
				(0.59)	(0.37)				(0.84)				(0.39)
Cigarettes	5	23	72	-7.39	-0.18	0.03	-0.18	707	213.43	0.03	-0.35	707	241.96
				(0.02)	(0.09)				(0.00)				(0.00)
Coffee	8	22	72	-6.76	-0.42	0.15	-0.49	108	121.05	0.10	-0.95	295	73.29
				(0.06)	(0.05)				(0.00)				(0.00)
Cornflakes	12	21	55	-5.48	-0.53	0.12	-0.68	77	89.03	0.12	-1.93	77	63.61
				(0.28)	(0.23)				(0.02)				(0.00)
Game	8	24	43	-6.24	-0.42	0.15	-0.45	101	110.37	0.13	-1.05	142	112.95
				(0.11)	(0.00)				(0.00)				(0.00)
Jeans	12	24	43	-0.45	-0.58	0.09	-0.60	277	125.89	0.09	-1.15	277	74.78
				(1.00)	(0.77)				(0.00)				(0.00)
Liquor	6	22	72	-7.04	-0.25	0.07	-0.28	291	315.45	0.08	-0.58	256	290.87
				(0.05)	(0.00)				(0.00)				(0.00)
Shirt	12	23	43	-0.81	-0.60	0.10	-0.65	270	70.24	0.09	-1.23	314	37.49
				(1.00)	(0.16)				(0.00)				(0.00)
Orange Juice	7	20	72	-6.70	-0.48	0.15	-0.52	148	136.32	0.13	-1.17	205	110.55
				(0.05)	(0.00)				(0.00)				(0.00)
Peaches	12	23	72	-7.02	-0.50	0.10	-0.59	218	377.95	0.10	-1.18	218	352.62
				(0.02)	(0.00)				(0.00)				(0.00)
Shampoo	9	23	43	-2.07	-0.63	0.07	-0.66	330	211.20	0.07	-1.05	330	169.94
				(1.00)	(1.00)				(0.00)				(0.00)
Shortening	12	23	72	-6.42	-0.40	0.11	-0.44	227	164.73	0.10	-1.01	301	147.77
				(0.11)	(0.06)				(0.00)				(0.00)
Soda	9	22	72	-6.15	-0.56	0.21	-0.63	181	154.50	0.20	-1.52	213	117.61
				(0.21)	(0.08)				(0.00)				(0.00)
Sugar	7	20	55	-6.98	-0.45	0.16	-0.45	57	117.72	0.16	-0.88	57	89.79
				(0.02)	(0.10)				(0.45)				(0.01)
Tennis	12	23	43	0.34	-0.46	0.19	-0.50	67	87.13	0.19	-1.58	67	86.80
				(1.00)	(0.86)				(0.37)				(0.01)
Tissue	8	22	72	-8.75	-0.43	0.13	-0.46	263	686.90	0.13	-0.80	263	671.81
				(0.00)	(0.00)				(0.00)				(0.00)
Toothpaste	11	23	43	-1.00	-0.48	0.18	-0.55	49	186.48	0.18	-1.52	49	174.83
				(1.00)	(0.57)				(0.69)				(0.14)
Tuna	11	23	43	0.02	-0.63	0.11	-0.62	213	128.04	0.11	-1.40	213	125.49
				(1.00)	(1.00)				(0.00)				(0.00)
Underwear	12	23	43	-4.90	-0.47	0.20	-0.53	60	66.55	0.20	-1.89	60	61.21
				(0.44)	(0.18)				(0.14)				(0.00)
Detergent	10	21	72	-5.83	-0.42	0.05	-0.41	767	120.02	0.17	-1.55	97	131.09
				(0.22)	(0.02)				(0.00)				(0.00)
Wine	12	23	43	-1.61	-0.60	0.12	-0.65	199	181.21	0.12	-1.71	199	188.40
				(1.00)	(0.95)				(0.00)				(0.00)

Results from estimation of $AR(1)$, EQ-TAR and $BAND$-TAR panel models where relative prices are grouped by commodity type. p is the order of serial correlation selected for the disturbance DGP, N is the number of data series in each panel, T is the number of quarterly observations in each panel, $t_{\rho_{GLS}}$ is the test statistic for the GLS panel unit root test, $\hat{\rho}_{OLS}$ is the OLS estimate of ρ for the full panel, and for each of the TAR models, a is the estimated threshold, $\hat{\rho}_1$ is the OLS estimate of ρ_1 outside the threshold, N_1 is the number of observations outside the threshold, and $LLR \equiv 2\ln(\mathcal{L}/\mathcal{L}_{AR})$ is the likelihood ratio statistic for that model against the $AR(1)$ alternative. p-values derived by parametric bootstrap for each test statistic appear in parentheses.

Table 2
Results for traded goods grouped by location, 1975–1992:4

	Full-panel $AR(1)$					*EQ-TAR* model				*BAND-TAR* model			
Location	p	N	T	$t_{\rho_{GLS}}$	$\hat{\rho}_{OLS}$	a	$\hat{\rho}_1$	N_1	LLR_{EQ}	a	$\hat{\rho}_1$	N_1	LLR_{BA}
Mobile, AL	10	8	72	-4.10	-0.56	0.17	-0.60	73	289.48	0.17	-1.08	73	266.29
				(0.18)	(0.00)				(0.12)				(0.01)
Blythe, CA	6	7	72	-3.70	-0.42	0.23	-0.68	26	84.96	0.11	-1.00	157	86.71
				(0.31)	(0.14)				(0.32)				(0.00)
Indio, CA	8	7	72	-3.32	-0.37	0.21	-0.39	75	125.68	0.20	-0.79	81	122.12
				(0.46)	(0.25)				(0.20)				(0.04)
Denver, CO	8	8	72	-4.45	-0.33	0.21	-0.39	65	85.55	0.23	-0.77	59	58.78
				(0.10)	(0.18)				(0.02)				(0.01)
Indianapolis, IN	9	8	72	-4.18	-0.53	0.26	-0.74	29	126.86	0.25	-2.36	30	116.77
				(0.15)	(0.00)				(0.84)				(0.23)
Cedar Rap., IA	10	7	72	-4.61	-0.54	0.15	-0.58	76	178.99	0.15	-1.00	76	161.58
				(0.04)	(0.01)				(0.06)				(0.01)
Lexington, KY	9	8	72	-4.11	-0.42	0.07	-0.43	274	128.77	0.07	-0.64	274	132.75
				(0.18)	(0.04)				(0.00)				(0.00)
Louisville, KY	5	8	72	-5.13	-0.49	0.05	-0.50	305	77.72	0.14	-1.11	136	79.23
				(0.02)	(0.01)				(0.00)				(0.00)
St. Louis, MO	8	8	72	-4.19	-0.42	0.17	-0.53	100	167.99	0.10	-0.85	224	163.61
				(0.18)	(0.00)				(0.01)				(0.00)
Hastings, NE	9	8	72	-4.06	-0.43	0.22	-0.55	48	109.42	0.22	-1.63	48	109.50
				(0.20)	(0.01)				(0.13)				(0.01)
Omaha, NE	6	7	72	-5.41	-0.55	0.16	-0.57	84	119.05	0.16	-1.07	84	103.06
				(0.00)	(0.00)				(0.13)				(0.02)
Rapid City, SD	9	8	72	-3.61	-0.36	0.11	-0.37	193	123.70	0.11	-0.57	193	122.26
				(0.40)	(0.53)				(0.00)				(0.00)
Vermillion, SD	7	8	72	-3.74	-0.54	0.16	-0.61	74	80.63	0.16	-1.28	74	58.95
				(0.38)	(0.01)				(0.11)				(0.01)
Chattanooga, TN	6	8	72	-4.05	-0.44	0.19	-0.52	82	139.75	0.06	-0.66	327	127.55
				(0.23)	(0.00)				(0.10)				(0.00)
El Paso, TX	8	7	72	-4.03	-0.27	0.17	-0.32	69	128.03	0.17	-0.57	69	122.70
				(0.15)	(0.45)				(0.04)				(0.01)
Houston, TX	7	8	72	-4.58	-0.37	0.16	-0.40	110	146.79	0.16	-0.64	110	116.81
				(0.09)	(0.47)				(0.13)				(0.02)
Lubbox, TX	9	7	72	-3.97	-0.37	0.27	-0.40	26	132.51	0.27	-0.85	26	124.58
				(0.16)	(0.14)				(0.74)				(0.27)
S. Lake City, UT	8	8	72	-4.38	-0.33	0.09	-0.34	231	186.22	0.09	-0.54	231	195.80
				(0.12)	(0.17)				(0.00)				(0.00)
Appleton, WI	6	8	72	-3.23	-0.38	0.15	-0.44	105	120.88	0.15	-0.97	105	127.66
				(0.62)	(0.01)				(0.00)				(0.00)
Casper, WY	9	8	72	-4.03	-0.40	0.18	-0.49	70	116.40	0.10	-0.75	180	125.22
				(0.20)	(0.32)				(0.06)				(0.00)

Results from estimation of $AR(1)$, *EQ-TAR* and *BAND-TAR* panel models where relative prices are grouped by location. p is the order of serial correlation selected for the disturbance DGP, N is the number of data series in each panel, T is the number of quarterly observations in each panel, $t_{\rho_{GLS}}$ is the test statistic for the *GLS* panel unit root test, $\hat{\rho}_{OLS}$ is the *OLS* estimate of ρ for the full panel, and for each of the *TAR* models, a is the estimated threshold, $\hat{\rho}_1$ is the *OLS* estimate of ρ_1 outside the threshold, N_1 is the number of observations outside the threshold, and $LLR \equiv 2\ln(\mathcal{L}/\mathcal{L}_{AR})$ is the likelihood ratio statistic for that model against the $AR(1)$ alternative. p-values derived by parametric bootstrap for each test statistic appear in parentheses.

Table 3
Results for perishable goods grouped by commodity type, 1975–1992:4

	Full-panel $AR(1)$					$EQ\text{-}TAR$ model				$BAND\text{-}TAR$ model			
Good	p	N	T	$t_{\rho_{GLS}}$	$\hat{\rho}_{OLS}$	a	$\hat{\rho}_1$	N_1	LLR_{EQ}	a	$\hat{\rho}_1$	N_1	LLR_{BA}
Bacon	8	24	67	-8.35	-0.71	0.19	-0.75	160	82.99	0.04	-0.89	1217	6.08
				(0.00)	(0.00)				(0.00)				(0.00)
Bananas	11	23	72	-8.21	-0.70	0.22	-0.72	114	24.05	0.02	-0.78	1430	12.69
				(0.00)	(0.01)				(0.00)				(0.00)
Bread	10	22	72	-6.04	-0.52	0.29	-0.64	109	76.64	0.09	-0.85	799	44.62
				(0.24)	(0.06)				(0.00)				(0.00)
Cheese	12	21	43	-4.16	-0.38	0.04	-0.37	304	58.30	0.04	-0.64	304	32.71
				(0.70)	(0.93)				(0.00)				(0.00)
Eggs	11	23	72	-4.61	-0.32	0.18	-0.24	183	186.90	0.18	-0.51	183	167.64
				(0.83)	(0.98)				(0.00)				(0.00)
Minced steak	9	23	72	-6.79	-0.58	0.11	-0.59	401	127.69	0.17	-1.58	156	24.91
				(0.07)	(0.04)				(0.00)				(0.00)
Lettuce	12	23	72	-7.47	-0.80	0.22	-0.79	261	57.96	0.01	-0.84	1556	0.79
				(0.01)	(0.12)				(0.00)				(0.00)
Margarine	9	23	72	-6.63	-0.51	0.22	-0.57	128	101.16	0.25	-1.59	83	59.14
				(0.09)	(0.00)				(0.00)				(0.00)
Milk	7	23	72	-5.52	-0.34	0.06	-0.36	400	337.79	0.06	-0.67	400	338.10
				(0.53)	(0.00)				(0.00)				(0.00)
Potatoes	11	23	72	-6.75	-0.73	0.30	-0.74	90	29.91	0.01	-0.77	1553	6.81
				(0.07)	(0.04)				(0.00)				(0.00)
Steak	11	22	72	-7.51	-0.52	0.11	-0.53	409	141.65	0.11	-1.11	409	98.63
				(0.00)	(0.26)				(0.00)				(0.00)
Chicken	8	23	72	-7.69	-0.61	0.13	-0.65	369	42.66	0.02	-0.74	1354	37.81
				(0.01)	(0.04)				(0.00)				(0.00)
Fr. chicken	11	23	43	-4.59	-0.49	0.14	-0.58	111	350.16	0.14	-1.34	111	343.19
				(0.40)	(0.02)				(0.99)				(0.90)
McDonalds	12	23	43	-0.33	-0.59	0.02	-0.61	519	338.34	0.02	-0.82	519	373.68
				(1.00)	(0.64)				(0.32)				(0.24)
Pizza	12	23	43	-0.93	-0.32	0.06	-0.34	283	308.43	0.06	-0.56	283	302.28
				(1.00)	(0.12)				(0.00)				(0.00)

Results from estimation of $AR(1)$, $EQ\text{-}TAR$ and $BAND\text{-}TAR$ panel models where relative prices are grouped by commodity type. p is the order of serial correlation selected for the disturbance DGP, N is the number of data series in each panel, T is the number of quarterly observations in each panel, $t_{\rho_{GLS}}$ is the test statistic for the GLS panel unit root test, $\hat{\rho}_{OLS}$ is the OLS estimate of ρ for the full panel, and for each of the TAR models, a is the estimated threshold, $\hat{\rho}_1$ is the OLS estimate of ρ_1 outside the threshold, N_1 is the number of observations outside the threshold, and $LLR \equiv 2\ln(\mathcal{L}/\mathcal{L}_{AR})$ is the likelihood ratio statistic for that model against the $AR(1)$ alternative. p-values derived by parametric bootstrap for each test statistic appear in parentheses.

Table 4
Results for perishable goods grouped by location, 1975–1992:4

	Full-panel $AR(1)$					$EQ\text{-}TAR$ model				$BAND\text{-}TAR$ model			
Location	p	N	T	$t_{\rho_{GLS}}$	$\hat{\rho}_{OLS}$	a	$\hat{\rho}_1$	N_1	LLR_{EQ}	a	$\hat{\rho}_1$	N_1	LLR_{BA}
Mobile, AL	9	8	72	-4.69	-0.82	0.09	-0.82	273	66.23	0.03	-0.96	469	22.77
				(0.06)	(0.00)				(0.00)				(0.00)
Blythe, CA	7	8	72	-4.01	-0.50	0.05	-0.51	431	56.12	0.05	-0.62	431	52.82
				(0.24)	(0.05)				(0.00)				(0.00)
Indio, CA	7	8	72	-3.82	-0.48	0.16	-0.48	212	29.78	0.16	-0.91	212	26.84
				(0.33)	(0.08)				(0.00)				(0.00)
Denver, CO	4	8	72	-5.23	-0.66	0.15	-0.70	181	73.41	0.07	-0.94	357	51.79
				(0.02)	(0.00)				(0.00)				(0.00)
Indianapolis, IN	9	8	72	-4.88	-0.72	0.10	-0.73	270	22.06	0.01	-0.76	542	2.69
				(0.04)	(0.01)				(0.00)				(0.00)
Cedar Rap., IA	11	8	72	-2.64	-0.77	0.04	-0.78	416	57.56	0.04	-0.98	416	73.77
				(0.81)	(0.07)				(0.00)				(0.00)
Lexington, KY	9	8	72	-3.65	-0.67	0.27	-0.61	40	31.52	0.27	-1.76	40	13.02
				(0.39)	(0.03)				(0.30)				(0.02)
Louisville, KY	11	8	72	-3.52	-0.68	0.08	-0.68	300	76.45	0.08	-1.03	300	85.69
				(0.41)	(0.09)				(0.00)				(0.00)
St. Louis, MO	11	8	72	-4.02	-0.59	0.21	-0.66	103	33.54	0.05	-0.74	420	16.55
				(0.21)	(0.00)				(0.00)				(0.00)
Hastings, NE	11	8	72	-3.88	-0.84	0.06	-0.84	335	34.18	0.00	-0.86	543	10.72
				(0.27)	(0.01)				(0.00)				(0.00)
Omaha, NE	10	8	72	-4.39	-0.79	0.18	-0.87	127	51.03	0.10	-1.36	258	70.30
				(0.09)	(0.00)				(0.00)				(0.00)
Rapid City, SD	9	8	72	-4.50	-0.77	0.14	-0.78	222	67.13	0.05	-0.94	405	35.94
				(0.09)	(0.01)				(0.00)				(0.00)
Vermillion, SD	11	8	72	-3.40	-0.70	0.16	-0.76	141	56.03	0.07	-0.98	350	27.26
				(0.48)	(0.01)				(0.00)				(0.00)
Chattanooga, TN	10	8	72	-5.25	-0.67	0.11	-0.67	246	54.53	0.11	-1.09	246	27.74
				(0.01)	(0.00)				(0.00)				(0.00)
El Paso, TX	11	8	72	-4.54	-0.59	0.11	-0.60	236	60.04	0.11	-0.95	236	50.63
				(0.06)	(0.22)				(0.00)				(0.00)
Houston, TX	7	8	72	-4.62	-0.64	0.09	-0.64	289	43.64	0.08	-0.93	316	19.57
				(0.07)	(0.00)				(0.00)				(0.00)
Lubbox, TX	9	8	72	-4.12	-0.69	0.15	-0.69	174	91.76	0.15	-1.31	174	86.08
				(0.19)	(0.00)				(0.00)				(0.00)
S. Lake City, UT	7	8	72	-4.95	-0.59	0.25	-0.64	88	115.84	0.13	-0.93	217	96.41
				(0.04)	(0.01)				(0.22)				(0.00)
Appleton, WI	7	8	72	-4.28	-0.75	0.24	-0.80	67	60.32	0.30	-2.27	37	33.74
				(0.16)	(0.00)				(0.52)				(0.16)
Casper, WY	11	8	72	-4.20	-0.76	0.16	-0.81	181	56.46	0.06	-1.03	371	41.06
				(0.14)	(0.00)				(0.00)				(0.00)

Results from estimation of $AR(1)$, $EQ\text{-}TAR$ and $BAND\text{-}TAR$ panel models where relative prices are grouped by location. p is the order of serial correlation selected for the disturbance DGP, N is the number of data series in each panel, T is the number of quarterly observations in each panel, $t_{\rho_{GLS}}$ is the test statistic for the GLS panel unit root test, $\hat{\rho}_{OLS}$ is the OLS estimate of ρ for the full panel, and for each of the TAR models, a is the estimated threshold, $\hat{\rho}_1$ is the OLS estimate of ρ_1 outside the threshold, N_1 is the number of observations outside the threshold, and $LLR \equiv 2\ln(\mathcal{L}/\mathcal{L}_{AR})$ is the likelihood ratio statistic for that model against the $AR(1)$ alternative. p-values derived by parametric bootstrap for each test statistic appear in parentheses.

Table 5
Results for services grouped by commodity type, 1975–1992:4

	Full-panel $AR(1)$					EQ-TAR model				$BAND$-TAR model			
Good	p	N	T	$t_{\rho_{GLS}}$	$\hat{\rho}_{OLS}$	a	$\hat{\rho}_1$	N_1	LLR_{EQ}	a	$\hat{\rho}_1$	N_1	LLR_{BA}
App. repair	8	23	72	-6.78 (0.07)	-0.25 (0.00)	0.10	-0.26	561	424.42 (0.00)	0.16	-0.63	248	404.64 (0.00)
Auto maint.	12	22	55	-5.20 (0.42)	-0.43 (0.00)	0.10	-0.47	256	375.94 (0.01)	0.10	-0.94	256	368.03 (0.00)
Beauty	12	24	43	-0.50 (1.00)	-0.35 (0.00)	0.06	-0.35	537	250.47 (0.00)	0.06	-0.51	537	253.15 (0.00)
Bowling	12	22	72	-5.35 (0.26)	-0.35 (0.00)	0.16	-0.46	146	277.81 (0.83)	0.16	-1.16	146	304.91 (0.22)
Dentist	7	23	72	-6.20 (0.24)	-0.28 (0.00)	0.14	-0.30	327	271.15 (0.00)	0.14	-0.68	327	254.25 (0.00)
Doctor	7	23	72	-5.74 (0.42)	-0.24 (0.00)	0.20	-0.28	132	189.31 (0.18)	0.20	-0.79	132	158.91 (0.02)
Dryclean	5	23	72	-6.73 (0.11)	-0.24 (0.00)	0.13	-0.30	94	620.44 (0.85)	0.13	-0.71	94	612.34 (0.40)
Hospital	12	23	72	-4.06 (0.75)	-0.13 (0.01)	0.16	-0.18	78	309.31 (0.80)	0.16	-0.49	78	306.79 (0.49)
Haircut	7	23	72	-7.15 (0.03)	-0.28 (0.00)	0.19	-0.40	100	405.55 (1.00)	0.07	-0.47	671	429.34 (0.00)
Movie	7	23	72	-7.04 (0.03)	-0.26 (0.00)	0.17	-0.29	136	1523.11 (1.00)	0.17	-0.51	136	1513.24 (1.00)

Results from estimation of $AR(1)$, EQ-TAR and $BAND$-TAR panel models where relative prices are grouped by commodity type. p is the order of serial correlation selected for the disturbance DGP, N is the number of data series in each panel, T is the number of quarterly observations in each panel, $t_{\rho_{GLS}}$ is the test statistic for the GLS panel unit root test, $\hat{\rho}_{OLS}$ is the OLS estimate of ρ for the full panel, and for each of the TAR models, a is the estimated threshold, $\hat{\rho}_1$ is the OLS estimate of ρ_1 outside the threshold, N_1 is the number of observations outside the threshold, and $LLR \equiv 2\ln(\mathcal{L}/\mathcal{L}_{AR})$ is the likelihood ratio statistic for that model against the $AR(1)$ alternative. p-values derived by parametric bootstrap for each test statistic appear in parentheses.

Table 6
Results for services grouped by location, 1975–1992:4

	Full-panel $AR(1)$					*EQ-TAR* model				*BAND-TAR* model			
Location	p	N	T	$t_{\rho_{GLS}}$	$\hat{\rho}_{OLS}$	a	$\hat{\rho}_1$	N_1	LLR_{EQ}	a	$\hat{\rho}_1$	N_1	LLR_{BA}
Mobile, AL	9	8	72	-1.59 (0.97)	-0.24 (0.13)	0.24	-0.37	32	97.81 (0.55)	0.24	-0.97	32	97.05 (0.21)
Blythe, CA	7	7	72	-3.96 (0.16)	-0.35 (0.01)	0.15	-0.38	106	210.34 (0.43)	0.15	-0.66	106	200.42 (0.31)
Indio, CA	7	7	72	-4.34 (0.08)	-0.32 (0.08)	0.12	-0.37	141	113.17 (0.00)	0.12	-0.73	141	103.99 (0.00)
Denver, CO	4	8	72	-5.32 (0.01)	-0.38 (0.00)	0.16	-0.44	112	110.24 (0.05)	0.16	-0.87	112	106.59 (0.01)
Indianapolis, IN	9	8	72	-3.39 (0.49)	-0.23 (0.29)	0.10	-0.24	179	130.37 (0.00)	0.08	-0.35	229	117.50 (0.00)
Cedar Rap., IA	11	7	72	-2.71 (0.67)	-0.16 (0.54)	0.17	-0.17	127	65.26 (0.00)	0.22	-0.46	84	62.95 (0.00)
Lexington, KY	9	8	72	-2.51 (0.80)	-0.20 (0.14)	0.08	-0.19	285	25.51 (0.00)	0.08	-0.31	285	14.45 (0.00)
Louisville, KY	11	8	72	-3.45 (0.39)	-0.22 (0.65)	0.24	-0.19	38	72.27 (0.20)	0.24	-0.57	38	68.98 (0.05)
St. Louis, MO	11	8	72	-4.41 (0.07)	-0.18 (0.18)	0.12	-0.19	176	43.62 (0.00)	0.12	-0.42	176	40.08 (0.00)
Hastings, NE	11	8	72	-3.05 (0.60)	-0.24 (0.40)	0.24	-0.24	31	58.91 (0.95)	0.24	-0.38	31	46.19 (0.75)
Omaha, NE	10	8	72	-4.28 (0.13)	-0.33 (0.02)	0.10	-0.34	216	50.78 (0.00)	0.10	-0.56	216	35.33 (0.00)
Rapid City, SD	9	8	72	-3.89 (0.23)	-0.26 (0.25)	0.27	-0.38	50	84.73 (0.38)	0.25	-1.26	55	87.63 (0.09)
Vermillion, SD	11	8	72	-3.25 (0.48)	-0.21 (0.33)	0.18	-0.20	92	52.20 (0.00)	0.18	-0.40	92	40.63 (0.00)
Chattanooga, TN	10	7	72	-4.90 (0.01)	-0.30 (0.01)	0.07	-0.30	277	109.42 (0.00)	0.07	-0.43	277	109.30 (0.00)
El Paso, TX	11	7	72	-3.86 (0.11)	-0.23 (0.01)	0.20	-0.30	57	69.58 (0.01)	0.19	-0.75	67	70.78 (0.00)
Houston, TX	7	8	72	-4.43 (0.10)	-0.23 (0.02)	0.10	-0.24	173	107.66 (0.00)	0.10	-0.38	173	98.19 (0.00)
Lubbox, TX	9	7	72	-3.72 (0.24)	-0.27 (0.09)	0.13	-0.30	157	39.53 (0.00)	0.12	-0.61	174	30.76 (0.00)
S. Lake City, UT	7	8	72	-4.67 (0.06)	-0.24 (0.02)	0.24	-0.28	58	72.68 (0.23)	0.23	-0.86	62	65.34 (0.08)
Appleton, WI	7	8	72	-3.89 (0.25)	-0.19 (0.25)	0.11	-0.21	210	70.48 (0.00)	0.11	-0.33	210	51.03 (0.00)
Casper, WY	11	8	72	-4.77 (0.03)	-0.28 (0.31)	0.12	-0.27	157	109.91 (0.00)	0.12	-0.45	157	102.17 (0.00)

Results from estimation of $AR(1)$, *EQ-TAR* and *BAND-TAR* panel models where relative prices are grouped by location. p is the order of serial correlation selected for the disturbance DGP, N is the number of data series in each panel, T is the number of quarterly observations in each panel, $t_{\rho_{GLS}}$ is the test statistic for the GLS panel unit root test, $\hat{\rho}_{OLS}$ is the OLS estimate of ρ for the full panel, and for each of the TAR models, a is the estimated threshold, $\hat{\rho}_1$ is the OLS estimate of ρ_1 outside the threshold, N_1 is the number of observations outside the threshold, and $LLR \equiv 2\ln(\mathcal{L}/\mathcal{L}_{AR})$ is the likelihood ratio statistic for that model against the $AR(1)$ alternative. p-values derived by parametric bootstrap for each test statistic appear in parentheses.

Table A.1
Detailed description of traded goods data

Good	Start	Description
Aspirin	82:2	Bayer; 325mg tablets, 100 count
Babyfood	75:1	Jar strained vegetables; 4.5 oz.
Beer	82:2	Miller Lite or Budweiser; 12 oz., 6 pack
Cigarettes	75:1	Winston, king-size, carton
Coffee	75:1	Maxwell House, Hills Brothers or Folgers; 2lbs, 1lb, or 13oz.
Cornflakes	79:2	Kellogg's or Post Toasties; 18 oz.
Game	82:2	Monopoly; standard (No. 9) edition
Jeans	82:2	Levi's; straight leg, 501s or 505s
Liquor	75:1	Seagrams 7 Crown or A&B Scotch; 750ml
Shirt	82:2	Arrow or Van Heusen; white, long sleeve, cotton-poly blend
Orange Juice	75:1	Can, 6 oz. or 12 oz.
Peaches	75:1	Del Monte or Libby's; #2.5 can (29 oz.), halves or slices
Shampoo	82:2	Johnson's or Alberto VO5; bottle, 11 oz. or 15oz.
Shortening	75:1	Crisco; all vegetable, 3lb. can
Soda	75:1	Coca-Cola; 1 quart or 2 litre
Sugar	79:2	Cane or beet, 4lbs. or 5lbs.
Tennis	82:2	Wilson or Penn; can, yellow, heavy duty, 3 count
Tissue	75:1	Kleenex; 1 roll, 4 roll or box, 175 count
Toothpaste	82.2	Crest or Colgate; 6oz. or 7oz.
Tuna	82.2	Starkist or Chicken of the Sea; in oil, can 6.5oz.
Underwear	82:2	Package of 3 briefs
Detergent	75:1	Giant Tide, Bold or Cheer; 42oz. or 49 oz.
Wine	82.2	Paul Masson Chablis or Gallo Sauvignon Blanc or Gallo Chablis Blanc; 750ml or 1.5 litre

Description of American Chamber of Commerce data as published in *Cost of Living Index.*

Table A.2
Detailed description of perishable goods data

Good	Start	Description
Bacon	75:1	1lb package
Bananas	75:1	1lb
Bread	75:1	20oz. or 24oz.
Cheese	82:2	Kraft; Parmesan, grated, canister, 8oz.
Eggs	75:1	Grade A, 1 dozen
Minced steak	75:1	1lb.
Lettuce	75:1	1 head
Margarine	75:1	1lb.
Milk	75:1	Half-gallon
Potatoes	75:1	White or red, 10lbs.
Steak	75:1	Round steak or T-bone; USDA choice, 1lb.
Chicken	75:1	Grade A frying, 1lb.
Fr. chicken	82:2	Kentucky Fried Chicken or Church's; breast and drumstick
McDonalds	82:2	Patty or patty with cheese, pickle, onion mustard and ketchup; 0.5lb.
Pizza	82:2	Pizza Hut or Pizza Inn; 12"-13" crust, regular cheese

Description of American Chamber of Commerce data as published in *Cost of Living Index.*

Table A.3
Detailed description of services goods data

Good	Start	Description
App. repair	75:1	Service call for color TV or washing machine; excluding parts
Auto maint.	79:2	Balancing; 1 or 2 front wheels, computer or spin balance
Beauty	82:2	Shampoo, trim and blow-dry; women's visit
Bowling	75:1	Evening price; per line
Dentist	75:1	Office visit; cleaning and inspection, no X-ray or flouride treatment
Doctor	75:1	Office visit; routine exam of existing patient
Dryclean	75:1	Man's suit; 2-piece
Hospital	75:1	Hospital room; semiprivate cost, per day
Haircut	75:1	No styling; man's
Movie	75:1	First run; indoor evening price

Description of American Chamber of Commerce data as published in *Cost of Living Index.*