NBER WORKING PAPER SERIES

NONLINEAR ASPECTS OF GOODS-MARKET ARBITRAGE AND ADJUSTMENT: HECKSCHER'S COMMODITY POINTS REVISITED

Maurice Obstfeld
Alan M. Taylor

Working Paper 6053

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

Forthcoming in *The Journal of the Japanese and International Economies*. Paper presented at the NBER-TCER-CEPR conference, "Purchasing Power Parity Revisited: The Exchange Rate and Price Movements, Theory and Evidence," Tokyo, December 20-21, 1996. We thank Charles Engel and John Rogers for providing the data from Engel and Rogers (1995), and thank Lawrence Officer for bringing Heckscher's discussion of commodity import and export points to our attention. We also acknowledge helpful comments and criticism from Menzie Chinn, Richard Clarida, Mitsuhiro Fukao, Takatoshi Ito, Colin McKenzie, Paul O'Connell, Axel Weber, an anonymous referee, and participants at the Tokyo conference and at the March 1997 meeting of the NBER International Finance and Macroeconomics group. Research support was provided by a grant from the National Science Foundation to the National Bureau of Economic Research. This paper is part of NBER's research program in International Finance and Macroeconomics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1997 by Maurice Obstfeld and Alan M. Taylor. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Nonlinear Aspects of Goods-Market Arbitrage
and Adjustment: Heckscher's Commodity
Points Revisited
Maurice Obstfeld and Alan M. Taylor
NBER Working Paper No. 6053
June 1997
JEL Nos. F31, F41
International Finance and Macroeconomics

ABSTRACT

We propose that analysis of purchasing power parity (PPP) and the law of one price (LOOP) should explicitly take into account the possibility of "commodity points" —thresholds delineating a region of no central tendency among relative prices, possibly due to lack of perfect arbitrage in the presence of transaction costs and uncertainty. More than eighty years ago, Heckscher stressed the importance of such incomplete arbitrage in the empirical application of PPP. We devise an econometric method to identify commodity points. Price adjustment is treated as a nonlinear process, and a threshold autoregression (TAR) offers a parsimonious specification within which both thresholds and adjustment speeds are estimated by maximum likelihood methods. Our model performs well using post-1980 data, and yields parameter estimates that appear quite reasonable: adjustment outside the thresholds might imply half-lives of price deviations measured in months rather than years, and the thresholds correspond to popular rough estimates as to the order of magnitude of actual transport costs. The estimated commodity points appear to be positively related to objective measures of market segmentation, notably nominal exchange rate volatility.

Maurice Obstfeld
Department of Economics
University of California
Berkeley, CA 94720
and NBER
obstfeld@econ.berkeley.edu

Alan M. Taylor
Department of Economics
Northwestern University
Evanston, IL 60208
and NBER
amt@nwu.edu

1 Introduction

It is a commonplace of the recent floating exchange rate era that both the law of one price (LOOP) and purchasing power parity (PPP) fail dramatically in the short run (see Rogoff 1996 for a comprehensive survey). Most recent research therefore has focused on the usefulness of PPP or LOOP as *long-run* propositions. Interpretation of the empirical record has proven controversial, despite the deployment in the last few years of powerful panel estimation methods designed to exploit simultaneously information from many national experiences. Even the studies most favorable to long-run PPP or LOOP, however, suggest an extremely slow decay rate for international price differentials. Estimated half-lives for PPP deviations, for example, tend to be somewhere on the order of four to five years.

In this paper we present empirical evidence that price differentials net of transaction costs may be substantially eliminated much more quickly than these latter estimates show—in months rather than in years—but that costs of international trade result in bands within which relative international prices can fluctuate with no central tendency.[1] We argue that existing price convergence studies reach the conclusion that convergence is slow by pooling data from two separate regimes: one of rapid convergence when price differences exceed transaction costs, and one of slow or nonconvergence when price differences are relatively small.[2]

[1]In earlier empirical work, De Grauwe, Janssens, and Leliaert (1985) and Davutyan and J. Pippenger (1990) have emphasized the capacity of tranaction costs to mask the force of international goods-market arbitrage. After completing a first draft of this paper, we found two other independent investigations of nonlinear adjustment to PPP. O'Connell (1996b) allows for the possibility of bands with different adjustment speeds, but his approach differs from ours in focusing on unit-root testing and power considerations. He does not utilize optimization (a search alogorithm) to locate thresholds as we do, but imposes them *a priori*. Another recent related piece is by Michael, Nobay, and Peel (1997), but theirs is a model of smooth transitions between adjustment regimes, unlike the threshold model presented here. Both of these latter works also differ from ours in focusing on adjustments toward PPP using broad price aggregates, and do not feature the disaggregated data for commodity categories which we employ.

[2]Hegwood and Papell (1996) show that convergence speeds appear more rapid once

The balance of this paper is organized as follows. Section 2 discusses the recent evidence on convergence to PPP and LOOP and sketches a theoretical framework that suggests an alternative mode of empirical investigation. Our approach begins with Heckscher's (1916) insight that international transaction costs should create some scope for deviations from LOOP and PPP, "commodity points" analogous to the gold points regulating specie flows under the gold standard. Advances in the theory of investment under uncertainty imply that the commodity points should be interpreted as resulting not only from concrete shipping costs and trade barriers, but also from sunk costs of international arbitrage and the resulting tendency of traders to respond only to sufficiently big price differentials. Section 3 specifies and estimates a simple nonlinear price adjustment model meant to capture the essentials of international price arbitrage subject to "bands of inaction." The model is applied both to aggregate CPI indices and to CPIs for disaggregated tradables in various cities and countries. In general, the model yields very reasonable estimates of commodity points, as well as rapid adjustment speeds outside of the implied bands. Our estimates of the commodity-point widths are positively related to distance, to aggregate tariff levels, and, very significantly, to nominal exchange-rate volatility. However, they appear unrelated to our (admittedly crude) measures of aggregate nontariff barriers. Section 4 summarizes and suggests directions for further research.

one allows for periodic discrete structural shifts in real exchange rates. The idea is related to Engel's (1996) point that real exchange rates may contain nonstationary as well as stationary components. Many of the distinct real exchange rate regimes Hegwood and Papell discern, however, do not obviously correspond to known structural shifts in the economies they study. Here instead, our idea is basically to classify PPP deviations according to whether they push relative prices beyond effective arbitrage points. In line with our findings below, Hakkio (1992) reports that PPP provides a useful guide to the direction of the U.S. dollar's future short-term movement when (and only when) the dollar is unusually far from a PPP benchmark level.

2 Recent Evidence and a Suggested Theoretical Framework

Initial empirical research on various measures of real exchange rates over the post-1973 floating exchange rate was based on pure time-series methods. It suggested near random-walk behavior in industrial-country real exchange rates, with little or no discernible tendency toward mean reversion.[3] A commonly (if often reluctantly) drawn inference was that exchange-rate theories based on PPP might be of limited use for understanding even the long-run behavior of exchange rates. However, the time-series work was plagued by a relatively short sample length, which offered little power to reject a unit-root null against near unit-root alternatives.

More recently, researchers have tried to exploit the potentially greater power of panel data sets to sharpen estimates of the speed of convergence to PPP. The first wave of these studies (including Wei and Parsley 1995; Chinn and Johnston 1996; Frankel and Rose 1996; Jorion and Sweeney 1996; Oh 1996; Wu 1996; and Lothian 1997) has found considerable evidence against unit roots in real exchange rates, but also claims relatively long half-lives for deviations from PPP, with consensus estimates clustering in the range of four to five years. But controversy remains, and the latest work continues to cast doubt on the long-run stationarity of international relative prices, even using wide panels and even when attention is restricted to supposedly tradable goods (O'Connell 1996a; Engel 1996; Canzoneri, Cumby, and Diba 1996). The most robust evidence in favor of long-run PPP and LOOP still appears to come from long time-series samples of a century or more in length (among the newer studies are Froot, Kim, and Rogoff 1995; Lothian and M. P. Taylor 1996; and A. M. Taylor 1996). As Froot and Rogoff (1995) stress, though, these long-run estimates blend data from very different nominal exchange-rate regimes, and thus do not illuminate the process of international price

[3] A contrary result for Japan is, however, reported by Obstfeld (1993).

arbitrage under the post-1973 floating exchange rate arrangements.

Much of the recent empirical work proceeds by augmenting the statistical power of a standard atheoretical regression framework. There has been relatively little attempt, however, to model explicitly the market forces determining the nature of convergence to LOOP or PPP. The literature on spatially separated commodity markets suggests, as Heckscher (1916) observed, that transport costs and other trade barriers play a key role in limiting intermarket price differentials for very similar commodities.

When arbitrage is immediate as in Williams and Wright (1991), for example, transport costs strictly delimit the range of price fluctuations. In the more recent model of Coleman (1995), time elapses while goods move between markets, allowing a period during which intermarket price differentials can exceed the cost of transport. Prices do eventually converge to eliminate static arbitrage profits, but this convergence need not imply price equality, only that price differentials decay until they no longer fall outside the commodity shipment points.[4] Thus, large price differentials converge toward transport-cost differentials, not toward absolute parity, and may spend time outside the bands transport costs would delimit in a world of instantaneous trade. Convergence is rapid in Coleman's model, but in more complex market settings one can conceive of the process being slowed by convexities in the shipment technology, lags in information transmission, and so on. With similar but differentiated tradables, one can also imagine customer-market effects (à la Phelps and Winter 1970 or Rotemberg and Woodford 1991) inducing gradual convergence.

It is this last type of model that informs the econometric methodology we outline in the next section. In essence, we assume that price differentials may exhibit no central tendency when they reside within the commodity points, but that larger differentials are arbitraged away and hence decay in

[4]Coleman (1995) applies his model to the U. S. slaughter cattle market, and finds some empirical support.

expectation according to a stable autoregressive process. We use maximum-likelihood methods to estimate both the commodity points and the decay rate for large price differentials.

In interpreting the commodity points we estimate below, it is important to observe that they may be wider than measurable transport costs and trade restrictions alone would imply. In the presence of sunk costs of arbitrage—for example, costs of setting up or expanding foreign retail distribution networks—uncertainty as to the permanence of the shocks causing relative price changes will widen the "bands of inaction" within which price differentials can fluctuate before arbitrage commences; see, for example, Dixit (1989), Dumas (1992), and Krugman (1989). Indeed, such effects are the most likely explanation for our finding in the next section that our estimated commodity points and nominal exchange-rate volatility are strongly linked.[5]

3 Empirics

3.1 Models of Price Adjustment

We first operationalize Heckscher's notion of commodity points in a simple form amenable to empirical application. Our starting point is the *standard model*, one used in many empirical analyses of purchasing power parity (PPP) and the law of one price (LOOP), namely the AR1 model. Let p_t^1 and p_t^2 be the log price levels of a good (or composite good, or basket of goods) in two locations at time t. Adjustment models are concerned with the dynamic behavior of the price gap $z_t = p_t^2 - p_t^1$.

If empirical data always considered identical, purely tradable goods at all locations and at all times, the above z_t would be a legitimate object of

[5]Another example of a "band of inaction" in a PPP model was provided by Sercu, Uppal, and Van Hulle (1995), but their model implies that the real exchange rate is confined within the band (unlike the gradual adjustment in our model outside the band) and their focus is on exploring the determinants of exchange rate adjustment within the band, for example, as a function of monetary policy in a cash-in-advance model.

study. In general, we need to consider the possibility of long-run trends in *measured* price differences arising from aggregation in price indices or from bundled nontradable components or quality differences. Indeed a large literature is concerned with the analysis of trend components in the real exchange rate (for example Marston 1987; Obstfeld 1993; Asea and Mendoza 1994; Chinn and Johnston 1996; Canzoneri, Cumby, and Diba 1996). For the present purpose, we are not concerned with the *long-run* trend behavior of the equilibrium price difference, only the *short-run* properties of adjustment toward that equilibrium, however it may be determined. Accordingly, we will admit the possibility of a long-run trend in our relative prices. Thus, we define x_t to be the detrended component of the price difference z_t, given by $z_t = \alpha + \beta t + x_t$, where x_t may be estimated as an OLS residual. (We also allow for a constant since we work with CPIs rather than absolute prices.) We might consider this x_t to be the candidate for the "equilibrium error" in the price differences. In practice, for our 1980–1995 data, we also used an identical model without initially detrending the data, and the results were qualitatively similar (see Appendix B).

In the standard model x_t is assumed to follow an AR1 process

$$\Delta x_t = \lambda x_{t-1} + e_t, \tag{1}$$

where e_t is $N(0, \sigma^2)$ and λ, expected to be between zero and minus one, is called the *convergence speed.* Note that x_t is already detrended and demeaned, so there is no constant term in (1). Thus, price differentials are diminished by a fraction λ in each period, plus an error term. This type of model has been used countless times in the analysis of PPP and LOOP. The convergence speed is usually interpreted as a measure of the integration of markets or the efficiency of arbitrage between spatially separate locations, and is expected to depend on the good or composite goods under consideration, the nature of transaction and transportation costs for these goods, and other aspects of economic distance between locations. Several recent panel studies have offered new estimates of λ of around -0.15 per annum, suggesting that deviations from PPP have a half life $\ln(0.5)/\ln(1+\lambda)$ of four to five years (see Rogoff 1996). But, as we have noted, the statistical significance of these estimates has not gone unchallenged.

It is natural to seek a way to insert Heckscher's idea of commodity points, a notion of nonlinear adjustment, into the contemporary methodology used to study LOOP and PPP. The simplest way to do this is to modify the standard model (1) from its AR1 form, to yield a threshold autoregression (TAR) model with thresholds corresponding to the commodity points c following Prakash's (1996) study of grain markets, or more generally, to "bands of inaction"—price limits within which arbitrage does not yield an expected net gain. In the TAR model, the process for x_t now exhibits two types of behavior. Inside the commodity points, when $|x_t| < c$, arbitrage does not operate and the price gap shows no central tendency; we model this behavior as a random walk. Outside the commodity points, when $x_t > c$, arbitrage operates as per the standard model, except that the price will now revert to the commodity point threshold, and *not* to 0. Thus, we may write the process as

$$\Delta x_t = \begin{cases} \lambda^{out}(x_{t-1} - c) + e_t^{out} & \text{if } x_{t-1} > c; \\ \lambda^{in} x_{t-1} + e_t^{in} & \text{if } c \geq x_{t-1} \geq -c; \\ \lambda^{out}(x_{t-1} + c) + e_t^{out} & \text{if } -c > x_{t-1}; \end{cases} \tag{2}$$

where e_t^{out} is $N(0, \sigma^{out^2})$, e_t^{in} is $N(0, \sigma^{in^2})$, $\lambda^{in} = 0$, and λ^{out} is the convergence speed outside the commodity points. In this model, equilibrium in prices obtains whenever x_t is within the commodity points, so we still speak of convergence to equilibrium. Only now, equilibrium for x_t is anywhere in a band or interval $[-c, +c]$, and not just at the point 0. Within that band there is no "error correction" force at work, because there is no error to correct.

Note that a looser specification would leave λ^{in} unrestricted. Clearly, within the no-arbitrage band, prices at each location follow a process determined by shifts in local excess demand functions, and the price difference will follow a process linked to differences in excess demand shocks. Those shocks could follow a totally arbitrary time series properties, for example, random walk, drift, stationary serial correlation, and so forth. In practice, for our 1980–1995 data, we also tested an identical model without the assumption of a unit root within the band, and the results were qualitatively similar (see Appendix C). We also investigated whether the possibility of serial corrleation might affect our threshold estimates, and here too the results did not greatly differ (Appendix D). [6]

The proposed model is one of a large family of TAR(p;k,d) models, which may be characterized by an arbitrary autoregressive length p, an arbitrary number of thresholds k, and an arbitrary delay parameter d (which determines the order of lag used to determine within which threshold regime the current observation falls). The simple model we have proposed is a TAR(1;2,1) with some restrictions on the coefficients for symmetry about $x = 0$. In order to test whether this kind of nonlinear specification is justified, we will shortly evaluate some specification tests developed for testing general TAR(p;k,d) alternatives against AR(p) nulls. In practice, we will also employ Monte Carlo simulations to test our specific TAR(1;2,1) against the AR1 alternative that is the standard model.[7]

[6]Concerning specification, our choice of null is consistent with our maintained assumption that the price difference z_t contains a deterministic trend which has been filtered out to generate a detrended, demeaned series x_t. It would be possible to test the TAR model against other nulls, such as a pure random walk for z_t, but with a widening menu of specification choices model selection is not simple. For example, it is clear that for given thresholds the benchmark TAR and AR models are non-nested. They are each nested models within a much broader class of TAR models. As an additional check we employed the standard J-test (Davidson and MacKinnon 1981) for non-nested models, but it proved unable to distinguish the two non-nested alternatives from each other. The objective of this paper is an illustration of the TAR technique, with a focus on fit (that is, a best fit description of the data process) and the implications of this for notions of equilibrium, and point estimates of convergence speed and half lives. Problems of inference remain, as already noted, given the need for simulation analysis, and given the non-nested structure of several candidate nulls and alternatives. For an analysis which is less concerned with fit and more concerned with inference and power see O'Connell (1996b).

[7]One extension would be to explore higher order autoregressive TAR(p;2;d) models with varying delays to see whether our results are sensitive to the particular chosen. Tsay (1989) offers a search algorithm to estimate p and d. The focus here is merely to contrast

Balke and Fomby (1997) call our model a "Band-TAR" model. We must emphasize that it is not the only conceivable model of threshold equilibrium adjustment. Alternative models might include their "equilibrium-TAR" (EQ-TAR) where reversion is toward the *center* of the band, not its edge:

$$\Delta x_t = \begin{cases} \lambda x_{t-1} + e_t & \text{if } x_{t-1} > c; \\ e_t & \text{if } c \geq x_{t-1} \geq -c; \\ \lambda x_{t-1} + e_t & \text{if } -c > x_{t-1}; \end{cases} \tag{3}$$

where $\lambda < 0$. Another possibility is the "returning-drift-TAR" (RD-TAR) where reversion is in the form of a random walk with inward drift outside the thresholds:

$$\Delta x_t = \begin{cases} -\mu + e_t & \text{if } x_{t-1} > c; \\ e_t & \text{if } c \geq x_{t-1} \geq -c; \\ \mu + e_t & \text{if } -c > x_{t-1}; \end{cases} \tag{4}$$

where $\mu > 0$. Despite local random walk behavior within the band, Balke and Fomby (1997) show that all of these processes are globally stationary but with different short-run adjustment dynamics. The RD-TAR is the most persistent, for example, and the equilibrium-TAR the least persistent.[8]

our estimates with the prevailing AR(1) paradigm in the PPP literature, so we maintain the assumption $p = 1$ for consistency across specifications.

[8]It is the EQ-TAR form that is used by O'Connell (1996b) to estimate a PPP reversion process with nonlinear, or price-differential specific, adjustment speeds. However, he does not impose a unit root within the arbitrage bands and his threshold choice is postulated a priori rather than being determined by a search, as we propose. If there are commodity points in the Heckscher sense, we would expect a band for which arbitrage profits are negative or zero, and within which there is no necessary tendency for further price convergence. In such a setting, the Band-TAR would seem to be the more natural specification. As noted above, in one set of tests we run below (see Appendix C), we allow for an unrestricted convergence speed within the band. Usually, estimates of λ^{in} are statistically insignificant and it is typical to find $\lambda^{out} < \lambda^{in}$, implying faster convergence for relatively large price differentials. Such results are also consistent with Michael, Nobay, and Peel (1997), who also find faster convergence for larger deviations. We have chosen to employ a Band-TAR, in which the expected, or predictable, change in x_t, $E(\Delta x_t) = f(x_{t-1})$, is a continuous function of x_{t-1} (unlike the EQ-TAR, where there is a discontinuity at the threshold). Continuity of this autoregressive function has inference implications. If such continuity is required at the threshold, then $\lambda^{in}=\lambda^{out}$ implies a pure AR model, and so specification tests with an unrestricted model are also tests of whether the convergence speeds differ in the two regimes.

In the next sections we will explore the implications of the using the TAR model (2) instead of the standard AR1 model (1). We first show, through simulation, how an incorrect choice of the standard model might lead to a biased underestimate of the convergence speed. We then show how to estimate and test the TAR model. Finally, we illustrate the method with an application to monthly data on a variety of goods for a cross section of cities and countries over the 1980–1995 period.

3.2 Model Simulation: TAR versus AR

Intuitively, there is good reason to suspect that the incorrect choice of the standard model instead of a true TAR model might lead to a biased estimate of λ. If the true process is the TAR in eq. (2), then applying the standard model amounts to an unjustified pooling of two types of data: observations outside the commodity points where the process does indeed exhibit reversion properties ($\lambda < 0$); and observations within the commodity points where the process is a pure random walk with no reversion ($\lambda = 0$). Thus, an incorrect choice of the standard model could bias our estimate of the convergence speed downward. It might also thus encourage a false acceptance of the hypothesis that x_t follows a random walk, especially if the process is relatively slow to reach the edges of the band, so that many random walk observations are observed.[9]

To illustrate these possibilities, we simulated a true TAR model according to equation (2) with some arbitrary parameters. The threshold level was chosen at $c = 1.0$, and the adjustment spread outside the threshold was set

[9]O'Connell (1996b), following M. K. Pippenger and Goering (1993), employs simulations and a power analysis to show that this is indeed the case. Power is shown to depend on the convergence speed parameter λ, and on the ratio of the standard deviation of the error to the threshold size σ/c. Power of the unit root tests is weakest as λ goes to zero, possibly the relevant range for PPP adjustment processes. But note again, that these calculations are for the EQ-TAR specification, not the Band-TAR used here. If the Band-TAR is the true model, then these problems should, intuitively, only worsen: the Band-TAR is even less persistent than the EQ-TAR, after all.

at $\lambda = -0.1$ (a half life of about 6.5 periods). The error process was specified to be the same inside and outside the thresholds, with $\sigma^{out} = \sigma^{in} = 0.5$.[10] Simulation proceeded with an initial value of x_t equal to zero, and the process was run for 2,100 periods, discarding the first 100 values. The final 2,000 periods of simulated data were chosen for analysis. The simulated data are shown in Figure 1. Of the 2,000 observations, 1,097 (55%) fell outside the threshold band.

Knowing the true process to be TAR, we then estimated the standard AR1 model on the simulated data. The result was the following OLS equation:

$$\begin{array}{ccccc} \Delta x_t & = & -0.017 & - & 0.053 \quad x_{t-1} \\ & & (0.053) & & (0.007) \end{array}$$

Thus, incorrect use of the standard model yields an estimate of the convergence speed roughly one half the true size, 5 percent per period (versus 10 percent in the true model), and an implied half-life roughly twice the true size, about 13 periods (versus about 6.5).

To show that this is not a freak occurrence, we repeated this simulation experiment 500 times and collected the resulting OLS estimates of λ from the standard model. The results were as follows:

Observations	500
Median	-0.052
Sample Mean	-0.052
Standard Error	0.005

Minimum	-0.067	Maximum	-0.037
01-%ile	-0.066	99-%ile	-0.040
05-%ile	-0.061	95-%ile	-0.044
10-%ile	-0.059	90-%ile	-0.045
25-%ile	-0.056	75-%ile	-0.049

[10] We have attempted to set up this simulation to correspond to a plausible "real world" scenario. This ratio of threshold c to the standard deviation of the error term σ ($c/\sigma = 2$) corresponds to O'Connell's (1996b) rough approximation of the ratio of typical observed real exchange rate variability (20%) to rough estimates of the typical ratio of transport costs to trade (about 10%) given by FOB to CIF comparisons (for example, Rogoff 1996, p. 653.)

It can be seen that the distribution of the OLS coefficient is tightly clustered around the value -0.052, with some 90% of observations falling in the interval $[-0.061, -0.044]$. This range is very far from the true value of $\lambda = -0.1$ outside the band.

It will be shown later, with reference to actual, not simulated, data, that these kinds of biases in the estimate of convergence speeds are not unusual in the application of this method to contemporary international price data. What is clear is that if we stop pooling the data, we could have adequate power to identify the true adjustment parameters. Figure 2 displays first the pooled regression reported above for the simulation data in Figure 1. The next two figures show the AR1 regression run on split samples: the first within the band, the second outside the band, with the adjustment assumed to be to the edge of the band. The estimated coefficients and scatterplots now accord very closely with the true model. Inside the band, the scatterplot if a diffuse array of dots, and the slope (of change in x versus lagged x) is not significant. Outside the band, the scatterplot shows a negative correlation, and the slope is significant (in a regression of change in x versus x minus the nearest threshold). Moreover, the outside band point estimate is very close to -0.1, the true adjustment speed in that regime. This exercise suggest a natural way to estimate the parameters of the model: for any given threshold value we split the sample in two, and estimate the parameters individually inside and outside the band. It then will prove straightforward to search for the best estimate of the threshold value by aiming to maximize some measure of fit.

3.3 Estimating and Testing the TAR Model

The problem with the standard model is that it ignores the thresholds and pools all the data as if they were generated by a uniform process. We therefore seek a method that allows us to seek the best-fit TAR model, properly partitioning the data into observations inside and outside the thresholds. We also seek tests that will allow us to ascertain more closely the right specification, and thereby distinguish between the standard model's AR1 process with no thresholds, and the TAR model.

A very general specification test is provided by Tsay (1989) for TAR alternatives against an AR null. It is a simple nonparametric test for exactly the kind of nonlinear adjustment process we have in mind here. It has already been applied to the analysis of price processes in the study of arbitrage and market integration (Coleman 1996; Prakash 1996; Prakash and Taylor 1997). The method is as follows. Following Tsay, our TAR model is of the general form

$$x_t = a_0^i + a_1^i x_{t-1} + e_t^i \quad \text{if } r(i) > x_{t-1} \geq r(i-1); \tag{5}$$

where the $r(i)$ for $i = 1, ..., k-1$ are nontrivial threshold values, with $k \geq 2, r(0) = -\infty, r(k) = +\infty$. We call (x_t, x_{t-1}) a *case* of data for $t = 1, ..., T$. We next order the cases according to the delayed level x_{t-1}, the lag of x that decides between which thresholds each case is located. Tsay's test is based on recursive residuals from the arranged AR1 autoregression of this case data. The intuition is as follows. Suppose we order the cases in increasing fashion, and the first m cases fall below the first threshold. Then, the recursive AR1 residuals will be orthogonal to the dependent variables of the regression. However, because of the regime change at the threshold level, the recursive residuals for cases after the mth will no longer have this same orthogonality. This fact suggests a simple F test for the recursive residuals.

In full, let the ordered case data be (x_{t_s}, x_{t_s-1}) for $s = 1, ..., T$. We perform an arranged AR1 autoregression on this ordered data and generate recursive residuals e_{t_s} for each case. We then do an OLS regression of these predictive residuals on the dependent variables,

$$e_{t_s} = \omega_0 + \omega_1 x_{t_s-1} + u_{t_s}, \tag{6}$$

and the orthogonality test amounts to calculating the conventional associated F statistic for this regression, which follows an asymptotic F distribution for large T. One final item is to note that, especially in small samples, the case data may not fall in all of the regimes delineated by every threshold value. Thus, as a practical matter, the Tsay test should be run with both increasing and decreasing ordering of the arranged regression. We follow that practice in what follows, and we report only the most significant p-values for each of the two F tests.

We now have a test for nonlinearity, but Tsay's test is nonparametric, and depends neither on the number of thresholds, $k-1$, nor their values $r(1), \ldots, r(k-1)$. However, from the standpoint of interpretation, if we find evidence of a TAR model in a PPP or LOOP context, it is natural to look first at the simple, symmetric, TAR specification (2), which incorporates two nontrivial thresholds corresponding to Heckscher's commodity points. In such a model, as in the standard model (1), important issues of interpretation rest on the speed of convergence (λ), and the spread of the commodity points (c), and how these parameters might relate to prior or independent estimates of fixed costs of transactions, transportation, and the like.

Thus, we need next a parametric method of estimating λ and c. Following Fanizza (1990), Balke and Fomby (1997), and Prakash (1996), we use a best-fit grid-search on the threshold parameter c. Such a search requires an objective function, and some choices are available. Balke and Fomby (1997) search over c and minimize the residual sum of squares. Fanizza (1990) and Prakash (1996) maximize a likelihood function. We follow the latter approach here.

Let $L_n(\lambda, \sigma)$ be the estimated log likelihood function of the standard (null) AR1 model(1), and let $L_a(\lambda, \sigma^{out}, \sigma^{in}; c)$ be the estimated log likelihood function of the (alternative) TAR model (2) for a given c. Estimation proceeds via a grid search on c which maximizes the log likelihood ratio $LLR = 2(L_a - L_n)$.[11] This is computationally simple, since for any given

[11] In principle, we could also allow for different upper and lower thresholds too. However, this appears unattractive for two reasons: in practical terms the computational cost of the search rises from order N to order N^2, as the search grid becomes two dimensional;

c, maximum likelihood (ML) estimation of the TAR model amounts to an OLS estimation on partitioned samples, namely, sets of observations with x_{t-1} wholly inside or wholly outside the thresholds.[12]

Finally, the likelihood ratio also provides a specification test of our chosen nonlinear model. This is a test of the specific two-threshold commodity-point TAR against the AR1 null, as compared with Tsay's general nonparametric test of AR1 versus alternatives with an indeterminate number of thresholds. However, a price is paid: care is needed since the parameter c is not identified under the AR1 null. Thus standard inference is invalid, and we must proceed using Monte Carlo methods, as in Prakash and Taylor (1997) and Fanizza (1990).[13] Empirical estimates of the distribution of LLR under the AR1 null can be simulated, and used to test whether the null should be rejected in favor of our specific TAR alternative. (See Appendix A.)

3.4 An Empirical Application

In this section we apply the above techniques to search for nonlinear adjustment in price levels. We consider disaggregated as well as aggregated CPIs for 32 city and country locations at monthly frequency from 1980 to 1995. The data are drawn from Engel and Rogers (1995). They include four cities in the U.S. and Canada, and national price levels for the U.S., Canada, Mexico, 15 countries in Europe, four countries in Asia, plus New Zealand and South Africa. The full set of indices available covers clothing, food, fuel, health, household, equipment, housing, recreation, transportation, and a basket of all consumer goods. (See Appendix E.) However, we focus only on the basket

and on theoretical grounds, we expect tests of PPP to satisfy some kind of base-country invariance principle, which, here, would imply symmetric upper and lower thresholds. The latter point might also be expected to imply symmetric speeds of convergence above and below the band. Still, such looser specifications as these are feasible extensions to the empirical methodology we propose.

[12]For practical purposes, choices of c with few observations in either partition are rejected (we chose 30 observations as a minimum).

[13]As shown by Balke and Fomby (1997), their "sup-Wald" statistic may be used to test for nonlinearity also, but again using Monte Carlo methods for the same reasons.

of all goods (for a PPP type test) and the first three quasi-tradable goods baskets, clothing, food, and fuel (for LOOP-type tests).[14] These comparisons are highly relevant to broader notions of PPP, however, because variations in relative tradables prices have been found to play an important—some would say, a dominant—role in the determination of real exchange rates (see, for example, Marston 1987; Engel 1995; and Canzoneri, Cumby, and Diba 1996). Even though we focus on putatively tradable goods, we are concerned that even these last three categories may contain substantial nontradable components in the form of distribution and marketing services, etc. In all cases we constructed x_t from the detrended and demeaned component of $z_t = p_t^1 - p_t^2$ for each location pair and each good by calculating the difference of demeaned and detrended price series for each location.

To start, we apply the threshold model to PPP using the basket of all goods. We chose the United States as the base country (series p_t^2) in each case. We first ask whether the price differential process shows any nonlinearities using Tsay's test. We found that the maximal Tsay F-statistic is highly significant for many locations (Appendix Table 1, final page, has the details). We next estimate the two-threshold TAR model (2), using a grid-search on c with the maximum likelihood method. The log likelihood ratio of the TAR model relative to the AR1 standard model is shown, along with the estimated threshold level c. For this c, we show the TAR estimates of the parameter λ, the absolute t statistic of λ, and the implied half-life. For comparison, we also show the same estimates for the standard AR1 model (Appendix Table 1 supplies detailed results here too).

Table 1 provides an encapsulated summary of our findings on thresholds and convergence speeds for various subsamples of our set of locations. The

[14] We ran our test on all nine goods categories, but the results for the other five mostly-nontradable baskets were very poor, which was not surprising. In general, one probably should not put much weight on the tests using the aggregate CPIs, as their intrepretation is problematic. Different commodities in principle are subject to different arbitrage costs. Moreover, the behavior of a broad price index might mask the positions of individual commodity price differentials relative to the transaction-cost band.

results accord with intuition, and with the findings of our earlier simulation method. Consider first the PPP-type results for the entire price index (CPI-All). The typical half-life of price differences in the AR1 model is of the order of 17 months, and in the range 20 to 40 months for locations outside of North America vis-à-vis the U.S.[15] In contrast, the typical half-life in the TAR model is only about 8 months, and roughly 12 months for locations outside North America vis-à-vis the U.S.[16] Thus, the "new consensus" of short half-lives of PPP deviations may well understate the case. Estimates of convergence speeds based on linear autoregressive models may indicate a half-life of several years, but estimates based on a threshold approach suggest that deviations have half lives of only a year or so. This is not far from the convergence speed that Parsley and Wei (1996) report for intra-U.S. price differentials for items other than services. It is also close to the convergence speed Cumby (1996) reports for the *Economist*'s international panel of Big Mac prices.[17] Table 1 shows that these basic findings are also repeated using baskets of disaggregated tradable goods in a LOOP-type analysis; indeed half-lives of deviations appear to be generally far below 12 months for the disaggregated indices. In all cases, the adjustment speed is much higher once a threshold is introduced.[18]

[15]It is clear that even our simple AR half-lives are somewhat shorter than the conventional wisdom suggests, being about two to three years on average rather than four to five. In addition to our new TAR approach, this result partly explains why we find such small half-lives in this paper. We think this result is mostly due to temporal-aggregation biases in most studies. Using quarterly or annual data (panel or single-country) these studies are working at an unfortunate sampling frequency which cannot hope to detect high-frequency mean reversion tendencies on the order of months or weeks.

[16]"All") are simple unweighted averages of the entries in Appendix Table 1. The "typical" half-lives, however, are calculated as $\ln(0.5)/\ln(1+\bar{\lambda})$, where $\bar{\lambda}$ is the appropriate typical λ.

[17]The results for Canada indicating implausibly slow convergence seem very puzzling and cast some doubt in our minds on the reliability of the Canadian data. The latter are collected from Statistics Canada, while for the U.S. the data are taken from the Bureau of Labor Statistics, and otherwise from Datastream. See Engel and Rogers (1995).

[18]As Evans and Savin (1981) point out, there could be small-sample biases in estimates of the convergence speed in AR or TAR models. This would be of the order (2+31)/T or about −0.01 in our sample. This could materially affect our longer half-life estiamtes:

It is also worth noting the implied threshold value: essentially a data-based estimate of transaction cost or, more generally, of "band of inaction" spreads. It ranges between 7.1% and 9.8% across the different types of goods (all-location averages), which is a very reasonable figure given the usual estimate of transaction costs derived from CIF-to-FOB ratios. Within each class of goods, the threshold shows predictable variation given the choice of a U.S. benchmark. Thresholds are below average in the U.S. and Canada (between 0.5% and 8%) and higher elsewhere. Thresholds are lower between the U.S. and Asia (2%–8%) than between the U.S. and Europe (9%–19%), which may come as no surprise given the increasingly Asia direction of U.S. trade in recent years and the general sense of Asia as a more open trading region with lower barriers to arbitrage.

Given that we are tempted already to relate our threshold estimates to measures of economic distance such as geographical distance between locations and transportation costs including trade barriers, Table 2(a) (upper panel) formally tests whether our implied thresholds relate to cross-sectional variation in selected measures of economic distance. These consist of distance (great circle in thousands of miles), exchange rate volatility (standard deviation of first-difference in log nominal dollar exchange rate, monthly), and summary data on tariff and nontariff barriers (fraction of value terms, from Lee and Swagel 1994).[19] Scatterplots summarize each of these bivariate regressions in Figures 3–6.

The results are somewhat encouraging. Almost always, measures of economic distance are positively related to threshold value. An exception in all

a half life of 23 months ($\lambda = -0.03$) would then be corrected to 34 months ($\lambda = -0.02$). However, this correction would matter less at the smaller half lives we typically find. And it is also worth pointing out that the reference point for our study is a PPP literature where this correction is almost never undertaken, though at smaller convergence speeds with smaller samples it could matter much more. For comparative purposes, we therefore present plain AR and TAR estimates in a form commensurate with the prevalent methodology.

[19]For this purpose, the U.S. price benchmarking requires a geographical choice of datum: quite arbitrarily, we measure distance from Chicago. See Appendix Table 5.

cases is the NTB (nontariff barrier) data which are known to be suspect.[20] The more robust of these correlations appear to be the relationships with distance and exchange-rate volatility. These findings are consistent with the Engel and Rogers' (1995) finding that the variability of deviations from PPP is positively related to distance and to exchange-rate volatility. Table 2(a) (lower panel) presents the results of multivariate cross-sectional regressions of estimated thresholds on the four market segmentation measures, but these results are less clear-cut, probably reflecting, to some extent, the well-known collinearity between distance and exchange-rate volatility, and a small sample size with several data clusters (see scatterplots). In this case, again, the tariff data provide little explanatory power.

We should note that, in the TAR model, adjustment is characterized by two parameters, the threshold and the convergence speed, and it is interesting to ask whether, like the threshold, convergence speed might also be related to measures of economic distance between location pairs. This is examined in Table 2(b) and Figures 7–10, following the same methodology used for thresholds above. Again, distance and exchange-rate volatility, but now also tariffs, show a consistent (inverse) relationship with convergence speed. As before, the multivariate results are less satisfactory.

However, we would not wish to overemphasize these findings, since a large degree of the correlation between thresholds and economic distance derives from a few cluster corrleations, as can be seen in Figures 3–6. North America has low thresholds and distance (relative to the U.S.), followed by most of East Asia with the next lowest thresholds and distances, then Europe lies beyond, with the higher thresholds and distances. Other countries like Japan, South Africa, New Zealand, and Mexico lie scattered around these clusters. Thus, further empirical work is needed to verify the thershold-

[20] Observe that all data on trade barriers are aggregative, and therefore far from ideal for studying the disaggregated price indices. Furthermore, as Lee and Swagel (1994) discuss, the NTB measures are "coverage ratios" that do not reflect the height of barriers, only their breadth.

distance relationship, and a larger sample of countries, or country pairs, would be desirable for that end, so as to fill in the gaps between the clusters.[21]

The exchange-rate volatility effect may well be due in part to reverse feedback. Not only does increased volatility make potential arbitrageurs less responsive to exchange-rate changes, as in models of trade with sunk costs mentioned earlier. In addition, low responsiveness itself may raise the amplitude of equilibrium exchange rate fluctuations, as suggested by Krugman (1989). Similarly, higher transaction costs probably increase the scope for exchange-rate fluctuations by impeding goods-market adjustment and promoting overshooting. These complex linkages deserve further study.

Specification tests based on Monte Carlo simulation (see Appendix A and Appendix Table 1, last page) sometimes reject the AR model in favor of the TAR model, but in most cases do not. This is perhaps not surprising, given the low power of relatively short univariate time series to distinguish among near unit root alternatives.

4 Conclusion

In this paper we have proposed that analysis of PPP and LOOP should take into account the possibility of "commodity points"—a region of no central tendency among relative prices, possibly due to an absence of arbitrage in the face of fixed transaction costs. This idea was indeed part of the refinement of PPP advanced by Heckscher (1916) (see Yeager 1976, chap. 11, especially pp. 217–218). To implement Heckscher's notion empirically, we devise an econometric method for estimating the commodity points and the speed of relative price adjustment outside the bands they define. Price adjustment is treated now as a nonlinear process. A threshold autoregression (TAR) offers an ele-

[21]However, a simple expedient such as trying all country pairs for the current sample seems unattractive for the reasons noted by Engel, Hendrickson, and Rogers (1996), and O'Connell (1996a): linearly dependent combinations of real exchange rates so derived contain no additional econometric information, and so tests of PPP such not be conducted with the use of such artifical data.

gant and economical specification for this purpose. Using TAR methods we can identify both thresholds and adjustment speeds by maximum likelihood methods. Nonparametric and Monte Carlo tests can be used to compare this TAR specification to the more standard AR specification used in empirical price adjustment analysis. Our TAR model yields empirical estimates of threshold and convergence speeds that seem quite reasonable. Adjustment outside the thresholds might imply half-lives of price deviations measured in months rather than years. And the thresholds themselves correspond to our intuitive or rough guesses as to the order of magnitude of transaction costs.

An empirical framework such as ours, based on costs to arbitraging tradable goods prices, is attractive for several reasons. It dovetails well with the observation of pricing-to-market behavior by international firms (see, for example, Knetter 1993). It also may help explain the related phenomenon of large swings in the international relative prices of even similar tradables. Recent theory suggests that if international arbitrage involves even relatively minor sunk investments, unhedgeable uncertainty can be levered up into relatively large and persistent international price differentials. Clearly the next step is to confront the data with more detailed models of profit-maximizing cross-border trade and arbitrage.

Transactions costs might also throw light on the "purchasing power parity puzzle" stressed by Rogoff (1996): How can international price convergence be so slow if the bulk of short-run PPP deviations is due to monetary shocks coupled with presumably transient price stickiness? Our results suggests that convergence is indeed rapid for large price discrepancies but that relatively small market frictions dramatically impede the final stages of relative price adjustment. To reconcile this type of story within a framework of long-run price flexibility clearly requires a general-equilibrium model, probably one based on imperfect competition and menu costs as well as costs of international trade.[22]

[22] A recent example of research in this mode is offered by Lapham (1995), who analyzes deviations from the law of one price for German manufacturing industries in a model with

We have also found some evidence of a positive correlation between our estimated "commodity points" and nominal exchange rate volatility. That result, if shown to be robust in other samples, is also suggestive of a sticky-price explanation. More generally, the finding that exchange rate volatility impedes international arbitrage is potentially strong evidence that floating exchange rates have indeed hampered international trade and an efficient international allocation of productive resources, as critics of market-determined exchange rates have long charged. Most studies of the effect of currency volatility on trade regress trade volume data on volatility (as well as on other explanatory variables; see, for example, Frankel and Wei 1993). A complementary and perhaps less ambiguous methodology could be based on international price differentials. Once again, however, reliable conclusions depend on the development of more detailed models of international commodity flows.

internationally segmented markets but flexible prices. A resolution of the PPP puzzle based on explicit modeling of the role of real and nominal shocks at short versus long horizons is discussed in Obstfeld (1997).

Appendices

Appendix A
General TAR Estimation Method

This appendix describes the basic TAR estimation method used in the paper. The object is to model an equation of the form

$$\Delta x_t = \begin{cases} \lambda^{out}(x_{t-1} - c) + e_t^{out} & \text{if } x_{t-1} > c; \\ \lambda^{in} x_{t-1} + e_t^{in} & \text{if } c \geq x_{t-1} \geq -c; \\ \lambda^{out}(x_{t-1} + c) + e_t^{out} & \text{if } -c > x_{t-1}; \end{cases} \tag{7}$$

where e_t^{out} is $N(0, {\sigma^{out}}^2)$, e_t^{in} is $N(0, {\sigma^{in}}^2)$, and we initially impose the restriction $\lambda^{in} = 0$.

Consider the likelihood function for the above TAR model as an alternative to an AR1 null,

$$\begin{aligned} L_a &= L_a(\lambda^{out}, \lambda^{in}, \sigma^{out}, \sigma^{in}; c) = \\ &- \textstyle\sum_{I_{in}(x_{t-1})=1} \frac{1}{2}(\log(2\pi) + \log({\sigma^{in}}^2) + {e_t^{in}}^2/{\sigma^{in}}^2) \\ &- \textstyle\sum_{I_{out}(x_{t-1})=1} \frac{1}{2}(\log(2\pi) + \log({\sigma^{out}}^2) + {e_t^{out}}^2/{\sigma^{out}}^2), \end{aligned} \tag{8}$$

where $I_{in}(x_{t-1}) = I(|x_{t-1}| \leq c)$, and $I_{out}(x_{t-1}) = I(|x_{t-1}| > c)$ are indicator functions which depend on the position of x_{t-1} being inside or outside the band.

The null AR1 model is

$$\Delta x_t = \lambda x_{t-1} + e_t, \tag{9}$$

where e_t is $N(0, \sigma^2)$, with a likelihood function

$$L_n = -\sum_t \frac{1}{2}(\log(2\pi) + \log(\sigma^2) + {e_t}^2/\sigma^2). \tag{10}$$

We define the likelihood ratio as $LLR = 2(L_a - L_n)$, and our objective is to maximize this ratio. We follow a search algorithm:

1. Find the 10th and 90th percentiles of $|x_t|$. Divide the intervening interval into steps of 0.001 width, marked by candidate thresholds $c_k; i = 1, ..., M$. These form candidate thresholds and, when the sample size is approximately $T = 180$, we are eliminating partitions with 18 or fewer observations in either part of the partition. Pick $k = 1$.

2. Chose c equal to kth highest value of c_k.

3. Partition the sample into observations inside and outside the band. Construct indicator values $I_{in}(x_{t-1})$ and $I_{out}(x_{t-1})$.

4. Calculate L_t as above, either by maximum likelihood, or, equivalently, by OLS on the partitioned samples. Then calculate $LLR = 2(L_a - L_n)$.

5. Increase k by one and go to step 2.

6. Locate the choice of k and c_k that maximizes LLR. Note the associated $(\lambda^{out}, \lambda^{in}, \sigma^{out}, \sigma^{in})$ estimated at step 3 for that choice of k. This is a maximum likelihood estimate of $(\lambda^{out}, \lambda^{in}, \sigma^{out}, \sigma^{in}, c)$.

Since c is not identified under the null, standard inference is invalid, and LLR does not follow the usual χ^2 distribution of likelihood ratio tests. An alternative is to follow Monte Carlo simulation methods. We proceed as follows:

1. Estimate the AR1 null model (9) on the actual data $(x_1, ..., x_T)$.

2. Generate $i = 1, ..., n$ simulations of this model. We chose $n = 600$. Start each with $x_{-b} = 0$, end at x_T, and discard the first b values $x_{-b}, ..., x_{-1}$ to avoid initial value bias. We chose $b = 50$.

3. For each simulation, estimate the TAR model as above and calculate the simulated LLR_i.

4. Calculate the empirical distribution of the LLR_i, and use this as the basis for inference in judging the alternative TAR model (7) against the AR1 null (9).

The TAR estimation results on our post-1980 monthly data are shown in full in Appendix Table 1. The last page of Appendix Table 1 (second set of columns) shows the Monte Carlo test results (in the form of p-values) for the TAR model against the AR1 null.

Appendix B
Sensitivity to Detrending

Note that our basic method described in Appendix A was applied to demeaned and detrended data (see main text). To investigate the sensitivity of our result to the preliminary detrending, we also ran the same procedure without initially detrending the data. The results were qualitatively similar, and are summarized in Appendix Table 2, which may be compared to Table 1 in the main text.

Appendix C
Sensitivity to Unit-Root Assumption

Note that our basic method described in Appendix A also imposes a unit root on the process inside the band. To investigate the sensitivity of our result to this assumption, we also ran the same procedure with λ^{in} not restricted to be zero in equation (2), with corresponding adjustments to the estimation procedure outlined above. The results were qualitatively similar, and are summarized in Appendix Table 3, which may be compared to Table 1 in the main text.

Appendix D
Sensitivity to Serial Correlation Assumption

In our models we have assumed that the error process is normal i.i.d. However, it is important to consider whether serial correlation in the residuals might change our threshold estimations dramatically. However, even applying the simple AR(1) model is not straightforward with serial correlation in the residuals: the true dynamics are an AR(2) process, but the serial correlation parameters are not uniquely identifiable.

A simple way to implement a TAR versus AR analysis in this context would be to just apply the models in the text with an AR correction to the residuals in each regime, but noting the above caveats. We tried this approach, and Appendix Table 4 reports the estimated thresholds for comparison with the baseline case. The usual Cochrane-Orcutt procedure was not available for the TAR case, of course, because changes of regime entail missing observations throughout the samples. Accordingly a grid search on the autoregressive parameter was employed, adding further econometric complexity. We wanted to know whether serial correlation would affect our threshold estimates in the TAR with and without the correction. No clear pattern emerged in the results. And as the summary Appendix Table 4 shows, the thresholds in each region looked very similar to those obtained without the serial-correlation correction.

Appendix E
Data

We are indebted to Charles Engel and John Rogers, who kindly provided the data from Engel and Rogers (1995) for our empirical analysis. The data consist of price series and exchange rates for 32 countries, covering nine categories of goods, at a monthly frequency from 1980 to 1995. There are missing observations. Our analysis only focuses on four goods categories that may be deemed "quasi-tradable": the composite price series we use correspond to CPI components for: (1) all goods; (2) clothing; (3) food; and, (4) fuel. We attempted our tests on the other "mostly-nontradable" CPI components in the dataset (health, household equipment, housing, recreation and education, and transportation) but the results were poor, and are not reported.

The data sources used by Engel and Rogers were as follows:

Prices from Datastream, except United States from Bureau of Labor Statistics, and Canada from Statistics Canada. All are index numbers.

Nominal exchange rates from IMF, *International Financial Statistics.* All are foreign currency units per U.S. dollar.

Having estimated our TAR model as in Appendix A, we also compared our threshold estimates to measures of economic distance, using data also compiled by Engel and Rogers. Appendix Table 5 reports these data. The thresholds are from Appendix Table 1. The monthly exchange rate volatility $\sigma(s)$ is derived from the above series. The other data were compiled by Engel and Rogers. Distance from Chicago is great circle distance, to capital city (except to specific U.S. and Canadian cities). Tariff and nontariff barriers (TB, NTB) are from Lee and Swagel (1994). For our purposes we estimate the total barriers for each pair (Chicago and location X) as being the sum of the component at each location, written $TB_{12} = TB_1 + TB_2$, and $NTB_{12} = NTB_1 + NTB_2$. These data form the basis of Tables 2 and 3 and Figures 3–6.

References

Asea, P. K., and E. G. Mendoza. 1994. The Balassa-Samuelson Model: A General-Equilibrium Appraisal. *Review of International Economics* 2 (October): 244–267.

Balke, N. S., and T. B. Fomby. 1997. Threshold Cointegration. *International Economic Review*. Forthcoming.

Canzoneri, M. B., R. E. Cumby, and B. Diba. 1996. Relative Labor Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel of OECD Countries. Working Paper Series no. 5676, National Bureau of Economic Research (July).

Chinn, M., and L. Johnston. 1996. Real Exchange Rate Levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries. Working Paper Series no. 5709, National Bureau of Economic Research (August).

Coleman, A. M. G. 1995. Arbitrage, Storage, and the 'Law of One Price': New Theory for the Time Series Analysis of an Old Problem. Princeton University (September). Photocopy.

Cumby, R. E. 1996. Forecasting Exchange Rates and Relative Prices with the Hamburger Standard: Is What You Want What You Get With McParity? Working Paper Series no. 5675, National Bureau of Economic Research (July).

Davidson, R., and J. G. MacKinnon. 1981. Several Tests for Model Specification in the Presence of Alternative Hypotheses. *Econometrica* 49 (May): 781–793.

Davutyan, N., and J. Pippenger. 1990. Testing Purchasing Power Parity: Some Evidence of the Effects of Transaction Costs. *Econometric Reviews* 9: 211–240.

De Grauwe, P., M. Janssens, and H. Leliaert. 1985. Real Exchange-Rate Variability from 1920 to 1926 and 1973 to 1982. Princeton Studies in International Finance no. 56 (September).

Dixit, A. K. 1989. Hysteresis, Import Penetration, and Exchange Rate Pass-Through. *Quarterly Journal of Economics* 104 (May): 205–228.

Dumas, B. 1992. Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World. *Review of Financial Studies* 5: 153–180.

Engel, C. 1995. Accounting for U.S. Real Exchange Rate Changes. Working Paper Series no. 5394, National Bureau of Economic Research (December).

Engel, C. 1996. Long-Run PPP May Not Hold After All. Working Paper Series no. 5646, National Bureau of Economic Research (January).

Engel, C., M. K. Hendrickson, and J. H. Rogers. 1996. Intra-National, Intra-Continental, and Intra-Planetary PPP. University of Washington (October). Photocopy.

Engel, C., and J. H. Rogers. 1995. Regional Patterns in the Law of One Price: The Roles of Geography vs. Currencies. Working Paper Series no. 5395, National Bureau of Economic Research (December).

Evans, G. B. A., and N. E. Savin. 1981. Testing for Unit Roots: 1. *Econometrica*

49 (May): 753–779.

Fanizza, D. G. 1990. Multiple Steady States and Coordination Failures in Search Equilibrium: New Approaches to the Business Cycle. Unpublished doctoral dissertation, Northwestern University.

Frankel, J. A., and A. K. Rose. 1996. A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries. *Journal of International Economics* 40 (February): 209–224.

Frankel, J. A., and S.-J. Wei. 1993. Trade Blocs and Currency Blocs. In *The Monetary Future of Europe*, edited by G. de la Dehesa, A. Giovannini, M. Gutián, and R. Portes. London: Centre for Economic Policy Research.

Froot, K. A., M. Kim, and K. Rogoff. 1995. The Law of One Price over 700 Years. Working Paper Series no. 5132, National Bureau of Economic Research (May).

Froot, K. A., and K. Rogoff. 1995. Perspectives on PPP and Long-Run Real Exchange Rates. In *Handbook of International Economics*, vol. 3, edited by G. Grossman and K. Rogoff. Amsterdam: Elsevier.

Hakkio, C. S. 1992. Is Purchasing Power Parity a Useful Guide to the Dollar? *Federal Reserve Bank of Kansas City Economic Review* 77 (Third Quarter): 37–51.

Heckscher, E. F. 1916. Växelkursens Grundval vid Pappersmyntfot. *Ekonomisk Tidskrift* 18 (October): 309–312.

Hegwood, N. D., and D. H. Papell. 1996. Quasi Purchasing Power Parity. University of Houston (August). Photocopy.

Jorion, P., and R. J. Sweeney. 1996. Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting. *Journal of International Money and Finance* 15 (August): 535–550.

Knetter, M. M. 1993. International Comparisons of Pricing-to-Market Behavior. *American Economic Review* 83 (June): 473–486.

Krugman, P. R. 1989. *Exchange Rate Instability.* Cambridge, MA: MIT Press.

Lapham, B. J. 1995. A Dynamic General Equilibrium Analysis of Deviations from the Laws of One Price. *Journal of Economic Dynamics and Control* 19 (November): 1355–1389.

Lee, J.-W., and P. Swagel. 1994. Trade Barriers and Trade Flows across Countries and Industries. Working Paper Series no. 4799, National Bureau of Economic Research (July).

Lothian, J. R. 1997. Multi-Country Evidence on the Behavior of Purchasing Power Parity under the Current Float. *Journal of International Money and Finance* 16 (February): 19–35.

Lothian, J. R., and M. P. Taylor. 1996. Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries. *Journal of Political Economy* 104 (June): 488–509.

Marston, R. C. 1987. Real Exchange Rates and Productivity Growth in the United States and Japan. In *Real-Financial Linkages among Open Economies*, edited by S. W. Arndt and J. D. Richardson. Cambridge, MA: MIT Press.

Michael, P., A. R. Nobay, and D. A. Peel. 1997. Transaction Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation. *Journal of Political Economy*. Forthcoming.

Obstfeld, M. 1993. Model Trending Real Exchange Rates. Working Paper no. C93–011, Center for International and Development Economics Research, University of California, Berkeley (February).

Obstfeld, M. 1997. Open-Economy Macroeconomics: Developments in Theory and Policy. University of California, Berkeley (May). Photocopy.

O'Connell, P. G. J. 1996a. The Overvaluation of Purchasing Power Parity. *Journal of International Economics*. Forthcoming.

O'Connell, P. G. J. 1996b. Market Frictions and Relative Traded Goods Prices. Harvard University (November). Photocopy.

Oh, K.-Y. 1996. Purchasing Power Parity and Unit Root Tests Using Panel Data. *Journal of International Money and Finance* 15 (June): 405–418.

Parsley, D. C., and S.-J. Wei. 1996. Convergence to the Law of One Price without Trade Barriers or Currency Fluctuations. *Quarterly Journal of Economics* 111 (November): 1211–1236.

Phelps, E. S., and S. G. Winter. 1970. Optimal Price Policy under Atomistic Competition. In *Microeconomic Foundations of Employment and Inflation Theory*, edited by E. S. Phelps. New York: Norton

Pippenger, M. K., and G. E. Goering. 1993. A Note on the Empirical Power of Unit Root Tests Under Threshold Processes. *Oxford Bulletin of Economics and Statistics* 55: 473–481.

Prakash, G. 1996. Pace of Market Integration. Northwestern University (September). Photocopy.

Prakash, G., and A. M. Taylor. 1997. Measuring Market Integration: Models of Arbitrage with an Econometric Application to the Gold Standard. Northwestern University (January). Photocopy.

Rogoff, K. 1996. The Purchasing Power Parity Puzzle. *Journal of Economic Literature* 34 (June): 647–668.

Rotemberg, J. J., and M. Woodford. 1991. Markups and the Business Cycle. *NBER Macroeconomics Annual* 6: 63–129.

Sercu, P., R. Uppal, and C. Van Hulle. 1995. The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity. *Journal of Finance* 50 (September): 1309–1319.

Taylor, A. M. 1996. International Capital Mobility in History: Purchasing Power Parity in the Long-Run. Working Paper Series no. 5742, National Bureau of Economic Research (August).

Tsay, R. S. 1989. Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association* 84 (March): 231–240.

Williams, J. C., and B. D. Wright. 1991. *Storage and Commodity Markets.* Cambridge, U. K.: Cambridge University Press.

Wei, S.-J., and D. C. Parsley. 1995. Purchasing Power *Dis*-Parity during the Floating Rate Period. Harvard University and Vanderbilt University (May). Photocopy.

Wu, Y. 1996. Are Real Exchange Rates Nonstationary? Evidence from a Panel-Data Test. *Journal of Money, Credit, and Banking* 28 (February): 54–63.

Yeager, L. B. 1976. *International Monetary Relations: Theory, History, and Policy.* 2d ed. New York: Harper and Row.

Table 1 **Summary Results for Threshold Model: Baseline Case**

	TAR c	TAR λ	TAR half-life	AR λ	AR half-life
CPI-All					
U.S. cities	0.005	-0.242	2.5	-0.132	4.9
Canada	0.045	-0.053	12.8	-0.003	252.3
Mexico	0.040	-0.036	19.1	-0.032	21.4
Europe	0.089	-0.054	12.6	-0.032	21.4
Asia	0.072	-0.061	11.1	-0.017	41.1
All	0.071	-0.083	8.0	-0.038	17.8
CPI-Clothing					
U.S. cities	0.011	-0.316	1.8	-0.245	2.5
Canada	0.041	-0.042	16.2	-0.008	82.8
Mexico	0.016	-0.056	11.9	-0.052	13.0
Europe	0.186	-0.213	2.9	-0.045	15.2
Asia	0.066	-0.113	5.8	-0.045	14.9
All	0.078	-0.157	4.1	-0.077	8.7
CPI-Food					
U.S. cities	0.021	-0.203	3.1	-0.056	12.0
Canada	0.080	-0.060	11.1	-0.004	167.6
Mexico	0.038	-0.041	16.8	-0.036	18.8
Europe	0.120	-0.088	7.6	-0.030	22.9
Asia	0.078	-0.196	3.2	-0.026	26.4
All	0.098	-0.116	5.6	-0.028	24.3
CPI-Fuel					
U.S. cities	0.030	-0.410	1.3	-0.228	2.7
Canada	0.042	-0.138	4.7	-0.054	12.4
Mexico					
Europe	0.105	-0.150	4.3	-0.040	16.8
Asia	0.022	-0.035	19.5	-0.030	23.0
All	0.072	-0.191	3.3	-0.079	8.4

Notes: The benchmark is versus the U.S. national average price (or four-city average, if no national price is available). See text and Appendix Table 1. Region data is simple average, except half lives, which are computed from region average convergence speed (λ) by the formula half-life = ln(0.5)/ln(1+λ).
Source: Appendix Table 1.

Table 2 **Estimated Thresholds and Convergence Speeds Versus Measures of Economic Distance**

a. Thresholds (c)

Bivariate regression: (c) on	Slope coefficient (absolute t-statistic) Number of Observations			
	All	Clothing	Food	Fuel
Distance (/1000, miles)	0.008 (1.9)	0.012 (1.9)	0.007 (1.6)	0.013 (1.8)
	29	19	30	21
Exchange-rate Volatility ($\sigma(\Delta s)$)	1.8 (2.3)	2.1 (2.0)	1.8 (1.9)	2.6 (2.5)
	30	20	31	21
Tariff Barriers (TB12)	0.005 (1.2)	0.002 (0.3)	−0.002 (0.4)	−0.011 (1.4)
	29	19	29	21
Non-Tariff Barriers (NTB12)	−0.000 (0.1)	−0.000 (0.4)	−0.002 (3.1)	−0.000 (0.6)
	29	19	29	21
Multivariate regression: (c) on	**Slope coefficient (absolute t-statistic)**			
	All	Clothing	Food	Fuel
Distance (/1000, miles)	0.009 (2.0)	0.011 (1.5)	0.007 (1.2)	−0.034 (1.8)
Exchange-rate Volatility ($\sigma(\Delta s)$)	0.64 (0.7)	1.6 (1.1)	0.3 (0.3)	7.8 (2.5)
Tariff Barriers (TB12)	0.007 (1.5)	0.004 (0.4)	0.003 (0.4)	−0.013 (1.9)
Non-Tariff Barriers (NTB12)	0.000 (0.0)	0.000 (0.2)	−0.002 (1.7)	0.001 (0.9)
R^2	.27	.31	.21	.43
Number of Observations	29	19	29	21

b. Convergence Speed ($-\lambda$)

Bivariate regression: ($-\lambda$) on	Slope coefficient (absolute t-statistic) Number of Observations			
	All	Clothing	Food	Fuel
Distance (/1000, miles)	0.013 (2.2)	0.003 (0.2)	0.008 (0.9)	0.033 (1.3)
	29	19	30	21
Exchange-rate Volatility ($\sigma(\Delta s)$)	2.9 (2.9)	2.0 (1.1)	2.5 (1.6)	5.8 (1.6)
	30	20	31	21
Tariff Barriers (TB12)	0.001 (0.1)	0.008 (0.6)	0.022 (3.0)	0.050 (2.1)
	29	19	29	21
Non-Tariff Barriers (NTB12)	−0.004 (5.1)	−0.006 (3.5)	−0.000 (0.2)	−0.007 (2.0)
	29	19	29	21
Multivariate regression: ($-\lambda$) on	**Slope coefficient (absolute t-statistic)**			
	All	Clothing	Food	Fuel
Distance (/1000, miles)	0.012 (2.8)	0.009 (0.9)	−0.009 (1.0)	0.091 (1.5)
Exchange-rate Volatility ($\sigma(\Delta s)$)	1.0 (1.2)	−2.5 (1.2)	2.9 (1.7)	−9.5 (1.0)
Tariff Barriers (TB12)	0.010 (2.4)	0.024 (2.0)	0.017 (1.9)	0.059 (2.6)
Non-Tariff Barriers (NTB12)	−0.004 (5.6)	−0.007 (3.9)	−0.001 (0.3)	−0.007 (1.9)
R^2	.73	.55	.36	.46
Number of Observations	29	19	29	21

Notes: Chicago is the benchmark location for distances. For U.S. and Canada we use New York and Ottawa as locations for distance measurement. TB12 and NTB12 are sums of TB and NTB (respectively) at Chicago and the other location.
Source: See Appendix Table 4.

Appendix Table 1 Detailed Results for Threshold Model

CPI-All

	T	LLR	TAR c	TAR λ	TAR t	AR λ	AR t	TAR half-life	AR half-life
U.S.									
Chicago	180	12.0	0.001	-0.332	5.5	-0.310	5.7	1.7	1.9
Los Angeles	180	13.9	0.010	-0.230	3.1	-0.053	2.1	2.6	12.8
New York	180	13.0	0.008	-0.261	3.1	-0.078	2.8	2.3	8.6
Philadelphia	180	6.8	0.003	-0.146	3.5	-0.085	2.8	4.4	7.8
Canada	179	4.8	0.012	-0.001	0.1	-0.002	0.1	737.7	408.7
Ottawa	180	6.5	0.009	-0.001	0.1	-0.002	0.1	589.6	407.1
Toronto	180	6.6	0.114	-0.156	1.5	0.003	0.3	4.1	-223.4
Vancouver	180	5.0	0.069	-0.099	1.5	-0.009	0.6	6.6	80.8
Winnipeg	180	5.1	0.020	-0.005	0.3	-0.005	0.4	127.3	141.6
Mexico	180	34.2	0.040	-0.036	2.1	-0.032	1.8	19.1	21.4
Austria	179	7.5	0.059	-0.050	2.3	-0.039	2.7	13.6	17.3
Belgium	180	3.9	0.082	-0.048	2.7	-0.035	2.8	14.0	19.3
Denmark	179	3.1	0.090	-0.049	2.2	-0.033	2.4	13.8	20.9
Finland	179	4.3	0.158	-0.071	1.7	-0.019	1.4	9.5	36.7
France	179	4.8	0.035	-0.039	2.4	-0.034	2.5	17.4	20.1
Germany	180	6.7	0.115	-0.063	2.3	-0.040	2.8	10.6	16.8
Greece	180	5.2	0.026	-0.041	2.4	-0.038	2.5	16.5	18.0
Italy	178	5.5	0.175	-0.107	1.9	-0.027	2.0	6.1	25.1
Netherlands	179	6.0	0.102	-0.057	2.2	-0.038	2.6	11.9	17.9
Norway	179	3.0	0.083	-0.048	1.7	-0.026	1.8	14.0	26.3
Portugal	180	5.6	0.142	-0.057	1.7	-0.028	2.1	11.9	24.7
Spain	179	4.4	0.023	-0.031	2.4	-0.027	2.4	22.0	25.0
Sweden	179	7.4	0.026	-0.027	2.0	-0.023	2.0	25.2	29.6
Switzerland	180	4.1	0.057	-0.057	2.3	-0.045	2.7	11.8	15.0
U.K.	179	8.0	0.169	-0.058	1.1	-0.026	1.7	11.5	26.8
Hongkong	180	21.7	0.099	-0.053	2.6	-0.013	1.5	12.6	54.0
Japan	179	2.9	0.033	-0.041	1.8	-0.031	1.8	16.4	21.9
Singapore	180	3.0	0.012	-0.008	0.5	-0.007	0.6	86.7	94.3
Taiwan	180	5.8	0.143	-0.140	1.8	-0.016	1.3	4.6	43.6
New Zealand									
South Africa	180	76.9	0.225	-0.187	1.6	-0.027	1.6	3.4	25.3
Panel	5387	77.4	0.029	-0.073	12.5	-0.058	13.4	9.1	11.6

Appendix Table 1 Detailed Results for Threshold Model

CPI-Clothing

	T	LLR	TAR c	TAR λ	TAR t	AR λ	AR t	TAR half-life	AR half-life
U.S.									
Chicago	180	1.0	0.007	-0.328	5.5	-0.283	5.4	1.7	2.1
Los Angeles	180	14.0	0.008	-0.232	4.3	-0.179	4.1	2.6	3.5
New York	180	10.1	0.022	-0.431	4.6	-0.266	5.1	1.2	2.2
Philadelphia	180	2.8	0.007	-0.274	4.9	-0.250	5.0	2.2	2.4
Canada	180	8.4	0.086	-0.143	1.6	-0.009	0.8	4.5	74.4
Ottawa	180	12.1	0.025	-0.007	0.4	-0.005	0.4	94.8	136.5
Toronto	180	5.8	0.019	-0.005	0.3	-0.004	0.3	133.2	181.9
Vancouver	180	10.9	0.058	-0.036	1.1	-0.010	0.7	18.8	66.0
Winnipeg	180	17.5	0.016	-0.018	1.0	-0.013	0.8	39.0	52.7
Mexico	180	29.5	0.016	-0.056	2.5	-0.052	2.2	11.9	13.0
Austria									
Belgium									
Denmark									
Finland									
France									
Germany									
Greece									
Italy	180	5.1	0.238	-0.394	2.2	-0.020	1.4	1.4	34.2
Netherlands	179	0.6	0.242	-0.169	1.5	-0.072	2.7	3.7	9.3
Norway	180	9.7	0.114	-0.112	1.6	-0.057	2.3	5.8	11.8
Portugal									
Spain									
Sweden	168	11.8	0.167	-0.343	3.6	-0.050	2.3	1.6	13.5
Switzerland									
U.K.	180	2.6	0.169	-0.044	1.1	-0.024	1.7	15.3	28.2
Hongkong	180	11.2	0.025	-0.074	2.7	-0.054	2.3	9.0	12.5
Japan	180	3.0	0.032	-0.094	2.6	-0.075	2.7	7.0	8.8
Singapore	180	6.1	0.127	-0.111	0.7	-0.008	0.7	5.9	84.7
Taiwan	180	7.9	0.080	-0.173	2.1	-0.044	2.0	3.6	15.3
New Zealand									
South Africa	180	77.6	0.098	-0.085	1.7	-0.053	2.2	7.8	12.6
Panel	3587	231.1	0.080	-0.163	9.2	-0.081	12.4	3.9	8.2

Appendix Table 1 Detailed Results for Threshold Model

CPI-Food

	T	LLR	TAR c	TAR λ	TAR t	AR λ	AR t	TAR half-life	AR half-life
U.S.									
Chicago	180	4.1	0.006	-0.075	2.4	-0.055	2.3	8.9	12.4
Los Angeles	180	5.0	0.016	-0.187	3.7	-0.084	3.3	3.3	7.9
New York	180	6.7	0.042	-0.387	2.4	-0.028	1.6	1.4	24.7
Philadelphia	180	8.2	0.020	-0.161	3.3	-0.057	2.8	3.9	11.7
Canada	179	8.7	0.041	-0.005	0.2	-0.003	0.2	138.9	262.2
Ottawa	180	7.7	0.095	-0.015	0.3	-0.006	0.4	45.7	110.8
Toronto	180	6.4	0.124	0.022	0.4	0.003	0.2	-31.7	-254.9
Vancouver	180	4.2	0.030	-0.017	0.7	-0.009	0.6	39.8	80.7
Winnipeg	180	5.0	0.110	-0.287	1.6	-0.006	0.4	2.1	116.5
Mexico	180	50.1	0.038	-0.041	1.6	-0.036	1.8	16.8	18.8
Austria	179	7.8	0.077	-0.054	2.0	-0.038	2.4	12.4	18.0
Belgium	180	4.3	0.112	-0.058	2.0	-0.035	2.4	11.6	19.6
Denmark	179	2.0	0.078	-0.053	2.0	-0.033	2.1	12.8	20.5
Finland	179	4.8	0.190	-0.085	1.5	-0.014	1.0	7.8	48.6
France	179	3.0	0.015	-0.034	2.2	-0.031	2.1	19.9	21.7
Germany	180	3.8	0.071	-0.049	2.0	-0.038	2.5	13.7	18.1
Greece	180	6.8	0.146	-0.111	2.1	-0.053	2.4	5.9	12.8
Italy	178	5.0	0.190	-0.140	1.8	-0.023	1.6	4.6	29.9
Netherlands	179	4.5	0.093	-0.052	1.9	-0.036	2.3	12.9	19.1
Norway	179	5.7	0.183	-0.271	1.8	-0.020	1.3	2.2	33.8
Portugal	179	6.2	0.056	-0.042	1.8	-0.028	1.8	16.1	24.3
Spain	179	5.1	0.249	-0.170	2.5	-0.023	1.9	3.7	29.8
Sweden	179	10.9	0.163	-0.091	2.2	-0.013	1.1	7.3	51.7
Switzerland	179	5.3	0.024	-0.045	1.9	-0.038	2.0	15.1	17.7
U.K.	179	5.3	0.156	-0.057	1.1	-0.025	1.6	11.7	27.9
Hongkong	180	22.7	0.144	-0.294	3.9	-0.011	1.0	2.0	63.1
Japan	179	1.7	0.036	-0.053	2.0	-0.039	1.9	12.8	17.4
Singapore	180	4.7	0.104	-0.389	1.8	-0.013	0.8	1.4	51.0
Taiwan	180	7.2	0.026	-0.050	1.8	-0.040	1.9	13.7	17.0
New Zealand	179	30.5	0.107	-0.068	1.7	-0.023	1.4	9.8	30.3
South Africa	180	58.8	0.306	-0.289	1.9	-0.018	1.2	2.0	37.2
Panel	5564	39.5	0.019	-0.069	13	-0.058	13.5	9.8	11.5

Appendix Table 1 Detailed Results for Threshold Model

CPI-Fuel

	T	LLR	TAR c	TAR λ	TAR t	AR λ	AR t	TAR half-life	AR half-life
U.S.									
Chicago	179	17.1	0.052	-0.727	3.5	-0.385	6.5	0.5	1.4
Los Angeles	179	16.1	0.034	-0.538	3.6	-0.261	5.1	0.9	2.3
New York	179	7.9	0.026	-0.118	1.8	-0.050	2.0	5.5	13.5
Philadelphia	179	4.2	0.007	-0.257	4.7	-0.218	4.6	2.3	2.8
Canada	179	11.2	0.071	-0.146	2.4	-0.037	1.7	4.4	18.6
Ottawa	179	10.0	0.081	-0.236	3.1	-0.046	1.8	2.6	14.7
Toronto	179	13.9	0.007	-0.066	2.3	-0.060	2.2	10.2	11.2
Vancouver	179	7.1	0.029	-0.181	3.6	-0.083	2.7	3.5	8.0
Winnipeg	179	7.7	0.023	-0.063	2.2	-0.046	2.0	10.7	14.7
Mexico									
Austria	179	5.0	0.052	-0.054	2.2	-0.040	2.3	12.6	16.8
Belgium	179	5.4	0.066	-0.061	2.3	-0.043	2.5	11.1	15.8
Denmark	179	5.6	0.033	-0.029	1.8	-0.025	1.8	23.4	26.9
Finland	179	4.1	0.030	-0.025	1.6	-0.022	1.6	27.1	31.4
France	179	6.2	0.053	-0.050	2.4	-0.040	2.6	13.5	17.1
Germany	179	6.7	0.043	-0.057	2.7	-0.047	2.9	11.7	14.5
Greece									
Italy	178	13.2	0.178	-0.131	1.7	-0.043	2.5	4.9	15.9
Netherlands	179	5.2	0.198	-0.516	3.1	-0.059	2.6	1.0	11.4
Norway	179	1.4	0.122	-0.049	1.8	-0.025	2.0	13.7	26.9
Portugal									
Spain									
Sweden									
Switzerland	179	8.4	0.250	-0.627	3.5	-0.075	3.3	0.7	8.9
U.K.	179	7.0	0.133	-0.047	1.2	-0.024	1.6	14.4	28.1
Hongkong									
Japan	179	12.8	0.022	-0.035	1.8	-0.030	1.6	19.5	23.0
Singapore									
Taiwan									
New Zealand									
South Africa									
Panel	3758	29.3	0.077	-0.164	11.3	-0.081	13.2	3.9	8.2

Appendix Table 1 Detailed Results for Threshold Model: Specification Tests

	Tsay's Nonlinearity Test TAR versus AR minimal p-value of F statistic (ascending & descending orders)				Monte Carlo Simulation TAR versus AR empirical p-value of LLR statistic (600 draws)			
	All	Clothing	Food	Fuel	All	Clothing	Food	Fuel
U.S.								
Chicago	1.00	0.63	0.56	0.12	0.11	0.91	0.71	0.11
Los Angeles	0.12	0.37	0.77	0.15	0.01	0.11	0.51	0.11
New York	0.01	0.80	0.51	0.39	0.01	0.11	0.41	0.31
Philadelphia	0.71	0.55	0.23	0.55	0.31	0.71	0.21	0.51
Canada	1.00	0.75	0.77	0.87	0.71	0.21	0.31	0.11
Ottawa	0.97	0.20	0.81	0.54	0.51	0.11	0.31	0.11
Toronto	0.77	1.00	0.47	0.90	0.61	0.61	0.51	0.01
Vancouver	0.81	0.20	1.00	0.28	0.61	0.11	0.81	0.31
Winnipeg	0.76	0.18	0.36	0.47	0.71	0.01	0.61	0.31
Mexico	0.29	0.00	0.09		0.01	0.01	0.01	
Austria	0.05		0.06	0.04	0.31		0.31	0.61
Belgium	0.05		0.34	0.38	0.71		0.71	0.51
Denmark	0.17		0.15	0.49	0.81		1.01	0.51
Finland	0.74		0.98	0.52	0.71		0.71	0.71
France	0.05		0.09	0.47	0.61		0.81	0.41
Germany	0.04		0.03	0.27	0.41		0.71	0.41
Greece	0.06		0.72		0.61		0.41	
Italy	0.29	0.23	0.17	0.98	0.61	0.61	0.61	0.11
Netherlands	0.07	0.99	0.19	0.22	0.51	1.01	0.71	0.51
Norway	0.54	0.44	0.63	0.45	0.91	0.11	0.51	1.01
Portugal	0.70		0.95		0.51		0.51	
Spain	0.18		0.21		0.71		0.61	
Sweden	0.53	0.99	0.49		0.31	0.11	0.11	
Switzerland	0.01		0.09	0.17	0.71		0.61	0.21
U.K.	0.17	0.22	0.19	0.13	0.31	0.91	0.61	0.41
Hongkong	0.17	0.33	0.44		0.01	0.11	0.01	
Japan	0.80	0.91	1.00	0.89	0.91	0.81	0.91	0.01
Singapore	0.69	0.78	0.41		0.91	0.51	0.71	
Taiwan	0.86	0.81	0.33		0.51	0.31	0.41	
New Zealand			0.70				0.01	
South Africa	0.31	0.53	0.66		0.01	0.01	0.01	
Panel	0.02	0.01	0.17	0.01	0.21	0.21	0.21	0.21

Source: Data are described Appendix E.
Notes: See text. See Appendix A on estimation of TAR model. For the panel estiamtes the data were prefiltered by regression on time dummies to control for cross-sectional dependence.

Appendix Table 2 Summary Results for Threshold Model: No Preliminary Detrending of Data

	TAR c	TAR λ	TAR half-life	AR λ	AR half-life
CPI-All					
U.S. cities	0.016	-0.194	3.2	-0.075	8.9
Canada	0.091	-0.186	3.4	-0.007	93.0
Mexico	0.033	-0.034	19.9	-0.030	22.7
Europe	0.191	-0.140	4.6	-0.017	40.2
Asia	0.152	-0.058	11.7	0.000	13408.4
All	0.139	-0.138	4.7	-0.022	31.5
CPI-Clothing					
U.S. cities	0.078	-0.276	2.1	-0.094	7.0
Canada	0.097	-0.086	7.7	-0.010	67.8
Mexico	0.040	-0.047	14.3	-0.037	18.2
Europe	0.251	-0.290	2.0	-0.020	34.7
Asia	0.252	-0.046	14.7	0.000	2932.4
All	0.170	-0.175	3.6	-0.030	22.7
CPI-Food					
U.S. cities	0.141	-0.031	21.9	-0.005	139.7
Canada	0.109	-0.028	24.5	-0.010	66.9
Mexico	0.129	-0.041	16.6	-0.020	34.4
Europe	0.349	-0.213	2.9	-0.003	223.0
Asia	0.194	-0.035	19.6	0.001	-1064.9
All	0.253	-0.123	5.3	-0.005	147.3
CPI-Fuel					
U.S. cities	0.029	-0.200	3.1	-0.160	4.0
Canada	0.071	-0.066	10.1	-0.024	28.4
Mexico					
Europe	0.166	-0.143	4.5	-0.032	21.1
Asia	0.216	-0.038	17.7	-0.010	71.3
All	0.120	-0.131	4.9	-0.054	12.6

Notes: See Appendix B.

Appendix Table 3 Summary Results for Threshold Model: No Unit Root Restriction Inside Band

	TAR c	TAR λ	TAR half-life	AR λ	AR half-life
CPI-All					
U.S. cities	0.006	-0.262	2.3	-0.132	4.9
Canada	0.034	-0.035	19.7	-0.003	252.3
Mexico	0.040	-0.036	19.1	-0.032	21.4
Europe	0.127	-0.081	8.2	-0.032	21.4
Asia	0.077	-0.062	10.8	-0.017	41.1
All	0.088	-0.096	6.9	-0.038	17.8
CPI-Clothing					
U.S. cities	0.023	-0.370	1.5	-0.245	2.5
Canada	0.042	-0.042	16.2	-0.008	82.8
Mexico	0.016	-0.056	11.9	-0.052	13.0
Europe	0.161	-0.210	2.9	-0.045	15.2
Asia	0.068	-0.115	5.7	-0.045	14.9
All	0.074	-0.167	3.8	-0.077	8.7
CPI-Food					
U.S. cities	0.021	-0.203	3.1	-0.056	12.0
Canada	0.068	-0.021	33.3	-0.004	167.6
Mexico	0.038	-0.041	16.8	-0.036	18.8
Europe	0.120	-0.086	7.7	-0.030	22.9
Asia	0.077	-0.197	3.2	-0.026	26.4
All	0.096	-0.109	6.0	-0.028	24.3
CPI-Fuel					
U.S. cities	0.034	-0.443	1.2	-0.228	2.7
Canada	0.043	-0.142	4.5	-0.054	12.4
Mexico					
Europe	0.101	-0.125	5.2	-0.040	16.8
Asia	0.022	-0.035	19.5	-0.030	23.0
All	0.071	-0.185	3.4	-0.079	8.4

Notes: See Appendix C.

Appendix Table 4 Summary Results for Threshold Model: With and without Serial Correlation Correction

	Baseline TAR c	Corrected TAR c
CPI-All		
U.S. cities	0.006	0.005
Canada	0.034	0.045
Mexico	0.040	0.040
Europe	0.127	0.089
Asia	0.077	0.072
All	0.088	0.071
CPI-Clothing		
U.S. cities	0.023	0.011
Canada	0.042	0.041
Mexico	0.016	0.016
Europe	0.161	0.186
Asia	0.068	0.066
All	0.074	0.078
CPI-Food		
U.S. cities	0.021	0.021
Canada	0.068	0.080
Mexico	0.038	0.038
Europe	0.120	0.120
Asia	0.077	0.078
All	0.096	0.098
CPI-Fuel		
U.S. cities	0.034	0.030
Canada	0.043	0.042
Mexico	—	—
Europe	0.101	0.105
Asia	0.022	0.022
All	0.071	0.072

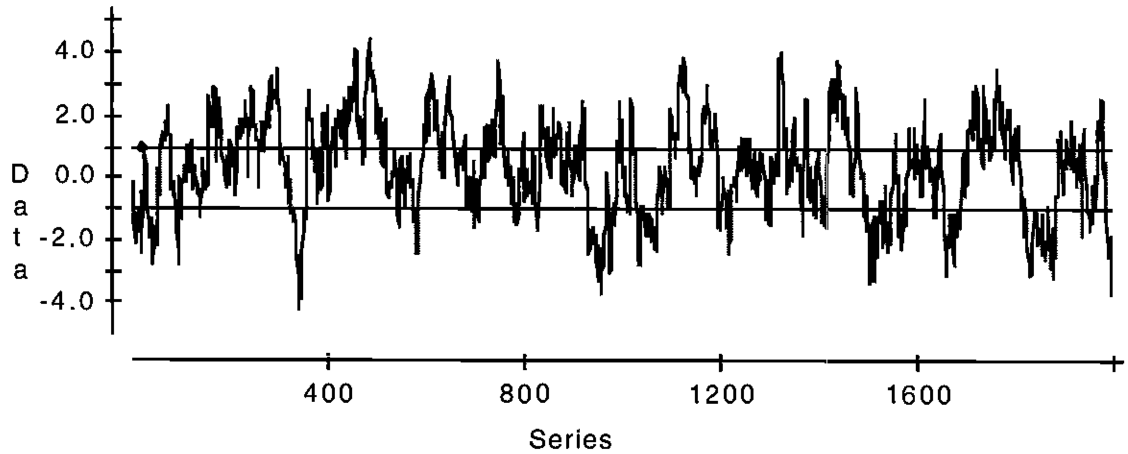
Notes: See Appendix D.

Appendix Table 5 Measures of Economic Distance and Estimated Thresholds

Location	Distance	$\sigma(\Delta s)$	TB	NTB	c All	c Clothing	c Food	c Fuel
U.S.	0	0.000	4.9	37.5				
Chicago	0	0.000	4.9	37.5	0.001	0.007	0.006	0.052
Los Angeles	1736	0.000	4.9	37.5	0.010	0.008	0.016	0.034
New York	716	0.000	4.9	37.5	0.008	0.022	0.042	0.026
Philadelphia	668	0.000	4.9	37.5	0.003	0.007	0.020	0.007
Canada	649	0.010	7.4	4.1	0.012	0.086	0.041	0.071
Ottawa	649	0.010	7.4	4.1	0.009	0.025	0.095	0.081
Toronto	441	0.010	7.4	4.1	0.114	0.019	0.124	0.007
Vancouver	1762	0.010	7.4	4.1	0.069	0.058	0.030	0.029
Winnipeg	716	0.010	7.4	4.1	0.020	0.016	0.110	0.023
Mexico	1686	0.058	9.65	7.38	0.040	0.016	0.038	
Austria	4688	0.030	4.3	7.7	0.059		0.077	0.052
Belgium	4143	0.030	7	19.6	0.082		0.112	0.066
Denmark	4254	0.029	7.1	18.2	0.090		0.078	0.033
Finland	4434	0.028	4.3	7.7	0.158		0.190	0.030
France	4135	0.029	7.4	18.4	0.035		0.015	0.053
Germany	4249	0.030	7.4	22.3	0.115		0.071	0.043
Greece	5435	0.028	7	25.5	0.026		0.146	
Italy	4814	0.029	7.6	20.9	0.175	0.238	0.190	0.178
Netherlands	4096	0.029	7.1	20.6	0.102	0.242	0.093	0.198
Norway	4042	0.025	4.3	7.7	0.083	0.114	0.183	0.122
Portugal	3995	0.029	7.1	19.1	0.142		0.056	
Spain	4182	0.028	6.8	13.9	0.023		0.249	
Sweden	4278	0.027	4.3	7.7	0.026	0.167	0.163	
Switzerland	4405	0.033		7.7	0.057		0.024	0.250
U.K.	3950	0.029	7.1	17.9	0.169	0.169	0.156	0.133
Hongkong	7786	0.012	0	0	0.099	0.025	0.144	
Japan	6293	0.029	4.9	11.3	0.033	0.032	0.036	0.022
Singapore	9364	0.011	0.9	1.1	0.012	0.127	0.104	
Taiwan	7443	0.009	10.6	45.7	0.143	0.080	0.026	
New Zealand	8678	0.027					0.107	
South Africa		0.036			0.225	0.098	0.306	

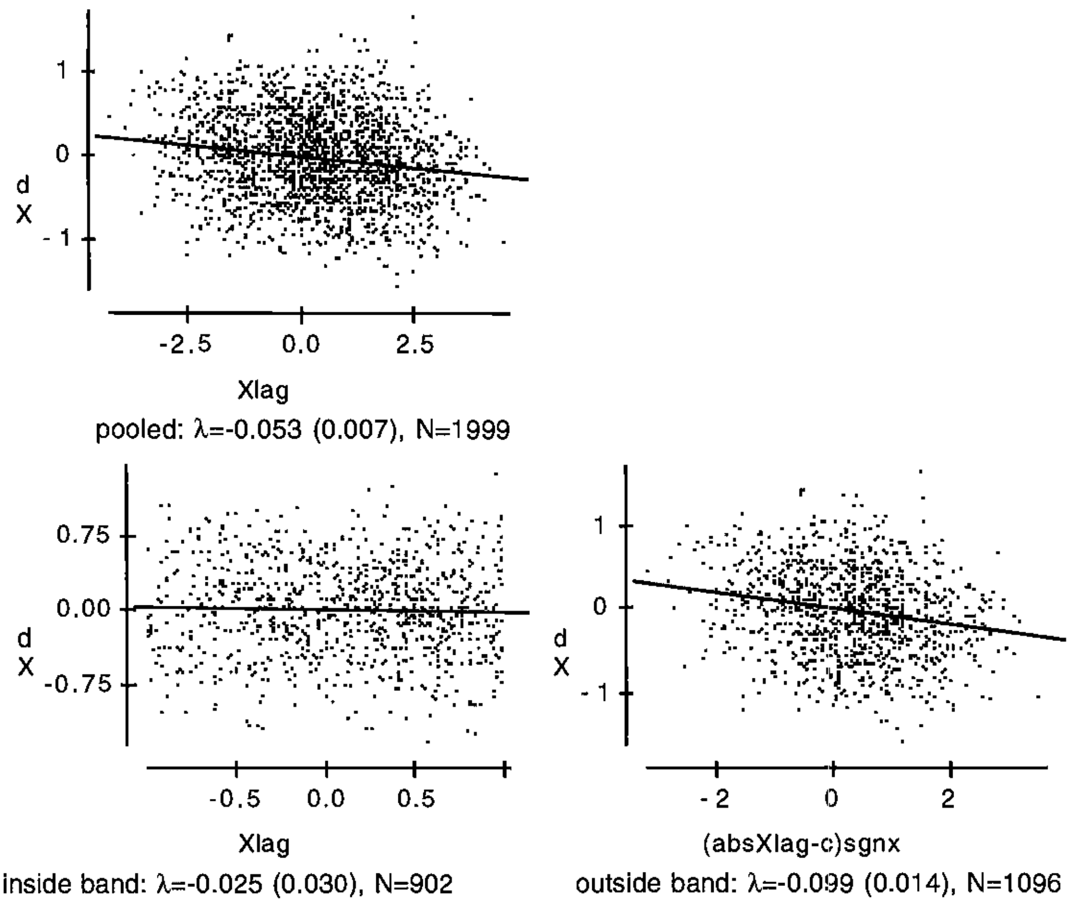
Notes: Distance is relative to Chicago. Exchange rate volatility $\sigma(\Delta s)$ is standard deviation of differenced log nominal exchange rate relative to U.S. dollar. In main tables and figures, a composite measure of tariff barriers between locations 1 and 2 is constructed as TB12 equal to the sum of TB in each location, and similarly for NTB12. The last four columns show estimated thresholds relative to a U.S. benchmark for the baseline specification. See text.
Source: Engel and Rogers (1995).

Figure 1 **Simulated Data for a Threshold Autoregressive Model**



Notes: See text.

Figure 2 **1-Regime Pooling versus 2-Regime Model with Thresholds**



Notes: See text.

Figure 3 Estimated Thresholds versus Distance

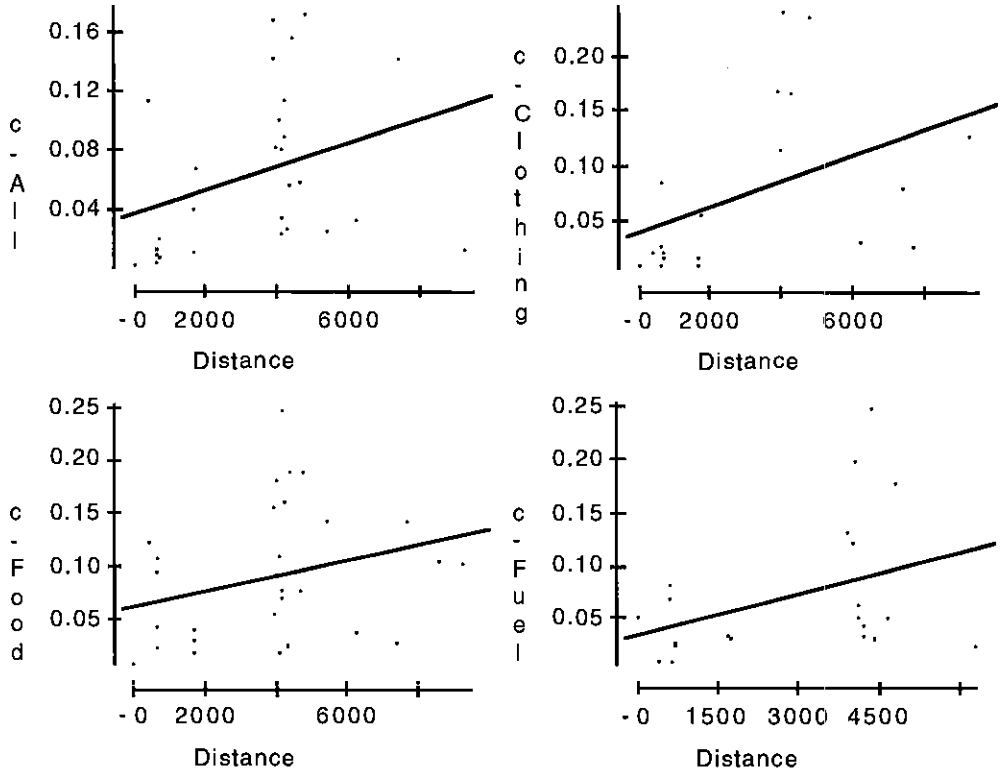


Figure 4 Estimated Thresholds versus Exchange-Rate Volatility

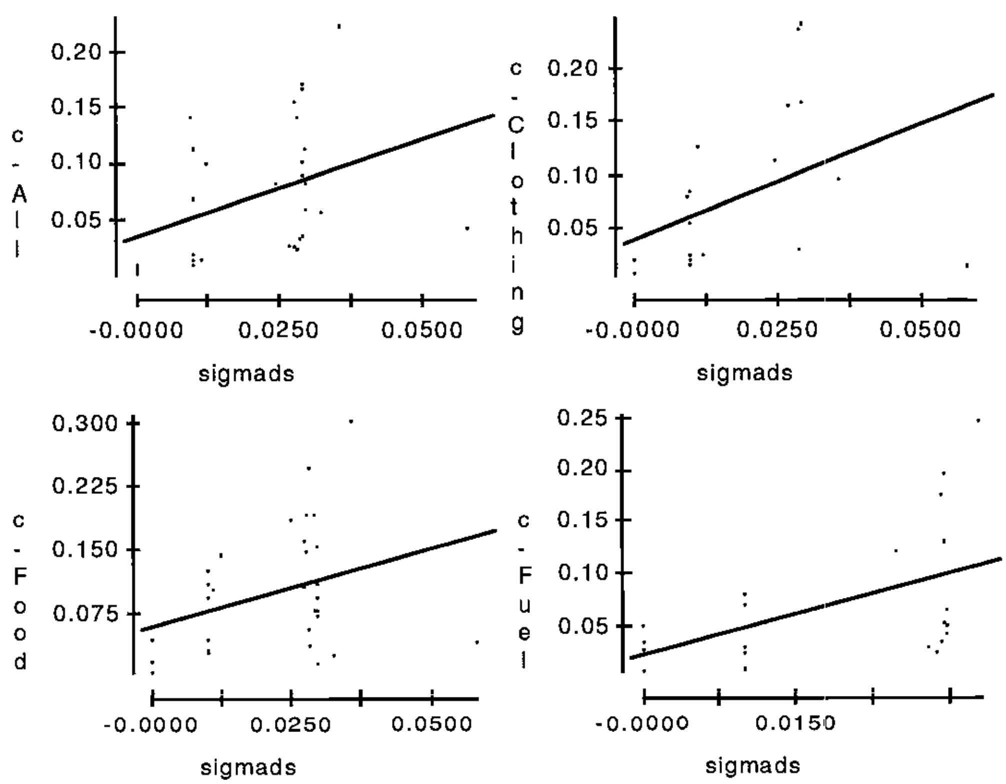


Figure 5 **Estimated Thresholds versus Tariff Barriers**

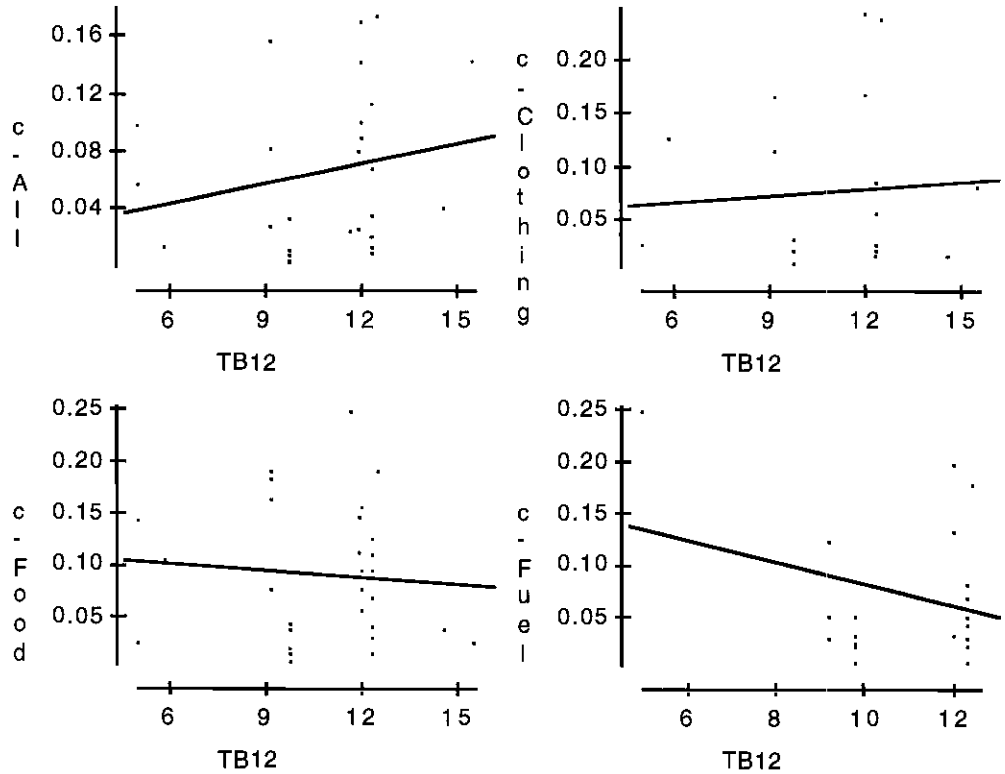


Figure 6 **Estimated Thresholds versus Non-Tariff Barriers**

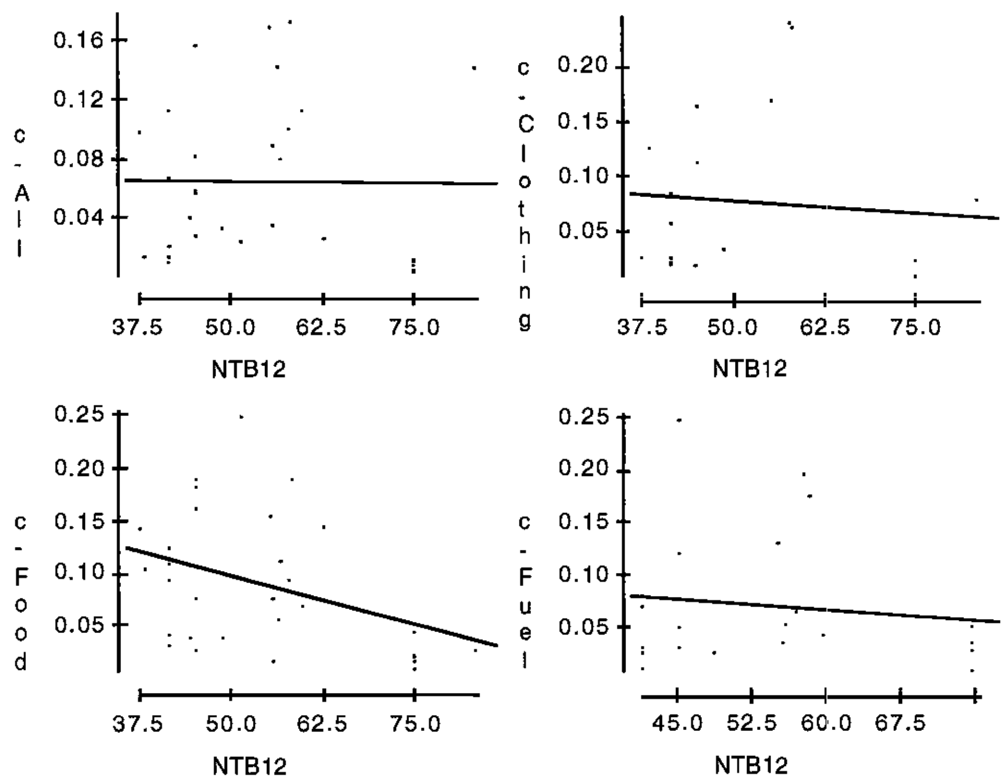


Figure 7 **Estimated Convergence Speeds versus Distance**

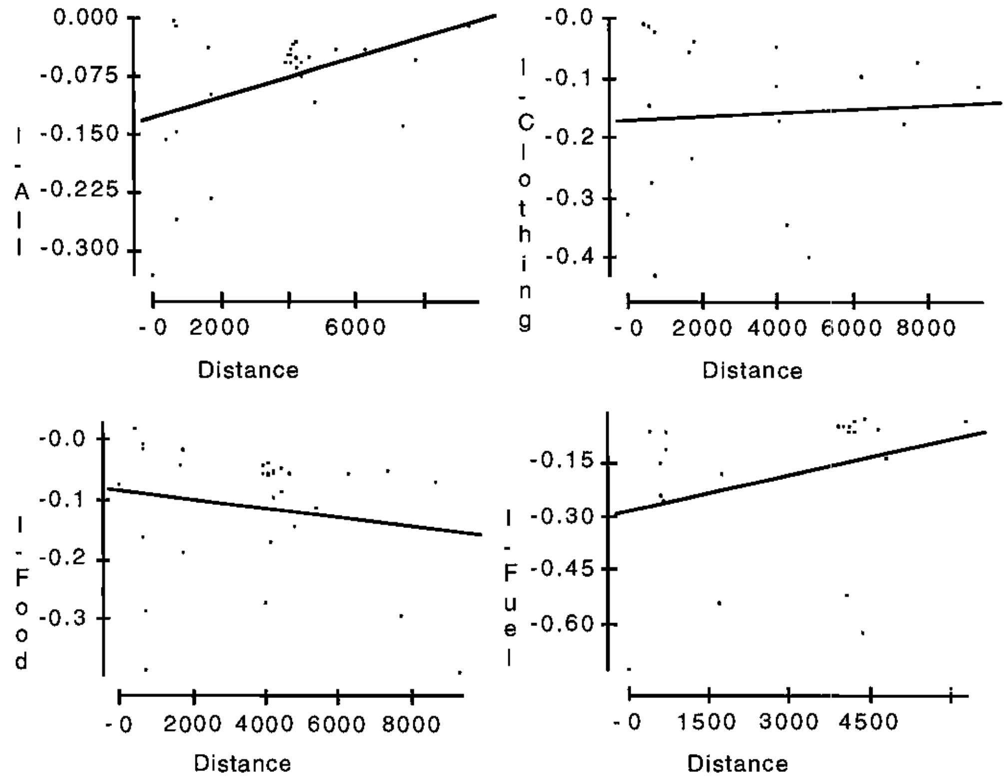


Figure 8 **Estimated Convergence Speeds versus Exchange-Rate Volatility**

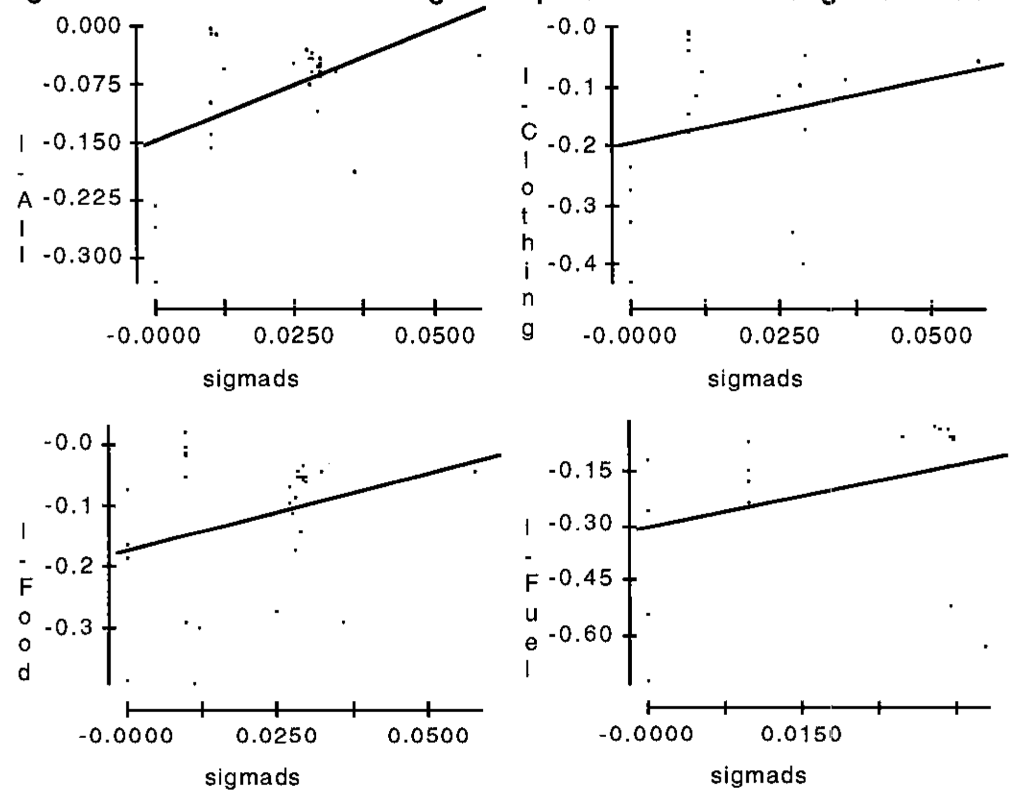


Figure 9 **Estimated Convergence Speeds versus Tariff Barriers**

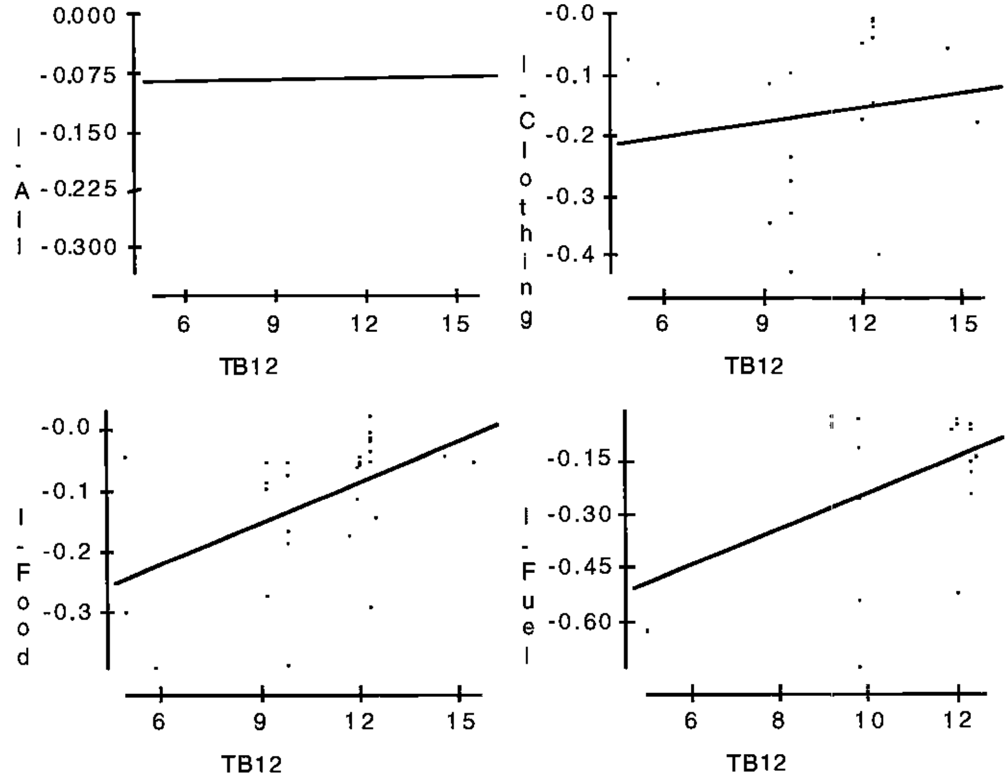


Figure 10 **Estimated Convergence Speeds versus Non-Tariff Barriers**