

Preliminary and not for quotation

The Impact of Monetary Union on Trade Prices*

By

Robert Anderton (ECB), Richard E. Baldwin (Graduate Institute for International Studies) and Daria Taglioni (Graduate Institute for International Studies)

August 2002

Abstract

This paper presents a model in which firms choose the extent to which they can bilaterally ‘price to market’ by investing in strategies that increase the cost of cross-border price arbitrage. As Friberg (2001) showed, the degree of bilateral exchange rate volatility is an important determinant of how profitable such segmentation investments would be. Our model goes beyond this by allowing firms a continuous choice of progressively more expensive segmentation strategies that result in their being able to maintain correspondingly larger price gaps. A straightforward prediction of our model – and one that is empirically testable – is that the no-arbitrage bands should be wider for bilateral trading relationships where firms anticipate a higher level of exchange rate volatility. To test this prediction we rely on a ‘natural experiment’ from Europe. While most intra-European bilateral exchange rates were fairly volatile in the 1980s and 1990s, one group of countries – the DM bloc – consistently maintained very narrow margins of fluctuation via a variety of practices. We therefore estimate separate threshold autoregressive (TAR) processes for intra-DM bloc trade and Germany’s trade with other EU nations. Our results, namely that the intra-DM bloc band is narrower, are consistent with the theoretical predictions. If the mechanism we model is empirically significant, then formation of a monetary union should induce changes that tend to homogenise price movements among monetary union members via the import and export price channels. The implication is that monetary union could produce changes in corporate strategies that result in faster cross-border transmission of price movements which, in turn, would tend to homogenise price movements across the member states of a monetary union.

*The views expressed in this paper are the author’s and do not necessarily correspond to those of the European Central Bank or the Graduate Institute for International Studies.

1. Introduction

Economic integration can help to harmonize price movements across member states of a monetary union. Trade is one of the primary channels for harmonisation of prices since price movements in one member state tend to spread to others in the union via traded goods prices.

Recent work on the trade effects of monetary union (Rose 2000, 2001) shows that monetary union may greatly increase the amount of trade among members, so intra-area trade prices could become more important in national price indices and therefore strengthen the impact of trade on price harmonisation. Unfortunately, standard estimates of the speed of price transmission across nations (Rogoff 1996, Goldberg and Knetter 1997) suggest that the price-homogenising effect of trade prices will operate very slowly. Indeed, quite a number of studies that focus on the gap between domestic prices and prices for similar goods on export markets find that the gap follows a random walk – implying that imported and exported inflation will not have an important homogenising impact on prices even if the degree of intra-union openness rises.

Some new empirical evidence (Parsley and Wei 2001, and Asplund and Friberg 2001) suggests that a reduction in exchange rate variability reduces the variability of international price differences. Moreover, the effect seems to be highly non-linear with monetary union having an effect even controlling for exchange rate volatility.

This arrangement of seemingly unconnected empirical results suggests an intriguing mechanism – formation of a monetary union among a group of nations may induce changes that help the process of price homogenisation. Indeed, the effect might be large if the elimination of exchange rate volatility simultaneously leads to a big increase in intra-union trade *and* a big increase in the speed at which price shocks are transmitted across members' goods markets.

This paper is a first attempt to piece together part of this mechanism, namely the impact of monetary union (and exchange rate volatility more generally) on the international transmission of price shocks via the imported/exported inflation channel. In doing this we generate specific testable hypotheses and confront these with a number of data sets on European prices.

In our simple model, a change in exchange rate volatility – formation of a monetary union, for example – induces trading firms to change their behaviour in a way that tends to harmonize aggregate price movements between nations. Our model combines elements from new work by Friberg (2001) with elements drawn from the old sunk-cost hysteresis literature (Baldwin 1988, Baldwin and Krugman 1989, Dixit 1990), and especially its application to European monetary union (Baldwin 1991).

Main message of the paper

If the mechanism we model is empirically significant, then formation of a monetary union should induce changes that tend to homogenise price movements among monetary union members via the imported and exported inflation channels.

To see our point in more detail, compare two extremes of the trade price transmission mechanism. If trade prices are unrelated to domestic prices (the random walk hypothesis),

then positive and negative price shocks in a given nation will not be transmitted among nations except perhaps by very indirect channels. In such a world, national inflation rates could remain quite different for long periods. At the other extreme, if all goods were traded and the price of domestically sold and exported goods were identical, national price shocks would be rapidly transmitted around the monetary union. This imported and exported inflation channel would keep the monetary union's price movements quite homogenous across members. The point of our paper is to argue that by eliminating exchange rate volatility, a monetary union tends to move member economies away from the random walk extreme and towards the full transmission extreme.

Organisation of the Paper

The paper begins with a formal presentation of the mechanism linking exchange rate volatility and trade price pass-through in Section 2. This is done in the context of a model that is kept as simple as possible. As part of this, Section 2 studies pricing behaviour under the two usual extreme assumptions of perfectly segmented markets and perfectly integrated markets. It then proceeds to endogenise the degree of segmentation and to study the implications of this for pass-through. Section 3 takes the model's main predictions to the data. The final section presents our concluding remarks.

2. Theoretical underpinnings

Firms engaged in international trade frequently set different prices for their goods in different markets. Actual and potential arbitrage, however, tends to limit the extent of such price gaps. If prices diverge sufficiently – and the goods sold in the different markets are sufficiently good substitutes – third parties will be tempted to arbitrage price gaps by re-exporting, re-importing goods, or by simply re-directing exports before they reach their intended destination. What all this goes to say is that unless firms explicitly engage in strategies to make price arbitrage expensive or illegal, their ability to price discrimination will naturally be limited.¹ Plainly, such market-segmenting strategies are likely to be expensive for the firm. For example, they may entail reduced scale economies if they involve producing different varieties for each market. Or, they may involve expensive lobbying that leads to laws, regulation and administrative practices that boost market segmentation.

The theoretical framework in this paper studies the optimal degree of market segmentation from the perspective of trading firms.² To this end, we consider the cost and benefits of 'buying' a given degree of market segmentation, where this is measured by the maximum price gap that cannot be arbitrated.

As it turns out, exchange rate volatility has an important impact on the degree of optimal market segmentation (Friberg 2001). The logic involves three steps. First, discriminatory pricing is always profitable when demand elasticities vary across markets, but exchange rate variation is surely a far more important real-world motive for wanting to price discriminate. When the exchange rate gets a long way from its steady state value, firms will have a very large profit-incentive to 'price to market', i.e. to charge very different prices in different markets. Since this is true for both appreciations and depreciations, the value of being able to price discriminate is a U-shaped function of the level of the exchange rate, with

¹ This is especially true in the European Union (EU) since hindering intra-EU trade is normally illegal.

² This is new since most of the literature on trade pricing simply assumes that markets are either perfectly segmented, or perfectly integrated.

the nadir of the U at the long-run exchange rate. The second step is to note that the expected value of this U-shaped function changes with the volatility of the exchange rate. In essence, a more volatile exchange rate means that one more frequently observes exchange rates where it would be very profitable to price discriminate. The final step is to consider the cost of market segmentation. Assuming that a higher degree of market segmentation involves corporate strategies that are progressively more expensive, profit maximising firms will choose a degree of segmentation (i.e. a maximum price gap) where the marginal benefit of further segmentation just equals its marginal cost. Because reduced exchange rate volatility lowers the marginal benefit of segmentation, firms will find it optimal to lower the degree of market segmentation in response to dampened exchange rate variability.

In our simple model, a change in exchange rate volatility – for example, the move to a monetary union – induces trading firms to change their behaviour in a way that tends to harmonize aggregate price movements between nations.

2.1. Basic Model

To illustrate the main economic logic of our analytic framework, we start with the simplest possible partial equilibrium model. We suppose that there are only two nations and we consider a market where there is a single firm that sells in both markets. The firm is located in the ‘home’ country and acts as a monopolist. The market in each nation is a priori identical with demand in the two markets given by:

$$(1) \quad q = a - p; \quad q^* = a - p^*$$

where p and q represent the price and quantity sold respectively; variables without an “*” are home-nation variables, while those with an “*” are foreign nation variables.³ The firm’s production technology involves constant marginal costs such that:

$$(2) \quad c[w, q + q^*] = c(q + q^*); \quad c \equiv wa_q$$

where c is the constant marginal cost, which depends on w , the home nation wage, and a_q , the unit input labour requirement, both of which are taken as constant.

For simplicity’s sake, we ignore all forms of trade costs (natural and manmade), so the firm’s two-market objective function is:

$$(3) \quad \max_{p, p^*} pq + \frac{p^* q^*}{s} - c(q + q^*)$$

where ‘ s ’ is the exchange rate (s is a mnemonic for spot rate) defined as foreign currency units per unit of home money. Note that a rise in ‘ s ’ raises the competitiveness of home-nation goods in the foreign market, i.e. a fall in s is a depreciation of the home currency.

2.2. Polar segmentation assumptions

Trade theory typically views international markets as integrated – prices must be identical in all market, i.e. third-degree price discrimination is not possible – or segmented, i.e. firms can set prices independently. To fix ideas, we first work out the firm’s problem assuming perfectly segmented markets, that is to say, assuming that there is no constraint on the choice of p and p^* .

³ Without loss of generality, units are chosen so that the slopes of the identical demand curves are unity.

2.2.1 Perfectly segmented markets

When markets are perfectly segmented, the firm is free to choose a price in each market. The solution to (3) thus involves two first order conditions, which solve to:

$$(4) \quad p = \frac{a+c}{2}; \quad p^* = \frac{a+cs}{2}$$

Since (3) is strictly concave in p and p^* , the first order conditions are necessary and sufficient, so (4) gives the profit maximisation prices with perfectly segmented markets. Notice that a sufficiently high 's' will yield a p^* that involves zero foreign sales. Specifically, the choke-off 's' equals a/c ; for spot rates higher than this, optimal foreign sales are zero. Using these optimal prices in the objective function, the profit function under the perfectly segmented markets hypothesis, which we denote as Π^{PSM} , is:

$$(5) \quad \Pi^{PSM} = \left(\frac{a-c}{2}\right)^2 + \frac{1}{s} \left(\frac{a-cs}{2}\right)^2; \quad 0 < s \leq \frac{a}{c}$$

for $s > a/c$, the profit function consists only of the first term in the expression since no foreign sales are made; 'PSM' is a mnemonic for perfectly segmented markets.

Intrinsic asymmetric of the profit function

Expression (5) shows that the profit function is inherently asymmetric with respect to the exchange rate, s . As the spot rate rises, the home-based firm loses competitiveness in the foreign market and eventually stops exporting.⁴ But as the spot rate declines, the profitability of exports sales increases without bound in terms of home country currency. This has nothing to do with functional forms because it is driven by the translation of foreign currency earnings into local currency units.

2.2.2 Perfectly integrated markets

The other extreme assumption typically made in the literature is that of perfectly integrated markets, i.e. where firms are constrained to choose a single price for sales to the two markets – presumably for reasons having to do with potential resale among consumers.

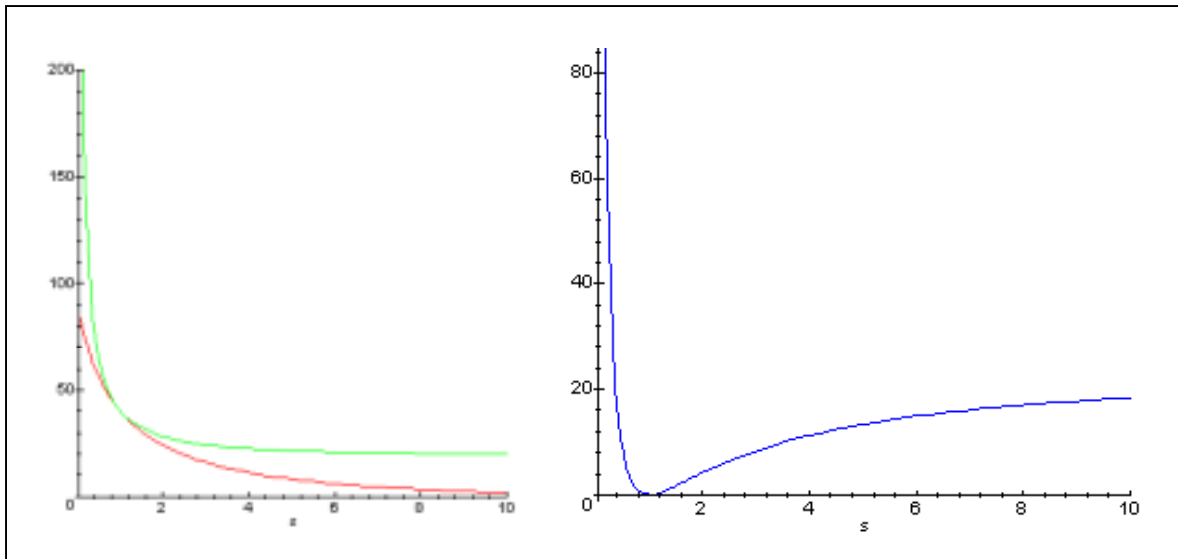
The firm's problem in this case is identical to that of the previous sections with the sole addition of the integrated market constraint, namely $p=p^*$. Solving (3) subject to $p=p^*$ yields optimal home and foreign prices (which differ only by s). Note that as 's' falls, the foreign market becomes increasingly attractive, so eventually the firm sets a price that implies zero home sales. The value of this choke-off s is $c/(2a-c)$. Plugging the optimal prices back into (3), the profit function under the perfectly integrated markets hypothesis, which we denote as Π^{PIM} , is:

$$(6) \quad \Pi^{PIM} = \frac{(2a-c(1+s))^2}{4(1+s)}; \quad \frac{c}{2a-c} \leq s$$

for s below the choke-off level, the firm sells only to foreign and chooses the optimal price for foreign sales as given in (4).

⁴ With linear demand this happens at a finite s , but with general demand, exports limit to zero as $s \rightarrow \infty$.

Figure 1: Difference between profits under PSM and PIM assumptions



2.2.3 Friberg's insight

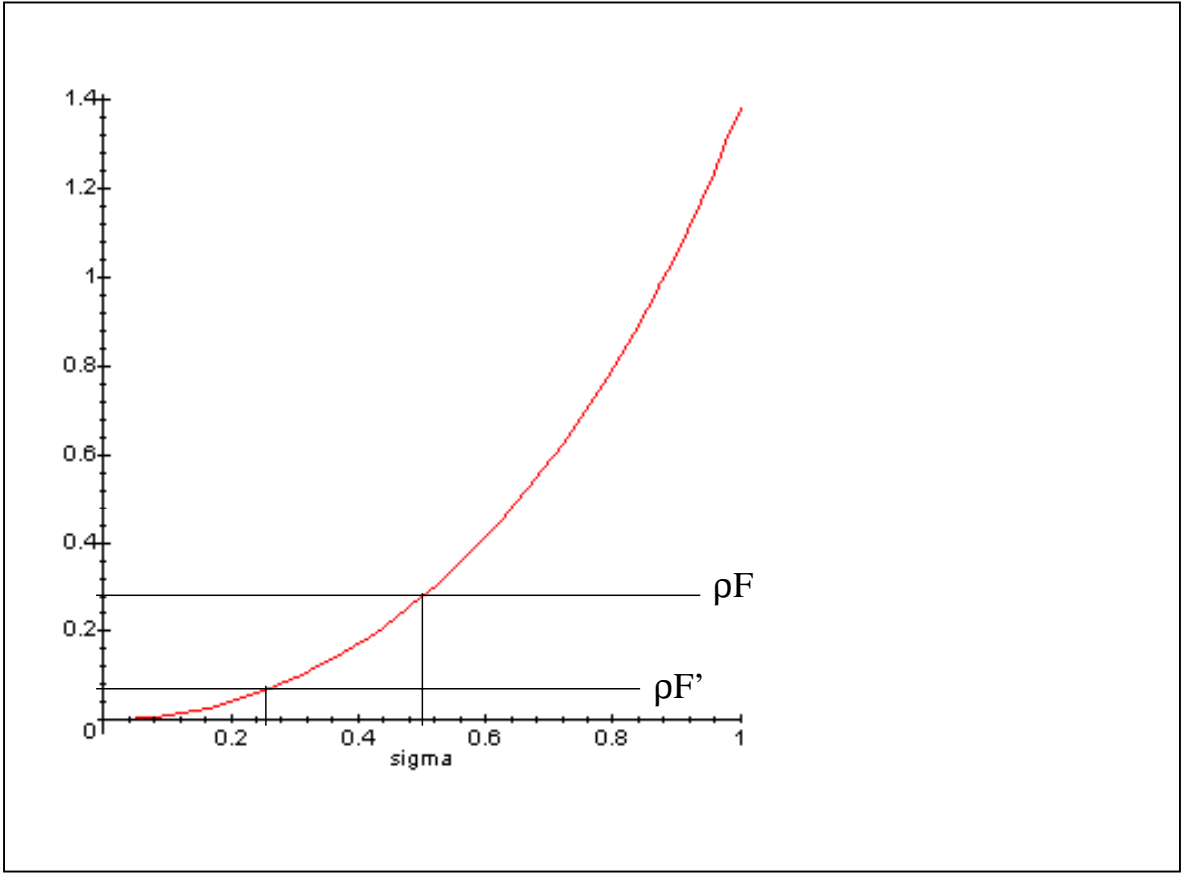
We turn now to showing the insight, due to Friberg (2001), that the reward to being able to engage in segmented-market pricing versus integrated-market pricing depends in a large part on the exchange rate's probability distribution. Using our specific functional forms, this general result can be directly illustrated. Simplifying, the difference between (5) and (6) becomes:

$$(7) \quad \Pi^{PSM} - \Pi^{PIM} = \frac{a^2(1-s)^2}{4s(1+s)}; \quad \frac{c}{2a-c} \leq s \leq \frac{a}{c}$$

While this is not everywhere convex – because of the asymmetric effects of appreciation and depreciation – it is convex in the neighbourhood of the long-run exchange rate $s=1$. Figure 1 shows, in the left panel, the plots of the two profit functions against s ; the higher curve is Π^{PSM} and the lower curve is Π^{PIM} .⁵ Clearly the wedge between the two, and thus the value of being able to price discriminate, increases as 's' deviates from its long-run value (which is here normalised to unity). The right panel plots the $\Pi^{PSM} - \Pi^{PIM}$ directly against s ; this clearly shows the intrinsic asymmetry involved in the PSM-vs-PIM wedge.

⁵ For the parameter values use to draw the figure ($a=10, c=1$), the choke off points are $1/19$ and 10 .

Figure 2: Volatility and the expected value of the PSM-vs-PIM wedge



The next step is to show that the expected value of the PSM-vs-PIM wedge is increasing in the volatility of the exchange rate. If the wedge were a strictly convex function, this result would be a straightforward application of Jensen's inequality. Given the asymmetry, however, the wedge will not in general be strictly convex, so Friberg (2001) is unable to prove a relationship between exchange rate volatility and the value of market segmentation. Nevertheless he conjectures that such a relationship exists (see his Corollary 2 and discussion).

Using our simple functional forms and an additional simple assumption on the stochastic process driving the exchange rate, we can establish a link between volatility and the reward to segmentation by direct calculation.

Assuming that the exchange is uniformly distributed on a support $[1-\sigma/2, 1+\sigma/2]$, the expected value of the PSM-vs-PIM wedge is:

$$(8) \quad \int_{-1+\sigma/2}^{1+\sigma/2} (\Pi^{PSM} - \Pi^{PIM}) f(s) ds = \frac{a^2}{4\sigma} \left(\ln\left(\frac{(\sigma+4)^4(2-\sigma)}{(4-\sigma)^4(2+\sigma)}\right) - \sigma \right); \quad f(s) = 1/\sigma$$

where σ measures exchange rate volatility. Note that $\sigma \leq 2$, since $s \geq 0$. Inspection of (8) proves the following result.

Result 1: The expected value of having perfectly segmented markets as opposed to perfectly integrated markets increases as exchange rate volatility increases.

A more complete characterisation of the link between volatility and expected value of the PSM-vs-PIM wedge is shown in Figure 2 (here $a=10$ and $c=1$). Notice that the relationship

is highly non-linear; the impact of volatility reduction on the value of having perfect segmentation gets weaker as the level of volatility falls.

We turn now to an extension of this framework that allows us to endogenously determine the degree of market segmentation.

2.3. Variably segmented markets

The classification of markets into perfectly segmented or perfectly integrated is theoretically convenient, but not terribly realistic. More generally, one expects that firms can engage in a whole array of strategies that make cross-market re-sale increasingly difficult.

Market Segmentation Technology

The key aspect of segmentation is the inability of consumers to re-sell goods across markets. Otherwise, arbitrageurs would exploit price differences up to the point where the price differences were eliminated. In the European Union, firms are not allowed to forbid re-sale within the EU, so segmentation requires some sort of product differentiation perhaps combined with government regulation. For example, a German food company that wishes to price discriminate between the German and French markets, but not between the German and Dutch markets, could produce one version of the good with a German-and-Dutch-language label and another with a French-language-only label. This labelling would make it expensive to arbitrage price differences, but not impossible unless this differentiation were combined with a government regulation on the sale of food labelled in the ‘wrong’ language’ and a ban on third party re-labelling. Implementing this strategy may involve a fixed cost in adjusting the production line, but the investment may be worth it if the exchange rate is sufficiently volatile.

In particular, suppose the segmentation technology involves a one-time sunk cost – call it F – that does not depreciate, so a firm will sink the cost if the expected value of doing so exceeds the sunk cost. Analysis of this sort of investment decision under a variety of stochastic processes is by now well understood. To illustrate our main points as simply as possible, assume that the exchange rate process over time is independently and identically distributed (iid). In this case, the firm undertakes the segmenting investment if and only if ρF exceeds the one-period expected value of the PSM-vs-PIM wedge.

Continuing the analysis with this simple market segmentation technology, we can see, with the help of Figure 2, that reducing exchange rate volatility would make it less likely that firms find segmentation profitable. For example, if the sunk cost facing a particular firm is F , then our measure of volatility would have to fall below 0.5 to make segmentation unattractive to the firm. Volatility would have to fall below 0.25 to make it unprofitable when the fixed cost was F' .

Notice that the impact of reduced volatility on segmentation is non-linear here, but the form of the non-linearity indicates that there would be little difference between very low volatility and a true monetary union. Since this seems to be counterfactual, given the empirical studies discussed above, we must go beyond the Friberg framework that relies on sunk costs.

Variably Segmented Markets Assumptions

To model firms’ choice over variable segmentation, we suppose that firms can choose a strategy that allows them to charge a cross-market price gap of up to “ G ” without fear of re-sale; to be concrete, we assume that G is fixed in foreign currency terms (G is a mnemonic for

gap). More specifically, for a given G , the firm acts as if markets are perfectly segmented when 's' is such that they want to charge prices within the following range:

$$(9) \quad sp + G \leq p^*, \quad p^* \leq sp - G$$

The left-hand inequality is relevant when s is low, so the home price looks very attractive to foreign consumers when translated into foreign currency terms – attractive enough to make it worthwhile paying the arbitrage cost of G per unit bought. The right-hand inequality is relevant when s is high enough to make home residents consider re-importing the good from foreign despite the arbitrage cost.

We refer to this framework as “variably segmented markets” and note that it will allow us to endogenously determine the extent of market segmentation.

The segmentation technology is such that the firm faces a rising cost of implementing strategies that result in higher G 's. To be concrete and to simplify calculations, we assume the segmentation cost function is:

$$(10) \quad C[G] = \alpha G^2 / 2$$

This implies that the marginal cost of a higher G is positive and rising in G . The idea here is that natural barriers always provide firms with very limited ability to segment but it becomes increasingly costly for a firm to institute strategies that permit it to charge progressively wider gaps.

2.3.1 Modelling variable market segmentation

The assumption of variably segmented markets (VSM) yields a richer problem for the firm. To explore this we consider the expected value of having a particular level of G .

The firm will face three types of exchange rates. If s is close enough to its long-run level (unity), then it can act as if the market were perfectly segmented, since its profit maximising prices imply a cross-market price gap that is too small to arbitrage. In this case, its problem and solution is identical to the one in (5). If, however, s is sufficiently high, the solution to its profit maximisation problem will involve the constraint that $p^* = sp - G$, and if s is sufficiently low, the constraint will be $p^* = sp + G$.

The critical values of s are easily determined, since when the inequalities in (9) hold with equality, the firm will be pricing according to (4). Using this fact, we find the critical values are:

$$(11) \quad s_{hi} = \frac{1}{1 - 2G/a}, \quad s_{lo} = 1 - 2G/a$$

This implies that the width of the hysteresis band, i.e. the band where arbitrage provide zero pressure for price convergence, is:

$$(12) \quad \text{band width} = \frac{4G(a - G)}{(a - 2G)a}$$

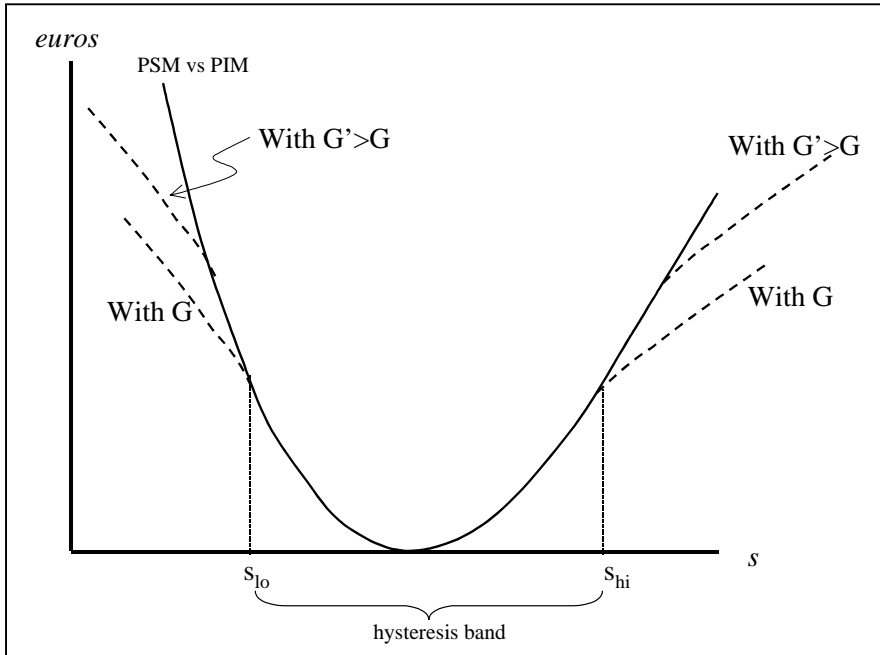
notice that this hysteresis band widens with G at an increase rate, i.e. the band is convex in G .

Optimal pricing in the two regions of s where arbitrage constraint binds are:

$$(13) \quad p_{hi} = \frac{2(a + sG) + c(1 + s)}{2(1 + s)}, p_{lo} = \frac{2(a - G) + c(1 + s)}{2(1 + s)}$$

The optimal prices in the band are given by (4). Plugging these optimal prices into the definition of profits yields three different profit functions which we do not report since they are not particularly revealing.

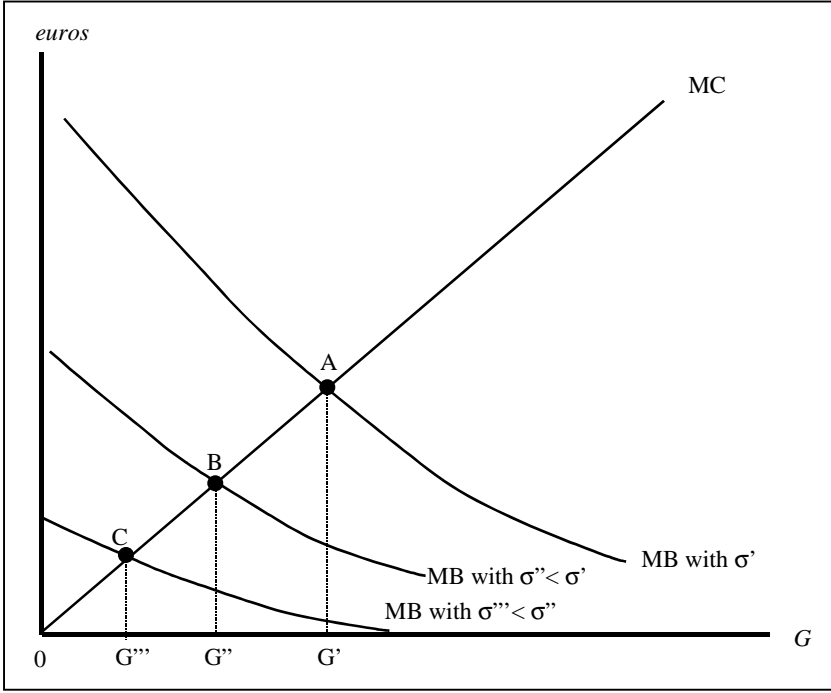
Figure 3: VSM vs. PIM profit functions



We can illustrate the impact of variably segmented markets (VSM) with Figure 3. For a given G , the high and low bounds on s will be fixed and for realisations of the exchange rate within this range, the difference between VSM and PIM is identical to the difference between PSM and PIM as in the right panel of Figure 1. When s is higher than s_{hi} , the firm is able to do somewhat better under VSM than it would with perfectly integrated markets, but not as well as if it had perfectly segmented markets. This is shown with the bottom dashed line on the right side of the diagram. Similarly, when s is below s_{lo} , the firm does better than PIM but worse than PSM, as shown by the lower right-side dashed line.

If the firm 'buys' a higher level of segmentation, as measured by G' , then its profits will unambiguously be higher or equal to those with G . The diagram shows this with the upper dashed lines on the right and left sides of the diagram. Notice that the benefit of higher G shows up only for exchange rates that are sufficiently far from the long-run level.

Figure 4: Marginal benefit and cost of G



2.4. Endogenous, variable market segmentation

Since the exchange rate is a random variable, optimising firms will choose the optimal G by considering the expected value of having a given level of market segmentation. The expected value is the sum of three parts. For s between $1-\sigma/2$ (the minimum realisation of s) and s_{lo} , the relevant profit function will be profits maximised subject to the left-hand price constraint in (9). For s between s_{lo} and s_{hi} , the relevant profit function is (5), and for s between s_{hi} and $1+\sigma/2$ (maximum realisation of s) the profit function is profits maximised subject to the right-hand price constraint in (9).

Integrating over the three ranges and subtracting the result from the expected value of perfectly integrated markets (which is assumed to apply when $G=0$ is chosen), we get:

$$(14) \quad EV = \frac{G}{\sigma} \ln\left(\frac{s^a (1+s)^G}{s(1+s)^{2s}}\right) \Bigg|_{1-\frac{\sigma}{2}}^{s_{lo}} + \frac{1}{4\sigma} \left(s - \ln\left(\frac{(1+s)^4}{s}\right)\right) \Bigg|_{s_{lo}}^{s_{hi}} + \frac{G}{\sigma} \left(s(a-G) + \ln\left(\frac{(1+s)^G}{(1+s)^{2s}}\right)\right) \Bigg|_{s_{hi}}^{1+\frac{\sigma}{2}}$$

where s_{lo} and s_{hi} are defined in (11).

It is important to note that this calculation assumes that G is such that the ordering of integration limits is correct, namely that $1-\sigma/2 \geq s_{lo}$ and $1+\sigma/2 \geq s_{hi}$. Since s_{lo} and s_{hi} depend upon G , this puts restrictions on the permissible G . Intuitively, the point is that no firm would ever find it profitable to buy a G that was so high that no price constraint was ever strictly binding. The specific restriction is that optimal G 's will be less than $a\sigma/2(2+\sigma)$.

Volatility and segmentation in the MB MC diagram

To illustrate the optimal choice of G , and thus the optimal hysteresis bandwidth, we plot the marginal benefit of G – i.e. the partial of (14) with respect to G – and the marginal costs in Figure 4.

As per (10), MC is rising. The MB curve is more involved. For a given level of exchange rate volatility, an increase in G widens the hysteresis band. Figure 3 helps us to evaluate the impact of this on EV. As the band widens, some realisations of s will correspond to a high profit level (the solid PSM vs PIM line instead of the dashed line), so a higher G will clearly raise EV. However, each successive increase in G affects a progressively narrower range of s . As a consequence, the marginal benefit of raising G falls with the level of G , as shown in Figure 4.

The optimal G corresponds to the intersection of the MC and MB curves. For sufficiently high exchange rate volatility, for example σ' , the firm will choose a positive G' . As the figure shows, lowering σ to, say σ'' , will result in a progressively lower optimal G , viz. G'' . Importantly, this model predicts segmentation only reaches zero when the level of is zero.

2.5. Testable implications

The analysis in the preceding theoretical sections can be summarised as follows:

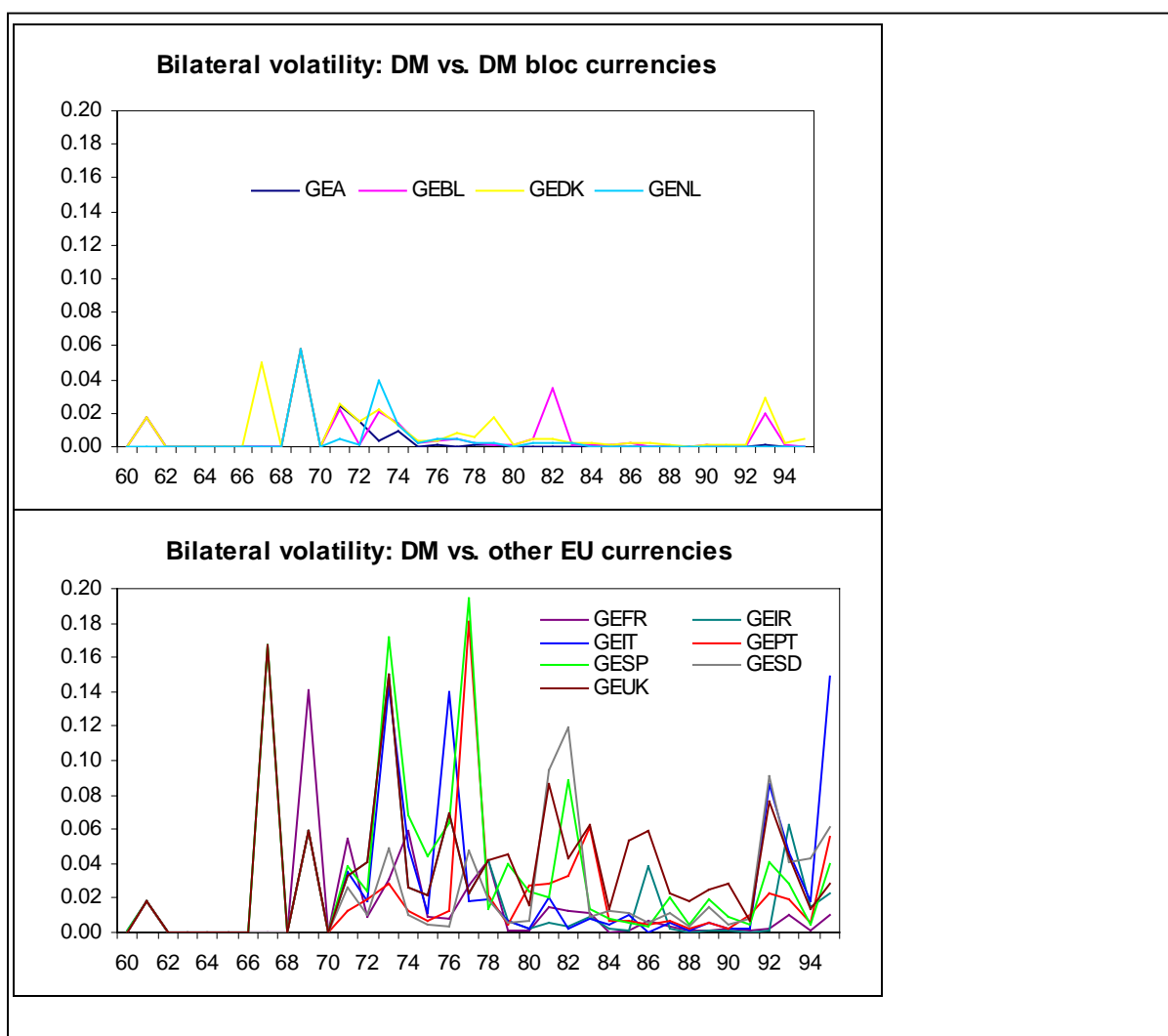
Result 2: The degree of endogenous market segmentation – as measured by the width of the hysteresis band – falls as the level of exchange rate volatility falls.

The testable implication of this result is clear. The estimated bandwidth should be narrower when using data for nations that have experienced relatively little bilateral exchange rate variation.

3. A Natural Experiment: The DM Bloc

The theoretical model suggests that pass-through should be linked to bilateral exchange rate volatility, but bringing this prediction to the data is problematic. The logic of our model suggests that pass-through is the result of conscious investments by firms. Since some of these investments are presumably long-lasting – for example, the decision to produce differentiated products for export markets, or lobbying for regulatory segmentation – it is somewhat difficult to know how to measure volatility. Firms make the investment today in expectation of tomorrow's volatility, so what we need is a measure of how volatile firms expected the exchange rate to be in the next few years. While there are a variety of ways of dealing with this econometrically, we choose instead to rely on a natural experiment.

Figure 5: Bilateral exchange rate volatility, DM bloc and others, 1960-1994



With the breakdown of the Bretton Woods exchange rate system in the 1970s, Europe adopted a string of managed exchange rate systems. European bilateral exchange rates nevertheless fluctuated widely with one important exception. Right from the beginning of post-Bretton Woods period, a group of countries – the deutsche mark (DM) bloc – maintained quite tight bilateral rates. For example, DM-guilder bilateral rate was not realigned between 1983 and the start of the monetary union and the bilateral exchange rate fluctuated within a very narrow band during this period, even narrower than the maximum allowed under the European Monetary System’s Exchange Rate Mechanism (ERM). Similarly, the Danish crown, Belgium franc, and Austrian schilling have had very stable relationships with the DM since the late 1980s.

Importantly, this arrangement was stable and widely expected to continue. As Figure 5 shows, the DM bloc exchange rates stayed close throughout the long sequence of exchange rate crises in the 1970s, 1980s and 1990s. Firms engaged in intra-DM bloc trading should, therefore, have had sufficient time and incentives to adjust their segmenting strategies to the lower degree of exchange rate volatility. What this means is that we should observe a difference in trade-pricing behaviour between Germany and these nations and between Germany and other nations using historical data.

3.1. The Null hypothesis: Linear model and equal bilateral pass through

Our theoretical model generates a number of testable predictions but the most obvious and the most robust is that pass-through behaviour should depend upon bilateral exchange rate volatility. Thus as a first cut, we test whether we can reject the null hypothesis that pass-through behaviour is unrelated to exchange rate volatility.

3.1.1 The naïve empirical model

To test the null hypothesis we need an empirical model and here we adopt the “linear model” that has been used countless times in purchasing power parity (PPP) and law of one price (LOOP) literature (see Rogoff 1996 for an overview). This model focuses on the gap between prices in different nations measured in the same currency. Taking ‘z’ as price gap, the model is usually written as:

$$(15) \quad \Delta z_t = \lambda z_{t-1} + e_t$$

where e_t is an iid error term is iid. Thus, we assume that z_t follows an AR1 process. For convenience, the model is not specified in terms of the underlying AR1 parameter ρ but rather the speed-of-convergence parameter, $\lambda \equiv \rho - 1$. That is, λ obeys $-1 \leq \lambda \leq 0$ and it is the convergence speed in the sense that price differentials are reduced each period by a fraction λ plus an error term. A λ equal to zero implies perfect persistence (random walk); a λ equal to minus one indicates zero persistence (iid behaviour).

Our data

We test this simple AR1 specification on a panel of German import price data from 17 industrialised trade partners covering 10 manufacturing (2-digit) sectors in the period 1975-1995 (see data appendix for details). Specifically, we denote p_t^1 and p_t^2 as the log price levels of a good in two locations at time t . In our dataset, p_t^1 is the German import price of goods imported in country 1 from country 2, and p_t^2 is the producer price in country 2 (the exporting country) both measured in the exporter’s currency.

In general one needs to consider the possibility of long term trends in measured prices differences, especially since our ‘prices’ are actually indices. One does observe trends in the relative prices of traded and non-traded goods and our producer price indices inevitably contain some non-traded goods. Moreover, systematic changes in quality can also produce trends in price-index gaps even in a world with perfect arbitrage. To test for this we perform a non-parametric testing based on Cuzick 1985. This test fully rejects the presence of a trend in the data.

Even without a trend, however, the mismatch of indices may produce a constant gap even in a world of perfect arbitrage, so we demean the data. Specifically, we generate the price gap, z_t , as the OLS residual from the regression of $\ln(p_t^1 - p_t^2)$ on a constant. (The appendix shows that qualitatively identical results hold for non-demeaned data series.)

Results

Following the logic of the DM bloc natural experiment, we distinguished between import suppliers belonging to the so called DM-block and other EU countries traditionally displaying highly volatile bilateral exchange rates vis-à-vis Germany. We grouped data together for imports from Austria, Belgium (and Luxemburg), Denmark, and the Netherlands in one sub-sample, and data on imports from France, Greece, Ireland, Italy, Portugal, Spain and United Kingdom in another.

As Table 1 shows, the difference in coefficients for the two groups is striking. The price gaps for intra-DM bloc trade are much less persistent (i.e. nearer to minus one) than the gaps for German imports from other EU nations. Specifically, it takes about a year and a half for half of a price shock to be reabsorbed (half life of 1.52 periods) on intra-DM bloc trade, while it takes more than 20 months for the ones coming from the rest of the EU. More formally, a log likelihood ratio test shows that the difference is statistically significant at 5% level of confidence.

Table 1: Testing the naïve linear model

<i>Linear AR1 specification</i>		
	<u>Lambda</u>	<u>AR1 half-life</u>
All	-0.328 (0.006)***	1.74
DM Block	-0.367 (0.007)***	1.52
Other EU	-0.337 (0.017)***	1.69

Notes: Standard errors in parentheses. *** indicates statistically significant at the 1% level.

Of course the fact that the two different datasets produce different estimates does not allow us to conclude that the difference is due to exchange rate volatility since, for example, intra-DM bloc might involve a substantially different set of goods. We can say, however, that the fact that adjustment is quicker in the less volatile environment is certainly consistent with the basic notion that firms choose to invest less in segmenting markets where the bilateral exchange rate is less volatile.

Additional hints can be found by splitting the whole sample of import data in two sets of observations, one displaying low volatility and the second high volatility. In this exercise, we include data on imports from Canada, Finland, Japan, Sweden, Switzerland and United States as well as the DM bloc and other EU nations. The resulting split is not country-specific since same supplier may be classed as low-volatility in some years and high volatility in other years. Running the simple linear model on these two dataset also produces significantly different coefficients with price gaps in the more volatile dataset displaying greater persistence (see appendix).

3.2. Testing for pass-through hysteresis bands

Confirmation that pass-through was quicker for trade flows marked by lower volatility provides a very rough check on our model, but the model makes much more precise predictions. We turn now to testing these.

The bedrock of our theoretical model is the existence of a band where price gaps are much harder to arbitrage. The simple model we worked with went further by assuming that arbitrage is perfect and instantaneous outside the band but impossible inside the band. Roughly speaking this would correspond to a price gap that followed a random walk with reflecting barriers. Plainly, instantaneous price changes are not the norm due to all sorts of adjustment costs. This leads us to introduce a new element, namely slow adjustment of the price gap outside of the band. In particular, our hypothesis is that when price gaps are small enough, there is either very slow or no convergence, while price differentials larger than a certain threshold are arbitrated away and hence decay in expectation according to a stable AR process.

In essence, the model allows for non-linear adjustment to long-run equilibrium. It predicts a sort of regime switching, where the applicable regime depends upon the size of the gap. In the statistical literature this sort of adjustment process is called a threshold autoregressive (TAR) process. Recent research shows that multiple-regimes threshold models are more powerful explanations for modelling purchasing power parity (PPP) in presence of international trade and arbitrage costs. Bec, Ben-Salem and Carrasco (2001) show that a TAR with unit-root in the middle regime is still stationary. Mark Taylor, Peel and Sarno (2001) demonstrate the very low power of standard univariate unit-root tests to reject a false null hypothesis of unit root behaviour when the true model is non-linearly mean reverting.

Before proceeding to a structural estimation of the TAR process, we start off with non-parametric tests based on one of the most fundamental predictions of our band-hypothesis, namely that the price-gap adjustment process should be non-linear, especially when trade costs are high.

3.2.1 Tsay test for non-linearity

Tsay (1989) proposes a very general specification test that compares a generic TAR alternative against a linear AR null. This test works with ‘case data’ which is defined as every pair of changes in the price gaps, e.g. $x_t - x_{t-1}$ is one case and $x_{t-1} - x_{t-2}$ is another. The case data are ordered according to their lagged value, first in ascending and then in descending order.⁶ The test goes as follows: we run an AR1 on the ordered data and generate a series of recursive residuals. We then regress the predictive recursive residuals on the dependent variable (the contemporaneous price gap). If these residuals are orthogonal to the depend variable, then we can reject non-linearity in the adjustment process. The orthogonality test amounts to calculating the conventional F statistic for this regression, which for large N approximates a χ^2 random distribution. The intuition of the test is that if there is a regime switching, only the first ‘n’ cases (falling under the first regime) will show orthogonality properties, While the remaining (N-n) cases will follow a different behaviour revealing the point of switching which should be related to the size of the price gap.

In performing this test, we find that the maximal χ^2 are all significant (the DM-block at 10 percent), so we conclude that we can reject the linearity hypothesis⁷. Our findings confirm the hypothesis that data are non-linear in presence of higher trading costs.

⁶ The reason why the Tsay test should be run with both increasing and decreasing ordering of the arranged case data is of practical nature and lies in the fact that, especially in small samples, the case data may not fall in all of the regimes delineated by every threshold value. In what follows, we adopt that practice and we report only the most significant p-values for each of the two (descending and ascending) F-tests.

⁷ Maximal F and χ^2 statistics are: (i) entire dataset $\chi^2(1)= 8.78$ with prob $>\chi^2 = 0.003$ (ii) DM Block $\chi^2(1)= 2.71$ with prob $>\chi^2 = 0.099$ and (iii) other EU countries $\chi^2(1)=76.88$ with prob $>\chi^2 = 0.000$

3.3. TAR estimation

We now proceed to estimate a TAR model on our non-linear data. While in general the adjustment process may display many regimes, we start with the parsimonious specification that assumes just one, symmetric band with one speed of convergence λ inside the band and another outside the band. Thus our empirical model is now:

$$(16) \quad \Delta x_t = \begin{cases} \lambda^{out} (x_{t-1} + b_{lo}) + \varepsilon_t^{out}, & x_{t-1} < b_{lo} \\ \lambda^{in} x_{t-1} + \varepsilon_t^{in}, & b_{lo} \leq x_{t-1} \leq b_{hi} \\ \lambda^{out} (x_{t-1} - b_{hi}) + \varepsilon_t^{out}, & x_{t-1} > b_{hi} \end{cases}$$

where λ^{in} and λ^{out} are the speeds of convergence in and out of the band and b 's define the band edges. To keep things simple we assume that the band is symmetric around a zero price gap, so we have only three parameters to estimate: the two λ 's and b .

Conceptually, estimation of the TAR is simple if one knows the 'b' a priori. After all, the TAR essentially says that the adjustment is driven by two separate models – one for gaps inside the band and one for gaps outside the band. If we had the b , we could partition the data into two populations and run a simple AR1 on the separate sample to obtain consistent estimates of the λ 's. We do not, of course, have the 'b' in hand, so we must estimate it. A standard way of approaching this is to order the price gap data in increasing size, assume that the band encompasses only data that falls in the, say, 1% smallest percentile and then estimate the λ 's on this provisional partition of the data. We then re-do the exercise for a slightly wider band.

To decide which partition is closer to the true model, we need an objective function. We follow Taylor (1997) in adopting a likelihood function. Specifically, we define $L_{AR}(\lambda, \sigma)$ as the estimated log likelihood function for the null AR1 linear model, and define $L_{TAR}(\lambda^{in}, \lambda^{out}, \sigma^{in}, \sigma^{out}, b)$ as the estimated log likelihood function of the alternative TAR model for a given b . Estimation proceeds via a grid search on b which maximises the log likelihood ratio $LLR = 2(L_{TAR} - L_{AR})$. Since the TAR model is locally linear, computationally, for any given b , the maximum likelihood estimation of the TAR corresponds to an OLS estimation on partitioned samples of case data wholly inside or wholly outside the threshold. The only limitation to our search is that a commodity point is unlikely to be close to the 0th or 100th percentile. For these extreme points there are not sufficient observations to provide an efficient estimate. Therefore we select an interval stretching from the 10th to the 90th percentile within which to look for the candidate thresholds.

In keeping with the natural experiment logic, we estimate a TAR for trade prices from all EU trade partners and a separate TAR for Germany's trade with other EU nations.

We expect "all EU" band to be smaller than EU vol because it contains data from DM Block countries.

Table 2: Summary results for TAR estimation (fixed effect panel)

	b	λ^{out}	<i>half-life out of band</i>	λ^{in}	<i>half-life in band</i>
DM Block	0.460	-0.392 (0.011)***	1.394	-0.391 (0.014)***	1.399
Other EU	0.734	-0.338 (0.036)***	1.683	-0.313 (0.016)***	1.845

Note: Estimation of TAR(2,2,1) using a maximum likelihood grid search over b .

The results of this procedure are reproduced in Table 2. These results are consistent with the main predictions of our empirical model. In particular, we find that the band width is greater for trade marked by more volatile exchange rate movements. Moreover, we estimate that the adjustment process is faster outside the band for the ‘Other EU’ sample. Inside the band the half-life is 1.8 years while outside it is 1.7 and this difference is statistically different. For the DM bloc trade, however, the two adjustment processes are basically identical with the λ ’s being -0.392 and -0.391 .

4. Concluding Remarks

This paper presents a model in which firms choose the extent to which they can bilaterally ‘price to market’ by investing in strategies that increase the cost of cross-border price arbitrage. Examples of such strategies include differentiating products by market, or lobbying for regulator controls that make third-party arbitrage expensive. As Friberg (2001) showed, the degree of bilateral exchange rate volatility is an important determinant of how profitable such segmentation investments would be. Our model goes beyond this by allowing firms a continuous choice of progressively more expensive segmentation strategies that result in their being able to maintain correspondingly larger price gaps. A straightforward prediction of our model – and one that is empirically testable – is that the no-arbitrage bands should be wider for bilateral trading relationships where firms anticipate a higher level of exchange rate volatility.

To test this prediction we rely on a ‘natural experiment’ from Europe. While most intra-European bilateral exchange rates were fairly volatile in the 1980s and 1990s, one group of countries – the DM bloc – consistently maintained very narrow margins of fluctuation via a variety of practices. We therefore estimate separate threshold autoregressive (TAR) processes for intra-DM bloc trade and Germany’s trade with other EU nations. Our results, namely that the intra-DM bloc band is narrower, are consistent with the theoretical predictions. The implication is that monetary union could produce changes in corporate strategies that result in faster cross-border transmission of price movements which, in turn, would tend to homogenise price movements across member countries of a monetary union.

References

- Asplund, M. and R. Friberg, 2001, "The law of one price in Scandinavian duty-free stores," **American Economic Review**, 91, 4, pp 1072-1083.
- Baldwin, R., 1991, "On the Microeconomics of the European Monetary Union," **European Economy**, Special Issue No. 1, "One market one money".
- Baldwin, R., 1988, "Hysteresis in Import Prices: The Beachhead Effect," **American Economic Review**, 78, 4, pp 773-785.
- Baldwin, R. and P. Krugman, "Persistent Trade Effects of Large Exchange Rate Shocks," with Paul Krugman, **Quarterly Journal of Economics**, pp 635-654, No. 419.
- Bec, F., M. Ben Salem and M. Carrasco, 2001, "Tests for unit-root versus Threshold specification with an application to the PPP", mimeo, University of Rochester.
- Friberg, Richard (2001), "Two monies, two markets? Variability and the option to segment" **Journal of International Economics**, 55, pp 317-327.
- Goldberg, P. and M. Knetter, 1997, "Goods prices and exchange rates: What have we learned?" **Journal of Economic Literature**, XXXV, pp 1243-1272.
- Parsley D. C. and S.J. Wei, 2001. "Limiting Currency Volatility to Stimulate Goods Market Integration: A Price Based Approach," **NBER Working Papers** 8468.
- Rogoff, K., 1996, "The purchasing power parity puzzle," **Journal of Economic Literature**, XXXIV, pp 647-668.
- Rose, A. 2001 "One Money, One Market: Estimating the Effect of Common Currencies on Trade" **Economic Policy**.
- Taylor, A., 2001, "Potential Pitfalls for the PPP Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the LOOP" **Econometrica**, 69, pp. 473-498.
- Obstfeld M. and A.M. Taylor "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited", **Journal of the Japanese and International Economies** Vol. 11, No. 4, December 1997 pp. 441-479.
- Taylor, M., D. Peel and L.Sarno "Non-linear mean reversion in real exchange rates: Towards a solution to the purchasing power parity puzzles" , **International Economic Review**, 42, 4, pp 1015-1042.
- Tsay (1989) "Testing and modeling Threshold Autoregressive Processes", **Journal of American Statistical Association**, vol. 84, no.405.

4.1. Appendix 1: Data description

When considering possible alternative measures, a number of dimensions were borne in mind. First, a long time-series sample period was desirable to allow an econometric investigation of the economic behaviour of trade prices. Second, sectoral information was preferred to allow for sector specific responses to economy-wide shocks -as the pass-through literature advocates- and third, accuracy of the data was a concern, as input into policy making was required. These considerations gave rise to a set of data being observed consisting of disaggregated annual producer and import price data covering the period 1975-1995. We calculated price gaps as the difference, in absolute values, between the natural logarithms of import and producer price indices (base = 1990).

Import prices are from 52 (4-digit ISIC rev.2) German manufacturing sectors and producer prices from 17 industrialized trade partners (remaining EU countries, Canada, Japan and US) relative to 8 (2-digit ISIC rev.2) sectors. The full set of indices available covers products ranging from foodstuff to chemicals and to various manufactured goods mainly classified by material (ref. Table 3 below).

Table 3

PRODUCER AND IMPORT PRICES MATCHING TABLE

Original sector description for import prices	SITC Rev 3 code	IMPORT PRICES (1990Q1-2001Q4)	ISIC rev.2	PRODUCER PRICES (1990Q1-2001Q4)	ISIC rev.3
FOOD AND LIVE ANIMALS CHIEFLY FOR FOOD (0)	0+1	FOOD	31	FOOD AND BEVERAGES (15) TOBACCO (16)	15_16
BEVERAGES AND TOBACCO(1)		BEVERAGES AND TOBACCO			
CHEMICALS AND RELATED PRODUCTS, N.E.S.	5	Chemicals and Chemical products	351.2	Chemicals and Chemical products	24
MANUFACTURED GOODS CLASSIFIED CHIEFLY BY MATERIAL (6):					
Leather and leather manufactures	61	Leather products	323	Leather products	19
Cork and wood manufactures (excluding furniture)	63	Wood and Wood Products, Including Furniture	33	Wood and Wood Products, except Furniture	20
Paper , paperboard and related products	64	Paper and Paper Products, Printing and Publishing	34	Paper and Paper Products	21
Textile yarns, fabrics and related	65	Textiles (321)	321	Textiles	17
Non metallic mineral manufactures	66	Non-metallic mineral manufactures	36	Non-metallic mineral manufactures	26
Iron and steel	67	Iron and Steel	371	Basic metals	27
Non ferrous metals	68	Non ferrous metals	372	Mining of metal ores	13
Manufactures of metals	69	Manufactures of metals	381	Manufactures of metals	28
MACHINERY AND TRANSPORT EQUIPMENT	7	Metal Products Machinery and Equipment	38	Fabricated Metal Products Machinery and Equipment	28/35

The lack of highly disaggregated producer prices constrained us to combine 4-digit import prices with industry averages of producer price data. Import prices have been taken from the International Trade by Commodities Statistics (ITCS) and have been transformed

from SITC rev.3 to ISIC rev.2 by Paul Brenton. Producer prices have been downloaded from the OECD Indicators of Industry and Services database.

All data are measured in national currency of the exporter. Exchange rate volatility is measured as the standard deviation of the first difference of the end of period monthly natural log of the nominal exchange rate (IFS line ae) in the year preceding period t.

4.2. Appendix 2: Summary results for threshold model

Table 4

Summary results for Threshold Model			
(simple OLS regressions)			
all		lambda	half-life
	naive AR1	0.042 (-0.002)***	16.065
	TARoff_band	-0.042 (-0.002)***	16.11
	TARin_band	-0.048 (0.009)***	14.23
DM Block		lambda	half-life
	naive AR1	-0.014 (0.002)***	48.69
	TARoff_band	-0.014 (0.002)***	48.42
	TARin_band	0.001 not signif	895.08
Other EU		lambda	half-life
	naive AR1	-0.050 (0.003)***	13.39
	TARoff_band	-0.014 (0.003)***	48.57
	TARin_band	-0.013 not signif	54.77

Table 5

Summary results for TAR on country groups and data partitioned according to degree of volatility

(Panel, fixed effect on non demeaned data)

all		c	lambda	half-life	all	c	lambda	half-life
	AR1_fe	0.46 (0.009)	-0.33 (0.006)	1.74	AR1_fe	0.46 (0.009)	-0.33 (0.006)	1.74
	TARoff_fe	0.79 (0.017)	-0.40 (0.008)	1.36	TARoff_fe	0.79 (0.017)	-0.40 (0.008)	1.36
	TARin_fe	-0.19 (0.014)	-0.39 (0.019)	1.41	TARin_fe	-0.19 (0.014)	-0.39 (0.019)	1.41
DM		c	lambda	half-life	low_vol	c	lambda	half-life
	AR1_fe	0.46 (0.014)	-0.37 (0.011)	1.52	AR1_fe	0.53 (0.013)	-0.39 (0.009)	1.40
	TARoff_fe	0.82 (0.040)	-0.36 (0.017)	1.56	TARoff_fe	0.72 (0.023)	-0.38 (0.011)	1.46
	TARin_fe	-0.11 (0.010)	-0.36 (0.019)	1.56	TARin_fe	-0.20 (0.024)	-0.33 (0.027)	1.70
euvol		c	lambda	half-life	hi_vol	c	lambda	half-life
	AR1_fe	0.42 (0.014)	-0.29 (0.008)	2.02	AR1_fe	0.31 (0.017)	-0.21 (0.010)	2.90
	TARoff_fe	1.20 (0.134)	-0.32 (0.032)	1.81	TARoff_fe	1.01 (0.038)	-0.47 (0.017)	1.09
	TARin_fe	0.30 (0.012)	-0.31 (0.010)	1.89	TARin_fe	-0.08 (0.022)	-0.41 (0.04)	1.31