NBER WORKING PAPER SERIES

EXCHANGE RATE HYSTERESIS:
THE REAL EFFECTS OF LARGE VS SMALL POLICY MISALIGNMENTS

Richard Baldwin

Richard Lyons

Working Paper No. 2828

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1989

We gratefully acknowledge the helpful comments and suggestions of Ricardo Caballero, Mike Gavin, Rick Mishkin, Alberto Giovannini and workshop participants at Dartmouth University. This paper is part of NBER's research program in International Studies. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #2828
January 1989

EXCHANGE RATE HYSTERESIS:
THE REAL EFFECTS OF LARGE VS SMALL POLICY MISALIGNMENTS

ABSTRACT

Using the sticky price monetary model of exchange rate determination and the sunk cost model of trade hysteresis, we show that a sufficiently large policy misalignment can induce hysteresis in the trade balance and thereby alter the steady–state real exchange rate. Thus in our model exchange rate dynamics are path dependent, PPP need not hold and money need not be neutral even in the very long run.

We present only positive analysis but conjecture that the results have strong welfare, policy, and econometric implications. Since hysteresis in our model can entail industrial dislocation and the scrappage of sunk assets, we suggest that these factors may constitute a welfare cost of large policy misalignments that have not been formally considered. On the policy side, one could sensibly argue against the dollar volatility of the 1980s without at the same time arguing for a return to a formal exchange rate regime (because 1980s–size swings may involve welfare costs that 1970s–size swings do not). Lastly, since the long–run exchange rate is path dependent, standard empirical tests of exchange rate models may be misspecified.

Richard Baldwin
Columbia Business School
Columbia University
New York, NY 10027

Richard Lyons
Columbia Business School
Columbia University
New York, NY 10027

aggregate demand (equation 2.2). Given our focus on external balance, as well as for the sake of clarity, we streamline the price adjustment specification to include only the net export component of aggregate demand, which we model as a linear function of both the real exchange rate and the current market structure. In the steady state of the system net exports equal zero (trade is balanced). Firms can determine the current value of K (using (1.1)), and use it in making their entry/exit decisions based upon current and expected future real exchange rates.

Per equation (2.4), uncovered interest parity holds; however, there is a non–standard component to the expectation of e due to the presence of the hysteresis band. The band implies that the expected change in K need not be zero, even though the money stock innovations have mean zero. Indeed, the closer the system is to an edge of the band, the more likely it is that next period's realization of ϵ will induce a change in K. Since any change in K induces additional exchange rate movement, rational agents in the foreign exchange market would include this in their forecasts. The first component of the right–hand side of (2.5), ${}_{t}e_{t+1}$, captures the standard endogenous dynamics conditional on a constant K. The second component (which includes d_t) is a linear representation of the role of the expected change of K. The variable d_t tracks the center of the inaction band of real exchange rates (recall $s_t \equiv e_t - w_t$). The lower the real exchange rate is relative to the middle of the inaction band, the greater the chance that K will rise (firms will enter) in the following period. Since a rise in K induces an additional rise in e, depreciation of the home currency conditional on a constant K should be slightly less than the interest differential.[6,7] An interesting implication of this necessary change in the standard model is that the possibility of market structure change exerts a stabilizing effect on the exchange rate within the $s_t^I - s_t^O$ range (see also Krugman (1988a)).[8]

It should be noted that an innovation that changes market structure will also necessarily change the value of d_t according to (1.12). Since a change in d_t also has an impact effect on e, albeit a slight one, the coefficient θ in the linear representation must capture this too. Nonetheless, since both effects occur simultaneously and are monotonic, this added condition hardly seems to make the original linearity assumption any more egregious (and the linearity was

The extreme movements of the US dollar in the 1970s and 1980s have led many to question the market–knows–best philosophy that has dominated the fixed versus flexible exchange rate debate. Indeed, several serious proposals have been made to peg formally the major currencies to a commodity or currency basket. Many economists remain skeptical of such proposals since, inter alia, the arguments against floating rates are typically made without the aid of anything even vaguely approaching welfare analysis (volatility is just taken as an evil in itself). This absence is particularly detrimental in light of the fact that much of the exchange rate fluctuation is attributed to a lack of international policy coordination. Yet, the various policies of the 70s and 80s were chosen by governments facing the difficult economic and political tradeoffs of the period. It is conceivable that no formal exchange rate regime would have significantly altered the tradeoffs that were ultimately responsible for most of the observed exchange rate movement.

Despite this counter–argument, a general uneasiness with the performance of floating rates in the 1980s remains. The exchange rate volatility of the 1970s, although quite high relative to the previous decades under Bretton Woods, did not result in the extensive industrial dislocation and persistent trade imbalances observed in the 1980s. It seems therefore that we need to distinguish between the two types of volatility: the 1970s' "gyrations", and the 1980s' prolonged cycle. Conventional models, however, ignore this distinction since they assume that temporary exchange rate swings—irrespective of their size—can have only temporary real effects (apart from changes in external debt).

This paper provides a theoretical basis for the distinction between the volatility of the 1970s and 1980s. We show that sufficiently large policy misalignments can lead to hysteresis in trade and the long–run exchange rate while small misalignments cannot.[1] Therefore, a complete comparison of managed and floating rate regimes (which we take as a shorthand for policy coordination regimes) cannot ignore the possibility that large and small exchange rate movements have qualitatively different effects. While our paper presents only positive analysis, we conjecture that the results have strong welfare and policy implications.

In the model, hysteresis can entail industrial dislocation and the scrappage of sunk assets.

Our exploratory model is not rich enough to gauge the welfare impact of this, but it suggests that the sunk asset wastage involved in industrial dislocation may be an important cost of prolonged currency swings. On the policy side, our analysis suggests that one can sensibly argue against the enormous dollar movements of the 1980s without arguing for a return to a formal fixed rate regime. The reason is that at least one type of welfare cost is due to 1980s–size policy misalignments rather than those of 1970s–size.

From the perspective of exchange rate theory, our results show that the long–run equilibrium real exchange rate is not unique.[2] Indeed, we show that it is path–dependent, implying that standard empirical tests of exchange rate models may be misspecified. Additionally, our results contribute an item to the list of reasons why purchasing power parity should not hold even in the long run. This added item, however, differs greatly from the others in that it is a nominal (as opposed to real) change that ultimately shifts the long–run rate. To put it a bit differently, the possibility of exchange rate hysteresis provides another channel for the long–run non–neutrality of money.

Our model combines a simplified sticky price monetary model of exchange rate determination (Dornbusch 1976, Mussa 1982) with a simple sunk cost model of the trade sector (Baldwin 1986). The basic economics of the combined model is accordingly simple. A sufficiently large overvaluation can induce additional foreign firms to enter the domestic market. If market entry costs are sunk, a return of the exchange rate to its pre–overvalued level will not force out all the new foreign entrants. This hysteretic shift in market structure (number of competitors) would cause a hysteretic shift in the relationship between the real exchange rate and trade flows. If any sort of intertemporal budget constraint is to hold at the country level, a temporary shock cannot lead to a permanent trade imbalance; clearly, hysteresis in trade will feed back into hysteresis in the long–run equilibrium exchange rate. Consequently, a large temporary shock could, in the absence of future compensating shocks, lead to a hysteretic reallocation of resources in the traded goods sector and the abandonment of sunk assets.

The sticky price monetary model has been criticized for the fact that it is not derived from maximizing behavior. This of course detracts from its intellectual elegance, but more importantly implies that the model's parameters may not be structural. That is, some of the parameters may actually vary systematically with the variables in the model. The first best solution to this problem would be to work out a maximizing model of exchange rate determination starting with preferences and technology. However, to date such efforts (e.g., Lucas 1982, Stulz 1984, Svensson 1985) have failed to provide models that are capable of accounting for the short– and medium–run dynamics of exchange rates in the 1970s and 1980s.

Here we take a second–best approach. There is a broad spectrum of industry and macro level empirical evidence suggesting that the relationship between the real dollar and US trade flows has shifted significantly in the 1980s (Mann 1986, 1987a, 1987b; Baldwin 1988a, 1988b; Feinberg 1986, 1987; Froot and Klemperer 1987). This implies that the ad hoc nature of the sticky price model is particularly irksome with respect to the link between net exports and the exchange rate. We develop a fully rational, maximizing model to investigate the structural relationship between trade flows and the real exchange rate. We then integrate this model into a highly stylized sticky price monetary model of exchange rate determination. The purpose of this exercise is not to present a realistic model of the world economy. Rather it is intended to clarify our thinking on (1) the feedback effects that sunk cost hysteresis in trade has on short– and long–run exchange rate dynamics in the context of a standard exchange rate model, (2) whether endogenizing exchange rate dynamics might in some way eliminate the possibility of hysteresis in trade, and (3) the importance of distinguishing between large and small policy misalignments.

We are fully aware that the mis–match in the degree of microeconomic justification of the two components of our model may appear awkward to some readers. However, we view this mis–match as one of the strengths of the approach: the industrial organization model allows us to clean up an empirically questionable aspect of the sticky price model, without fundamentally altering its appealing dynamic properties. We choose the sticky price model because the two standard competing models of exchange rate determination are not appropriate given the questions at hand: the flexible price monetary model imposes purchasing power parity and therefore is useless for investigating pronounced real exchange rate variation; the portfolio balance model emphasizes issues that are not our primary concern.

The idea that large exchange rate shocks can have qualitatively different trade effects than small shocks dates back at least to the quantum effect of Orcutt (1950). The sunk cost model of hysteresis in trade was introduced by Baldwin (1986). That paper, and a later version Baldwin (1988a), defined the basic sunk cost model and demonstrated the possibility of hysteresis in trade price and volume. Kemp and Wan (1974) showed that the Heckscher–Ohlin model is subject to hysteresis in the long–run trade pattern when factor prices are not equalized instantaneously and labor is internationally mobile. Their model does not involve sunk costs of any type. The intellectual history of the sticky price monetary model (Dornbusch 1976 and extensions) is well–known.

The feedback of sunk cost hysteresis into exchange rate determination is examined in Baldwin and Krugman (1986) and Krugman (1988b). Both papers use the income–elasticity approach to the balance of payments as a model of exchange rate determination. Essentially, net capital inflow is taken as an exogenous process and the exchange rate is determined in the goods market. That is, the exchange rate is assumed to jump to whatever point is necessary to move the balance of trade to the level where it just offsets this period's exogenous capital inflow. While these papers help define the basic issues, they do not truly work out the feedback effects of hysteresis. The values of the world's major currencies are generally thought to be determined by asset market conditions; therefore, net capital flows and exchange rates should be co–determined by more elemental policy or real shocks.

The paper has four sections. The first examines an industry–level model of trade hysteresis. The second examines the dynamics of a bare–bones sticky price monetary model allowing for the possibility of trade hysteresis. The third presents a number of thought–experiments to spotlight the policy implications of the model. The last section presents a summary and concluding remarks.

I. Hysteresis in the Trade Balance

The trade flows of a country are often modeled as depending on real exchange rates and national incomes. Our simplified exchange rate model in section II takes national outputs as constant, allowing us to highlight the exchange rate–trade balance link. In this section we first work out the evolution of the number of competitors in an import sector marked by sunk market–entry costs (the hysteretic sector), showing that hysteresis in imports is a possibility. Then we aggregate the hysteretic sector and a normal (non–hysteretic) sector to show that hysteresis in the trade balance can occur.

A. Hysteresis and The Evolution of Market Structure

Consider a representative foreign firm located in the foreign country which has the possibility of selling to the home country. In order to sell to the home market the firm must incur a sunk cost F; to continue selling during subsequent periods, it must incur a fixed maintenance cost M. If M is not expended the firm is said to exit. Re–entry then requires F to be re–incurred. We assume that $F > M$ and both are fixed in home currency terms. If the firm is active in the home market in period t, it earns operating profits π_t (in home currency units) which are related to the log of the real exchange rate, s (price home goods/price foreign goods), and the number of foreign firms competing in the market, n, by the function $\pi[s_t,n_t]$. The function is assumed to be bounded above and below, and non–increasing in both s and n. We assume all firms are identical and entry is not restricted. Also, we ignore the integer constraint on n. A special case which would lead to such a profit function is that of period–by–period Cournot competition among symmetric firms producing a homogeneous product. It may be useful to think of F as the cost of setting up a distribution and after–sales service network, and M as the cost of maintaining these assets. The timing is as follows: at the beginning of the period the firm observes the current period's monetary shock and then decides whether to be active on the basis of the implied real exchange rate (s and n are thus jointly–determined).

Given the function, π, we re–phrase the firm's entry–exit decision in terms of an active–inactive decision: Each period the firm decides whether or not to buy a "license" to be active. If the firm buys the license, it receives $\pi[s_t, n_t]$. If it does not, it receives nothing. Introducing a binary state variable A ($A_t = 1$ if the firm was active last period, $A_t = 0$ otherwise), the cost of the activity license is $F + DA_t$ where $D \equiv M - F$.

Optimal Entry–Exit Strategy

Since the derivation of the representative firm's optimal entry–exit strategy is rather involved, we first provide an intuitive explanation of the optimal strategy. To fix ideas, consider for the moment the standard case of a firm facing a fixed cost in a static setting. The optimal strategy in this case is that the firm should choose to be active only if this period's profits can at least cover the fixed cost; that is, only if $\pi_t \geq G$, where G is the fixed cost. In a multi–period setup, this entry condition must be modified when at least some of the entry cost is sunk. With sunk costs, the firm finds it cheaper to stay "in" the market than it is to get into the market. Consequently, being active this period may provide the firm with an advantage during future periods. We refer to the expected value of this advantage as the "incumbency" premium and denote it as Ψ. Thus the entry condition with sunk cost can be viewed as a modification of the standard entry condition: the firm should choose to be active only if the sum of the incumbency premium and this period's profits can at least cover the costs of being active (which will be F or M); that is, $\pi_t + \delta\Psi > F + DA_t$, where δ is a constant discount factor.

More formally, the state variables are s_{t-1}, n_{t-1}, A_t, and ϵ_t. For notational convenience, the first three are grouped in the vector $x_t = (s_{t-1}, n_{t-1}, A_t)$. The control variable is U_t ($U_t = 1$ if the firm chooses to be active, $U_t = 0$ otherwise). The discrete–time system function f,

(1.1) $\quad x_{t+1} = f[x_t, U_t, \epsilon_t]$, conveniently summarizes the laws of motion:

(1.2) $$A_{t+1} = U_t,$$

(1.3) $$s_t = h[s_{t-1}, n_{t-1}, \epsilon_t], \text{ and}$$

(1.4) $$n_t = n\left[s_t, n_{t-1}\right],$$

where the function h is implicitly defined by (2.11) and (1.4), and ϵ is an independently and identically distributed (i.i.d.) random variable related to the i.i.d. money supply disturbance η in (2.8). We denote the random variable as ϵ and its realizations as ϵ_t.

The typical firm chooses an optimal entry–exit strategy, μ, to maximize the discounted cash flows (μ states U_t as a function of x_t and ϵ_t):

$$(1.5) \qquad V^{\mu}[x_0,\epsilon_0] = E\{ \sum_{t=0}^{\infty} \delta^t g\left[x_t,\mu[x_t,\epsilon_t],\epsilon_t\right] \}, \quad \text{where}$$

$$g\left[x_t,\mu[x_t,\epsilon_t],\epsilon_t\right] = \begin{cases} \pi[s_t,n_t] - F - DA_t, & \text{if } \mu[x_t,\epsilon_t] = 1 \\ 0, & \text{if } \mu[x_t,\epsilon_t] = 0 \end{cases}$$

and

$$x_{t+1} = f\left[x_t,\mu[x_t,\epsilon_t],\epsilon_t\right].$$

Here the expectations are over all ϵ_t. The two primary goals of this subsection are to characterize the optimal entry–exit strategy, μ, and to characterize the evolution of the market structure implicit in (1.4).

The solution to (1.5) is characterized using Bellman's optimality principle. That is, we obtain a characterization of μ by asking what is optimal this period given that we already know the optimal strategy from next period onward. The apparent circularity of this approach can be avoided by first characterizing the optimal strategy for a similar finite horizon problem (working backward from the last period), and then studying how the strategy changes as the horizon tends toward infinity. However, standard dynamic programming techniques allow us to obtain a characterization more parsimoniously. To this end, we suppose for the moment that the value function $V[x_t,\epsilon_t]$ (which depends on μ) is known and the firm can use this together with its knowledge of the i.i.d. shock to form an expectation of its optimal value in period t+1 conditional on information available in period t. This allows us to study μ by solving the much simpler sub–problem: choose U_t to maximize,

$$(1.6)\qquad g[x_t,U_t,\epsilon_t] + \delta E\{V\Big[f[x_t,U_t,\epsilon_t],\epsilon\Big]\},$$

where the expectation is over the period t+1 realization of ϵ.

The Decision of a Potential Entrant

A firm that was inactive last period can enter or remain inactive. An optimizing firm calculates the value of (1.6) under each option, compares the two values and chooses the option that yields the highest. Consequently, its optimal value in period t is:

$$(1.7)\qquad V[x_t,\epsilon_t] \equiv \text{Max}\left\{\pi[s_t,n_t] - F - DA_t + \delta EV\Big[f[x_t,1,\epsilon_t],\epsilon\Big],\ \delta EV\Big[f[x_t,0,\epsilon_t],\epsilon\Big]\right\},$$

where again the expectations are over the realizations of ϵ in t+1, and s_t and n_t are related to x_t and ϵ_t by (1.1). This equation implicitly defines the function $V[x_t,\epsilon_t]$. It also implicitly defines the decision rule as (recall that we are considering the case where $A_t = 0$): Be active only if ϵ_t is such that:

$$(1.8)\qquad \pi\left[h[s_{t-1},n_{t-1},\epsilon_t],n\Big[h[s_{t-1},n_{t-1},\epsilon_t],n_{t-1}\Big]\right] + \delta\Omega[x_t,\epsilon_t] > F + DA_t, \text{ where}$$

$$\Omega[x_t,\epsilon_t] \equiv EV\left[f[(s_{t-1},n_{t-1},A_t),1,\epsilon_t],\epsilon\right] - EV\left[f[(s_{t-1},n_{t-1},A_t),0,\epsilon_t],\epsilon\right].$$

The expectations are over the period t+1 realization of ϵ. The function Ω gives the incumbency premium (which is a conditional expectation) as a function of the period t state variables.

To get a handle on the decision rule implicit in (1.8), we focus on the value of ϵ_t that makes all potential entrants just indifferent to entry. This borderline value partitions the state space into two decision regions, and is equal to the ξ_t that solves (recall $A_t = 0$):

$$(1.9)\qquad \pi\left[h[s_{t-1},n_{t-1},\xi_t],n\Big[h[s_{t-1},n_{t-1},\xi_t],n_{t-1}\Big]\right] + \delta\Omega[x_t,\xi_t] = F + DA_t,$$

By assumption, s_{t-1}, n_{t-1}, A_t, and ϵ_t are observed by the firm prior to its activity–inactivity decision, so the most mathematically elegant solution to (1.5) is to write μ in terms of these four variables (as is done implicitly in (1.9)). However, the economics of the problem comes out more clearly when we make the mathematical detour of writing an equivalent decision rule in terms of s_t, n_t and A_t (we call this rule $\varphi[s_t,n_t,A_t]$). This detour is possible since (1) firms can infer s_t and n_t from x_t and ϵ_t using (1.1), and (2) the triplet (s_t,n_t,A_t) constitutes a sufficient statistic for the state of the system. In other words, we can express the period t value of the firm as a function, $J[s_t,n_t,A_t]$, such that $J[s_t,n_t,A_t] = V[x_t,\epsilon_t]$ for all x_t and ϵ_t. Heuristically, this is due to the single lags in (2.11) and (1.4) so that s_t and n_t summarizes all effects of history. Formally this can be shown by construction, using (2.11), (2.3), and the definition of g in (1.5). It may seem strange that a firm can determine n_t before any firms (including itself) have decided to be active or not. However, given our assumption of free–entry and no integer constraint, the firm's own activity decision has no effect on the number of firms in the market. If entry looks advantageous, firms will enter up to the point where the marginal firm is just indifferent to entry. This will be true regardless of whether any particular firm is among the entrants. Incidentally, since all firms are symmetric, if the marginal entrant is indifferent to entry, so are all other non–incumbents.

The equivalent decision rule expressed in terms of s_t, n_t and A_t is: Be active only if:

$$\text{(1.10)} \qquad s_t < \alpha_t, \quad \text{such that}$$

$$\pi[\alpha_t,n_t] + \delta\Psi[\alpha_t,n_t] = F + DA_t, \text{ and}$$

$$\Psi[\alpha_t,n_t] \equiv EJ\left[h[\alpha_t,n_t,\epsilon],n\left[h[\alpha_t,n_t,\epsilon],n_t\right],1\right] - EJ\left[h[\alpha_t,n_t,\epsilon],n\left[h[\alpha_t,n_t,\epsilon],n_t\right],0\right].$$

Here we could have used the fact that if $s_t = \alpha_t$, no firms enter so $n_t = n_{t-1}$. Notice that as long as n_t does not change, the critical value α_t depends only on A_t. The importance of this is that the entry–exit strategy is time–invariant during intervals when n_t is stable.

Figure 1 depicts the optimal decision rule and α_t. We refer to α_t when $A_t = 0$ as s_t^I since this is the s_t which makes entrants just indifferent to coming "in". The NC line shows the net cost of being active this period (F net of the incumbency premium). It is positively sloped since high (unfavorable) realizations of s_t make high realizations of s more likely in all future periods. This reduces the incumbency premium since it makes it less likely that the firm will benefit in future periods from being active in t. Any realization of s_t that is greater than or equal to s_t^I does not induce previously inactive firms to enter. For s_t's more favorable (lower) than s_t^I, firms would enter up to the point where the marginal firm is just indifferent to entry. Thus n_t would be implicitly defined by the condition that (recall $A_t = 0$):

$$\pi[s_t, n_t] + \delta\Psi[s_t, n_t] = F + DA_t, \tag{1.11}$$

This relationship implicitly defines one part of the law of motion (1.4). Clearly, for $s_t < s_t^I$ the more favorable is s_t, the larger will be n_t.

The Decision Facing an Incumbent

Consider the problem of a typical firm that was active during the previous period (so $A_t = 1$). Its options are to exit or to stay active. Its value in period t is therefore given by (1.7) taking $A_t = 1$. The decision rule can be characterized by (1.9) and (1.10), taking $A_t = 1$. To pin down the optimal decision rule we focus on the borderline value, s_t^O, at which an incumbent is just indifferent between staying active and exiting. This value is defined by the α_t which solves (1.10) for $A_t = 1$. It can be seen graphically in figure 2 which is similar to figure 1 except that the NC line is drawn at $M - \delta\Psi[\cdot, n_{t-1}]$, instead of $F - \delta\Psi[\cdot, n_{t-1}]$. Any s_t less than or equal to s_t^O will not lead any incumbents to exit. Values of s_t greater than this will lead some firms to exit. Indeed, they will exit up to the point where the remaining number of firms is given implicitly by (1.11), taking $A_t = 1$. This characterizes another part of the law of motion (1.4). Clearly, for values of s_t greater than s_t^O, the larger is s_t, the smaller will be n_t.

Putting together the problems of the incumbents and potential entrants completes our characterization of $\varphi[s_t, n_t, A_t]$, and the law of motion for n implicitly given by (1.4). Figure 3, which combines figures 1 and 2, divides the realizations of s_t into three regions. If s_t is favorable enough (less than s_t^I), firms enter until (1.11) holds for $A_t = 0$. If s_t is bad enough (greater than s_t^O), firms exit until (1.11) holds for $A_t = 1$. Any s_t in the s_t^I–s_t^O range induces neither entry nor exit. Firms that were active last period remain active; firms that were inactive last period remain inactive. This is the no–entry–no–exit or hysteresis band. The decision rule is:

$$\text{(1.12)} \qquad \varphi[s_t,n_t,A_t] = \begin{cases} 1, & \text{if } s_t < \alpha_t \\ 0, & \text{if } s_t \geq \alpha_t \end{cases}, \qquad \text{such that}$$

α_t is the solution to (1.10). Equation (1.4) is:

$$\text{(1.13)} \qquad n_t = \begin{cases} \pi[s_t,n_t] + \delta\Psi[s_t,n_t] = F, & \text{for } s_t < s_t^I \\ n_{t-1}, & \text{for } s_t^I \leq s_t \leq s_t^O \\ \pi[s_t,n_t] + \delta\Psi[s_t,n_t] = M, & \text{for } s_t > s_t^O \end{cases}.$$

Definition of the Expectations

Up to this point we have simply assumed that the value functions, V and J, and their expectations exist. The proof that they do, and that (1.9) and (1.10) indeed describe optimal decision rules, is a straightforward application of standard textbook results (e.g., Bertsekas 1976, Bellman 1957). The proof is therefore omitted (a sketch of the proof can be found in Baldwin 1988c). The important assumptions are that π is bounded, $0 < \delta < 1$, and the underlying disturbance is i.i.d.. Nonetheless, to build intuition we examine these expectations more closely.

The expected value (at time t) of the firm in t+1 can be found by determining what the value in t+1 will actually be for any realization of s_{t+1} and then integrating over all possibilities weighted by the probability measure of s_{t+1}. Consider first the expectation of a firm that chooses to be active in t, so that it faces the incumbent's decision rule in t+1: For $s_{t+1} \leq s_{t+1}^O$ it stays in, so its value is $\pi[s_{t+1},n_{t+1}] - M + \delta EJ[s_{t+2},n_{t+2},1]$; for $s_{t+1} > s_{t+1}^O$ the firm exits, so it is worth

$\delta EJ[s_{t+2},n_{t+2},0]$. Integrating over these possibilities (using the laws of motion to express s_{t+2} and n_{t+2} in terms of s_t and n_t):

$$(1.14)\quad \int_{z=-\infty}^{s^O_{t+1}} \{\pi\left[z,n[z,n_t]\right]-M+\delta EJ\left[h\left[z,n[z,n_t],\epsilon\right],n\left[h\left[z,n[z,n_t],\epsilon\right],n[z,n_t]\right],1\right]dP_s[z|s_t,n_t]$$

$$+\int_{z=s^O_{t+1}}^{\infty} \delta EJ\left[h\left[z,n[z,n_t],\epsilon\right],n\left[h\left[z,n[z,n_t],\epsilon\right],n[z,n_t]\right],0\right]dP_s[z|s_t,n_t],$$

where z is the variable of integration and $P_s[\cdot|s_t,n_t]$ is the conditional distribution function of s_{t+1}. Equation (1.14) expresses the expectation (conditioned on s_t and n_t) of the firm's value in t+1 when it chooses to be active in period t. Similarly, the conditional expectation of a firm's value in t+1 when it chooses to be inactive in period t is:

$$(1.15)\quad \int_{z=-\infty}^{s^I_{t+1}} \{\pi\left[z,n[z,n_t]\right]-F+\delta EJ\left[h[z,n[z,n_t],\epsilon],n\left[h\left[z,n[z,n_t],\epsilon\right],n[z,n_t]\right],1\right]dP_s[z|s_t,n_t]$$

$$+\int_{z=s^I_{t+1}}^{\infty} \delta EJ\left[h[z,n[z,n_t],\epsilon],n\left[h\left[z,n[z,n_t],\epsilon\right],n[z,n_t]\right],0\right]dP_s[z|s_t,n_t],$$

Actually, (1.15) is zero for all t. A firm that chose to be out in period t may enter in t+1 if s_{t+1} turns out to be favorable enough. However, so will other firms, driving the expected value of the firm to zero, as pointed out in (1.11). If the realization of s_{t+1} is greater than the critical value s^I_{t+1}, the firm stays out. Since if the firm ever does enter in the future, it will have zero value, it is obvious that $EJ[s_{t+1},n_{t+1},0]$ is zero.

The expected value of V in (1.9) could be constructed in a similar manner, using ξ^I_{t+1} and

ξ^{O}_{t+1} as the limits of integration, and the density of ϵ.

As it turns out a good deal of intuition can be gathered by examining the value of the incumbency premium Ψ more closely. For realizations of s_{t+1} less than s^{I}_{t+1} firms will enter until the value of potential entrants is zero. Since the incumbents pay only M instead of F, the difference between the value of an incumbent firm and a non–incumbent in this situation is simply (F – M). For realizations of s_{t+1} between s^{I}_{t+1} and s^{O}_{t+1}, a previously active firm will remain active but no new firms will enter. The difference over this range will therefore be $\pi[s_{t+1},n_{t+1}] - M + \delta\Psi[s_{t+1},n_{t+1}]$. Lastly for s_{t+1} greater than s^{O}_{t+1} the difference is zero because for such realizations, firms either exit and are worth zero, or stay in and earn operating profits which exactly equal M net of the incumbency premium (as in (1.11)). Taking expectations over these outcomes, $\Psi[s_t,n_t]$ is:

$$(1.16) \qquad (F-M)\int_{z=-\infty}^{s^{I}_{t+1}} dP_s[z|s_t,n_t] + \int_{s^{I}_{t+1}}^{s^{O}_{t+1}} \left\{\pi\left[z,n[z,n_t]\right] - M + \delta\Psi\left[z,n[z,n_t]\right]\right\} dP_s[z|s_t,n_t]$$

B. Trade Balance Hysteresis

To study the feedback of import hysteresis to the exchange rate dynamics, we assume that there are two trade sectors: a hysteretic sector and a normal or non–hysteretic sector. For simplicity, the possibility of hysteresis in exports is ignored.[4] We assume the hysteretic sector's import level, H, is related to the real exchange rate and the number of firms in the market:

$$(1.17) \qquad H_t = H[s_t,n_t],\ H_s[\cdot,\cdot] < 0,\ \partial H[\cdot,\cdot]/\partial n_t > 0 \text{ (everywhere)},$$

In the normal sector, the relationship between the real exchange rate and the value of the trade balance is assumed to be invariant to exchange rate shocks. That is:

$$B^n[s_t] = X[s_t] - M[s_t], \;\; dB^n[\cdot]/ds_t > 0 \text{ (everywhere)} \tag{1.18}$$

where X and M are the real value (in terms of the import good) of exports and normal–sector imports respectively. Clearly the number of firms operating in the normal sector could affect the level of trade. However in absence of sunk costs, these firms do not face asymmetric entry and exit conditions. The number of firms would thus depend monotonically on the real exchange rate. Given this, we can describe the relationship between B^n and the s with a reduced form such as (1.18). Putting the two sectors together we have that the balance of trade, B_t, is:

$$B[s_t,n_t] = B^n[s_t] - H[s_t,n_t], \;\; \partial B[\cdot,\cdot]/\partial s_t > 0, \; \partial B[\cdot,\cdot]/\partial n_t < 0. \tag{1.19}$$

The dynamic behavior of B_t is implicitly controlled by the equations (1.1) and (1.12).

II. The Macro Implications of Hysteresis

Baldwin and Krugman (1986) model the macro effects of hysteresis using an elasticities approach to exchange rate determination. The driving force in their system is an exogenous, stochastic net capital flow each period which must be offset by trade balance adjustment induced by real exchange rate changes. As they are quick to point out, this approach to the relevant macro linkages clearly leaves much to be desired. Here we implement the sticky price monetary model of Dornbusch (1976). Given the state of the art in exchange rate determination, this model clearly dominates the two principal alternatives for the questions at hand: flexible price monetary models assume purchasing power parity (PPP) holds at all times—a difficult context in which to talk about the persistent effects of large real exchange rate shocks. Portfolio balance models, which do allow real exchange rate changes, emphasize considerations that are not of primary concern here.

Before launching in, however, a couple of the characteristics of the standard setup need to be considered. First, the system as typically specified implies a long–run steady state at PPP. Yet, hysteresis should involve a shift in the long–run exchange rate as market structure changes. That is, the system should have a "memory" of its past states that the standard system does not. We will refer to this property as path–dependent–PPP (PDPPP): the path of the real exchange rate will determine the market structure, which in turn determines the real exchange rate toward which the system regresses. And, second, we want the direct effect of real exchange rate changes on net exports to be governed by a constant parameter since, empirically, this elasticity appears stable in the face of adjustments that do not induce a change in market structure.

These characteristics are captured in the simplified, sticky price economy:

(2.1)	$m_t - w_t = -\lambda i_t$	LM Curve
(2.2)	$w_{t+1} - w_t = \gamma(e_t - w_t - K_t)$	Price Adjustment
(2.3)	$K_t = -a + \beta n_t$	Market Structure
(2.4)	$E[e_{t+1} \mid \Omega_t] - e_t = i_t - i_t^*$	Uncovered Parity
(2.5)	$E[e_{t+1} \mid \Omega_t] = {}_te_{t+1} + \theta[d_t - (e_t - w_t)]$	Expectations
(2.6)	$d_t \equiv s_t^I + (1/2)(s_t^O - s_t^I)$	Center of Band
(2.7)	$m_{t+1} = m_t + \eta_{t+1}$	Money Process
(2.8)	η ~ i.i.d. Uniform (mean zero)	Innovations

Here, m is the log of the nominal money stock; w is the log of the domestic price level; i is the nominal interest rate (asterisk denotes foreign); and e is the log of the nominal exchange rate (home/foreign).[5] As in section I, we assume that output remains at the full employment level. The variable K captures the effect on the system of market structure change; the more competitive is the import sector (as indexed by the number of foreign firms n_t), the larger is K and the lower is net exports.

As is standard with sticky price monetary models, price adjustment is tied to the level of aggregate demand (equation 2.2). Given our focus on external balance, as well as for the sake of clarity, we streamline the price adjustment specification to include only the net export component of aggregate demand, which we model as a linear function of both the real exchange rate and the current market structure. In the steady state of the system net exports equal zero (trade is balanced). Firms can determine the current value of K (using (1.1)), and use it in making their entry/exit decisions based upon current and expected future real exchange rates.

Per equation (2.4), uncovered interest parity holds; however, there is a non–standard component to the expectation of e due to the presence of the hysteresis band. The band implies that the expected change in K need not be zero, even though the money stock innovations have mean zero. Indeed, the closer the system is to an edge of the band, the more likely it is that next period's realization of ϵ will induce a change in K. Since any change in K induces additional exchange rate movement, rational agents in the foreign exchange market would include this in their forecasts. The first component of the right–hand side of (2.5), ${}_te_{t+1}$, captures the standard endogenous dynamics conditional on a constant K. The second component (which includes d_t) is a linear representation of the role of the expected change of K. The variable d_t tracks the center of the inaction band of real exchange rates (recall $s_t \equiv e_t - w_t$). The lower the real exchange rate is relative to the middle of the inaction band, the greater the chance that K will rise (firms will enter) in the following period. Since a rise in K induces an additional rise in e, depreciation of the home currency conditional on a constant K should be slightly less than the interest differential.[6,7] An interesting implication of this necessary change in the standard model is that the possibility of market structure change exerts a stabilizing effect on the exchange rate within the $s_t^I - s_t^O$ range (see also Krugman (1988a)).[8]

It should be noted that an innovation that changes market structure will also necessarily change the value of d_t according to (1.12). Since a change in d_t also has an impact effect on e, albeit a slight one, the coefficient θ in the linear representation must capture this too. Nonetheless, since both effects occur simultaneously and are monotonic, this added condition hardly seems to make the original linearity assumption any more egregious (and the linearity was necessary to maintain the tractability of the model).

Normalizing $i^*=0$, the solution of the model generates the saddle–path equation:[9]

(2.9) $$e_t = Gw_t + Qm_t + K_t + Q\lambda\theta d_t$$

where $$G \equiv \frac{\rho+\gamma-1}{\gamma}$$

$$Q \equiv 1 - G$$

and $$\rho \equiv 1 + \left[\frac{\theta-\gamma}{2}\right]\left[1 + \left[1 + \frac{4\gamma}{\lambda(\theta-\gamma)^2}\right]^{1/2}\right] < 1.$$

Note that G+Q=1—money is neutral in the long run in the event there is no change in market structure. Now the role of market structure, as captured by K, is quite clear: the entry or exit of foreign firms induces a shift in the saddle–path. Moreover, any non–zero value of K implies that the steady–state value of the real exchange rate, (e_t-w_t), is no longer zero.

At any point in time, the expected evolution of the system is described by:

(2.10) $$\begin{bmatrix} e_t \\ w_t \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ \gamma(\rho+\gamma-1)^{-1} \end{bmatrix} (\rho)^t + \begin{bmatrix} m_t + \lambda\theta d_t + K_t \\ m_t + \lambda\theta d_t \end{bmatrix}$$

where C_1 is a constant determined by initial conditions. Note that since the role of d_t in the long–run values of e and w is the same, d_t plays no role in the determination of the long–run real exchange rate (e–w). Equation (2.10) implies that the real exchange rate follows an autoregressive process that admits a role for market structure:

(2.11) $$s_t = \rho s_{t-1} + (K_t - \rho K_{t-1}) + (Q\lambda\theta)(d_t - d_{t-1}) + \epsilon_t$$

where $\epsilon_t \equiv Q\eta_t$.

The response of the system to a negative monetary innovation is illustrated in figure 4. The initial equilibrium is at point A along the 45° line ($K_0=0$). The dotted lines represent "profitability contours" for the foreign firms, with higher profits in the northwesterly direction. If the shock is not large enough to induce a change in market structure then the adjustment is the standard jump to B then movement along SP_1 to steady state at C. If, on the other hand, the real appreciation is large enough to induce foreign firms to enter then the saddle–path SP_2 is appropriate and the steady state now must fall on the line along which e=w+K. Thus, the jump is to D with gradual adjustment to steady state at E.

The illustration makes another point clear. Baldwin and Krugman (1986) focus on large real exchange rate shocks. Yet, what matters to firms is which profitability contour the system is on, not how it got there. Consequently, a series of positively autocorrelated small shocks can also be a source of market structure change. This is particularly important in a context where agents are learning, and these correlated small "shocks" are their rational revisions over time [see Lewis (1988) and Lyons (1987)]. For example, one might argue that the upswing of the dollar cycle resulted from a series of revisions of agents' expectations of fiscal deficit decay. Such a series of revisions (autocorrelated small shocks) could push the real exchange rate to the point where it induced a change in market structure.

III. Exchange Rate Hysteresis

As we show in the previous section, the dynamics of exchange rate adjustment and the long–run equilibrium exchange rate depends upon the market structure (number of competitors) in the trade sector. Consequently, a shock that leads to a hysteretic change in the market structure will lead to exchange rate hysteresis.[10] In this section we show that hysteresis in the market structure is a possibility even allowing for feedback, and show that there is a qualitative difference between a policy regime which entails only small policy misalignments and one in which large misalignments are possible.

A. A Range of Steady–State Exchange Rates

To investigate the properties of the model, we first consider the benchmark case where the system is at steady state and is expected to remain there forever. Using "^" to indicate that the value of the variable is a steady–state value, the optimal expected value of a representative firm that is operating in the hysteretic–sector is:

$$\hat{J}' = \left[\pi[\hat{s}',\hat{n}'] - M\right]\left(\frac{1}{1-\delta}\right) \tag{3.1}$$

If $\hat{n}'$ is itself a steady–state value, then it must be that neither additional entry nor exit is optimal. This condition places the following bounds on the value of $\hat{V}'$:

$$0 \leq \hat{J}' \leq (F - M) \tag{3.2}$$

The bounds on $\hat{J}'$ can be directly translated into bounds on $\hat{n}'$ (assuming that the inverse of π with respect to n exists):

$$\pi_n^{-1}[\hat{s}', (F-M)(1-\delta)+M] < \hat{n}' < \pi_n^{-1}[\hat{s}', M] \tag{3.3}$$

Likewise the balanced trade condition (which is implicit in (2.2)) defines an implicit link between $\hat{s}'$ and $\hat{n}'$. That is the pair $\hat{s}'$, $\hat{n}'$ must satisfy the condition:

$$0 = B[\hat{s}, \hat{n}]. \tag{3.4}$$

Assuming $B[\cdot,\cdot]$ is monotonic and continuous, there exists a function,

$$\hat{s} = s[\hat{n}], \tag{3.5}$$

which relates the steady–state values of s and n.

Putting this all together, we see that, unlike the standard sticky price model, the model can now support a whole range of steady–state exchange rates and market structures. As we shall illustrate below, hysteresis is a possibility since a sufficiently large shock could shift the system from one steady–state equilibrium to another.

B. The Possibility of Exchange Rate Hysteresis

To demonstrate formally the possibility of hysteresis in market structure we take the usual circuitous route of showing that the contrapositive implies a contradiction. Assuming that hysteresis is not a possibility and that the system starts at a steady–state equilibrium, then even a shock large enough to induce some entry should be expected eventually to reverse itself in the absence of additional shocks. That is, while a shock may change the number of firms and the real exchange rate, these changes are only temporary. Our task is to show that this is not possible.

The argument is facilitated by first establishing three facts. A negative monetary innovation this period (call it period 1) will lead to an immediate real appreciation. The first fact is that a representative firm expects the real exchange rate to decay in a strict monotonic fashion back up to $\hat{s}$, and the probability density function (p.d.f.) of s to also shift monotonically back up to the p.d.f. of $\hat{s}$ as pictured in figure 4. This fact follows directly from the section II analysis. Second, the firm expects that expected operating profits π_t will decay monotonically (although not necessarily strictly monotonically) back down to $\pi[\hat{s},\hat{n}]$. This follows directly from fact 1, the law of motion for n_t given by (1.13), and the assumption that π is decreasing in both its arguments. Third, if a firm never exited its value decays monotonically back to $J[\hat{s},\hat{n},1]$. Since $J[s_\tau,n_\tau,1] \geq \bar{J}_\tau \equiv \sum_{t=\tau}^{\infty} \delta^{t-\tau}[\pi(s_t,n_t)-M]$ (for all s_τ, n_τ), this establishes a lower bound on V.

Now for the argument. For $\hat{s}$ and $\hat{n}$ to be steady–state values, it must be that for a representative firm in the market:

$$0 \leq J[\hat{s},\hat{n},1] \leq (F-M), \text{ where} \tag{3.6}$$

$$J[\hat{s},\hat{n},1] = \left[\pi[\hat{s},\hat{n}] - M\right]\left[\frac{1}{1-\delta}\right].$$

Consider a period one shock, $\bar{\epsilon}_1$, that induces entry. As was shown in section I, s_1 must satisfy:

$$\pi[s_1,\hat{n}] + \delta E\{J[s_2,n_2,1] \mid s_1,\hat{n}\} > F. \tag{3.7}$$

where $E\{J[s_2,n_2,1] \mid s_1,\hat{n}\}$ is the expectation of the firm's value in period 2 conditioned on s_1 and $\hat{n}$. Given $\bar{\epsilon}_1$, firms would enter until n_1 satisfied:

$$\pi[s_1,n_1] + \delta E\{J[s_2,n_2,1] \mid s_1,n_2\} = F. \tag{3.8}$$

Since all firms including the new entrants have the same value from period 2 onwards, the third fact implies that the new firms expect their value to decrease monotonically from its period 2 level. Furthermore, the representative firm's value is never less than $\bar{J}_\tau$, so it is also never less than $J[\hat{s},\hat{n},1]$. This is a contradiction. If the value of the firm is decreasing to the steady–state value, but the steady–state value is greater than zero, by (3.1), then none of the new entrants would ever exit, so n_1 would not return to $\hat{n}$.

Hysteresis with One–Time–Only Shocks

Although this constitutes a proof, we illustrate the possibility of hysteresis by working through a simple example. The equations (2.1)–(2.8) allow for the system to be continually subjected to shocks. However, one useful simplification is to ask: Starting from a steady–state position, how would the system react to a one–time unexpected shock in the absence of additional shocks? Given the standard assumptions on the eigenvalues of the system, the size of a one–time–only shock is a well–defined concept. That is, a negative shock of size ϵ^0 will lead to an adjustment path as depicted in figure 4 as SP_1. This path could be described by a function relating the level of the real exchange rate to the size of the one–time shock, ϵ, and time: $s_t = s[\epsilon,t]$. A slightly larger negative shock $\epsilon' < \epsilon^0$ will lead to the adjustment path SP_2 which is parallel to and everywhere to the left of SP_1. In other words we know that $s[\epsilon^O,t] < s[\epsilon',t]$ for all t This fact will allow us to divide once–and–forever ϵ's into those that cause hysteresis in trade and the exchange rate (large ϵ's) and those that do not (small ϵ's).

Consider first the case when the shock is such that it induces no entry or exit. The optimal expected value of a firm upon realization of the unexpected once–and–forever shock is (call the period in which it occurs period 1):

$$(3.9) \qquad J_1 = \sum_{t=1}^{\infty} \delta^t \left[\pi\left[s[\epsilon^O,t],\hat{n}\right] - M \right].$$

For any given $\hat{n}$ we define the range of ϵ's $[\epsilon^{min},\epsilon^{max}]$ that are small as those that satisfy:

$$(3.10) \qquad 0 = \sum_{t=1}^{\infty} \delta^t \left[\pi\left[s[\epsilon^{max},t],\hat{n}\right] - M \right], \text{ and}$$
$$(F-M) = \sum_{t=1}^{\infty} \delta^t \left[\pi\left[s[\epsilon^{min},t],\hat{n}\right] - M \right].$$

This range is depicted in figure 5. For any given $\hat{n}$, J_1 is a decreasing function of ϵ, so the above condition holds for values of ϵ between $\epsilon[\hat{n}]^{min}$ and $\epsilon[\hat{n}]^{max}$. If the benchmark number of firms is higher, say $\hat{n}''$, then the range is shifted down to $\epsilon[\hat{n}'']^{min}$ and $\epsilon[\hat{n}'']^{max}$.

For any value of ϵ outside this range, firms will either enter or exit in period 1. As was shown in Baldwin (1988a), such large shocks can themselves be classified into two types: those for which n changes only once (in period 1) and those in which n jumps in period 1 and then returns to a steady–state value that is different from the original value. We only consider the simpler first case here. The first category of large shocks consists of all ϵ that meet two conditions. The first is:

$$(3.11) \qquad \sum_{t=1}^{\infty} \delta^t \left[\pi\left[s[\epsilon,t],\hat{n}^O\right] - M \right] > (F-M) \quad \text{for appreciation or,}$$

$$\sum_{t=1}^{\infty} \delta^{t}\left[\pi\left[s[\epsilon,t],\hat{n}^{0}\right] - M \right] < 0 \qquad \text{for depreciation.}$$

If either condition holds, firms will enter or exit until the new steady state $\hat{n}'$ is given by:

$$(3.12) \qquad \sum_{t=1}^{\infty} \delta^{t}\left[\pi\left[s[\epsilon,t],\hat{n}'\right] - M \right] = (F-M) \text{ or}$$
$$\sum_{t=1}^{\infty} \delta^{t}\left[\pi\left[s[\epsilon,t],\hat{n}'\right] - M \right] = 0$$

To insure that n actually stays at $\hat{n}'$, the case we are considering here, it must be true that exit and entry are never optimal in future periods. This gives us the second condition:

$$(3.13) \qquad (F-M) > \sum_{\tau=0}^{\infty} \delta^{\tau}\left[\pi\left[s[\epsilon,t+\tau],\hat{n}'\right] - M \right] > 0,$$

for all t = 2,...,∞. Since s is monotonically depreciating over time toward its new steady–state value, we can easily rule out future entry and exit: any firm that found it optimal to enter in a future period would do better by entering in period 1.

In summary, small shocks as defined by (3.10) lead to no change in $\hat{n}$ and therefore no change in K in the macro model. Large shocks, on the other hand, lead to a hysteretic change in n and K. Consequently, hysteresis in the long–run equilibrium exchange rate can be induced by a sufficiently large ϵ.

C. Effects of Large versus Small Policy Misalignments

In this section we make the argument that a monetary regime that entails only small policy misalignments would be qualitatively different than one that entails large policy misalignments. In particular we show that the latter would involve hysteresis and therefore the possibility of the abandonment of sunk assets.

In our macro model the foreign interest rate is constant, so that any deviation of the domestic rate from the foreign rate can be thought of as a monetary policy misalignment. Consequently we can rigorously define a convenient (although certainly not a perfectly general) metric for the degree of policy coordination: The larger is the support of the distribution of ϵ in a given regime, the less coordinated is policy in that regime. Perfect coordination corresponds to a degenerate distribution of ϵ; the probability that $\epsilon_t = 0$ is unity and the probability that $\epsilon_t \neq 0$ is zero. As the bounds of ϵ's distribution widen, greater policy misalignments are possible.

We first examine the case of fixed exchange rates, showing that if the policies are perfectly coordinated then hysteresis cannot occur. Next we show that even if ϵ follows a non–degenerate, bounded distribution, then we can rule out hysteresis as long as the bounds are not too wide.

The case of fixed exchange rates is in fact a trivial extension of our benchmark analysis. We therefore know that hysteresis cannot occur when policies are perfectly coordinated. The second step is somewhat more involved since with a stochastic ϵ the value of the firm takes account of the possibility of future entry and exit. However, having assumed that the shocks are bounded, we know that the realizations of s are also bounded. Indeed it is straightforward to show that if K does not change:

$$(3.14) \qquad s^{max} = \hat{K} + \frac{1}{1-\rho}\epsilon^{max}, \text{ and} \qquad s^{min} = \hat{K} + \frac{1}{1-\rho}\epsilon^{min},$$

where $\hat{K}$ reflects any steady–state market structure. Consequently, we can relate bounds on the value of the firm to ϵ^{min} and ϵ^{max}. Specifically, the bounds are:

$$(3.15) \qquad J^{max} = \frac{1}{1-\delta}(\pi[s^{max},\hat{n}] - M), \text{ and} \qquad J^{min} = \frac{1}{1-\delta}(\pi[s^{min},\hat{n}] - M).$$

Now if J^{min} is greater than zero, firms would never exit, and if J^{max} is less than (F–M) then firms would never enter. For distributions of ϵ for which this holds, n and K are constant. This range of ϵ's is defined by:

$$0 = \frac{1}{1-\delta}\left[\pi\left[\hat{K} + \frac{\epsilon^{max}}{1-\rho}, \hat{n}\right] - M\right], \text{ and } \frac{1}{1-\delta}\left[\pi\left[\hat{K} + \frac{\epsilon^{min}}{1-\rho}, \hat{n}\right] - M\right] = (F-M)$$

Combining these arguments, we know that for regimes with a sufficiently narrow range of policy misalignments, there will be no hysteresis, since:

$$0 = \frac{1}{1-\delta}\left[\pi\left[\hat{K} + \frac{\epsilon^{max}}{1-\rho}, \hat{n}\right] - M\right] < EJ_{\tau} < \frac{1}{1-\delta}\left[\pi\left[\hat{K} + \frac{\epsilon^{min}}{1-\rho}, \hat{n}\right] - M\right] = (F-M).$$

where EJ_{τ} is the actual expected value of the firm in any period τ.

IV. Summary and Concluding Remarks

This paper shows that hysteresis in trade and the exchange rate can occur if monetary policies get sufficiently misaligned. That is, uncoordinated policies may lead to the hysteretic entry or exit of firms in the traded goods sector. Since exit here entails the abandonment of sunk costs, the possibility of hysteresis means that a regime which permits large policy misalignments will involve the scrappage of sunk assets. Furthermore we showed that to avoid such hysteresis would not require a return to a fixed exchange rate regime. Rather, it could be avoided by an exchange rate regime in which policies were never allowed to get too far out of line. We conjecture that these positive results have several interesting implications for exchange rate policy and welfare analysis of managed versus floating rates.

The policy implication is that it may not be necessary to return to a formal, tightly managed exchange rate system in order to avoid much of the undesirable effects of floating rates observed in the 1980s. The industry–level model in this paper is not rich enough to allow us to determine theoretically how large the policy misalignments can be without inducing hysteresis. However, the empirical studies discussed in the introduction support the hypothesis that the exchange rate fluctuations of the 1970s did not induce hysteresis in trade, while the 1980s dollar cycle did. Thus loosely speaking it seems that a hysteresis–inducing shock must be of the magnitude of the 1980–1987 dollar overvaluation. Dixit (1987a), using the basic sunk cost model, makes a number of highly special assumptions about the nature of the industry–level competition and the exchange rate process which allow him to relate the width of the no–entry–no–exit band to readily observable values. Plugging in reasonable estimates for these values he finds that the band is quite wide. In his central case, the real exchange rate would have to move to a point where an importer's marginal revenue is 48 percent above total costs in order to induce entry; to force exit would require that the marginal revenue is 31 percent below variable costs. Thus Dixit's reasoning seems to lend further credence to the conjecture that policy coordination need not be extremely tight in order to avoid hysteresis.

On the welfare side, the exploratory model in this paper is not rich enough to study the welfare effects of sunk asset abandonment. However it may be taken as an indication that such hysteretic industrial dislocation may be an important issue in the debate between fixed and floating exchange rate regimes. To pin down the welfare costs of the scrapping of assets would require a model in which the amount of available capital in the economy were in some way limited. To keep such models tractable (for instance Solow's growth model with vintage capital), it is usually necessary to make assumptions (e.g.,perfectly flexible prices and continuously clearing markets) which eliminate the appealing dynamic properties of the sticky price monetary model of exchange rate determination. A reasonable approach to gauging the welfare effects of hysteresis would therefore be a study of what happens to real incomes in a two–sector growth model (with sunk costs) when the economy is subjected to large versus small exogenous terms of trade shocks. However, as pointed out in Baldwin (1986), import hysteresis has pro–competitive effects following large appreciations and anti–competitive effects following large depreciations. Any complete welfare analysis would need to include these effects.

From the perspective of the exchange rate determination literature, our model shows that the long–run equilibrium exchange rate is not unique. In fact we show that the steady–state exchange rate is path dependent. This conclusion suggests that the standard tests of the sticky price monetary model are misspecified since they typically impose a unique steady–state exchange rate level. It also adds another item to the already long list of reasons why PPP should not be expected to hold even in the very long run. Yet, this reason differs greatly from the others in that it is a nominal (as opposed to real) change which shifts the long–run exchange rate. To put this in other words, the possibility of PDPPP implies that in our model money need not be neutral even in the very long run.

FOOTNOTES

[1] Hysteresis is said to occur when an external shock alters a system in such a manner that the system is not expected to return to its original state in the absence of future corrective shocks.

[2] Of course, a number of factors are known to shift long–run real rates such as productivity growth differentials, shifts in comparative advantage, and intertemporal budget constraints. In our model the shift is due to an effect outside of this traditional set of explanations.

[3] All we really need is that ϵ has a bounded support. The uniform distribution, however, reduces the complexity of the formal analysis.

[4] Allowing for hysteresis in exports would amplify the effect of a large exchange rate swing. For instance, a large appreciation could force home firms to exit the foreign market as well as induce additional foreign firms to enter the home market. Both effects work in the same direction as far as shifting the exchange rate–trade balance link is concerned.

[5] Augmenting the specification of real money supply by using a price index defined over both domestic and foreign goods, with appropriate weights, does not involve any qualitative difference in the results and for that reason is not included in this streamlined model.

[6] This asymmetry creates a type of peso problem. Note that deviations from uncovered parity would be serially–correlated in this model, which is in accord with empirical results.

[7] Our simple i.i.d. mean zero money process is not meant to capture reality here. Richer, more realistic specifications, such as those in Lyons (1988), would involve only quantitative rather than

qualitative differences.

[8] In the Krugman (1988a) model, when the intervention policy is credible, the exchange rate cannot go outside of the band and investors know it. Our model differs in that the exchange rate is free to go outside the $s_t^I - s_t^O$ range. Nevertheless, since crossing the boundary exerts a regressive effect on the nominal exchange rate through K, e does not respond to innovations as much as it otherwise would when within the range.

[9] We make the standard assumption that γ and λ are such that the system is saddle–path stable. We also assume that γ and λ are such that the eigenvalue of absolute value less than one, ρ, lies in the interval (0,1) in order to ensure monotonic convergence. Empirical magnitudes for these parameters suggest that this assumption on ρ is very weak.

[10] It should be noted at this juncture that this hysteresis does not imply that the exchange rate deviates from fundamentals (as specified in our models). What changes are the fundamentals, market structure in particular.

REFERENCES

Baldwin, Richard. "Hysteresis in Import Prices: The Beachhead Effect." American Economic Review, September 1988a, 78.

________. "Some Empirical Evidence on Hysteresis in Aggregate U.S. Import Prices." NBER Working Paper No. 2483, January 1988b.

________. "On The General Existence and Nature of Sunk Cost Hysteresis in Trade." NBER Working Paper No. ??? (1988c).

________. "Hysteresis in Trade." M.I.T. mimeo, prepared for the 1986 NBER Summer Institute, April 1986.

Baldwin, Richard and Paul Krugman. "Persistent Trade Effects of Large Exchange Rate Shocks." NBER Working Paper No. 2017, September 1986.

Baldwin, Richard and Richard Lyons. "The Mutual Amplification Effect of Exchange Rate Volatility and Unresponsive Trade Prices." NBER Working Paper No. 2677, August 1988.

Bellman, Richard. Dynamic Programming. Princeton, New Jersey: Princeton University Press, 1957.

Bertsekas, Dimitri P. Dynamic Programming and Stochastic Control. New York: Academic Press, 1976.

Dixit, Avanash. "Entry and Exit Decisions of Firms Under Fluctuating Real Exchange Rates." Mimeo, Princeton University, September 1987a.

________. "Hysteresis, Import Penetration, and Exchange Rate Pass–Through." Mimeo, Princeton University, November 1987b.

Dornbusch, Rudiger. "Expectations and Exchange Rate Dynamics." Journal of Political Economy, December 1976, 84, 357–371.

________. Open Economy Macroeconomics. New York: Basic Books, Inc., 1980.

Feinberg, Robert M. "The Effects of Foreign Exchange Movements on U.S. Domestic Prices." ITC mimeo, December 1986.

________. "Recent Foreign Exchange Movements, Prices, and U.S. Trade Flows." I.T.C. mimeo, March 1987.

Frankel, J. and Richard Meese. "Are Exchange Rates Excessively Variable?" In Stan Fischer (ed.) NBER Macroeconomics Annual 1987.

Froot, Kenneth A. and Paul C. Klemperer. "Exchange Rate Pass–Through When Market Share Matters." Mimeo, Sloan School of Management, M.I.T., December 1987.

Froot, Kenneth A. and Takatoshi Ito. "On The Consistency of Short–Run and Long–Run

Exchange Rate Expectations." NBER Working Paper No. 2577, May 1988.

Hooper, Peter and Catherine Mann. "The U.S. External Deficit: Causes and Persistence." International Finance Discussion Paper no. 316, November 1987.

Kemp, Murray C., and Henry Y. Wan, Jr. "Hysteresis of Long–Run Equilibrium From Realistic Adjustment Costs," in George Horwich and Paul Samuelson (eds.) Trade, Stability, and Macroeconomics. New York: Academic Press, 1974.

Krugman, Paul R. "Target Zones and Exchange Rate Dynamics." NBER Working Paper No. 2481, January 1988a.

_________. "Deindustrialization, Reindustrialization, and the Real Exchange Rate." NBER Working Paper No. 2586, May 1988b.

Lewis, Karen K. "The Persistence of the 'Peso Problem' When Policy is Noisy." Journal of International Money and Finance. 7 (March 1988): 5–22.

Lucas, Robert E., Jr. "Interest Rates and Currency Prices in a Two–Country World." Journal of Monetary Economics. 10 (November 1982): 335–359.

Lyons, Richard K. "Whence Overshooting: Money Stock or Flow?" Mimeo, Columbia Graduate School of Business, October 1987.

Mann, Catherine. "Prices, Profit Margins, and Exchange Rates." Federal Reserve Bulletin, June 1986.

_________. "Prices, Profit Margins, and Exchange Rates: After the Fall." DIF Discussion Paper, Federal Reserve Board, July 1987a.

_________. "The Puzzling Behavior of Non–oil Import Prices." Mimeo, November 1987b.

Mussa, Michael L. "A Model of Exchange Rate Dynamics." Journal of Political Economy. 90 (February 1982): 74–104.

Orcutt, Guy. "Measurement of Price Elasticities in International Trade." Review of Economics and Statistics. 32 (May 1950): 117–132.

Stulz, Rene M. "Currency Preferences, Purchasing Power Risks, and the Determination of Exchange Rates in an Optimizing Model." Journal of Money, Credit and Banking. 16 (August 1984): 302–316.

Svensson, Lars E. O. "Currency Prices, Terms of Trade, and Interest Rates: A General Equilibrium Asset–pricing Cash–in–Advance Approach." Journal of International Economics. 18 (February 1985): 17–41.

FIGURE 1

The Entry Decision

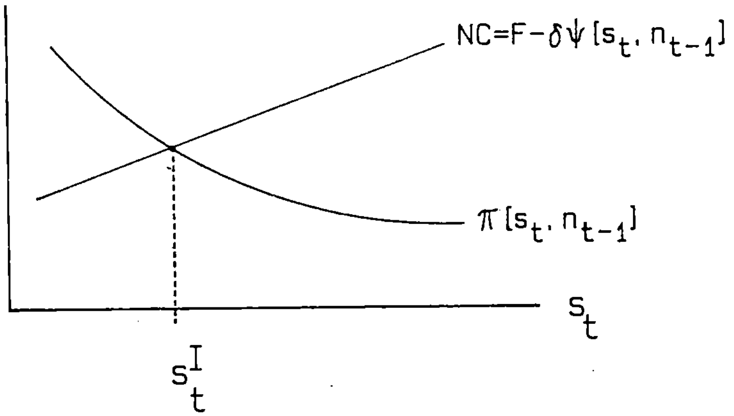

FIGURE 2

The Exit Decision

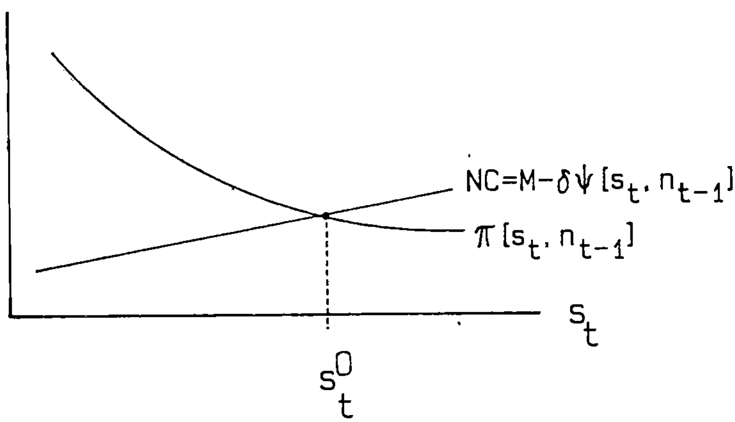

FIGURE 3

The Three Exchange Rate Regions

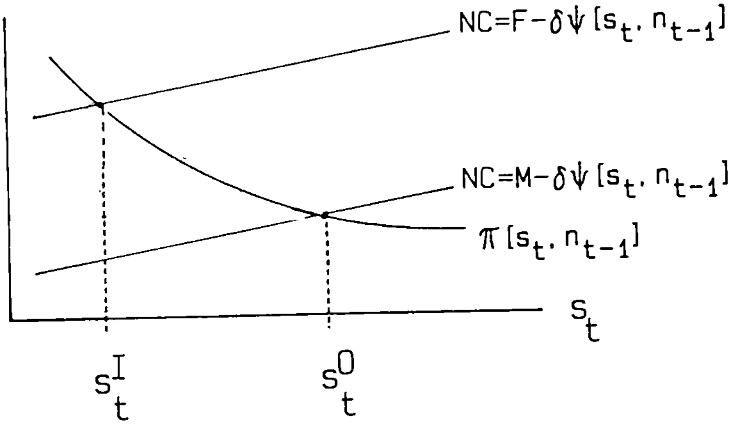

FIGURE 4

System Response to Different Shocks

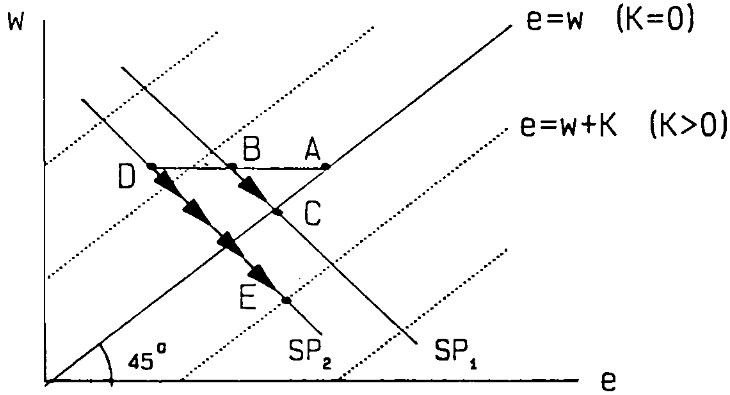

FIGURE 5

Large vs Small Shocks