NBER WORKING PAPER SERIES

PERSISTENT TRADE EFFECTS OF LARGE EXCHANGE RATE SHOCKS

Richard Baldwin

Paul R. Krugman

Working Paper No. 2017

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1986

The research reported here is part of the NBER's research program in International Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Working Paper #2017
September 1986

Persistent Trade Effects of Large Exchange Rate Shocks

ABSTRACT

This paper presents a theoretical basis for the argument that large exchange rate shocks - such as the rise of the dollar from 1980 to 1985 - may shift historical relationships between exchange rates and trade flows. We begin with partial models in which large exchange rate fluctuations lead to entry or exit decisions that are not reversed when the currency returns to its previous level. Then we develop a simple model of the feedback from "hysteresis" in trade to the exchange rate itself. Here we see that a large capital inflow, which leads to an initial appreciation, can result in a persistent reduction in the exchange rate consistent with trade balance.

Richard Baldwin
Graduate School of Business
611 Uris Hall
Columbia University
New York, NY 10027

Paul R. Krugman
Department of Economics
MIT
E52-383A
Cambridge, MA 02139

The extreme strength of the US dollar in the early 1980s has led to substantial losses of market position by US-based firms. Conventional economic wisdom, and conventional econometric estimates, suggest that a return of the dollar to more normal levels will, after some lag, reverse these losses. Many businessmen are, however, less sanguine. They argue that the extent of the dollar's overvaluation has been so large and so persistent that in many cases US firms have abandoned markets altogether, or foreign firms have entered markets that had previously been US preserves. A fall of the dollar to its 1980 level, they suggest, will not reverse these effects. Once foreign firms have invested in marketing, R&D, reputation, distribution networks, etc. they will find it profitable to remain in the market even at a lower exchange rate. Once US firms have abandoned markets, a mere return of the exchange rate to former levels will not be enough to make the expensive recapture of these markets worthwhile. Following the recent fashion in macroeconomics, we may describe this argument as the view that there is "hysteresis" in the response of trade to exchange rates (see in particular Blanchard and Summers 1986).

Now the conventional, non-hysteretic wisdom on the dynamics of trade response to exchange rates is based on surprisingly little systematic analysis. Distributed lags on the exchange rate are universal in empirical work, and such concepts as the J-curve are part of every international economist's vocabulary, yet little attempt has been made to provide microeconomic foundations for trade dynamics. So the pessimistic view about the effects of dollar decline cannot be dismissed out of hand. It is not difficult to imagine that in the presence of important sunk costs and increasing returns, a temporary shock to the exchange rate could have a long term impact on international trade. In particular, it seems intuitively plausible that very large shocks, such as the one we have just gone through, can have qualitatively different effects from smaller disturbances.

The purpose of this paper is to formalize the idea that large shocks to the exchange rate can have persistent effects on international trade. It builds on a previous paper by one of the authors (Baldwin 1986). In that paper a simple model was developed in which a foreign firm can enter a domestic market only by incurring a once-for-all sunk cost. The paper showed that in this case a temporary rise in the exchange rate, if sufficiently large, would induce permanent entry by the foreign firm. This entry would shift the subsequent relationship between imports and the exchange rate, so that even if the exchange rate returned to its previous level the trade pattern would not.

The present paper extends Baldwin (1986) in three ways. First, we replace the finite-horizon, perfect-foresight framework of the original paper with an indefinite-horizon, stochastic setup. In this setup we can impose stationarity on the exchange rate shocks, which has the advantage of helping to clarify the meaning of persistence in the behavior of trade.

Second, we examine the aggregated behavior of imports when there are many industries subject to potential foreign entry. This is important to verify that the results of the single-industry case do not get smoothed away in the aggregate. For example, one might wonder whether the distinction between large exchange rate shocks and small is meaningful when there are many sectors, since a strong dollar from one industry's point of view may seem weak from another's. What we can show is that the results do not get smoothed away: even in a multi-industry case, large exchange rate shocks have persistent effects in a way that small do not.

Finally, we examine the feedback from entry and exit decisions to the exchange rate itself. Clearly macroeconomic constraints imply that a temporary shock to the exchange rate cannot lead to a permanent trade surplus or deficit. The exchange rate must adjust so as to preserve intertemporal budget balance. But the nature of this adjustment needs examining. One view might be that a temporary overvaluation will automatically be followed by a corrective undervaluation that restores the intial market positions. We will show that this is not necessarily the case. In the particular model we consider, a temporary overvaluation is followed by a persistent reduction in the equilibrium exchange rate, one which is enough to restore trade balance but not enough to regain lost markets.

The models presented in this paper are, to say the least, highly simplified. Furthermore, the approach is resolutely partial equilibrium. Even where the feedback to the exchange rate is discussed, we use a quasi-partial, "elasticities" approach to the balance of payments. We commit these sins for the sake of tractability, and in the belief that the key insights would survive in a more satisfactory framework.

1. The single-industry model

Consider an industry in which there is a single foreign firm that is capable of supplying the domestic market. It would be possible to consider the case of oligopoly, but for simplicity we assume that if the foreign firm chooses to enter it will be a monopolist. The demand for its product in any period t will be represented in inverse form:

(1) $P_t = D(X_t)$

where X is deliveries to the market and P is the price in domestic currency.

The foreign firm is assumed to have a constant marginal cost in terms of its home currency, and to be concerned with profits measured in that currency. If the firm enters the US market, we can measure its operating profits as

(2) $Y_t = E_t P_t X_t - cX_t$

where Y is operating profits and E is the price of domestic currency in terms of foreign. If the foreign firm is in the market at all, it will choose X so as to maximize Y, so that we can represent the outcome of this conventional monopoly pricing problem as

(3) $Y_t = Y(E_t)$

Clearly Y will be an increasing function of E.

We next turn to the question of whether the firm will actually be in the market. What we will assume is that both getting into the market and staying there are costly. To get into the market, if the firm is not already there, requires an investment in marketing, reputation, distribution, and so on; we summarize all these costs as a single entry fee N. If the firm is already in the market, it will still have to spend something on all these areas to remain there; we summarize the cost of remaining in the market by a single maintenance cost M. We will assume that N>M: it costs more to enter a market than to stay there. The difference between N and M is the sunk cost aspect of the model, and as we will see is key to the results.

In any given period, the firm earns a net revenue that is equal to operating profits less entry or maintenance costs. We can define net revenue R by

(4) $R_t = 0$ if the firm chooses not to be in the market

$= Y_t - M_t$ if the firm was already in and stays there

$= \pi_t - N_t$ if the firm was out and gets in

Let us suppose that the firm is risk-neutral. Then its objective is to maximize the expected present value of net revenue. Assuming a constant discount rate δ, this objective is to maximize

$$(5) \quad W = E[\sum_t R_t \delta^t]$$

The firm's strategy depends on the behavior of the exchange rate E. Later in the paper we will introduce a model in which E's behavior and the strategies of firms are jointly determined. Initially, however, we will simply impose the assumption that E is a random variable, i.i.d. across periods. The value of E will be assumed to be revealed in each period before the firm makes its decision whether or not to be in the market.

The i.i.d. assumption is clearly problematic from an empirical standpoint. In reality exchange rate <u>levels</u> are clearly highly serially correlated over time, to such an extent that in monthly data they are not too far from a random walk. We therefore need to think of our periods as being quite long. We believe that the basic insights would go through as long as the exchange rate follows any stationary process, e.g. an autoregressive one; however, without the i.i.d. assumption the level of technical difficulty will rise sharply.

We are now prepared to consider the firm's decision problem. If N were no higher than M, this problem would be very simple: participate in the market if and only if the current Y(E) exceeds the entry cost. Because of the sunk cost aspect, however, entry now puts the firm in a favorable position in later periods, and this option value needs to be taken into account. On the other hand, the existence of a maintenance cost means that the entry decision is not irreversible: the firm must also take into account the possibility of leaving the market at some future date.

To take the various possibilities into account, we can treat this as a problem of dynamic programming. Consider first a firm that was in the market last period. It has two options. It can remain in the market: if it does so, it will have an expected present value

$$Y(E) - M + \delta V_I$$

where V_I is the expected present value of future revenues of a firm that was in the market the previous period, evaluated <u>before</u> we know the exchange rate. Given our i.i.d. assumption on the exchange rate, V_I will be a fixed number. Its value is however contingent on the strategy of the firm, and must be determined simultaneously with that strategy.

Alternatively, the firm can drop out of the market. In that case it realizes no current revenues. But it still has the option of entering later, so its expected present value is δV_0, where V_0 is the expected present value of a firm that was out of the market the previous period.

Similarly, a firm that was out last period will have the choice between entering and staying out. If it enters, its expected present value will be

$$Y(E) - N - \delta V_I$$

while if it remains out its expected present value will be δV_0.

The optimal strategy of the firm should now be apparent. If it was out of the market last period, it should enter if the current exchange rate exceeds a value we can label E_I. Otherwise it should stay out. If the firm was in the market last period, it should remain in unless the exchange rate falls below a value E_0. Otherwise the firm should drop out. The critical exchange rates E_I and E_0 are defined implicitly by the condition of indifference between staying and moving:

$$(6)\quad Y(E_I) - N - \delta V_I = \delta V_0$$

$$(7)\quad Y(E_0) - M - \delta V_I = \delta V_0$$

We still need to explain where V_I and V_0 come from. The answer is that they are the true ex ante expected present values of a firm following the strategy we have just described, when the firm starts out in or out respectively. Let f(E) be the density function of E. Then

$$(8)\quad V_I = \int_{E_0} Y(E)f(E)dE - \left[\int_{E_0} f(E)dE\right][\delta V_I - N] + \left[\int^{E_0} f(E)dE\right]\delta V_0$$

$$(9)\quad V_0 = \int_{E_I} Y(E)f(E)dE - \left[\int_{E_I} f(E)dE\right][\delta V_I - M] + \left[\int^{E_I} f(E)dE\right]\delta V_0$$

To characterize the firm's behavior, it is necessary to solve the recursive equations (8) and (9) simultaneously with (6) and (7) to derive the four variables E_I, E_0, V_I, and V_0. We have not come up with any specific examples that prove particularly enlightening. The main result, however, does not depend on deriving a closed-form solution. From (6) and (7), we see that

$$Y(E_I) - Y(E_0) = N - M > 0$$

Since Y(E) is increasing in E, this implies that $E_I > E_0$. <u>The exchange rate that induces entry is higher than the exchange rate that induces exit</u>.

The implications of this result can be illustrated using Figure 1. In the figure, we show import volume in this industry as a function of the exchange rate. The import schedule has two parts. If the firm is out of the market, imports are zero. This is represented by the horizontal line along the axis, OO. If the firm is in the market, imports will be an increasing function of the exchange rate. This is shown by the schedule II. Now the point is that there is a range of exchange rates, from E_0 to E_I, where either schedule could apply. If the firm is not in the market, then as long as the exchange rate does not go above E_I it will not enter. If it is in the market, then as long as the exchange rate does not go below E_0 it will stay in. Thus the level of imports will depend on history as well as the current exchange rate.

Suppose in particular that the distribution of E is such that it usually falls between E_0 and E_I, and only rarely lands outside that range. Then for long stretches imports will either be zero or fluctuate along II. Occasionally there will be a large exchange rate shock, leading the firm either to enter or to leave; this will shift the industry to the other segment of the schedule, leading to what will appear to be a structural change in the exchange rate-import relationship.

Notice that the persistence we find here is not captured by the simple notion of a lag in the effect of the exchange rate. To see this, imagine that the exchange rate were to rise from a level between E_0 and E_I to a level high enough to induce entry by the firm, then return to its original level and stay there forever. (With random E, we will of course not expect this to happen; thus this is purely a thought experiment). If persistence were simply a matter of a lag, the level of imports would eventually return to its original level as well. In this model, however, the level of imports will remain permanently higher.

We have now seen how large exchange rate shocks can have persistent effects in a single industry. The next question we address is whether the attribution of persistent effects to large shocks continues to make sense when we are concerned with the aggregate behavior of a large group of industries with different characteristics.

2. Many industries

When we discuss a single industry, it is natural to think of two discrete states: the foreign firm is either in or out. When we are concerned with aggregate imports, however, we must ask whether the conclusions are going to be softened by the presence of many industries with different characteristics. One might suppose that there will always be some industry with a foreign firm just on the margin of entry, and another industry with a foreign firm just on the margin of exit. If this were the case, any exchange rate shock would induce some entry or exit. This would vitiate our point that large shocks will produce persistent effects that small shocks will not. So it is important to ask whether the aggregation of many industries will smooth out the results we derived for the single industry case. What we will do is show that for one interesting special case the aggregate behavior will be similar to the behavior we have analyzed for a single industry. We will then argue that in the more general case the result will be only somewhat softened by aggregation.

As a starting point we need to specify the space across which industries are to be distributed. In the model of Section 1 we found that the dynamic behavior of an industry could be summarized by two values: E_I, the minimum exchange rate that will induce foreign entry, and E_0, the maximum exchange rate that will lead to foreign withdrawal. These values in turn depend in a complex way on underlying parameters, but for current purposes all that we need to do is classify industries by the result. Let us, then, represent each industry as a point in E_0, E_I space, where the coordinates indicate the critical values for that sector.

Now E_0 and E_I may be said loosely to vary across industries for two reasons. On one side is what we may think of as comparative advantage. For industries where foreign firms have low costs relative to domestic competitors, we might expect both E_0 and E_I to be low; for industries where foreign firms have high costs, both would tend to be high. To the extent that it is differences in comparative advantage that mostly dominate the spread among industries, we would expect to find that E_I and E_0 were positively correlated. On the other side, industries may vary in the degree to which their entry costs are sunk -- the effect captured by the difference N - M in our single industry example. If industries had similar comparative advantage positions but were very different in sunkness, we would expect to find E_0 and E_I negatively correlated: easy-in, easy-out "contestable" industries would have both high E_0 and low E_I.

What we will do is concentrate first on the case where comparative advantage is dominant. Specifically, we will assume that if one industry has a higher E_0 than another, it also has a higher E_I. This implies that it is possible to assign an index z to industries such that E_0 and E_I are both increasing functions of z. For simplicity, let us further assume that these functions are continuous.

The resulting situation is illustrated in Figure 2. The axes of the figure represent the E_0 and E_I specific to each industry. Given our special assumption, the distribution of industry characteristics lies along an upward-sloping line like ZZ. Each point on that line corresponds to a particular industry. Note that ZZ lies everywhere to the left of the 45 degree line. This reflects the fact that $E_I > E_0$ always, which in turn as we saw reflects the sunk cost assumption that $N > M$.

In each of these industries the foreign firm may be in or out, depending on the past history of the industry. There is, however, an equilibrium configuration which, once established, will be maintained over time. It is the following: in all industries with z less than some value $\bar{z}$ the foreign firm will be in, while in all industries with z greater than $\bar{z}$ the foreign firm will be out. Once this configuration is established, the effect of exchange rate shocks can be wholly summarized by shifts in $\bar{z}$. That is, the state of the whole import sector will be summarized by a single number representing the range of goods in which foreign firms are present.

To see that this configuration will in fact be self-replicating, we examine the geometry of Figure 2. Suppose that initially the borderline industry is represented by the point A; in all those industries corresponding to points on ZZ to the southwest of A the foreign firm is in the market, while in all the industries to the northeast of A the foreign firm is not in. Now let there be an actual realization of the exchange rate, which we can represent as a point on the 45 degree line.

Clearly there are three possibilities. First, suppose the exchange rate falls between B and C. Then E is less than E_I for industry A, and hence also less for all the industries to the northeast of A. So no foreign firms enter. At the same time, E exceeds E_0 for A, and hence for all industries to the southwest of A. So no foreign firms exit either.

Second, suppose that the realized E is above C, say at C'. Then additional foreign firms will enter, up to the industry corresponding to A'. No firms will exit. So the form of the configuration will remain the same, but the location of the marginal industry will have shifted.

Finally, suppose that the exchange rate is realized at a level below B. The case will be symmetrical to the second case: the form of the configuration will be retained, but the margin will shrink in.

What is important to note is that even though there is a continuous distribution of industries in terms of the exchange rates that will make foreign firms enter or exit, it is not the case that there is always a firm on the margin of entry or the margin of exit. On the contrary, there will always be a range of exchange rates for which no firm either enters or exits. Thus when the marginal firm is A, the exchange rate range BC leads to neither entry nor exit. Geometrically, this range of no change exists because ZZ lies to the left of the 45 degree line. This in turn, as we have already noted, reflects the presence of sunk costs.

We have now seen that the dynamic behavior of a group of industries can, under these special assumptions, be reduced to changes in the location of the marginal industry. The next question is whether the behavior of the aggregate of a group of industries will still show the same kind of persistence as each individual industry. As our measure of the aggregate, we will focus on the total value of imports from all industries.

It is clear that the total value of imports, which we will denote as T, will depend not only on the exchange rate but on the range of goods in which foreign firms are present. Thus

(10) $T = T(E,\bar{z})$

Depending on the demand functions, $\partial T/\partial E$ may be either positive or negative; for our diagrams we will draw it as positive. Clearly, however, T will be an increasing function of $\bar{z}$.

Now consider a case parallel to that which we used to examine persistence for a single industry. Suppose that in Figure 2 the exchange rate usually lies in the range B'C. Then if the marginal industry is A, the economy may go on for a time without any change in the range of industries in which foreign firms participate. As long as this is the case, there will appear to be a stable relationship between the value of imports and the exchange rate. This relationship is illustrated as AA in Figure 3.

But now suppose that one period's exchange rate happens to lie outside B'C, say at C'. We already know what happens: the marginal industry shifts to A', so that $\bar{z}$ rises. If the exchange rate then returns to B'C, this shift will not be reversed. The effect will be to shift the import-exchange rate relationship up, to A'A' in Figure 3. The economy may then fluctuate for a time along this new schedule. It will appear as if the economy has experienced a structural change.

This scenario is clearly very similar to what we saw in Figure 1. The main difference is that AA and A'A' are not the only two possible schedules in E,T space. There is in fact a continuum of schedules, each corresponding to a different $\bar{z}$. Some of these schedules will be sustainable only by exchange rates that occur only rarely, and the economy will not usually remain on these schedules more than one period. If sunk costs are large enough, however, there will be a wide range of $\bar{z}$, and correponding import schedules, that will tend to persist as long as the exchange rate falls in its normal range. The typical behavior of the economy will be to stay on any one schedule for a while, until a large exchange rate shock pushes it "off the edge" onto a different schedule. Thus the exchange rate-import relationship will seem to alternate periods of stability with abrupt structural changes following large shocks.

To derive this result we have relied on an assumption about the distribution of industries, namely, that E_0 and E_I move strictly together. To conclude this section, we ask whether aggregate imports would still show this kind of behavior with a more general distribution of industry characteristics.

Let us try to answer this by sneaking up on the issue. First, let's suppose that there are two groups of industries, 1 and 2, each of which can be represnted by an upward-sloping schedule in E_0, E_I space. In Figure 4 we show these groups of industries as Z_1Z_1 and Z_2Z_2. The fact that industry group 2's schedule lies further from the 45 degree line than industry group 1's indicates that sunk costs are more important for 2.

What can we say about the behavior of these two groups? Both will exhibit some persistence of effects of large exchange rate changes. This effect will be more pronounced for group 2, where larger shocks will be needed to shift the marginal industry. The difference will complicate the description of dynamics, since we need to keep track of two margins instead of one. Clearly, however, we will still be justified in saying that in the aggregate the whole import side exhibits persistent effects of large shocks.

If we can have two groups of industries, however, we can have more. Thus we can approximate a two-dimensional distribution of industries in E_0, E_I space as finely as we like by a series of parallel lines. The more lines, the more complex the detail of the dynamics, but the basic point will not change.

We have seen, then, that the peculiar behavior we have derived for a single industry will not in general get averaged away in the aggregate. When we begin to consider large aggregates, however, our assumption of an exogenous distribution for the exchange rate begins to become suspect. We would expect there to be some feedback from the entry and exit decisions of firms to the exchange rate itself -- if only because a country must sooner or later pay its way in international trade. The final step, then, must be to try to model this feedback.

3. Feedback to the exchange rate

To allow for feedback to the exchange rate it will be useful to simplify the assumed structure of the economy in a somewhat different way from that in the last section. We now assume that there are two kinds of sectors: normal sectors, where none of these dynamic issues apply, and hysteretic sectors, where they do. For the normal sectors there will be a static relationship between trade flows and the exchange rate. We will simply summarize everything that goes on in these sectors by a reduced form relationship between the exchange rate and the balance of trade,

(11) $B_t = B(E_t)$

We will assume that there is a large group of hysteretic import sectors that are all perfectly symmetrical, i.e., that all have the same entry and maintenance costs, face the same demand, and have the same marginal cost. (There is no reason why could not also have export sectors where domestic firms enter or leave; we concentrate on imports as an arbitrary choice among alternative simplifications). This allows us to concentrate on the analysis of a representative sector. It also lets us summarize the effects of past history by a single number f, the fraction of these industries in which the foreign firm is in the market. The total value of imports from these sectors will depend on both the exchange rate and f:

$$(12)\quad T_t = T(E_t, f_t)$$

In order to assess the feedback from trade to the exchange rate, we need a model of exchange rate determination. The approach we will take has two virtues: it is simple, and it forces the economy to balance its trade in the long run. The vices of the approach will be immediately apparent and no doubt infuriating to those who worry seriously about exchange rate determination. We look forward to the day when our points can be made in a less ad hoc way.

Our basic approach is a partial equilibrium, elasticities one, modified to take account of capital flows. Interest payments on past debt accumulations are ignored, so that the condition of balance of payments equilibrium is

$$(13)\quad B_t - T_t - K_t = 0$$

where K is net capital inflow.

Now comes the awful part: we will assume that K is an i.i.d. random variable. That is, random shocks to the capital account will be the forcing variable that generates exchange rate movement and the entry and exit of firms. The only justification for this approach is its usefulness, which will soon become apparent.

The exchange rate must move to balance payments. We can express the exchange rate as a function of net capital inflow and the fraction of foreign firms that have entered import markets:

$$(14)\quad E_t = E(K_t, f_t)$$

with the exchange rate increasing in both arguments.

Now let us consider the problem of a representative foreign firm. We assume the following timing: at the beginning of a period the size of net capital inflows K is revealed. Then all firms decide simultaneously whether to enter or exit. Finally, the exchange rate and everything else gets determined.

Suppose that a foreign firm is not in the market, and considers entry. In the current period it will earn Y(E) - N, where E depends on both K and the number of other firms that enter. Having entered, it will also have an expected present value next period. In contrast to the case we considered in Section 1, however, the expected present value V_I will not in general be constant over time, even though K is assumed i.i.d.. The reason is that the number of firms feeds back to the exchange rate, and thus the number of firms in the market by the end of this period affects the expected value of being in the market in the future. We must, then, think of a function $V_I(f)$. Correspondingly, a firm that is not in the market has an expected value $V_0(f)$.

Take the functions $V_I(f)$ and $V_0(f)$ as given for a moment. Then the decision problem of a firm may be written as follows. A firm that is out will enter if

$$Y(E(K,f)) - N + \delta V_I(f) > \delta V_0(f)$$

A firm that is in will exit if

$$Y(E(K,f)) - M + \delta V_I(f) < \delta V_0(f)$$

Now suppose that at the end of the previous period the number of firms was f_{t-1}. Clearly there is a range of current values of K that will lead no firms either to enter or exit. If K lies above this range, firms will enter until there is no incentive for more to enter; if K lies below this range, firms will exit until they are indifferent between staying or leaving. Let $K_I(f)$ and $K_0(f)$ be the critical values for entering and leaving the market. Notice that these are <u>functions</u>, not fixed values. These functions may be defined implicitly by the entry and exit conditions,

$$(15)\quad Y(E(K_I(f),f)) - N - \delta V_I(f) = \delta V_O(f)$$

$$(16)\quad Y(E(K_O(f),f)) - M - \delta V_I(f) = \delta V_O(f)$$

The law of motion for f can then be described as follows:

$$(17)\quad f_t = f_{t-1} \text{ if } K_O(f_{t-1}) < K_t < K_I(f_{t-1})$$

$$K_I(f_t) = K_t \text{ if } K_t > K_I(f_{t-1})$$

$$K_O(f_t) = K_t \text{ if } K_t < K_O(f_{t-1})$$

We will turn to the interpretation of (17) in a moment. First, however, we need to ask where the expected value functions $V_I(f)$ and $V_O(f)$ come from. The answer, of course, is that they must be determined simultaneously with the entry and exit functions $K_I(f)$ and $K_O(f)$. We can imagine the following computational procedure. Start with a guess at the functions $V_I(f)$ and $V_O(f)$. This will enable us to solve (15) and (16), and thus to compute the behavior described in (17). Once we have done this, however, we can ask what the expected present value of a firm would actually be for any given f. We can then use this computed function as a new guess at the expected present value functions, and repeat. An equilibrium will be a fixed point of this computational process. This is an equilibrium defined as a fixed point in <u>function</u> space rather than in price space -- a concept that has been widely used in the work of Robert Lucas (see Lucas (1981)). We will not attempt here to prove existence, let alone uniqueness. Instead, we simply suppose that the functions $V_I(f)$ and $V_0(f)$ can be determined, and that we therefore do in fact end up with the dynamics described by (17).

Now we turn to trying to make sense of these dynamics. Here we can use essentially the same trick we used in Section 2, in a somewhat different space. A particular value of f_{t-1} has associated with it values $K_I(f_{t-1})$ and $K_0(f_{t-1})$, the exchange rates that will induce entry and exit respectively. It seems obvious that both will be increasing in f_{t-1}: the more firms in the market, the lower the exchange rate for any given K and thus the higher the K needed to induce entry or deter exit. Thus we can represent the range of possible values of f_{t-1} as an upward-sloping line in K_0, K_I space. This is illustrated as FF in Figure 5. Each point on FF has a value of f_{t-1} associated with it. Also, FF must lie to the left of the 45 degree line, because $K_I > K_0$: a larger capital inflow is needed to induce entry than to deter exit. Finally, parallel to our technique in Section 2, we can represent a realization of the actual capital inflow as a point on the 45 degree line.

From this point the analysis is exactly parallel. Suppose that we start with a number of foreign firms in the market corresponding to point A in Figure 5. Then a capital inflow in the range BC will not lead either to entry or exit. A large capital inflow, however, such as C', will lead additional foreign firms to enter. By examining (17), we see that foreign firms will enter to exactly the extent that no more would enter if the capital inflow remained at exactly the same level; in other words, f will shift up to the level that corresponds to A'. If the capital inflow then falls to the range B'C, the new, increased number of foreign firms in the market will persist. We can thus imagine that for stretches of time there will be a stable number of foreign firms in the market, but that at intervals that number will be shifted by large capital inflows or outflows.

The interesting question, however, is what this says about the behavior of the exchange rate. As long as the number of firms in the market remains fixed, the exchange rate will have a static relationship with the capital inflow, fluctuating around some mean. A large inflow (outflow) will at first produce a large appreciation (depreciation). If the capital flow then returns to its initial value, however, the exchange rate will not. Instead, following a large capital inflow that provokes entry by foreign firms the exchange rate will tend to fall <u>below</u> its original level. We can see this by noting that f will rise, and this will tend to worsen the trade balance for any given exchange rate.

The implied behavior of the exchange rate is illustrated in Figure 6. For stretches of time the exchange rate will appear to fluctuate around a constant mean. Then a large capital inflow or outflow will produce a temporary movement of the exchange rate in one direction, followed by a shift of the mean in the opposite direction. It will seem as if the initial exchange rate shock has produced a structural change that permanently lowers or raises the normal level of the exchange rate -- until the next large shock.

The implications of this analysis, if it is at all relevant to the current US situation, are obvious. Massive capital inflows pushed the dollar to very high levels in the early 1980s. If this was indeed, as businessmen believe, a "large" shock that leads to entry by foreign firms (and by extension, exit by US firms), the dollar in subsequent years can be expected to fluctuate around a level that is persistently lower than that of the 1970s.

4. Concluding remarks

In this paper, as in Baldwin (1986), we have presented simple models (although with fancy equilibrium concepts) designed to clarify our thinking rather than to be realistic. We can see from the models that there is a reasonable case to be made for persistent trade effects of large exchange rate shocks in the presence of sunk fixed costs, and that this kind of trade persistence is not simply a kind of lag in the response to the exchange rate. We hope that even this step will help encourage economists to take the notion of hysteresis in international trade seriously. Clearly, however, the next step must be to make the analysis more operational. In our view, this will involve three main tasks.

The first task is to make the models more reasonable at a micro level. We would like to dispense with the i.i.d. assumption about the distribution of shocks, and in general sand away the rough edges of the models to the point where they look as though they might apply to actual industries. As recent work in the application of industrial organization models to trade has demonstrated (Dixit 1986; Baldwin and Krugman 1986; Venables and Smith 1986), this will not be easy. The truth is that the existing models of indutrial organization can be made to look like real industries only by a Procrustean effort of model modification. This is true even of static certainty models; it will be even more true of the dynamic uncertainty models that are vital here.

The second task is to get the macroeconomic linkages better specified. We are of course concerned about our ad hoc exchange rate approach, although we speculate that the results will not be too much changed by a better model here. Perhaps more important is the modelling of investment decisions and the cost of capital. In the models of this paper, entry and exit decisions are in effect invisible investment decisions. Properly speaking, a decision to enter a market is a kind of investment; a decision to abandon a market amounts to capital consumption. If it is wrong to ignore feedback from trade to the exchange rate, it is probably also wrong to ignore feedback to the cost of capital.

Finally, we need some idea of how important these effects really are in practice. Here the problem is one of both technique and data.

The dynamic effects we model are not captured by the usual econometric assumptions that behavior can be represented by continuous functions and a fixed structure of leads and lags. Thus unconventional statistical techniques. and perhaps a reliance on case-study-like evidence, may be necessary. Furthermore, it is at least suggested by our models that very large data sets may be required. Suppose that really big exchange rate shocks occur only once a generation. Then long time series may show a stable relationship between exchange rates and trade -- yet that relationship may change abruptly when a large shock does hit. When it seems plausible that this generation's one big shock has just happened, this is a disquieting thought.

In sum, then, the analysis presented here needs a great deal of extension. We hope, however, that we have shown that the businessmen's pessimism about the effects of a decline in the dollar is not necessarily bad economics, and that we had better not be complacent about the stability of econometric estimates of trade behavior.

References

Baldwin, R. (1986): "Hysteresis in trade", mimeo, MIT.

Baldwin, R. and Krugman, P. (1986): "Market access and competition: a simulation study of 16K random access memories", NBER working paper #

Blanchard, O.J. and Summers, L. (1986): "Hysteresis and the European unemployment problem," NBER Working Paper # 1950.

Dixit, A. (1986): "Optimal trade and industrial policy for the US automobile industry," mimeo, Princeton.

Lucas, R. (1981): Studies in Business Cycle Theory, MIT, Cambridge.

Venables, A. and Smith, M.A.M. (1986): "Trade and industrial policy under imperfect competition," Economic Policy, 3.

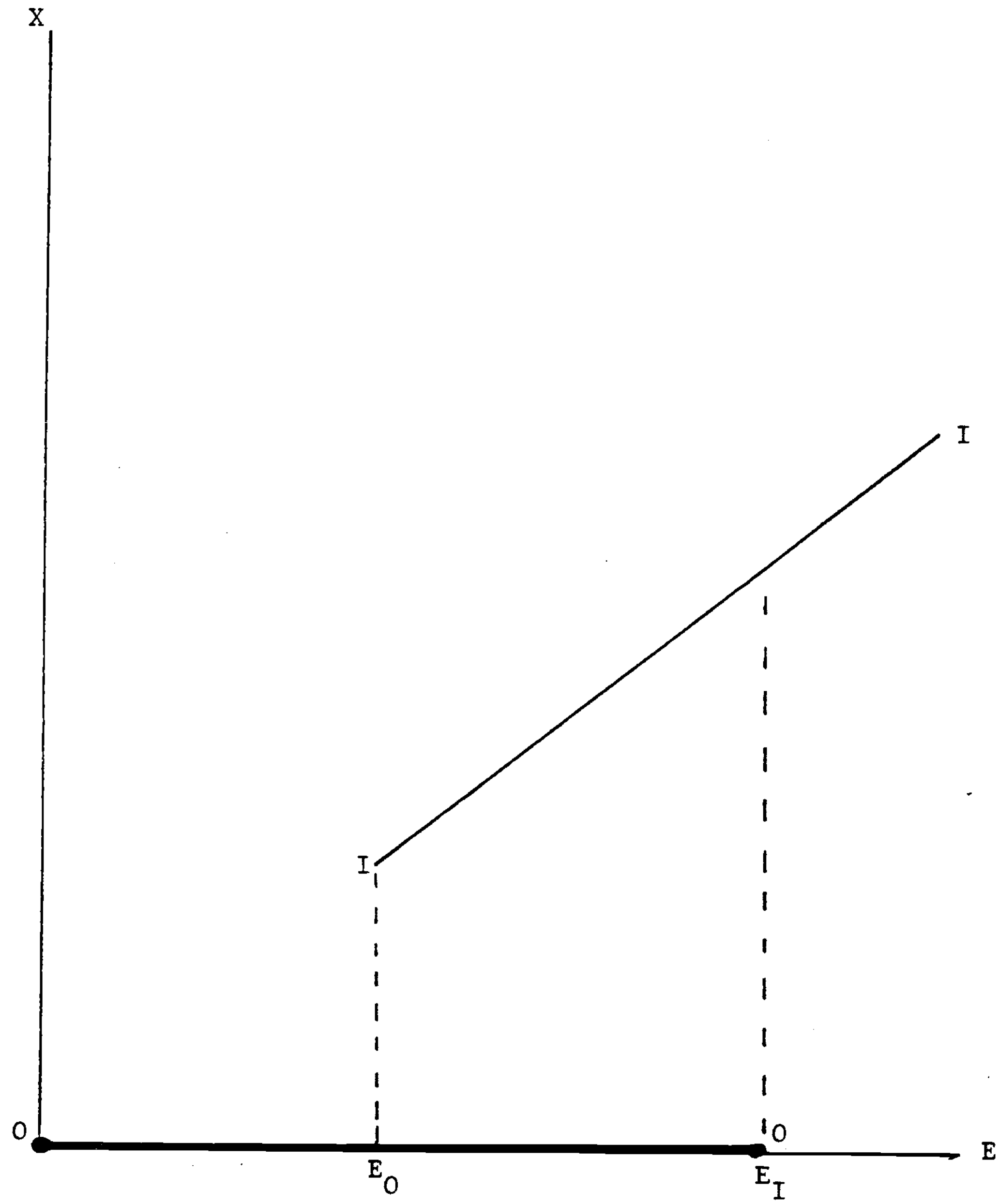

Figure 1

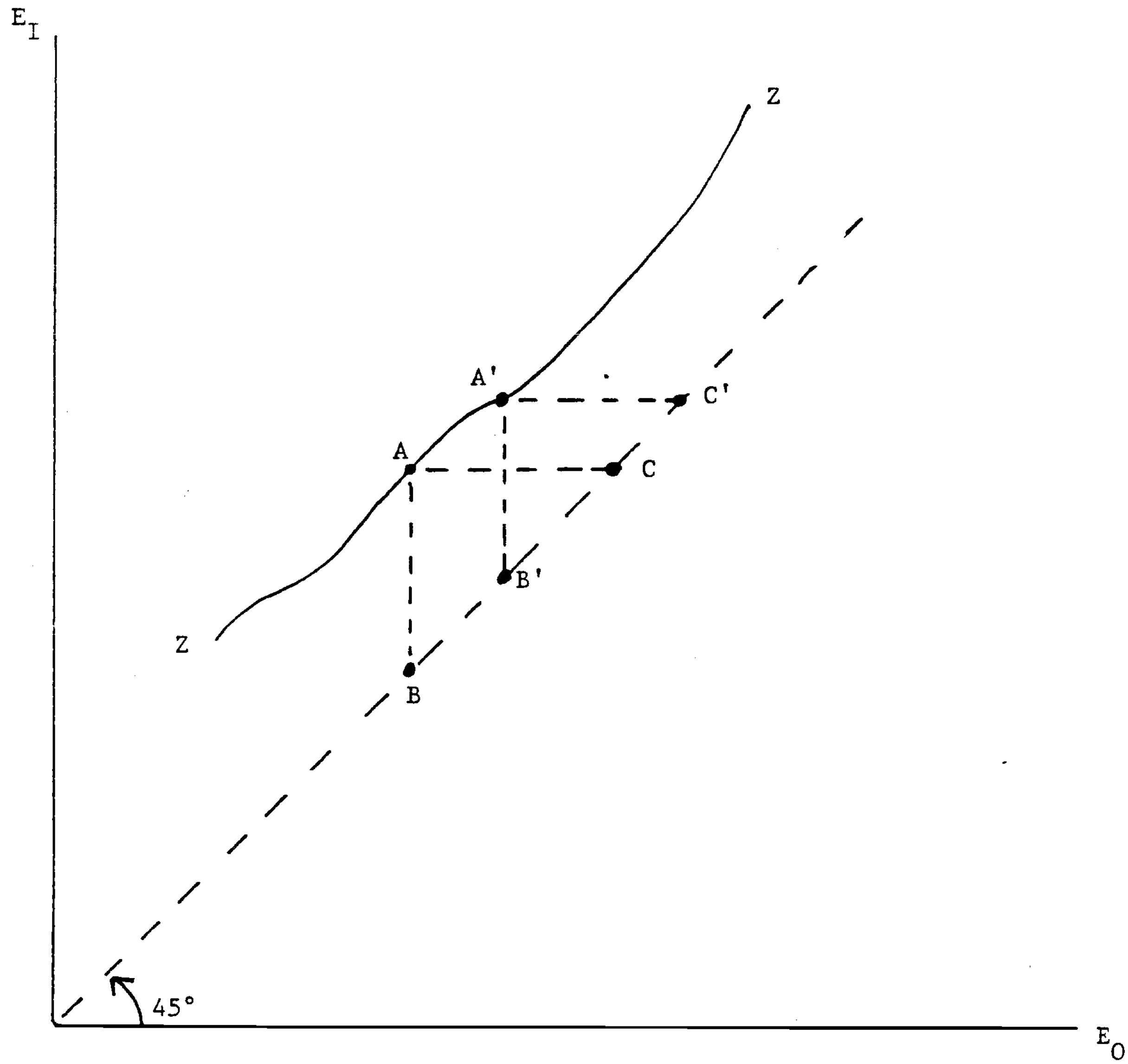

Figure 2

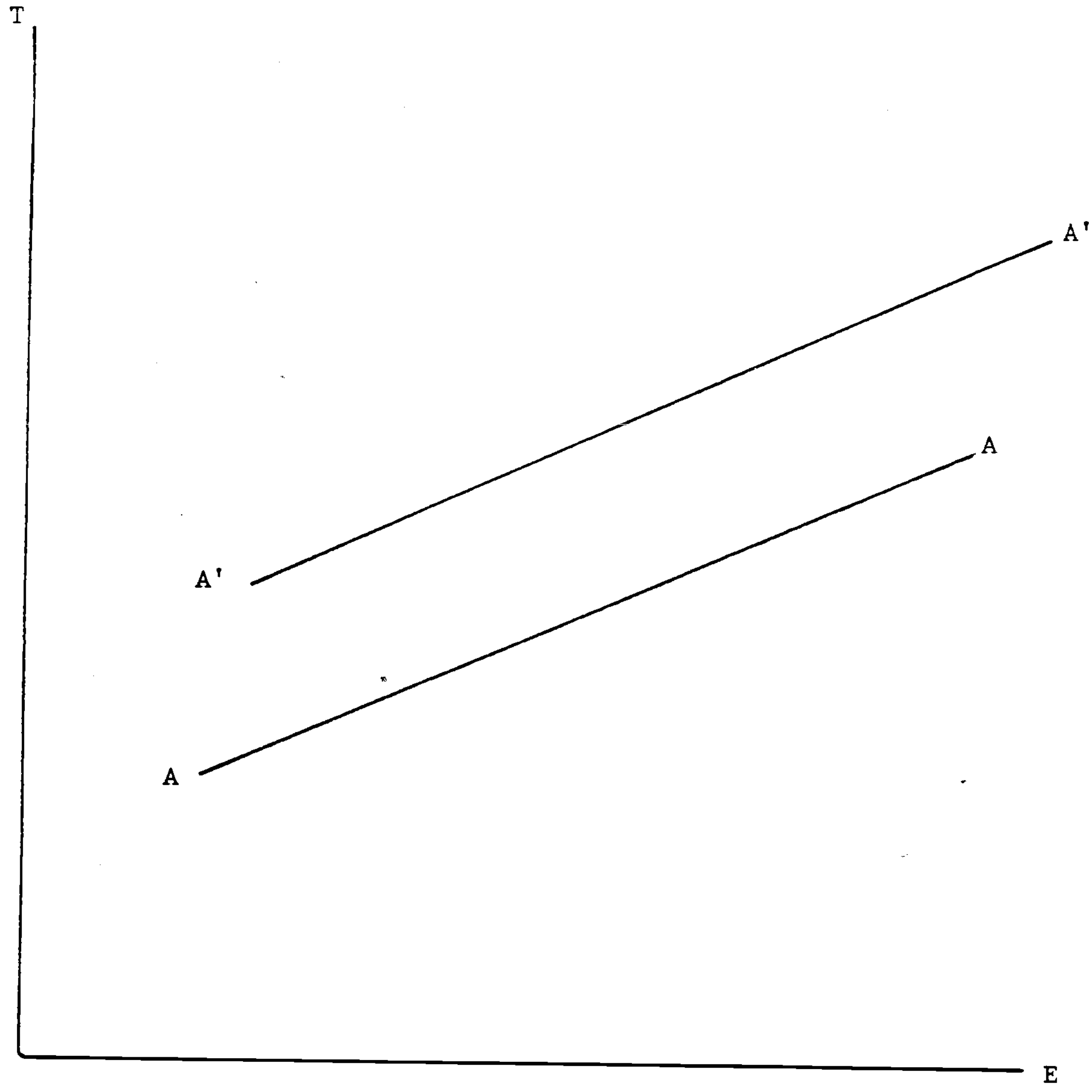

Figure 3

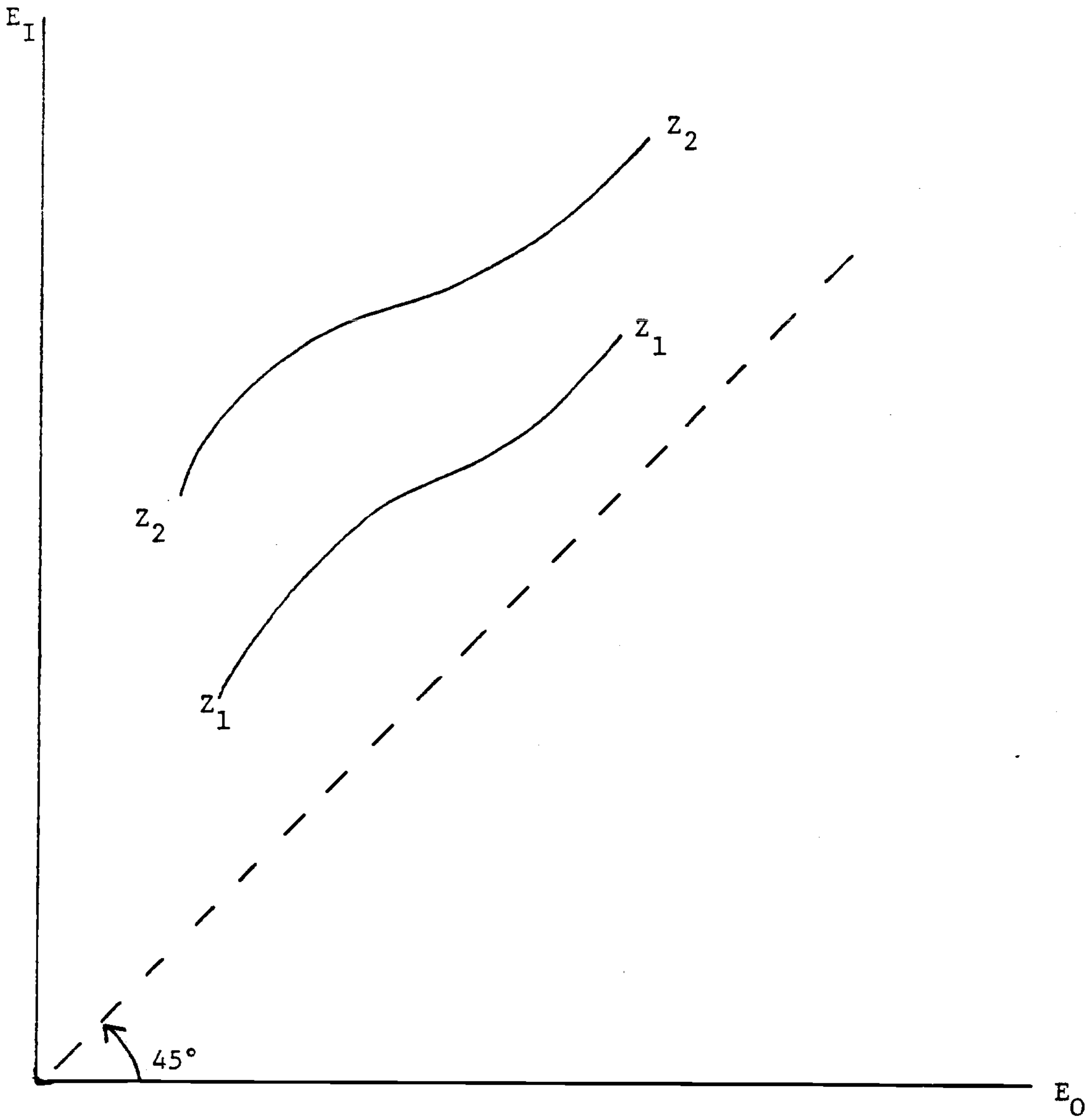

Figure 4

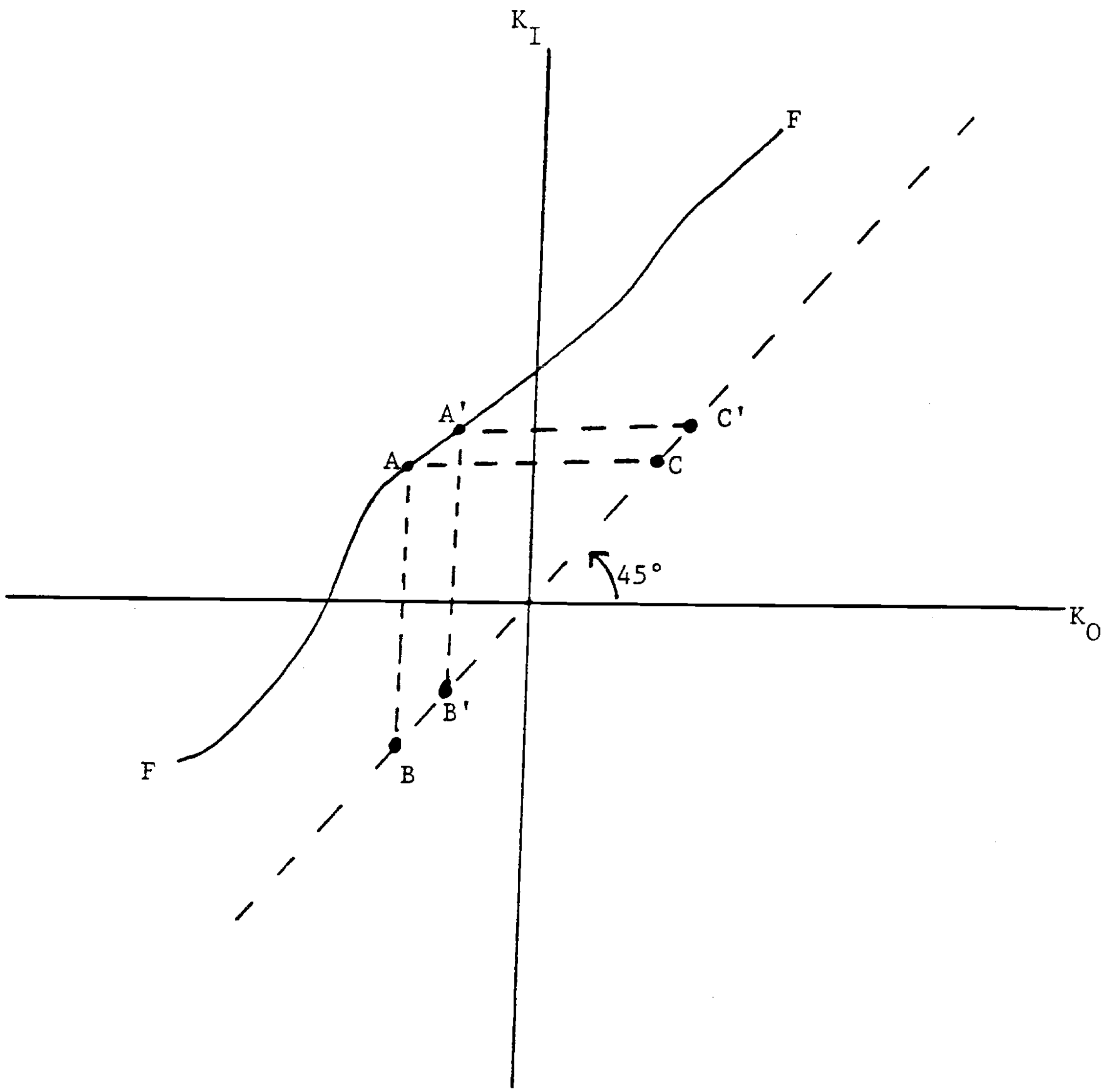

Figure 5

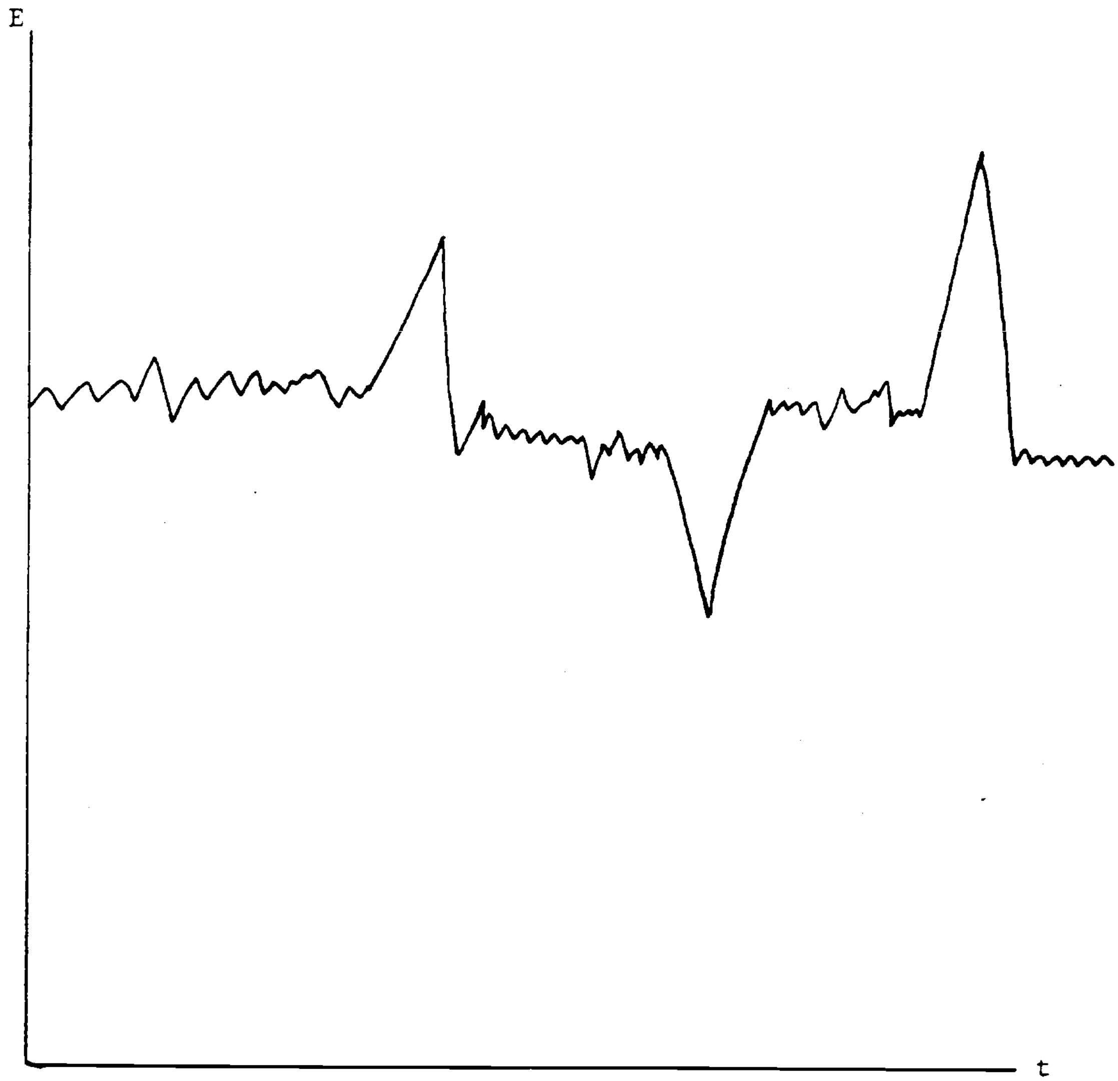

Figure 6