Are the traditional trade-exchange rate theorems relevant for developing countries facing entry costs in international markets?

Trond-Arne Borgersen

Høgskolen i Østfold
Arbeidsrapport 2008:5
Are the Traditional Trade-Exchange Rate Theorems Relevant for Developing Countries Facing Entry Costs in International Markets?

Trond-Arne Borgersen
Østfold University College
Faculty of Business, Social Sciences and Languages
Remmen
1757 Halden
Norway

November 2008
Abstract

This paper investigates the relevance of traditional trade-exchange rate theorems for developing countries facing sunk entry costs in international markets. First the theorems analysing pricing of tradable goods and the trade balance dynamics following exchange rate shocks are accounted for. Second the sunk cost hysteresis model of foreign trade is described, including the possibility for hysteresis both at the microeconomic and at the macroeconomic level. Third the implications of sunk cost hysteresis for the predictions of the traditional trade-exchange rate theorems are discussed, focusing on both pricing of tradable goods as well as short and long run trade balance dynamics following exchange rate shocks. The paper argues that the sunk cost model provides a microeconomic basis for trade dynamics that allows for non-linearities and regime switches, something often seen in empirical analysis. The predictions of structural adjustment programs are however drawn from the traditional theorems, lacking the possibility for non-linearity. The sunk cost model is argued to push both pricing rules and trade balance dynamics closer towards the empirical record, mainly by allowing for a state-dependent relationship between exchange rates and foreign trade. When it comes to policy implications the paper argues in favour of context specific policy interventions and against the one size fits all approach of structural adjustment programs.
1 Introduction

This paper looks more closely at the relationship between international trade flows and exchange rates. Trading goods internationally means that goods are sold in different national markets, and as the exchange rate measures the terms at which relative prices are compared internationally, it should also be a key factor when analyzing international trade. However, research on international trade is only to a minor extent concerned with exchange rates. In fact, despite having been under severe criticism since the mid 80s, a few well established theorems still dominate the reasoning.

The aim of this paper is to look more closely at these theorems, with a particular emphasis on evaluating their predictions in relation to how developing countries can expect exchange rates to affect their foreign trade flows, taking into account the irrecoverable entry costs they face in international markets. Export supply responses is of central concern to the World Bank and its client countries, being a key element in structural adjustment programs emphasizing commercial policy reforms and devaluations. Unfortunately, export supply responses are not well understood (Roberts and Tybout, 1997b). A part of the reason for this lack of understanding is the neglect of entry costs firms have to incur in international markets in trade models. In fact, Obstfeld and Rogoff (2001) argue that all the major puzzles in international macroeconomics in general are related to trade costs. Even so, the basic claim of Stiglitz et al (2006), criticizing the economic policies advocated by the IMF for developing countries, is precisely the lack of sensitivity the IMF shows to the particular circumstances these countries face. For developing countries exporters, entry costs in international markets are pronounced. Thus, for understanding export supply responses in developing countries the impact of market entry costs should be accounted for.

The paper starts by describing the traditional relationship between prices and exchange rates, as given by the law of one price (LOP) and the purchasing power parity (PPP) theorems, including their modifications, Pricing to Market (PTM) and Exchange Rate Pass-through (ERPT), relating nominal exchange rates to the pricing of tradable goods. Then, the elasticity approach is analyzed, including the J-curve, the Marshal-Lerner and the Harberger-Metzler-Laursen (H-M-L) condition, discussing short, medium and long run relationships between the (real) exchange rate and international trade flows. Third, income absorption, linking the relationship between exchange rates and the current account to the wider economy, is the focus of attention. In
the fourth part the sunk market entry cost hypothesis, where the irrecoverable entry costs firms have to incur in international markets is accounted for. Both the implications for prices and trade volumes are analyzed at the microeconomic level. At the aggregate level both "weak- and strong hysteresis" are described, and the latter is argued to be most relevant for developing countries. The resulting move from time- to state dependent relationship between exchange rates and trade flows following sunk entry costs in international markets is highlighted. The fifth part analyses the implications of sunk-cost hysteresis for the traditional theorems relating exchange rates to trade flows, including both pricing of tradable goods as well as short- and long run trade balance responses following exchange rate shocks. The last part concludes.

2 Pricing of Tradable Goods

2.1 The Law of One Price and Purchasing Power Parity

When goods are traded internationally, the terms at which these transactions take place are partly determined by the exchange rate. The bedrock assumption relating prices and exchange rates is still the law of one price (LOP), even if its overall empirical validity is limited.1

Basically, LOP states that if trade is costless, competition is perfect and markets integrated, identical products will sell for the same common currency price in different countries. Let for instance \( p_i \) represent the home currency price of good \( i \) measured in domestic currency, \( p_i^* \) the home currency price in country \( F \), and \( E \) the exchange rate of home currency per unit of \( F \)'s, then LOP can be expressed as:

\[
p_i = E p_i^*
\]

This condition ensures that an article sells for the same price in different markets, when measured in the same currency. If LOP holds, the relationship between exchange rates and prices is rather simple, as exchange rates adjust to make sure home country prices equal foreign country prices, when

---

1 For a thorough analysis of both LOP and PPP, see e.g. Froot and Rogoff (1995), Rogoff (1996), Goldberg and Knetter (1997), Obstfeld and Taylor (1997) or Taylor and Taylor (2004).
measured in the same currency. Exchange rate pass-through is now perfect, and exchange rates are completely passed onto domestic prices. When prices are equal, no further arbitrage will come about, making LOP an equilibrium condition in international markets.

A first extension to LOP is when all goods are traded internationally, and every country has the same consumption bundle, making LOP hold for all consumption goods. Then, LOP is extended to the so-called purchasing power parity (PPP) condition

\[ P = EP^* \] (2)

Now, the domestic price level \( P \) is related to the foreign price level \( P^* \) through the nominal exchange rate \( E \), and the exchange rate is assumed to adjust in order to make sure that PPP holds.\(^2\) Today, the most widely used application of PPP is found in the so-called Burgernomics, which is developed by The Economist. Burgernomics is based on the theory of purchasing-power parity, the notion that a dollar should buy the same amount in all countries. Thus in the long run, the exchange rate between two countries should move towards the rate that equalizes the prices of an identical basket of goods and services in each country. The Economist’s "basket" is a McDonald’s Big Mac, which is produced in about 120 countries. The Big Mac PPP is the exchange rate that would mean hamburgers cost the same in America as abroad. Burgernomics compares actual exchange rates with PPP to indicate whether a currency is under- or overvalued (The Economist, 2007).

The absolute versions of LOP and PPP are, however, based on a set of assumptions, including the absence of transportation, distribution and resale costs, neither of which can be argued to be realistic for a broad category of goods. The absolute versions of both LOP and PPP are often modified to allow for less than complete price equalization, but keeping price differentials stable between countries, and considered to depend on productivity differentials. The relative versions of LOP and PPP can be expressed as\(^3\)

\[ p_i = \alpha E p_i^* \] (3)

and

\[ P = \alpha E P^* \] (4)

\(^2\)The first application of PPP was presented by Cassel (1922), reproduced by Pitchford (1997), as an aggregate form of LOP.

\(^3\)It is implicitly assumed that all traded goods face the same transportation costs.
where $\alpha$ is the real product exchange rate, normally assumed constant over time. In addition to the assumptions highlighted above, an additional argument in favor of the relative versions is related to assumptions regarding identical goods. In any transaction a number of characteristics will differ, including the seller/producer of the good, the buyer, the location at which the transaction takes place, the physical characteristics of the product, in addition to other non-price characteristics such as date of delivery and invoicing currency (Goldberg and Knetterer, 1997). Such variations will induce differences in prices. However, despite rather weak empirical merits, LOP and PPP are still key assumptions when analyzing pricing of tradable goods. In fact, a number of approaches have been taken in order to revitalize the hypothesis. For instance, Goodwin et al (1990) argues that since trade takes time, the appropriate comparison is not between contemporaneous prices, but between a current market prices and the price expected to prevail in an alternative market at a later date, as this is the basis for arbitrage. Their analysis of 17 narrowly defined US traded primary products, based on rational expectations, is a version of LOP argued to strengthen LOP as a pricing rule. The acknowledgment that trade takes time is a core assumption when constraining LOP and PPP to the long run. As it takes time for markets to correct imbalances, short run relationships between exchange rates and prices are driven by other factors than arbitrage.

### 2.2 Pricing to Market and Exchange Rate Pass-through

Today most economists agree that LOP is restricted to situations where international markets are completely integrated, and is a long run relationship for pricing of tradable goods. Where markets are segmented, other relationships between prices and exchange rates dominate. The two dominating lines of reasoning, exchange rate pass-through (ERPT) and pricing to market (PTM) both came about during the economic turbulence of the 1980s. In general, ERPT refers to how import prices respond to exchange rates within different industry structures, while PTM is related to how price discrimination across different (national) markets implies different responses in different markets following an exchange rate shock.
2.2.1 Exchange Rate Pass-through

The ERPT approach focuses on how prices adjust to exchange rates for transactions between an exporting and an importing country, where exporters from different countries can have different ERPT in an importing country. Krenin (1977) is often used as a seminal reference on ERPT, using import prices from a different importer - one which relevant exchange rate did not change - as a control variable, and applying the change in import prices as an indicator for exchange rate pass-through. Krenin calculates exchange rate pass-through in different markets, estimating it to be only 50 percent for US imports, 60 percent for German, and about 100 percent for Italian imports. Following Krenin, a number of papers have used industrial organization models to show how exchange rate pass-through depends both on the nature - and the degree - of industry competition. For instance, Feenstra et al (1996) applies Bertrand competition in the case of differentiated products, and argues that the exchange rate pass-through depends on the extent of foreign concentration in the industry. The result follows the seminal Dornbusch (1987) paper, analyzing exchange rate pass-through within a number of different industrial organization models, showing the apparent asymmetry between the intrainsdustral international trade theory, and the literature on how exchange rates were assumed to affect the prices on tradable goods. Dornbusch argued that while intrainsdustral trade flows were analyzed within Cournot-, Bertrand- and Spence-Dixit-Stieglitz (S-D-S) models, the relationship between exchange rates and prices often was postulated according to LOP. Dornbusch analyzed short run exchange rate pass-through within a number of conventional industrial organization models, showing how exchange rate pass-through depends on industry characteristics, such as the degree of market concentration, the degree of product homogeneity and substitutability, in addition to domestic and foreign firms market shares.\footnote{Dornbusch emphasises the short run nature of the model, by focusing on the fact that wages are fixed, and that no industry relocation occurs (Dornbusch; 1987, p.104).}

Consider for instance the situation of Dornbusch (1987), where Cournot competition, linear demand and domestic and foreign firms producing a homogeneous good, the industry price $P$ equals

$$P = \left( \frac{nw + n^* e w^*}{N} \right) + \left( \frac{a}{bN} \right) \text{ where } N = n + n^* + 1 \quad (5)$$

Here, $n^*$ is the number of foreign firms, and $n$ the number of domestic firms in...
the industry, $e$ the exchange rate, while $w^*$ and $w$ are unit costs in foreign and domestic currency respectively. The degree of exchange rate pass-through depends on the relative number of foreign firms in an industry and the industry mark-up. The exchange rate pass-through elasticity ($\varphi$) equals:

$$\varphi = \frac{n^*}{N} \left( \frac{ew^*}{p} \right) \text{ where } N = n + n^* + 1$$

(6)

The model implications are particularly interesting as they encompass situations where exchange rate pass-through can stretch from the small country case of complete exchange rate pass-through, that is LOP, to situations where sales are dominated by a few domestic firms and exchange rates have virtually no effect on market prices. The model has a potential for explaining unaffected as well as steep price changes, following exchange rate shocks, and includes both market structure and type of competition.\textsuperscript{5}

2.2.2 Pricing to Market

The PTM model, initiated by Krugman (1987), derives a relationship between prices and exchange rates based on firms participation in a number of different export markets. The firm’s maximization problem is given by Goldberg and Knetterer (1997) as

$$\Pi(p_1, ..., p_n) = \sum_{i=1}^{n} p_i q_i(E_i p_i; v_i) - C \left( \sum_{i=1}^{n} q_i(E_i p_i; v_i) w \right)$$

(7)

where $p_i$ is the market price in exporter’s domestic currency, $q_i$ the quantity demanded in that market, $E$ the exchange rate, and $C(q, w)$ the cost function, where $w$ denotes input prices. The first order condition implies that the marginal revenue from all markets should be equal, making the export price in each market determined by a common marginal cost component, and a market (destination) specific mark-up:

$$p_i = C_q \left( \frac{-\eta_i}{-\eta_i + 1} \right), \forall i$$

(8)

\textsuperscript{5}Dornbusch (1987) also analyses exchange rate pass-through within a S-D-S model, extending it, by allowing for strategic interaction, creating an interesting industry structure, quite similar to a Bertrand model.
where the arguments of the marginal cost function $C_q$ are suppressed, and $\eta_i$ is the absolute value of the elasticity of demand in (foreign) markets. That is, PTM includes both LOP and potentially differing exchange rate pass-through between markets. Marston (1990) relates PTM to LOP by differentiating the first order condition, and shows that the manner export prices respond to exchange rates, depend on the convexity of the demand curve, and the marginal cost effect that might come about due to changes in output or higher import prices. The results are support by a number of papers, as e.g. Irandoust (1999), highlighting policy conclusions following PTM behavior by exporting firms.

2.3 Some Comments on the Pricing of Tradable Goods

Even though LOP and PPP have some long run support, their short term merits are weak. Rather, different exchange rate response between prices seems to be a consequence of third degree price discrimination (Goldberg and Knetterer; 1997, p. 1270). In effect, in the short run national markets seem to be better viewed as segmented than integrated, and allowing for more comprehensive pricing rules in international markets, encompassing both LOP and the more complicated pricing behavior of PTM and ERPT, seems to increase the understanding of the relationship between prices on tradable goods and exchange rates in the short run, even though conditions for arbitrage are important for pricing in the long run. In fact, even though the empirical merits are weak Obstfeld and Taylor (1997) have shown how deviations from PPP and LOP are bounded, and when taking these thresholds into account, mean reversion is fairly fast, and relative prices are kept within corridors determined by the cost of arbitrage.

3 Trade Balance Dynamics and Exchange Rate Shocks

3.1 The Elasticity Approach

In the relationship between exchange rates and the trade balance, the early approaches highlighted the role of price elasticities. The elasticity framework

\footnote{See for instance Gagnon and Knetterer (1995), or Giovannini (1988).}
was originally set out in a framework where the exchange rate was pegged and capital movements played a minor role (Pitchford, 1997), but has, despite these restrictive assumptions, maintained a key role in international economic theory. The elasticity approach states that if the trade balance is hit by a domestic currency depreciation, the question of whether this would improve the trade balance or not, depends on the elasticities of imports and exports and the exchange rate pass-through. The combination of differing short and long run price elasticities, makes the time horizon crucial for the trade balance response.

3.1.1 The J-Curve

The claimed difference between short and long run price elasticities is far from new. The rather extreme position that short run elasticities are negative was initially set out by Magee (1973), arguing the negative elasticities to induce negative trade balance responses in the short run following depreciations, even though the long run response is positive.

The textbook J-curve is derived from the assumption that exchange rates are passed on to domestic import prices completely and immediately, export prices in domestic currency are unaffected and imports and exports change slowly (Meade; 1988, p. 633-644). The initial trade balance response is therefore negative, basically since the quantity response is gradual. When quantities eventually respond, the trade balance improves, as imports fall and exports rise when domestic goods become relatively cheaper than foreign goods. The elasticity approach thus compass a dichotomy between export and import prices, as the latter is determined by LOP conditions, while the former is fixed in domestic currency value.

The textbook model of a J-curve is derived from a situation where markets are perfectly competitive, and prices only reflect costs. In such a situation, the exchange rate pass-through to import prices is complete, while export prices are unaffected. The export price response, reflects the partial nature of the model. At the same time quantities, due to lags in information, recognition, contract, order and delivery structures, respond slowly to price changes (Meade, 1988). We can specify the trade balance (TB) as

\[ TB_t = P_{Xt}X_t - P_{Mt}M_t \]

where \( P_{Xt} \) and \( P_{Mt} - (P_{Mt} = E_tP^*_M) \) are period (t) prices on exports and imports, while \( X_t \) and \( M_t \) are the volumes of exports and imports respec-
tively. The trade balance is now affected by the exchange rate through three channels: Directly through both import and export prices, and indirectly through the quantity responses following changes in relative prices.

**Figure 1 about here**

Magee (1973) provides the theoretical framework for the J-curve by classifying 3 periods, one which he refers to as the currency contract period, where the trade balance response is negative since existing contracts must be fulfilled. The second is a pass-through period, where new contracts are made on the basis of new exchange rate conditions, but where demand still is unaffected. Last, there is a quantity response period, where, as quantities finally respond, the trade balance improves. The quantity response is again slow, due to a number of lags in both supply and demand.\(^7\)\(^8\)

The analysis of Magee also investigates how the currency denomination of contracts affects the elasticities of supply and demand, and thus the trade balance, during the pass-through period. Magee shows how it is only in the case where import contracts are in foreign currency, that the initial negative J-curve period is inevitable, and how results otherwise are ambiguous. Magee shows how the ability of countries to contract imports and exports in domestic currency, which is related to their international market power and the extent of a domestic peso problem, affects the short run trade balance response following changes in exchange rates. The effect of currency denomination also implies that the trade balance response should differ both between developed and developing countries, as well as between small and large countries, as their ability to contract imports and exports in domestic currency differs. In most developing countries imports and exports are predominately invoiced in foreign currency, increasing both the value of imports and exports during the currency contract period. Therefore, the possibility of a J-curve is not as plausible for developing as for developed countries. Nonetheless, it may still be possible to observe a short term worsening of the

\(^7\)See Rhomberg and Junz (1973) for a general description of the lag structure in export supply. However, additional lags arising from the demand side of international transactions can easily be imagined. The switching cost literature of Klemperer (1987) derives demand lags from uncertainty regarding product quality, contractual or learning costs, creating stickiness in demand, and makes it respond less than perfectly to changes in prices.

\(^8\)Isaac (1995) refers to other sources of lags, related to distribution bottlenecks, adjustment costs in distinguishing between temporary and permanent changes in relative prices, intertemporal substitution effects and implicitly contracted long term trade relationships.
trade balance, depending on the initial trade balance position.\textsuperscript{9}

Furthermore, during the pass-through period the trade balance development depends on the short run elasticities in both export supply and import demand, and the results are again in general ambiguous. The worst case scenario is one where exchange rate pass-through in both imports and exports is complete. And last, what exactly happens to the trade balance in the quantity adjustment face is, even though the long run effect is positive, conditional on what happened during the pass-through phase.\textsuperscript{10}

Despite being rigorous in its form, the J-curve is a partial equilibrium process, taking national income, wages, interest rates as given, and must be interpreted as a process which only is valid when the economy has available resources. Noland (1989) argues that the partial equilibrium estimates of trade elasticities may differ substantially from the general equilibrium estimates, ignoring the exchange rate induced effects on domestic activity. However, Nolan also states that the trade balance results are not without interest, as they show the direction that policy induced changes affect real variables.

The existence of a J-curve is still in dispute, and Hsing and Savvides (1996), following Moffett (1989), provide a summary of some recent findings. Moffett (1989) questioned the Magee phases and sketched some of the alternative approaches that have been taken to the relationship between exchange rates and foreign trade. In addition, Moffett showed how the US experience gives rise to a sine-wave rather then a J-curve. A number of different approaches have been taken when analyzing the relationship between exchange rates and trade balance dynamics, questioning the J-curve as such. For instance, Backus et al (1994) simulates numerically a general equilibrium model, where the results resemble an S-curve, more than a J-curve. Their trade balance response is derived in a general equilibrium framework, taking the effect of exchange rates on the wider economy into account, relating it to the income-absorption approach. Likewise, Demirden and Pastine (1995) criticize the conventional empirical methodology underlying J-curve analysis, claiming these approaches not to be valid when the exchange rate is floating.

\textsuperscript{9}Since most trade is invoiced in US $, the US case might, however, be less relevant.

\textsuperscript{10}Gerlach (1989) derives an alternative J-curve by allowing for sticky prices, and two potentially offsetting effects in demand; a relative price effect between domestic and foreign goods, and an intertemporal price effect affecting the allocation of consumption over time. Now, a J-curve might emerge, if the intertemporal substitution effect dominates the relative price effect.
instead applying a VAR-approach to include exchange rate feedback to domestic income, savings, prices etc. Rose and Yellen (1989) and Rose (1991) both question the fundamental theoretical reasoning behind the J-curve, and support their criticism with empirical findings showing that there is no significant relationship between the real exchange rate and the trade balance in any of the major OECD-countries that resembles a J-curve, arguing that the J-curve is due to neglect of aggregation issues, and suffers from an empirical simultaneity problem.

3.1.2 The Marshal-Lerner/Bickerdyke-Metzler-Laursen condition

While the J-curve characterizes short-run adjustments, the Marshal-Lerner condition is viewed as the steady-state condition relying on long run price elasticities (Kongsted, 1996). The Marshal-Lerner condition provides the restrictions necessary for a depreciation of the domestic currency to improve the trade balance in the long run. The Marshal-Lerner condition is, when the economy is out of equilibrium, extended to the Bickerdyke-Metzler-Laursen condition.

The Marshal-Lerner and the Bickerdyke-Metzler-Laursen condition can be derived as follows. Consider a two-country model, where the countries produce goods which are imperfect substitutes. Demand depends conventionally upon income and relative prices, and import demand functions are specified as

\[ D_m = D_m(Y, p_m) = \alpha_y Y - \alpha_p p_m \]  
(10)

and

\[ D^*_m = D^*_m(Y, p^*_m) = \alpha^*_y Y^* - \alpha^*_p p^*_m \]  
(11)

where \( D_m(D^*_m) \) is the volume of goods imported by the domestic (foreign) country; \( Y(Y^*) \) the real income measured in domestic (foreign) output, \( p_m \) is the price of imported goods relative to domestically produced goods in the home country, both measured in home currency, while \( p^*_m \) is the relative price of imported to domestically produced goods abroad, when both prices are measured in foreign currency.

The supply side assumes that export supply depends on the relative price of exportables as:

\[ S_x = S_x(p_x) = \beta_p p_x \]  
(12)
and

\[ S^*_x = S^*_x(p_x) = \beta^*_p p^*_x \]  \hspace{1cm} (13)

where \( S^*_x(S^*_x) \) is the quantity of domestic (foreign) exportables. The relative price of exportables at home equals \( p_x \), defined as the ratio between the price of exportables in domestic currency \( P_X \), relative to the domestic price level \( P \), and similarly for \( p^*_x \). The real exchange rate equals \( q \equiv \frac{EP^*_x}{P} \). Given the supply and demand specifications, we can define the relative price of imported to domestically produced goods at home \( p_m \) and abroad \( p^*_m \) as:

\[ p_m = \frac{EP^*_x}{P} = q p^*_x \]  \hspace{1cm} (14)

and

\[ p^*_m = \frac{P_x}{EP^*_x} = \frac{p_x}{q} \]  \hspace{1cm} (15)

The assumption of imperfect substitutes implies that changes in relative prices equilibrate supply and demand both home and abroad:

\[ D^*_m = S^*_x \]  \hspace{1cm} (16)

and

\[ D^*_m = S_x \]  \hspace{1cm} (17)

The domestic real trade balance can now be given as

\[ RTB = \left( \frac{P_x D^*_m - EP^*_x D^*_m}{P} \right) = p_x D^*_m - q p^*_x D^*_m \]  \hspace{1cm} (18)

or again

\[ RTB = \frac{\beta^*_p (\alpha^*_y)^2 (Y^*_y)^2 q^2}{(\alpha^*_p + \beta^*_p q)^2} - \frac{\beta^*_p (\alpha^*_y)^2 Y^2 q}{(\beta^*_p + \alpha^*_p q)^2} \]  \hspace{1cm} (19)

which express the real trade balance as a function of domestic and foreign income, in addition to the real exchange rate. The partial derivative of the RTB with respect to the real exchange rate, is used to calculate the Bickerdyke-Robinson-Metzler condition, given as:
\[
\frac{\partial RTB}{\partial q} = (1 + \epsilon_S) \left( \frac{\epsilon^*_D}{\epsilon^*_D + \epsilon_S} \right) D^*_m p^*_x - D_m q p^*_x (1 - \epsilon_D) \left( \frac{\epsilon^*_S}{\epsilon^*_S + \epsilon_D} \right) \quad (20)
\]

where \(\epsilon_D(\epsilon^*_D)\) denotes the absolute value of the domestic (foreign) price elasticity of demand, and \(\epsilon_S(\epsilon^*_S)\) denotes the absolute value of the domestic (foreign) price elasticity of supply. The sign of this derivative is in general undetermined, making the effect of a higher real exchange rate on the real trade balance ambiguous. If foreign trade initially is in equilibrium, and the supply curves are perfectly elastic, the expression reduces itself to the Marshal-Lerner condition, where a real depreciation improves the trade balance if the sum of the (absolute values) of the two demand elasticities exceeds unity:

\[
(\epsilon^*_D + \epsilon_D) > 1 \quad (21)
\]

The Marshal-Lerner condition is considered to be a stability condition in international economics.\(^{11}\) The Marshal-Lerner condition is model specific, since it is derived within a one home good model which contains severe limitations due to its simplicity. Pitchford (1997) for instance shows how the trade balance always is improved by a depreciation, in a situation where both tradable and non tradable domestically produced goods are present, and despite its crucial stand in international economics, there are no theorems in consumer theory which imply that the Marshal-Lerner condition must hold.\(^{12}\) Empirically, however, there seems to exist a vast range of literature claiming that the Marshal-Lerner condition holds in real life.\(^{13}\)

---

\(^{11}\)See for instance Ethier (1988) splitting the Marshal-Lerner condition into domestic and foreign elasticities of substitution in both supply - and demand, as well as the propensities to import, and the level of foreign trade in an economy. Pitchford (1997) derives a similar stability condition as the conventional Marshal-Lerner condition within a setting where accumulation of foreign assets, or debt, balances trade in goods. Pitchford, however, questions the policy implications of exchange rate shocks normally argued on the basis of such analysis, as the exchange rate is endogenous.

\(^{12}\)The Marshal-Lerner condition implies that nominal prices are unaffected by exchange rates, and, thus, are kept fixed through open market operations for the situation to be valid.

\(^{13}\)The link between trade elasticities and exchange rates is analysed by Goldstein and Khan (1985) and Menon (1995b).
3.2 The Income-Absorption Approach

A number of authors claim the elasticity approach to be insufficient for describing the long run relationship between exchange rates and the trade balance, since elasticities by far determine the trade balance’s exchange rate response by themselves:

“Relative prices are not the only thing that matters. Whatever the price elasticities of demand and supply for imports, tradables and non-tradables, it is worth remembering the lesson of the absorption approach that an improvement in the current account requires a reduction in expenditure relative to income; likewise, the monetary approach reminds us that there is a capital account as well as a current account, and that an exchange rate induced excess demand for real money balance, if permitted to arise can be satisfied from either channel” (Goldstein and Khan, 1985, p.1096)

A number of papers, stemming from the seminal papers of Harberger (1950) and Alexander (1950), relate exchange rates to national income identities in order to show how the trade balance not only included pure trade decisions, but how the trade balance and the current account must be jointly determined with national expenditures, savings, income and investments. The total effect of exchange rates on the trade balance is derived from the total effect on all of these variables. At the same time as the relationship between excess expenditures and the current account is highlighted, the relationship between current account deficits and foreign debt is clarified, creating a natural link to the monetary approach.\(^\text{14}\) Within the income-absorption approach, the exchange rate is treated as an endogenous variable, determined together with other relative prices. Manipulation of the exchange rate for policy purposes thus requires special conditions, equal to those set out in the elasticity approach.

3.3 The Harberger-Metzler-Laursen Effect

The conventional relationship between domestic income and expenditures is the Harberger-Metzler-Laursen (H-M-L) condition, which states that a worsening in the terms of trade (TOT) will deteriorate the current account at each level of nominal income. The argument is that a TOT deterioration decreases real income, and that this reduces savings at all levels of income,

\(^\text{14}\) The monetary approach to the current account is often seen as a natural extension of the specie flow mechanism for the case of fixed exchange rates (Pitchford, 1995, p.30)
when measured in terms of exportables. Now, if investments are fixed, and there is no government deficit, the change in savings will equal the current account by definition. Thus, the H-M-L condition implies that the current account will deteriorate in response to a TOT deterioration.

Consider the reasoning of Alexander (1952), where the foreign balance $B$ is equal to the difference between the total production $Y$ and total absorption $A$

$$B = Y - A$$  \hspace{1cm} (22)

where changes in these variables are denoted by small letters, so that

$$b = y - a$$  \hspace{1cm} (23)

is a fundamental identity. The identity indicates that the change in foreign balance equals the difference between the change in domestic output and the change in domestic absorption. How a devaluation affects the trade balance depends on its effect on $y$ and $a$. The absorption of goods is assumed to depend partially on real income, which itself is equal to output. Absorption is also assumed to depend directly upon the price level, and other factors related to the devaluation, so that

$$a = cy - d$$  \hspace{1cm} (24)

where $c$ is the propensity to absorb, equal to the sum of the propensity to consume and invest. The term $d$ is the direct effect of the devaluation on absorption, reflecting whatever tendency there is for the devaluation to change the level of absorption. Combining these two equations gives us the relationship

$$b = (1 - c)y + d$$  \hspace{1cm} (25)

The total trade balance effect is now equal to the sum of the effect on income $(y)$, the effect of income on the level of absorption $(c)$, and the effect on the direct absorption $d$. Now, let $e$ denote the exchange rate and the H-M-L condition equals

$$\frac{db}{de} = \left(\frac{dy}{de}(1 - c) - y\frac{dc}{de} + d\frac{dd}{de}\right) > 1$$  \hspace{1cm} (26)
Being a general equilibrium condition, the criticism of H-M-L have arisen from a number of different strands. Both by arguing improper modelling of investments and savings, a claimed difference between temporary and permanent TOT effects, the neglect of substitution effect following the one good model structure, and the exogenous income assumption have all been addressed.

3.4 Modifications and extensions to H-M-L

The H-M-L effect was first criticized for fitting static savings behavior only, derived from a one good, open economy Keynesian model without capital mobility. The attempt to overcome the criticism related to static savings behavior was first (?) addressed by Obstfeld (1982), who derived a situation where a small economy consists of an infinitely lived representative consumer with an Uzawa (1982) type of utility function, where the rate of time preference is increasing in the level of utility. Now, the economy has a target level of wealth, and since a TOT deterioration implies a wealth reduction, it is necessary to accumulate wealth in order to reach the desired level, and the economy must thus increase savings. Therefore, a TOT deterioration increases savings, and improves the current account.\(^{15}\)

However, Obstfeld was criticized for the choice of utility function, which is crucial for his results.\(^{16}\) The H-M-L effect was again analyzed by applying a special type of utility function by Mansoorian (1993), applying the habit persistence model of Ryder and Heal (1973). Now, the H-M-L condition depends on whether the marginal utility of consumption is increasing or decreasing in the habitual standard of living.

A more conventional approach was taken by Svensson and Razin (1983), analyzing TOT effects within an intertemporal framework with conventional preferences, separating the trade balance effects of a TOT deterioration into three; a direct effect following the revaluation of the export vector, a wealth

\(^{15}\)Obstfeld (1980) questioned the H-M-L effect by considering the possibility for the economy to import intermediate goods. The effect of a real exchange rate appreciation is then most likely an improvement in the current account, in contradiction to the H-M-L effect.

\(^{16}\)Svensson and Razin (1983) argues that there are three reasons for Obstfeld’s result. First, his intertemporal utility function is simply a discounted sum of period utilities, implying that preferences are homothetically weakly separable. Second, he considers a permanent terms of trade deterioration leaving the interest rate unaffected. Third, the time rate of preferences are assumed to be increasing with the level of utility.
effect on consumption following the reduction in domestic wealth, and a pure substitution effect on consumption due to changes in relative prices both within and between periods. In general, the sum of these effects is ambiguous. However, even though Svensson and Razin (1983) are unable to make precise predictions about how a TOT shock affects the trade balance, they can make statements about the different effects following temporary and permanent deteriorations in TOT. A temporary deterioration generates a temporary fall in income, and a change in the real interest rate. The reduction in income leads both directly, and indirectly through the wealth effect in consumption, to a deterioration of the trade balance. At the same time the interest rate effect induces a substitution effect on spending, and if the interest rate falls following the TOT deterioration, this reinforces the two former effects, and the current account unambiguously deteriorates. On the other hand, a permanent TOT effect leaves the real interest rate unaffected, and the current account effect depends on whether the rate of time preference decreases or increases with welfare. Thus, the main result of Svensson and Razin (1983) is that the general H-M-L effect only is valid in case of a temporary TOT deterioration. The distinction between temporary and permanent shocks is in line with Edwards (1989), Gavin (1990) and Mendoza (1992), questioning how a TOT deterioration affects the trade balance in a forward looking framework.

In addition to highlighting the role of savings, and explicitly including conventional preferences by households, the other main line of extensions to the original H-M-L effect is found in the modelling of investments. In the original H-M-L theorem investments are fixed, and the current account is driven by the exogenous savings component. Several authors have analyzed the role of investments relating the current account to exchange rates. Persson and Svensson (1985) allowed for both savings and investment dynamics, and as there are cycles in the current account, they argue that the H-M-L effect on savings can be either positive or negative for all plausible parameter values. Without focusing explicitly on savings, Risager (1988) considers the effects of a devaluation on investments, and shows how the short run trade balance effect is uncertain, while the trade balance is unaffected in the long run. The short run trade balance effect is also analyzed by Bo Nielsen (1991), who analyses the effect of a devaluation on investments when wages are sticky.

---

17 See Persson and Svensson (1985) for a short survey on the modelling of investments and the effect on the current account.
The result depends on the length of the period in which nominal wage contracts are fixed. If the contract period is short, the devaluation improves the current account, while a devaluation deteriorates the current account the more, the longer the wage contract period and the smaller the adjustment costs of capital.

Even so, the original H-M-L encompassed only the absorption effect of a devaluation, leaving out the effect on income. Sen and Turnovsky (1989) analyzed a TOT shock within an infinite horizon model where employment is endogenous, households are allowed to make choices related to leisure and labor and where a q-theory of investments is applied, allowing for capital accumulation. Both permanent and temporary, as well as anticipated and unanticipated exchange rate shocks are analyzed, and the long run response is crucially depending on the long run response of capital. The long run trade balance response depends on a negative substitution effect and a positive income effect on consumption, and the H-M-L effect now depends on which of the two that dominates. However, no matter whether the TOT effect is temporary or permanent, if it is anticipated or not, if the more plausible substitution effect dominates the income effect, the current account initially goes into surplus following a depreciation, in opposition to the H-M-L prediction.

4 Hysteresis in Foreign Trade

Following the development in the US trade balance during the 1980s, several authors questioned the validity of the conventional international trade-exchange rate theorems

"why is it that in some countries and time periods, a given trade and exchange rate regime supports large-scale production for foreign markets, while in other countries and time periods, the same policies appear to induce a minimal export supply". (Roberts and Tybout, 1997, p.545)

The conventional theorems’ lack of ability to explain the empirical record, combined with a weak microeconomic foundation, lead to questions regarding the proper reasoning relating international trade flows to exchange rates.

18Sen and Turnovsky (1989) summarises other intertemporal models of the current account which includes labour supply and the q-theory models of investments.
Following the turbulence of the 1980s it was first argued by Richard Baldwin, Paul Krugman and Avinash Dixit that international trade would emerge as a natural framework for hysteresis, if one allows for a proper microeconomic foundation.\textsuperscript{19} They argued that by taking the irrecoverable entry cost firms face in international markets into account, and the non-linear relationship between exchange rates and foreign trade they induce, the relationship between international trade flows and exchange rates is at heart state dependent. Thus, exchange rate shocks of different magnitudes will entail different foreign trade responses, and even large temporary shocks to the exchange rate can lead to permanent effects in foreign trade flows.\textsuperscript{20} Even so, sunk cost hysteresis allows history to play a role in trade patterns and trade structures, making it highly relevant for developing countries, knowing the current problems of developing countries to enter the markets of developed countries (Roberts and Tybout, 1997).\textsuperscript{21}

4.1 Foreign Trade and Hysteresis at the Firm Level

The forerunners of supply side hysteresis are Baldwin (1988a, 1988b, 1990), Baldwin and Krugman (1989) and Dixit (1989a, 1989b). These papers derived dynamic versions of a supply side model for international trade where firms have to incur some irrecoverable entry costs in international markets, and where entry and exit is allowed for.\textsuperscript{22} This microeconomic structure implies that foreign trade is non-linear in nature, and that exports and imports should be analyzed as regular investment decisions within intertemporal

\textsuperscript{19}By hysteresis they mean that the system after being hit by an external shock failed to reverse itself as the underlying shock is removed. A temporary shock thus leaves behind a permanent effect (Dixit, 1989a, p. 622). The term hysteresis is often somewhat confusingly applied alongside terms as persistence, and path-dependance. See for instance Cross (1994).

\textsuperscript{20}The hysteresis effects in foreign trade come about both through supply- and demand side effects. The supply side effects include both the existence of sunk market entry costs (Baldwin, 1988a, 1988b), learning effects (Krugman, 1984), ”cold baths” (Harris, 1992) and the theories of vintage capital and the life cycle of firms (Harris, 1992). The demand side effects of Froot and Klemperer (1989) base their arguments on ”consumer switching costs” and consumer ”uncertainty regarding product quality”.

\textsuperscript{21}See for instance Schembri (1991), discussing the role of history in various foreign trade models containing hysteresis.

\textsuperscript{22}Baldwin, Baldwin and Krugman and Dixit make different assumptions regarding market structure and exchange rate determination, but the key factor of all the papers is the irrecoverable market entry costs firms have to incur.
frameworks. The market entry costs are firm specific, and include the costs necessary to set up a distribution and resale network, product launching, training of staff, product upgrading according to local health- and environmental standards, in addition to finance the foreign market expansion itself (Göecke, 1994, p. 558). The costs are at least to some extent sunk, and create a special situation for trade dynamics both at the microeconomic and at the macroeconomic level.\(^\text{23}\)

The reasoning of the model can easily been seen as indicated by figure 2, which is along the lines of Baldwin and Krugman (1989). Consider a situation where for simplicity only foreign firms are able to supply the domestic market, so that there is no domestic competition. Also, assume that an entry cost must be incurred in order to supply the domestic market. By examining the dynamic programming problem of a representative foreign firm, Baldwin and Krugman (1989) derive two values of the exchange rate (measured as the foreign price of domestic currency) - the minimum level of the exchange rate that will induce foreign entry \(e_I\), and the maximum exchange rate level that will induce exit \(e_O\), where \(e_I > e_O\). The asymmetry between the exchange rates that trigger entry and exit is due to the existence of sunk entry costs. If initially no foreign firms operate in the domestic market, imports are zero until \(e = e_I\). However, once the exchange rate appreciates beyond this value, foreign firms will enter and domestic imports (foreign exports) becomes positive. Even so, imports will remain positive if the exchange rate falls below \(e_I\) again, as long as it stays above the exit trigger, i.e \(e > e_O\). That is, for exchange rates between \(e_I\) and \(e_O\) no entry or exit occur, and domestic imports (and foreign exports) are determined by the number of exporters that already are active. For exchange rates within the band no exporters change positions, and those that are out of the market stay out, and those that are in, stay in. Entry or exit only comes about if the exchange rate passes either of the triggers. That is, market structures are affected and regime shifts in both pricing and trade volumes now comes about. The distance between the exchange rates that separates situations where regime shifts accompanying entry or exit comes about from where it does not, is referred to as the hysteresis band (Baldwin, 1988a).

\(\text{23}\)The existence of sunk costs also provides a potentially new and stronger reason for why exchange rate uncertainty might hamper trade flows. As in conventional investment theory the existence of sunk costs and uncertainty extends the band of inaction - that is the hysteresis band - where it is optimal for firms not to alter their position. Thus, increased exchange rate uncertainty unambiguously hampers foreign trade (Dixit 1989a, 1989b).
The multibranch structure of an individual firm’s export behavior is often referred to as a microeconomic hysteresis loop, since a firm’s export market position is indeterminate for exchange rates within the band, as active export positions can coexist with inactive positions for otherwise identical firms.\(^{24}\) The decision of whether to begin, or to cease exporting is similar to those of when to exercise an option, and "a wait and see strategy" is thus optimal (Dixit, 1989b). The sunk entry costs, makes exporting and importing equal to investments, only incurred if expected to be profitable.

The microeconomic hysteresis effect, where even temporary shocks in exchange rates can have persistent effects on foreign trade flows, is also easily seen from figure (2). Consider for instance the situation where the exchange rate initially lies between the entry and exit triggers, combined with a temporary exchange rate shock that moves the exchange rate outside the band. The shock makes some new firms enter, or some existing firms exit, and a mere reversal of the exchange rate is now not enough to restore the initial positions. A corrective shock in the opposite direction is due to the sunk entry costs necessary to bring about the initial situation, otherwise a permanent change in foreign trade structure comes about.\(^{25}\)

\(\text{Figure 2 about here.}\)

The sunk cost model have with its microeconomic foundation and logical consistency been the focus of significant theoretical interest. Lack of data on individual firm’s export participation patterns have however so far hampered extensive empirical investigation of the model. Sullivan (1996) and Roberts and Tybout (1997) are a few of the papers focusing directly on the implications of sunk costs, investigating entry and exit patterns in international markets using plant level data. They present a dynamic discrete choice model of export behavior that separates the role of profit heterogeneity from sunk costs, as sources of export persistence. By applying data from Morocco and Columbia respectively, both papers find considerable persistence in individual firms’ export participation patterns. They argue in favor of the sunk cost

\(^{24}\)The width of the 'hysteresis band' can be seen to be wider apart, the higher the level of entry costs, the higher the entry costs relative to the recurrent costs that must be incurred each period for maintenance purposes, the higher the degree of mean reversion in the exchange rate, the higher the degree of exchange rate uncertainty and the more entry / exit affects industry profits. (Dixit; 1989a).

\(^{25}\)Figure 2 is a duplication of Figure 1 in Baldwin and Krugman (1989) p. 641.
hypothesis by showing how prior export experience increase the probability of exporting today with as much as 60%. Roberts and Tybout (1997b), analyzing microeconomic foundations of industrial export booms in developing countries more generally, argue that sunk cost is a significant factor for deciding export participation patterns, and in particular to be relevant for developing countries. Second, based on the sunk cost model, strong policy conclusions are drawn. That is, countries undertaking export promotion policies should distinguish clearly between measures aimed at expanding export volumes from existing exporters, and from policies aimed at promoting exports from new entrants. If entering a new market is a significant hurdle, policies aimed at promoting entry might be more expanding than those aimed at stimulating exports from existing exporters through subsidies.

4.2 Foreign Trade and Hysteresis at the Aggregate Level

Even though the possibility of hysteresis seems obvious at the firm level, some caution might be in order when aggregating non-linearities, as there is always a possibility that aggregation will smooth away non-linearities. There are, however, a number of papers that have shown how non-linearities might survive aggregation. Krugman and Baldwin (1989) provides a first simple procedure for showing how hysteresis might not be overturned by aggregation within the elasticity approach, followed by Chow (1991), Chen (1991) and Han (1991).

Even so, the most powerful (?), yet simple, argument for why hysteresis will survive aggregation is by Dixit (1989), who argue that the only change in the results of the microeconomic sunk cost model by substituting the Brownian motion assumption with a mean reverting exchange rate, that is moving from a partial to a general equilibrium framework, is that the exchange rate triggers that defines the hysteresis band will be further apart. That is, as the mean reverting assumption reduces the expected future income stream from exporting, compared to the situation with a Brownian motion, exporting firms will now be less willing to change market positions. However, at some exchange rates, it is of course still profitable to change positions. That is, the hysteresis band is wider and the exchange rate triggers are further apart in the case of mean reverting exchange rates, but the basic non-linear structure remains.
Ljungquist (1994) develops a general equilibrium model where hysteresis in foreign trade might be present following temporary exchange rate shocks, and, in addition, equity values in the country facing a positive demand shock ultimately will fall below their pre-shock level. Hysteresis in trade is thus accompanied by hysteresis in equity values.\(^{26}\) Even so, Baldwin and Lyon (1994) integrate a sunk cost model with exchange rate overshooting as in Dornbusch (1976), and shows how hysteresis in foreign trade can be transferred back into the exchange rate.

In fact, a number of authors have argued that hysteresis in trade, in addition to feed-back into other variables and survive aggregation, in fact even might be strengthened by it. This approach was developed by Amable et al (1991, 1993), and Cross (1994). These papers combine a multitude of discontinuous microeconomic hysteresis loops of heterogenous firms into a macroeconomic hysteresis loop, where hysteresis effects occurs continuously with every change in the direction of the input path, as long as local extreme values are passed. Since persistence is reinforced by aggregation, this macroform of hysteresis is referred to as "strong hysteresis".\(^{27}\) Figure (3) illustrates the continuos macroeconomic "strong hysteresis" loop of Göcke (1994).

\textit{Figure 3 about here.}

A simplification of the original hysteresis loop is presented by Göcke (1993,1994) based on linear partial functions and a "loading-unloading" approach which may be applied for econometric analysis. The model is applied to the relationship between US imports from Japan and the dollar-yen exchange rate, and show indications of "strong hysteresis". The "strong hysteresis" approach is again applied by Borgersen and Göcke (2007) within an overshooting model along the lines of Dornbusch (1976), showing how sticky prices and exchange rate overshooting can induce foreign trade hysteresis at the aggregate level. Short run exchange rate overshooting can now, in addi-\(^{26}\)Christophe(1997) examines the value of the US multinational corporations following the dollar turbulence of the 80s and finds that firm values decrease, which is in accordance with the predictions of Ljungquist.

\(^{27}\)The term "strong hysteresis" was first used in an international trade context by Amable et al (1991), who used the reasoning of Cross (1994) and showed how the specific features of such a loading-unloading approach implies that all local extremas in the exchange rate gives a persistent effect on a countrys foreign trade flow. They also discuss the various ways in which the term hysteresis is used and is interpreted.
tion to affecting foreign trade structures, also create changes in the long run equilibrium exchange rate. The approach of Borgersen and Göecke (2007) extends that of Baldwin and Lyon (1994) as the former apply a "strong hysteresis" methodology for analyzing foreign trade structures, whereas the latter is based on a "weak hysteresis" approach. While only large shocks to the exchange rate induce hysteresis in the latter, it occurs with every change in direction of the input path, as long as the exchange rate reaches a local extreme value, in the former. In basic, the distinction between "weak hysteresis" and "strong hysteresis" can be related to the cost heterogeneity of firms. Borgersen (2007b) discusses how export supply responses might differ between the two, impacting on the optimal industrial structure of export processing free zones.

When it comes to hysteresis however, a number of varieties exist both at the microeconomic and at the macroeconomic level. In fact, even within supply side hysteresis, several versions appear at the aggregate level. For instance, Krugman (1991) sets out the basis for a related form of supply side hysteresis, the so-called "deep hysteresis". Again, aggregation strengthens the microeconomic non-linearity. Krugman relates "deep hysteresis" to deindustrialisation of the American economy, and the aggregate economic processes at work. In addition, Krugman (1988a) examines the long run impacts of the US $ appreciation during the mid 80s on the US current account and develops three models, where two rely on the existence of foreign debt and non tradables respectively, but the third in fact is a sunk-cost model, referred to as an invisible asset model. Here, firms have to incur some fixed costs in order to acquire a invisible market asset, such as brand recognition, implying market entry costs. Now, conventional non-linearities exists, and the foreign trade balance dynamics is state-dependent.

4.3 Some Comments on the Empirical Merits of Sunk Cost Hysteresis

At the microeconomic level, the empirical merits of sunk cost hysteresis in developing countries are - as discussed in section 4.1 - somewhat favorable, even though there is a need for more extensive research on the issue. At the aggregate level, the empirical merits are however weak. This is first of all due

---

28 A number of definitions and concepts are referred to as hysteresis, or path-dependence, see Schembri et al (1991) or Harris (1992) for a survey.
to lack of data for export patterns at the plant level in developing countries, making aggregate levels difficult to construct. Second, even when it comes to developed countries, the merits are blurred. This is in part due to methodology, as hysteresis is analyzed as a mix of sunk costs as such, and structural breaks more generally. In addition, applying both "weak hysteresis" and "strong hysteresis" as theoretical frameworks, and different interpretations of the hysteresis phenomenon as well, complicate the results. For instance, David and Papell (1997) investigates trade structures in the OECD area and find evidence of structural breaks in most countries. The paper is however not related to sunk cost and hysteresis, while others, emphasizing hysteresis more strongly as a theoretical framework, e.g. Parley and Wei (1993), analyzing bilateral trade flows between US and Japan at a disaggregated level, finds no support for hysteresis in the form of structural breaks. That is, similar approaches are differently related to sunk cost, and contain different results and interpretations.

Even so, a number of the papers find support for hysteresis in developed countries. Bean (1988) analyses export behavior in the UK from 1900 to 1986 in light of the possibility of having hysteresis in export supply, and finds support for the hypothesis. In fact, Bean argues the possibility of hysteresis in UK exports, to be situated both at the supply and at the demand side. Likewise, Anderton (1999) investigates UK export performance for the period 1971Q1-1992Q4, applying a model along the lines of Baldwin (1988b), finds hysteresis in both price elasticities and trade volumes. Giovanetti and Samiei (1996) develops an empirical model of trade account hysteresis which is able to distinguish between two types of hysteresis; that arising from changes in exchange rate pass-through, and that arising from regime switches in supply, that is a model structure along the lines of Dornbusch (1987), including entry and exit of foreign firms. The analysis is compared to a conventional time series approach that links hysteresis to non-stationarity and unit roots. In the latter case any non-specific shock can produce long-term effects. Applying the model to both German, Japanese and US data, Giovanetti and Samiei finds evidence in favour of hysteresis however only in the case of Japanese exports.

As hysteresis implies that waiting is optimal, the distinction to long adjustment lags is vague. The relationship between the two is illustrated by Gagnon (1989), analyzing adjustment lags in foreign trade structures, related to both adjustment costs and market entry costs. Gagnon states that exports and imports are not very responsive to relative prices in the short run, and
derive a framework for justifying the use of long lags. Gagnon argues that due to information- and transportation lags, as well as spatial contracting costs unique to international business relationships, trade flows are slow to respond to changing economic circumstances. The spatial contracting costs, e.g. language barriers, unfamiliarity with foreign business practices, the expense of transporting a negotiating team to the relevant market, etc., are exclusive to foreign firms and make international trade respond more slowly than domestic sales to changing market conditions. By applying data on US exports to Japan, Germany and the UK within a full information framework, the paper rejects the hypothesis of no adjustment costs. The paper does however not support the hysteresis approach, even though the model contains market entry costs, but instead specifies a time-dependent relationship with a long adjustment lag structure.

The empirical merits of hysteresis at the aggregate level in developed countries seems to be somewhat blurred by the fact that one analyses the implications of the sunk cost models, more than actually looking at sunk cost as such. Even so, with the exception of Göcke (1994), most papers apply "weak hysteresis" as their core frame of reference, even though the heterogeneity of exporters, the basic difference to "strong hysteresis" is established by a number of papers, e.g., Bernard and Jensen (2004) for US exports, Bernard and Wagner (1998) for German exports, and Campa (2004) for Spanish exports. That is, the methodological approach, highly partial in nature and often assuming "weak hysteresis", seems to constrain the empirical validity of hysteresis at the aggregate level in developed countries.

When it comes to developing countries, there are as mentioned only a few papers analyzing hysteresis explicitly. The most comprehensive analysis of sunk cost for developing countries is the above mentioned papers by Sullivan (1996), and Roberts and Tybout (1997, 1997b) discussed in section 4.1. The existence of entry costs in international markets for African countries is in addition discussed by Bigsten et al (2004), Gumede (2004), Söderbom and Teal (2000), pinpointing the relevance of international market entry costs, the heterogeneity of exporters, and the potential non-linear processes governing trade dynamics in African countries. In fact, even though less rigorous, the empirical analysis on sunk costs in developing countries export supply responses seems more to the point, than the empirical analysis on sunk costs in developed countries.
5 Sunk Cost Hysteresis and the Conventional Trade-Exchange Rate Theorems

5.1 Assessing The Pricing of Tradable Goods and Sunk Cost Hysteresis

The existence of market entry costs invalidates the core assumption underlying the absolute version of LOP, since the market entry cost by nature is a transportation cost. Market entry costs thus immediately lead the way for the relative version of LOP. However, entry costs simultaneously deter entry and restricts competition, leading the way for PTM or ERPT reasoning. Even so, the possibility of entry or exit shows how regime switches can come about, and why exchange rate pass-through is context specific.

In the case of "weak hysteresis" a direct implication of sunk costs is that small and large shocks to the exchange rate are passed on to industry prices differently. Small shocks that do not induce entry or exit of foreign firms, contain a fixed exchange rate pass-through and maintain a time-dependent relationship between exchange rates and prices. In the case of large shocks however, pass-through might change over time, and makes the relationship between exchange rates and prices less general. \(^{29}\) In fact, the sunk cost model shows how both pricing of tradable goods and exchange rate pass-through might vary both between countries and over time, which is in accordance with empirical observations (Goldberg and Knetterer, 1997).

The logic behind the situation, encompassing both large and small shocks to the exchange rate, can easily be seen by combining the industry price and the exchange rate pass-through of Dornbusch (1987), and the non-linear entry function of Krugman and Baldwin (1989). The industry price of Dornbusch is as given by equation (5)

\[
P = \frac{nw + n^*ew^*}{N} + \frac{a}{bN} \tag{27}
\]

while the exchange rate pass-through is as given in equation (6)

\[
\varphi = \left(\frac{n^*}{N}\right) \left(\frac{ew^*}{p}\right) \ \text{where} \ N = n + n^* + 1
\]

\(^{29}\) See for instance Dornbusch (1987) for an analysis on the relationship between exchange rates, prices the number of firms in an industry and the exchange rate pass-through.
Combined with the non-linear entry function described earlier

\[
n_t^* = \begin{cases} 
\tilde{n}_t & \text{if } e_t > e_t^I \\
n_{t-1} & \text{if } e \in (e_t^I, e_t^O) \\
\tilde{n}_t & \text{if } e_t < e_t^O 
\end{cases}
\]  

(28)

the system shows how entry and exit of exporting firms, and the accompanying regime switching, easily can come about in the case of large exchange rate shocks. That is, in the case of "weak hysteresis", an ERPT situation containing regime switches comes about.

The hysteresis band \((e_t^I, e_t^O)\) separates two regimes for relating exchange rates to pricing of tradable goods and exchange rate-pass through. Inside the band, the exchange rate affects industry prices through existing exporters pricing rules, and the exchange rate pass-through is determined by the industry mark-up and the degree of competition. Within the band the degree of mean reverting is strong, as price differences now mainly are determined by productivity differentials. Outside the band, an additional effect comes into play through entry or exit of foreign firms. The effect of changes in market structure impacts on the degree of mean reverting, and makes the total effect on both market prices and exchange rate pass-through state-dependent. The existence of such a band is claimed by Krugman and Baldwin (1989), and is in accordance with empirical observations by Obstfeld and Taylor (1997).

When it comes to "strong hysteresis", Borgersen and Göcke (2007), extending Göcke (1993), is one of the few papers directly analyzing the implications for prices and exchange rate pass-through. At the aggregate level, the effects of a monetary shock on the exchange rate and nominal prices are not proportional, as the long run equilibrium exchange rate is path-dependent. The PPP condition will thus not be fulfilled. The exchange rate pass-through is determined by current prices’ deviation from the long run equilibrium price level. As the equilibrium price level continuously is affected by changes in the exchange rate, the pass-through is completely idiosyncratic. In any given period the pass-through depends on the number of both foreign and domestic firms in an industry, the size of the entry cost, the exchange rate volatility, the exchange rate shocks itself, as well as the exchange rate history. That is, there is always some degree of mean reverting governing pass-through, but as the long run equilibrium price level is not unique, no clear predictions regarding size, or even the direction, of exchange rate pass-through, can be applied. For instance, as a short term depreciation is accompanied by a long term ap-
preciation, established exporters governed by different planning horizon can entail completely different exchange rate pass-through. While an exporter with a short horizon might respond to the initial depreciation by increasing market prices, an exporter with a longer horizon might reduce prices (Borgersen, 2007). That is, in the case of "strong hysteresis" no time-dependent prediction regarding short- or long run exchange rate pass-through can be made, as equilibrium prices are path-dependent.

5.2 Assessing The Elasticity Approach and Sunk Cost Hysteresis

Dixit (1994) and Backus (1994) were the first (?) two papers to discuss the relationship between the J-curve and sunk cost hysteresis, both applying the "weak hysteresis" framework. Dixit compares the two by stating that in the J-curve the lag that comes into play before the trade balance improves following an exchange rate shock, is a time-lag, while the lag in the hysteresis approach is explicitly denominated in the dimension of the economic state (Dixit; 1994, p.106). That is, in the case of hysteresis the distance is between the current level of the exchange rate and the action threshold defining the hysteresis band. The time it takes to travel this distance is random, and depends on the parameters of the exchange rate process, and the optimal decisions of exporters and importers. Dixit argues that the fixed rule of thumb for relating trade balance dynamics to exchange rates, as for instance the Obstfeld and Krugman (1991) statement that for most industrial countries sufficient elasticity response emerges within a year for the trade balance to improve following a depreciation, is likely to be misleading. If sunk costs are more pronounced for developing countries, such a statement is then even less relevant for developing countries. Context specific predictions about the likely duration are in general better for predicting the foreign trade dynamics following exchange rate shocks. In a hysteresis framework, both the exchange rate’s current deviation from its long run equilibrium, its trend as well as the short run volatility are important for assessing trade dynamics. Dixit applies the model to the US current account during the 80s, and is able to explain its behavior, arguing that the J-curve at heart is a state dependent process, without a fixed time lag. The large appreciation of the US dollar during the 80s, induced regime shifts in the US foreign trade structure along the lines of "weak hysteresis".
Dixit (1994) questions the theoretical assumptions underlying the J-curve, first of all regarding perfect short run export supply elasticities. Whether short run elasticities are small enough to generate a J-curve is in itself controversial. Even so, according to the sunk cost model short run price elasticities will differ between small and large exchange rate shocks, as the latter induce entry or exit, processes the former does not encompass. If entry occurs, and new exporters are able to increase exports immediately, even short run elasticities can be substantial. If entry takes time, long run elasticities can be expected to differ between situations where entry occurs from where it does not. If existing exporters not are able to increase exports neither in the short nor in the long run, different short and long run elasticities hinges on new entry, which in the case of sunk market entry costs, again depends on the size of the exchange rate shock.\footnote{However, Isaac (1995) argues that the restrictions regarding different short and long run elasticities is questionable, and in opposition to empirical facts: "The hysteresis in trade model..., but it does not accomodate the well documented difference between the long run and the short run trade elasticities" (p.342).} Thus, by questioning the short run elasticities Dixit (1994) questioned the J-curve as such, as the medium and long run trade balance improvements are conditional on short run developments.

When analyzing the implication of trade account hysteresis for the relationship between prices and exchange rates a conventional approach has been to look for structural breaks in the relationships determining prices, in particular import prices. Baldwin (1990) develops a theoretical model, referred to as "the beachhead model", focusing on the hysteretic effects following large exchange rate shocks. From the solution to the maximization problem faced by exporting firms, two testable implications arise, both with implications for the elasticity approach: The first is that industry prices will be inversely related to the number of firms. Thus, if hysteresis occurred as a result of the appreciation of the US dollar during the 80s, the number of exporting firms operating in the US market should have increased, and hence reduced equilibrium prices. Baldwin analyses import prices and exchange rate pass-through in which an index of US import prices is regressed on current and lagged proxies for foreign marginal costs (measured in domestic currency), and finds weak evidence of such a structural shift.\footnote{Although he finds weak evidence of such a shift he admits that it might be due to a number of other reasons besides hysteresis (Baldwin, 1988b, p. 784).} The second implication of the model, again directly related to the elasticity approach, and the reasoning of Dixit (1994), is that the price elasticity should
increase with the number of firms. That is, large shocks to the exchange rate should also impact on price elasticities. Following Baldwin (1988b), as well as the "bottleneck model" of Baldwin and Foster (1986), Baldwin (1990) finds evidence of a structural break in US import prices during the mid 80s. In contrast to the results of Baldwin, neither Hooper and Mann (1989) nor Chow (1991) however, find little support for a structural shift in the US import pass-through during the 80s, but instead somewhat less stable structural equations.

Feinberg (1992) tests another implication of the hysteresis model, related to that not only the volume of trade, but also its spread across markets should be affected by exchange rate shocks. Large depreciations (appreciations) should be followed by increased (decreased) spread in exports (imports), and decreased (increased) spread in imports (exports). Applying US export data at the 4-digit industry level the paper identifies structural breaks in the distribution of exports following large exchange rate shocks. The results differ between industries, ranked by the level of entry costs. Feinberg argues the changed spread in exports to be consistent with an important implication of the hysteresis hypothesis. Likewise, motivated by the sunk cost hypothesis, Knetterer (1994) analyzed how the degree of competition in U.S. product markets was affected by the dollar appreciations of the 80s, finding weak support for increased competition, where again, results varied between industries, ranked according to their market entry costs.

Göcke (1993) is however the only (?) paper directly analyzing "strong hysteresis" empirically, finding evidence for hysteresis in imports and exports volumes between Japan and the US. Applying a model with linear-partial functions Göcke (1993) argues that trade account elasticities are state dependent, and as "strong hysteresis" is present, all hypotheses regarding elasticities explicit in the dimension of time are invalid.

5.3 Assessing The Income-Absorption Approach and Sunk Cost Hysteresis

The income-absorption approach draws its predictions on trade balance dynamics following exchange rate shocks based on how both the generation of income and the ability to absorb that income is affected by exchange rates. Thus, a general equilibrium framework is applied. The relationship between hysteresis and the income absorption approach is discussed by Back-
hus (1994), arguing that since the basic sunk cost models are partial in nature, no direct link can easily be established between the two.

The first (?) comprehensive analysis of hysteresis and income-absorption is by Blecker (1992). Motivated by the debate following the US trade balance development during the 80s, the paper analyses whether there is a secular declining trend in US competitiveness during this period, or if the development simply was due to unfavorable income elasticities in the US impacting income and absorption differently. The findings support a hypothesis stating that there was a positive secular trend in imports, and hysteresis, in the sense that there was an increased trend in the growth rate of imports, in addition to unfavorable income elasticities. However, the paper also states that the best specification of foreign trade - that is whether one should allow for unequal income elasticities or hysteresis in econometric trade models - is uncertain. The problem of having both different income elasticities and different trends in the various foreign trade components, is illustrated by Blecker (1982) as follows:

Suppose that imports and exports are determined by the following constant elasticity reduced form functions, with log linear time trends representing changing supply conditions:

\[
M = \left( \frac{EP^*}{P} \right)^{\psi} Y^\mu e^{\sigma t} \tag{29}
\]

\[
X = \left( \frac{P}{EP^*} \right)^{\eta} Y^{*\varepsilon} e^{\sigma t} \tag{30}
\]

where \( E \) is the exchange rate, \( P \) the domestic market price, \( Y \) national income, \((*)\) denotes foreign variables, and \( \sigma \) and \( \rho \) are the trends in exports and imports respectively. Combining balanced trade, that is \( PX = EP^* M \), with the expressions for imports and exports, and taking logarithmic derivations with respect to time \( t \) the growth rate of national income equals

\[
y = \left[ \varepsilon y^* + (\sigma - \rho) - (1 + \psi + \eta)(e + p^* - p) \right]/\mu \tag{31}
\]

where lower letters represent growth rates, and by assuming that purchasing power parity holds in the long run \( (e + p^* - p = 0) \), a balance of payment constrained growth rate for domestic income is:32

32Krugman (1989) has denied that the measured income elasticities can be interpreted as an autonomous source of growth, rather than the other way around. See Blecker (1992)
\[ y_B = \left[ \varepsilon y^* + (\sigma - \rho) \right] / \mu \] (32)

Thus, even if the income elasticities are equal, domestic income must grow slower than foreign income \((y_B < y^*)\) in order to maintain balanced trade as long as \((\sigma - \rho) < 0\), that is, as long as the structural trends in the economy are unfavorable. That is, by interpreting hysteresis as structural shifts in the growth rate of imports, it can coexist with unfavorable income elasticities. Therefore, assessing the impact of hysteresis on the income-absorption relationship can be difficult.

Some temporary assessments can be made by applying the reasoning of Ljungquist (1994). Ljungquist argues that when a country faces a positive demand shock equity values at first increase, but that they ultimately will fall below their pre-shock level as the temporary increase in profits, by new attracting competitors, reduces future profitability. Hysteresis in trade is thus accompanied by hysteresis in equity values, with potential implications for both consumption possibilities and the incentives for investments. Christophe (1997) finds empirical support for hysteresis in equities. While the initial increase in equity values at first stimulates consumption, the reduction in long run equity values impacts consumption negatively (and savings positively) over time. The stronger the tendency for consumption smoothing, the smaller the effects on the current account accompanying such wealth effects. Even so, according to the H-M-L condition, for any given level of income, the trade balance initially deteriorates, but improves over time, following the hysteretic effects on equity values, stemming from the temporary depreciation and the accompanying permanent exchange rate appreciation.

If equity values also impact investments, and hence ultimately also income, the current account dynamics becomes even more uncertain. While a long run reduction in equity values is likely to deteriorate investments, the temporary increase might have opposite effects. Even though the former most likely will dominate the latter, the impact of investments on the current account becomes harder to predict. Complicated investment behavior is most likely the reason for why traditional theorems de-emphasize the role of investments when analyzing the current account effect of a devaluation, and as sunk entry cost also makes exports and imports into investments decisions, aggregate investment behavior is now most likely even more complicated, with accompanying effects on the trade balance.

for a wider discussion on the relationship between growth and income elasticities.
A paper focusing on both the effect of savings and investments and the resulting effects on the current account is Lin and Tseng (1993), deriving an intertemporal optimization model with monopolistic competition. The paper shows how a temporary exchange rate shock can yield long run effects even without relying on short run disequilibriums, as in Risager (1988) and Bo-Nielsen (1991). Lin and Tseng allows for a situation with monopolistic competition and increasing returns to scale, where firms have to incur a fixed entry cost in order to export, that is a sunk cost framework. The reasoning of "weak hysteresis" is applied as a devaluation affects the value of the post entry profits, thereby also entry in itself, and the number of exporters, inducing a new steady state number of firms. To accommodate a new steady state with a higher number of firms, consumers immediately reduce intertemporal consumption. The consumption effect reduces the demand for labor, and prevents domestic wages from rising even when new firms increase the competition for labor. Thus, the devaluation lowers the equilibrium real wage relative to the pre-shock level, making savings too small to meet the initial investment boom following the devaluation. The paper shows how a devaluation creates an immediate deterioration, which becomes even more negative over time, when sunk market entry costs are present. That is, hysteresis and current account dynamics is directly related to the nature of industrial organization.

Likewise, the permanent appreciation following the temporary depreciation of Borgersen and Göcke (2007) and the "strong hysteresis" approach, can induce short and long run effects on both income and absorption. And again, while the effect on savings reasonably can be related to intertemporal consumption smoothing, the reasoning on investment is harder, and will always be model specific. Even so, when temporary depreciations induce long term appreciations, and equilibriums are path-dependent, both the structure of investments and savings might be affected, inducing complex trade balance dynamics in the short- as well as in the long run. All in all, when it comes to income-absorption the impact of hysteresis is still not accounted for, neither in the form of "weak hysteresis" nor in the form of "strong hysteresis", and as the contemporary results seems highly partial in nature, the general criticism on the H-M-L condition is by far overcome.
6 Summary and Discussion

The conventional theorems relating exchange rates to international trade flows have been under pressure for the last two or three decades. Today the LOP is in general argued only to be valid in the long run, while the relevance of the J-curve is questioned as such, and the Marshal-Lerner condition - in addition to its lack of empirical support- is criticized for its weak microeconomic foundation. Within the income-absorption approach, the effect on the trade balance of an exchange rate shock, is first of all uncertain, and, second, highly model specific.

Sunk cost hysteresis provides a strong microeconomic foundation for the relationship between trade flows and exchange rates, inducing highly complex relationships at the aggregate level. In the case of heterogenous exporters, the microeconomic non-linearity following sunk costs will in fact be strengthened by aggregation. Even though empirical research provides some promising results matching the sunk cost model to real world issues, it still does not provide satisfactionary answers to all the problems of the conventional trade-exchange rate theorems. Empirical research on sunk cost hysteresis suffers both from methodological issues, as well as from lack of data at plant level regarding export behavior. In this sense, this paper is a progress report. Still, the paper shows how sunk cost hysteresis provides some important insights regarding both pricing of tradable goods as well as the dynamic relationship between exchange rates and trade flows. This insights might in particular prove useful for developing countries, being the ones facing entry costs in international markets most intensely, and experiencing the strongest policy interventions based on the traditional trade-exchange rate theorems.

First of all, the theoretical predictions of the model implies pricing to market, and in the case of "weak hysteresis" different exchange rate pass-through following small and large shocks to the exchange rate. That is, the hysteresis band separates two regimes regarding pricing of tradable goods and exchange rate pass-through. This model structure provides a basis for empirical findings related to having a band around the exchange rate, a band where the degree of mean reverting in export prices is high, and within the band exchange rate shocks only contain temporary effects on prices. Outside the hysteresis band mean reverting is weaker, as entry and exit affect equilibrium prices and the optimal degree of exchange rate pass-through. Thus, the implications of sunk cost resembles the commodity points of Hecksher, as discussed by Obstfeld and Taylor (1997). Even so, allowing for "strong
hysteresis” completely state-dependant pricing processes and exchange rate pass-throughs emerge. There is still pricing to market, and exchange rate pass-through is governed by mean reverting, but as the equilibrium is path-dependent, pass-through will be completely context specific.

Second, the sunk cost model provides theoretical support for econometric analysis which seems to constrain structural stability for the equations determining imports and exports to situations where exchange rate shocks are small. However, the theoretical implications regarding stability differ between "weak hysteresis" and "strong hysteresis". In the case of "weak hysteresis", stability will differ between small and large shocks to the exchange rate, a result directly in accordance with empirical findings. In the case of "strong hysteresis", stability is questioned as such, as equilibriums are path-dependent and shocks have completely idiosyncratic effects. Now, every change in the direction of the exchange rate shock will induce shifts in structural parameters, as long as local extreme values (of the exchange rate) are passed. That is, the concept of structural stability is only valid for a given direction in exchange rates, that is for any depreciation (appreciation) as long as the shock produces a local extreme value. Even though cost heterogeneity is a reasonable assumption for describing firms in developing countries, there is still a need for further empirical analysis regarding the existence of "strong hysteresis". However, no matter whether "weak-" or "strong hysteresis" is assumed, the trade balance dynamics accompanying exchange rate shocks is context specific, and in the case of "strong hysteresis", econometric analysis on structural stability can in fact easily be misleading.

Third, the context specific pricing behavior highlighted above invalidates the J-curve in itself, as the lags of the J-curve are denominated in the dimensions of time. In the case of hysteresis lags are state-dependent, and trade elasticities will vary both according to the size of the exchange rate shock, the initial number of exporters in the industry, the cost heterogeneity of exporters, as well as the persistence of the shock. In the case of homogenous exporters short- and long run elasticities will differ between small and large shocks to the exchange rate, while structural analysis of the elasticity approach can be questioned as such, if exporters are heterogenous. Different export supply responses and differing elasticities following exchange rate shocks, are also argued by Roberts and Tybout (1997) to be one of the most peculiar phenomenons in current international economics.

Fourth, the existence of sunk cost makes importing and exporting investment decisions, and the costs are only incurred if they are assumed to increase
potential future income. Hysteresis in exports and imports can thus induce changes in the long run equilibrium values of both investments, exchange rates, equity values, prices, savings, consumption and income, simultaneously affecting both the income- and the absorption side of the H-M-L condition no matter whether "weak- or strong hysteresis" is assumed. Still, no unconventional analysis regarding the general equilibrium effects exist (?). While awaiting such an analysis, a first initial assessment is that sunk cost hysteresis does not seem to make the trade balance response following exchange rate shocks any less controversial than the conventional H-M-L condition does.

Even though a number of issues remain to be addressed, and others only partially can be discussed within the existing frameworks, sunk entry costs in international markets obviously makes it crucial for developing countries to question the conventional wisdom regarding how exchange rates impact trade flows. As state-dependent relationships between the two seem paramount, context specific analyses should, in opposition to the one size fits all approach of structural adjustment programs, be the basis for policy interventions intended to improve the trade balance. State-dependent processes are often hard to accept for economists. Therefore "strong hysteresis" might be difficult to accept as such, even though sunk entry costs are acknowledged to be present in international markets. The distinction between "weak-" and "strong hysteresis" can however be reduced to an assumption regarding exporter heterogeneity. In the case of heterogenous exporters, aggregation strengthens non-linearities and induce path-dependent equilibriums. Dominated by smaller firms, differing in size as well as production technology, heterogeneity, already claimed important for developed countries, should be even more plausible when it comes to developing countries. In basic state dependency, and context specific relationships between exchange rates and foreign trade volumes might be more pronounced than what is proclaimed in conventional theorems governing the relationship between exchange rates and foreign trade dynamics supporting the structural adjustment programs.

References


42


Appendix

Figure 1: The J-curve

\[ TB_t \]

Time

Figure 2: Microeconomic non-linearity

state of activity of an exporting firm j

exporting

inactive

exit trigger

entry trigger

variable unit costs of firm j

exchange rate (price of FX)

\( q_{t} \)

exit costs

entry costs

\( \beta_j \)

\( \alpha_j \)

Exit for \( q_t < \beta_j \)

Entry for \( q_t > \alpha_j \)

\( j \)

\( \epsilon \)
Figure 3: Macroeconomic non-linearity