(4)

ESTIMATING ELASTICITIES OF DEMAND AND SUPPLY FOR SOUTH AFRICAN MANUFACTURED EXPORTS USING A VECTOR ERROR CORRECTION MODEL

PREPARED FOR: SCHOOL OF ECONOMICS
UNIVERSITY OF CAPE TOWN

PREPARED BY: ALBERTO BEHAR
IN PARTIAL FULFILMENT OF REQUIREMENTS FOR A MASTER OF BUSINESS SCIENCE DEGREE

SUPERVISOR: MR LAWRENCE EDWARDS

19 AUGUST 2002

TABLE OF CONTENTS

ABSTRACT	3
1. INTRODUCTION	4
2. SPECIFICATION ISSUES	7
3. ECONOMETRIC ISSUES	16
4. THE CVAR APPROACH	26
5. DATA ISSUES	34
6. ESTIMATIONS	52
7. CONCLUSION	80
BIBLIOGRAPHY AND DATA SOURCES	82

I would like to thank Professors Nicoli Nattrass and Murray Leibbrandt for their helpful comments. I particularly acknowledge Mr Lawrence Edwards for introducing the topic to me and thank him for his expert guidance throughout the course of this study. I however assume full responsibility for any errors and controversies contained in this dissertation.

Abstract

Elasticities of demand and supply for South African manufactured exports are estimated using the cointegrating vector autoregression / vector error correction model approach in order to address simultaneity and non-stationarity issues. Demand is highly price-elastic, ranging from –3 to –6. The price elasticity of supply is 1. Competitors' prices and world income are an important determinant of demand, but domestic capacity utilization is not an important determinant of export supply.

CHAPTER 1: INTRODUCTION

According to Abbott & De Vita (2002), many trade studies have tried to find the reason why some countries, particularly the newly-industrialised-countries, are successful exporters. The main issue is *"...whether manufactured exports ... are predominantly dependent upon the economic prosperity of [the countries'] trading partners or ... their ability to compete in export markets on the basis of price"* (Abbott & De Vita, 2002:1025).

After receiving the Nobel Prize in 1980, Lewis warned that the slow-down in industrial country growth rates would reduce the pace of development in the rest of the world. This has been subsequently challenged by assertions that developing countries face downward-sloping demand curves and can therefore expand their exports through price competition (Senhadji & Montenegro, 1999). The international debate continues.

South Africa's *Growth, Employment and Redistribution* (GEAR) policy document states that promoting export led growth requires measures designed to lower unit costs and enhance competitiveness (RSA, 1996). GEAR also speaks of the need for a competitive real exchange rate (Bhorat, 1998).

While a policy of pursuing competitive export prices / real exchange rates is certainly more active than one of simply hoping the world economy grows, it may not work. Restructuring to become more competitive has many social costs, so the benefits of such policies must be clear.

Another important consideration is the relationship between exports and the domestic business cycle. One view is that exporters export only when domestic demand is insufficient. This suggests a negative association between exports and growth (Tsikata, 1999). Other studies find a positive relationship (Goldstein & Khan, 1985), thus supporting the view that exports are an exogenous component of Keynesian-style aggregate expenditure.

This study contributes to these debates, but its main aim is simply to derive elasticities of demand and supply for manufactured exports using time series data. These can be used as inputs into other studies, especially in the growing computable general equilibrium model arena.

Following the generally accepted specification in Goldstein & Khan (1985), the underlying model is based on the standard laws of demand supply. However, the choice of which specific variables to use is fairly wide. Chapter 2 discusses this and motivates adding competitors' prices to the established framework as well as representing domestic income separately as potential output and capacity utilization.

There are two standard flaws in other studies. The first flaw is the estimation of a single equation (eg demand or supply) when a system of two equations is appropriate. Unless the other equation is perfectly price elastic (which should not be assumed) an estimate of a single equation produces biased estimates. The second flaw is a failure to account for non-stationary data, which may cause spurious regressions (Gujarati, 1995).

Chapter 3 presents these and other econometric issues in more detail. Chapter 4 proposes a method that addresses both flaws. This study uses the so-called cointegrating vector autoregression (CVAR) approach. A vector error correction model (VECM) explains changes in exports in terms of lagged changes in all the variables in the system and in terms of adjustment to long run equilibrium. The long run equilibrium is governed by a cointegrating regression. The elasticities are contained in this cointegrating relationship (Patterson, 2000).

The CVAR approach requires lots of observations over relatively long periods of time. While the specification of variables is fairly standard, this study runs numerous estimations with different combinations of data sets. The aim is to gauge the robustness of the estimates to different representations of a given variable. A substantial part of the research involved the sourcing, combination and construction of long data sets. This process is described in chapter 5, together with a visual inspection of some possible trends and relationships.

Chapter 6 applies the CVAR approach, finding that export demand is highly price elastic, ranging from –3 to –6, and that competitors' prices are important demand factors. The price elasticity of supply is about 1, but there seems to be no clear relationship between capacity utilization and exports. It is important to highlight the fact that many estimation combinations do not converge or they produce estimates that have signs and magnitudes that are grossly inconsistent with the theoretical model introduced in chapter 2. The findings are however based on complete regressions that are satisfactory in terms of standard demand and supply theory.

Chapter 7 provides a brief summary and interpretation and suggests avenues for extending the study.

This study consults a variety of South African and international sources for guidance and comparison. Goldstein & Khan (1985) has an extensive survey of the issues relating to international studies of price and income elasticities. Chapter 2 draws on this research and complements it with information from South African studies, namely those by Wood (1995), Fallon & Pereira de Silva (1994) and Tsikata (1999).

The South African studies only estimate reduced-form single equations. Bhorat (1998) uses a similar method to the CVAR technique for South Africa, thus dealing with non-stationarity, but only estimates the supply equation.

In an empirical study of Hong Kong's exports, Riedel (1988) attempts to deal with simultaneity issues but does not address non-stationarity. In later studies, Riedel and others debate the relevance of export prices using more sophisticated techniques adapted for simultaneity and non-stationarity. The method Abbott & De Vita (2002) employ most closely resembles the one used in this study.

The key contribution this study makes is the use of the CVAR approach to estimate elasticities of demand and supply for South African manufacturing exports simultaneously. This has not been done before. It gauges robustness by using many combinations of long time series data sets that are not readily available and also introduces additional variables.

CHAPTER 2. SPECIFICATION ISSUES

The models will consist of two equations – one for export demand and one for export supply – solving for two unknowns, South Africa's export price and export volume. Alternative representations of each variable in the two equations are discussed.

2.1 BASIC FRAMEWORK

This model's demand and supply curves for a country's exports are based on the conventional laws of supply and demand. Being an equilibrium model, export demand is assumed to equal export supply (Goldstein & Khan, 1985). Equation 2.1 below sets out the basic framework. The functional form is assumed to be logarithmic, mainly so the coefficients can be interpreted as elasticities.

$$\begin{aligned} X^S &= f(p^e; p^d; y^d) \\ X^d &= g(p^e; p^f; p^c; y^f) \end{aligned} \tag{2,1}$$

X^s is the volume of exports supplied

X^d is the volume of exports demanded

p^e is the export price

p^d is the domestic price

y^d represents measures of income, capacity or capacity utilization in the exporting country

p^f is the foreign price

p^c is competitors' prices

y^f is foreign income

2.1.1 Goods produced for export and goods produced for domestic consumption are modeled as imperfect substitutes.

Producers are seen as having the choice of producing to meet domestic demand or to meet foreign demand. Making products for export and making goods for domestic consumption are therefore production substitutes, which raises two issues. The first is whether exports are a direct stimulus to production, or whether exports are merely the residual after local demand has been satisfied. This important topic is deferred to section 2.4. The second issue is the extent to which exports and local consumption are substitutes. This drives the decision whether to model exports as perfect substitutes or as imperfect substitutes for domestic goods (Goldstein & Khan, 1985).

The perfect substitutes model treats the foreign country's import demand as the excess quantity demanded in that country's local market. Similarly, if there is excess quantity supplied in the domestic market, the surplus is export supply. Estimates of import demand and export supply are thus derived entirely from estimates of foreign demand and supply and domestic demand and supply (Goldstein & Khan, 1985).

In contrast, the imperfect substitutes framework has producers choosing whether to produce for export or for local consumption. This can be interpreted as a choice between selling a given tradable product at home or abroad on the one hand, or substituting between the production of tradable and non-tradable goods on the other hand (Goldstein & Khan, 1985).

This study uses an imperfect substitutes model. According to Goldstein & Khan (1985), most empirical work on price and income elasticities uses this model as well. The imperfect substitutes model applies when the goods in question are heterogeneous to some degree. Manufactures are differentiated to some extent, so this study uses the imperfect substitutes model.

2.1.2 Export supply is a function of the price of exports, the price of production substitutes and a variant of GDP.

Export volumes will rise as exporting becomes relatively more profitable. A higher price for exports raises profitability absolutely, and lower domestic prices promote exports in two ways. First, lower domestic prices mean it is less attractive to sell to domestic consumers and therefore relatively more attractive to export. Second, lower domestic prices entail lower input costs and higher export supply (Goldstein & Khan, 1985).

Goldstein and Khan (1985) do not include a variable related to the exporting country's GDP in their review. Nonetheless, such variables are included in most of the studies discussed in section 2.4. Notably, Goldstein and Khan (1978) themselves include a measure of full-employment income to proxy domestic production capacity, because it indicates an economy's ability to produce exports.

2.1.3 Demand for a country's exports is a function of the price of exports, the price of consumption substitutes and foreign income.

The foreign country's GDP is its budget constraint. Subject to this constraint, the country (consumer) must decide on its consumption bundle. For the foreign country, substitutes come in the form of (i) that foreign country's domestically produced alternatives and (ii) other countries' exports. For example, consumers in the United States must choose between South African exports, US products and, say, Mexican exports.

The demand function therefore includes foreign income and price variables for South African exports, competitors' exports and the foreign country's domestically produces substitutes.

2.1.4 Perfect elasticity of demand or supply must not be assumed.

Some studies assume the exporting country is a price taker, meaning the exporting country has no influence over its export price. The export price is equal to the international price. Bhorat (1998) and Fallon & Pereira de Silva (1994) estimate supply equations only. Bhorat justifies this, saying South Africa is a small open economy and therefore faces a perfectly elastic demand curve for its exports. While this is a plausible argument for homogenous commodities, it is less likely to hold in manufacturing. The uncertainty alone motivates an estimate of the price elasticity of demand.

Alternatively, Riedel (1988) suggests many studies have assumed perfectly elastic supply because the supply equation is hard to model. Chapter 3 will discuss the difficulties with estimating two simultaneous equations. This study addresses these difficulties and is therefore not forced to assume perfect price elasticity on either the demand or supply side.

2.2 PRICE VARIABLES

A correctly specified model should have four different price variables:

i. The price in the country being exported to or a weighted average of countries being exported to, or some international price, henceforth *foreign prices* (p^f)
ii. Competitors' export prices, henceforth *competitors' prices* (p^c)
iii. The price of goods made for domestic consumption, henceforth *domestic prices* (p^d)
iv. The exporting country's export price, henceforth *export prices* (p^e)

This can get confusing, as the prices are often highly correlated, and studies not using all four occasionally use the one as a proxy for another. The export price is the endogenous price variable.

Export demand is a positive function of foreign prices and competitors' prices and a negative function of export prices. Export supply is a positive function of export prices and a negative function of domestic prices.

2.2.1 Homogeneity assumptions are theoretically grounded and can be econometrically convenient

A function is said to be homogenous of degree zero if the same percentage change in all variables has no effect on its value, and is an important element of optimizing behaviour (Varian, 1992). Price homogeneity in a two-good demand function means the sum of the coefficients equals zero (Riedel, 1988). In the export supply function, this assumption would mean an equal percentage rise in domestic and export prices would have no effect.

Export demand contains foreign prices, export prices and competitors' prices, but the third price is often ignored. Attaching estimated coefficients to the three price variables in the demand equation:

$$\beta_1 \log P^e + \beta_2 \log P^f + \beta_3 \log P^c \qquad (2.2)$$

Imposing price homogeneity restrictions means $\beta_1 = -\beta_2 - \beta_3$, so the estimate is effectively:

$$\begin{aligned} &\beta_1 (\log P^e - \log P^f - \log P^c) \\ &= \beta_1 \log\left(\frac{P^e}{P^f P^c}\right) \end{aligned} \qquad (2.3)$$

While there are accepted theoretical grounds for price homogeneity, its use seems to be econometric in most cases. Estimating one coefficient instead of three can save many degrees of freedom in certain estimation procedures and can remedy multicollinearity between price variables (Senhadji & Montenegro, 1999).

Many studies have two price variables, so many researchers perform estimations on the ratio of prices as their variable. While this implicitly assumes homogeneity (Muscatelli, Stevenson & Montagna, 1995), not all studies mention it and fewer actually test for the assumption. Riedel (1988) and Abbott & De Vita (2002) do explicitly test for homogeneity. They list the price variables separately in their theoretical models, before estimating ratios / imposing homogeneity restrictions.

Others specify their theoretical models directly as ratios, even if the variables are not in log form (Goldstein & Khan, 1978; Bhorat, 1988; Fallon & Pereira de Silva, 1994). Tsikata (1999) only has the export price in her export supply model.

2.2.2 It is better to list export price variables separately as their individual significance can be tested and own-price elasticities can be estimated.

Section 2.2.1 presented reasons for imposing homogeneity assumptions and using price ratios. Not imposing a homogeneity assumption has the advantage of allowing one to test the significance of each price variable. While the homogeneity assumption implies no price variables or all price variables are significant, this study adds competitors' prices, so it is particularly important to test for individual coefficients. Homogeneity can always be tested econometrically afterwards.

2.2.3 Effective exchange rates are especially inappropriate ratios because they often over-emphasise trading partners at the expense of competing exporters.

Econometric issues, and assumptions convince the authors in section 2.2.1, except Goldstein and Khan (1978), to estimate single equations instead of separate demand and supply equations. This and data problems cause the variables used in estimation to differ from those in the theoretical models.

Tsikata (1999) uses the real effective exchange rate she specifies in her demand equation, but does not use the export price from her supply equation. Fallon & Pereira de Silva (1994) also use the real effective exchange rate. Bhorat (1998) keeps the relative prices from his supply equation. Wood (1995) uses the deviation of the exchange rate from purchasing power parity and the ratio of domestic prices (not export prices) to trading partners' prices.

Using real exchange rates is clearly justified when the question being asked is the effect of changes in the exchange rate. This study does not ask this question. Furthermore, only trading partners, not competitors, determine effective exchange rates (at least in the Pereira de Silva study; the others do not reveal the construction of their exchange rate variable). Future studies using future exchange rates may be more promising. The IMF is working on new real effective exchange rates that capture competition more accurately (Golub, 2000).

2.2.4While export prices and export volumes are the two separate endogenous variables, trade data is often recorded as a value, which is the product of volume and price.

"*... conventional trade models ... treat import (export) quantities or prices but not their product as the dependent variables. Trade data, however, are oblivious to this theoretical nicety and are most readily available in value terms.*" (Goldstein and Khan, 1985:1054)

Using any price deflator to separate the components is subject to numerous problems. Choosing current or base period weights, changes in quality and changes in relative product weights are problems common to many indices. Cross-country comparability problems are particularly relevant to this study (Goldstein and Khan, 1985).

An export price index based on actual export contracts or transactions is in principle the first choice (Goldstein & Khan, 1985). For industrial countries, such as those studied by Goldstein and Khan (1978), this option is likely to be available. Export price indices are usually not available for developing countries, long time periods and disaggregated data. Goldstein and Khan (1985) list two alternatives.

The first alternative deflator is a unit value index. These are constructed by dividing export values by export volumes. The main drawback in price indices of aggregated goods is that a change in the composition of exports in favour of higher-quality or higher-value goods results in higher unit values (Mahdavi, 2000). Shiells (1991) finds that unit values do not greatly alter import demand elasticity estimates that originally used import price indices. Hanninen & Topinen (1999) and Muscatelli, Srinivasan & Vines (1992) use unit values.

The second alternative is the domestic consumer price index (CPI) or producer price index (PPI). While perhaps more soundly constructed, it suffers from the serious drawback that it contains both tradable and non-tradable goods (Goldstein and Khan, 1985). One of the key elements of the model is relative export and domestic prices. The PPI already incorporates goods destined for export, so it is not a useful proxy. Wood (1995) uses sectoral producer prices.

Tsikata (1999) omits the export price variable from her study and Fallon & Pereira de Silva (1994) use the real effective exchange rate as a proxy. Golub (2000) studies South Africa's price competitiveness, and therefore devotes much space to alternative measures of competitiveness, including many alternative real exchange rate constructions and relative unit labour costs.

2.2.5 Wholesale price indices for tradable goods are the best variable for domestic prices.

The first issue is whether to use producer prices or consumer prices. Golub (2000) lists many drawbacks of consumer prices, notably their sensitivity to price controls and other distortions and their not capturing intermediate goods.

Goldstein & Khan (1985) argue that the index should exclude non-tradable goods, rendering the wholesale price index or GDP deflator sub-optimal. Wood (1995), and Bhorat (1998) use sectoral PPI. Fallon & Pereira de Silva capture all relative prices using the real effective exchange rate. Tsikata (1999) does not have a domestic price variable either. Goldstein and Khan (1978) use wholesale prices.

2.2.6 Foreign countries' import prices should be used instead of export prices or the real effective exchange rate.

Some authors use the foreign countries' export prices as the foreign country price variable (Goldstein & Khan, 1978; Bhorat, 1998). Wood (1995) uses sectoral producer prices in South Africa's most important trading partners. Others incorporate foreign prices by using the real effective exchange rate. As is case for the exporting country, foreign countries' export price indices are a better option when available.

However, foreign countries' import price indices should be used instead. After all, the products a foreign country imports and the domestic country exports are likely to be closer substitutes than both countries' exports. Furthermore, an import price index should track closer domestically produced substitutes for imports than the export price index.

2.2.7 No study reviewed explicitly includes a separate variable for competing exporters, although some real effective exchange rate measures may capture this variable.

Within a certain industry, trade theory predicts that the products a country imports from a variety of sources are distinct in some way from the products it produces domestically. Therefore, the products exported to a country by two or more rival exporters should be closer substitutes for each other than for products produced by the importing country. There is therefore a strong argument for including competitors' exports in the demand equation. No study reviewed explicitly does so.

The real effective exchange rate is usually trade-weighted, so competitors are underrepresented, although some recent constructions may correct for this (Golub, 2000).

2.3 EXPORT QUANTITIES

Because the quantity of exports demanded is restricted to equal the quantity of exports supplied, the same variable appears in both the demand and supply equations when two equations are estimated. Along with export price, this is the other endogenous variable.

Sometimes the data are available only in nominal values. Otherwise, measures of real exports should be used. These can be in a currency unit but at constant prices (Tsikata, 1999). Some studies use indices of unit volume (Goldstein & Khan, 1978). Wood's (1995) variation is to use South Africa's share of world exports while Fallon & Pereira de Silva (1994) use exports divided by gross output. Exports at constant prices are subject to the flaws of price deflators.

2.4 DOMESTIC INCOME OR PRODUCTION CAPACITY

There are two issues. The first is whether the variable should be actual GDP, potential GDP or capacity utilization. The second is the direction of causality between exports and the chosen variable. These separate issues are often confused.

2.4.1 Potential GDP represents the country's ability to produce exports.

The argument for using potential GDP or trend income is that, while exporters may be willing to respond to relative prices, they may not be able to. The higher a country's production capacity, the higher its export supply (Goldstein & Khan, 1985). This can also be represented by time or the capital stock (Muscatelli, Stevenson & Montagna, 1995).

2.4.2 Higher cyclical income or capacity utilization increases the profitability of selling domestically, but exports can be a major contributor to aggregate demand.

Goldstein & Khan (1985) reason that higher domestic demand pressures make selling in the domestic market more profitable, and that for some reason this effect is not completely absorbed by relative prices.

A related argument is the "vent-for-surplus" argument found in Bhorat (1998) for example. If producers cannot sell their products domestically, they will sell them externally. Alternatively, if producers cannot exhaust their production capacity with local demand, they will tilt their production towards foreign markets.

Furthermore, higher capacity utilization means the country's production ability is used up. This affects, for example, producers' access to inputs and hence their ability to increase export output. Again, relative prices may not capture this entirely.

Goldstein and Khan (1985) and the "vent-for-surplus" argument highlight the effect of cyclical income on incentives to export. These arguments suggest that (i) causality runs from domestic income to exports and (ii) higher income leads to lower exports.

However, there is a strong reason to believe that exports and cyclical income are positively correlated and that the causation is reversed. Simple Keynesian models list exports alongside domestic demand as one of the components of aggregate expenditure, where exports are determined by international factors, not domestic demand. The implication is that exports contribute to capacity utilization, not vice versa.

The conflicting arguments above have serious policy implications. A key feature of GEAR is export-led growth (RSA, 1996), not export-the-leftovers growth. Bhorat (1998) argues South African firms should seek export opportunities actively instead of being "residual" (pg 8) exporters. If the latter attitude prevails, mindsets will have to be changed for exports to be a growth driver.

2.4.3 Separate production capacity and capacity utilization variables should be used

Tsikata (1999) uses manufacturing capacity utilization. Fallon & Pereira de Silva (1994) represent cyclical income using the deviation of actual from potential output. They find this is highly correlated with manufacturing capacity utilization. Wood (1995) uses South African capacity utilization relative to that of her major trading partners. Riedel (1988) uses a time variable and Goldstein & Khan (1978) use an index of production capacity. Bhorat (1998) calls his variable an index of production capacity, but uses an index of the physical volume of production. He also includes a trend variable.

Using only one variable can only address one of the two issues. Using actual output, which is a combination of production capacity and capacity utilization, is unlikely to be informative. For example, the "vent-for-surplus" argument may dominate the Keynesian-type argument, but might be overwhelmed by the production capacity effect. The resulting positive coefficient is inconclusive. This motivates using both production potential and capacity utilization.

2.5 FOREIGN INCOME

Higher foreign income means that country consumes more goods, including South African goods. It also means more inputs are needed for its production process. This section presents the merits of distinguishing between cyclical and trend income and alternative proxies for world income.

2.5.1 A foreign country is more likely to import after a rise in cyclical income than trend income.

Section 2.4 built the case for separating trend from cyclical income in the exporting country. Strictly speaking, this applies to the foreign country too. Goldstein and Khan (1985) predict the elasticity of demand for imports is higher for cyclical variations than for changes in trend income and cite studies that support this. None of the studies consulted do so.

2.5.2 To achieve a satisfactory measure of world income, a number of large countries' or trading partners' income can be added.

Some studies avoid aggregation issues by including world imports directly in the model (Tsikata, 1999). Others sum the incomes of a handful of countries, be they the world's largest economies (Wood, 1995; Riedel, 1988) or the exporter's major trading partners (Bhorat, 1998). Goldstein & Khan (1978) sum the incomes of 18 countries, weighting them by their share of the given exporter's exports. Bhorat and Riedel (ibid.) also use some sort of weighting – presumably a trade-weighting.

2.5.3 Incomes are standardized for summation by converting into a common currency or by using purchasing power parity (PPP) GDPs.

When aggregating incomes for many countries, it is important to measure each country's GDP consistently. This is especially important if one sees import demand as derived demand, because one is interested in how many products the importing country makes and therefore how many inputs it demands, not necessarily the value its production. None of the studies reviewed here specifies how they aggregate their data.

Because the countries have different currencies, one option is to take each country's real GDP and convert it into a common currency at the exchange rate prevailing in that period. This has its critics, primarily because exchange rates seldom consist of direct price comparisons, which are necessary for achieving comparable levels of production volumes. Researchers should use purchasing power parity measures instead (Schreyer & Koechlin, 2002).

CHAPTER 3. ECONOMETRIC ISSUES

The econometric issues are introduced here because they are particularly relevant to the trade literature, explain why some of the methods in other studies are flawed or unsuitable, and motivate the use of the cointegrating VAR approach (CVAR). The two most important issues are simultaneity in multiple equations and non-stationarity.

3.1 AGGREGATION, LAGS AND MULTICOLLINEARITY

Using aggregated manufacturing trade data to calculate single elasticity estimates tends to bias them downwards, especially when manufacturing sub-sectors have diverse elasticities. However, disaggregated data is less reliable and the issue of how to combine the estimates has not been resolved (Goldstein & Khan, 1985).

There are good theoretical reasons to include lagged values of variables in a model. Exporters take time to adjust production and importers take time to adjust their habits. The question is not whether to have lags in trade models, but how (Goldstein & Khan, 1985).

Wood (1995) includes lagged versions of some of the explanatory variables in the regression. This model is a kind of distributed lag model. Autoregressive models include lagged values of the dependent variable as an explanatory variable. The biggest danger is (Gujarati, 1995).

occurs when the right-hand-side variables of a regression are linearly related. The key problem is that they may be measured individually insignificant when they are significant (Gujarati, 1995). is likely to arise because many of the variables, especially the price variables, are likely to be highly correlated and because lagged values of the variables are fundamental to the CVAR.

Riedel (1988) includes a lagged dependent variable as an explanatory variable. Fallon & Pereira de Silva (1994) include lagged values of both the dependent and explanatory variables in autoregressive distributed lag models (Gujarati, 1995).

3.2 SIMULTANEITY

There are two equations, export demand and export supply. This section will show why it is generally wrong to estimate only one of the equations and why estimating reduced-form equations is problematic.

3.2.1 Single equation estimates bias true demand and supply elasticities downwards, unless the assumption of perfect demand or supply elasticity is valid.

Orcutt (1950, in Goldstein & Khan, 1985) states that, because quantities and prices are related, single equation estimates of elasticities are biased downward. The following empirical export demand and supply equations, based loosely on Gujarati (1995), explain why.

$$Q^s = \alpha_1 + \beta_{11}P^e + \beta_{12}P^d + \beta_{13}Y^d + \mu_1 \qquad (3.1a)$$
$$Q^d = \alpha_2 + \beta_{21}P^e + \beta_{22}P^f + \beta_{23}P^c + \beta_{24}Y^f + \mu_2 \qquad (3.1b)$$
$$Q^s = Q^d \qquad (3.1c)$$

The precise meanings of the variables (given in chapter 2) are not important. Export price (P^e) appears in two equations. Two additions to chapter 2 are the equilibrium condition (3.1c) and the error terms. Assume that some exogenous change like sanctions reduces export *demand*. This will be reflected in μ_2.

Sanctions change export *supply* via changes in export price. Therefore, export price and export quantity are related. Furthermore, the fall in export price mitigates the negative effect of sanctions on export demand. Therefore, export price and the error term in the demand equation are related. This violates one of the assumptions required for ordinary least squares regression (OLS), resulting in biased estimates (Gujarati, 1995).

Simultaneity bias is not a problem for estimating equation 3.1a over time if the demand curve is perfectly elastic. Sanctions shift the horizontal demand curve down and export supply adjusts via export price. However, β_{21} is zero, so the OLS assumption is not violated. Studies have traditionally focussed on demand elasticities, assuming that supply elasticities are perfectly elastic (Goldstein & Khan, 1978; Senhadji & Montenegro, 1999). Bhorat (1998) does the opposite.

3.2.2 Using reduced form equations corrects for simultaneity but can prevent the extraction of elasticities.

Given that neither demand nor supply can be assumed to be perfectly elastic, there are two options. The first is to convert the structural equations (as in 2.1 or 3.1) into reduced form equations (Goldstein & Khan, 1985). The second option is to use simultaneous equation methods, introduced in 3.4. The reduced form equations have the following structure:

$$Q = \pi_0 + \pi_1 P^f + \pi_2 P^c + \pi_3 Y^f + \pi_4 Y^d + \upsilon \qquad (3.2a)$$

$$P^e = \lambda_0 + \lambda_1 P^f + \lambda_2 P^c + \lambda_3 Y^f + \lambda_4 Y^d + \omega \qquad (3.2b)$$

More specifically, using 3.1c, rewriting 3.1b in terms of P^e and substituting into 3.1a yields:

$$Q = \frac{\alpha_1 \beta_{21}}{\beta_{21} - \beta_{11}} + \frac{\beta_{11}}{\beta_{11} - \beta_{21}} \left(\alpha_2 + \beta_{22} P^f + \beta_{23} P^c + \beta_{24} Y^f + \mu_2 \right) + \frac{\beta_{21}}{\beta_{21} - \beta_{11}} \left(\beta_{13} Y^d + \mu_1 \right) \qquad (3.3a)$$

and therefore

$$P^e = \frac{\alpha_1}{\beta_{21} - \beta_{11}} + \left(\frac{1}{\beta_{11} - \beta_{21}} \right) \left(\alpha_2 + \beta_{22} P^f + \beta_{23} P^c + \beta_{24} Y^f - \beta_{13} Y^d - \mu_1 + \mu_2 \right) \qquad (3.3b)$$

There is no problem in estimating equations 3.3a and 3.3b separately. However, it may not be possible to extract values for elasticities for coefficients like β_{11} and β_{21} in equation 3.1, as discussed next.

When the coefficients of the structural equations (3.1) can be extracted from the reduced-form equations, the equations are said to be "just-identified". If there are too few variables in an equation, the terminology is "underidentified". If there are too many, the coefficients could take on more than one value and the equation is "overidentified" (Gujarati, 1995).

The necessary conditions for identifiability are based on the number of variables in the system that are *not* in a particular equation. For an equation to be just-identified in an *m* equation system, it must exclude *m-1* variables that are in the system. If it excludes more than *m-1* variables, it is over-identified (Gujarati, 1995). As Goldstein & Khan (1985) claim is usually the case, system 3.1 is overidentified, so the reduced-form approach will not work.

Tsikata (1999) uses a single reduced form equation. She specifies the structural equations in terms of p^e, sets them equal to each other and sets export supply equal to export demand to estimate an equation for a single export quantity. She acknowledges the coefficients are not strictly elasticities, but then interprets one of the coefficients as an elasticity. At best, one can say that a given percentage change in, say, capacity utilization led to a given percentage change in export quantities, but one can't say it leads to a given change in export supply or demand.

Wood (1995) directly presents a single equation model for South Africa's share of exports that includes both traditional supply and demand factors. Both authors answer interesting questions with their models, but don't give price elasticities. The aim of this study is to estimate price elasticities. Therefore, simultaneous equation methods for estimating the structural equations must be considered (Goldstein & Khan, 1985).

3.3 NON-STATIONARITY AND COINTEGRATION

This section reviews stationarity and the problems with estimations of non-stationary data. It will describe situations when two non-stationary series are said to be cointegrated, which means estimations can be performed after all.

3.3.1 A stationary time series has a constant mean, a constant variance and constant autocorrelations for a given lag

Figure 3.1 shows both manufacturing export prices and foreign income have risen over time. Both variables may have risen over time for independent reasons but it will seem like they move together and are correlated. The fact that the two variables may appear related when they are not is the main problem with estimations involving non-stationary data.

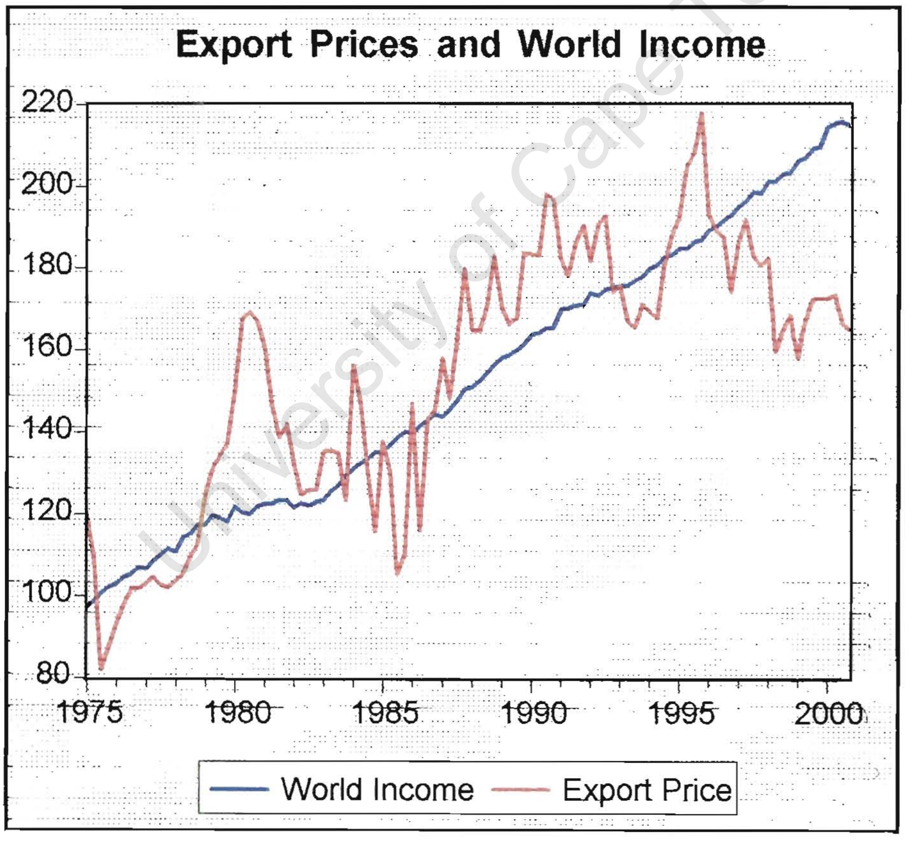

Figure 3.1: The fact that world GDP and export prices have both generally increased over time may suggest a correlation that is not there.

Figure 3.1 clearly shows that the means of both variables are not constant over time. Besides a constant mean, a series must have a constant variance over time to be stationary. Furthermore, autocorrelations for a given lag must be consistent over time for the series to be stationary (Gujarati, 1995).

3.3.2 Random walks are a common form of non-stationary series

Assume export supply is modelled as a function of export supply in the previous period, a constant (α) and the error term (ε).

$$Q_t^s = \alpha + \beta Q_{t-1}^s + \varepsilon \tag{3.4}$$

When $\beta=1$, exports in a given time period are expected to be α units higher than in the previous period, resulting in a non-constant mean. Furthermore, an exogenous shock in a given period would have an infinitely long impact on the quantity of exports supplied (Patterson, 2000).

3.3.3 Non-stationary series have one or more unit roots

Following Patterson (2000), re-arranging equation 3.4 yields

$$Q_t^s - \beta Q_{t-1}^s = \alpha + \varepsilon \tag{3.5}$$

Defining lag operator $L^j Q_t \equiv Q_{t-j}$ results in

$$Q_t^s (1 - \beta L) = \alpha + \varepsilon \tag{3.6}$$

The value of L that makes the left hand side of 3.5 equal to zero is $\frac{1}{\beta}$. If $\beta=1$, then $L=1$.

This is where the term unit root comes from. The case of equation 3.4 is trivial, but a unit root is not as easy to find in equation 3.7.

$$Q_t^s = \alpha + \beta_1 Q_{t-1}^s + \beta_2 Q_{t-2}^s + \varepsilon \tag{3.7}$$

This can be transformed to:

$$Q_t^s(1-\beta_1 L-\beta_2 L^2)=\alpha+\varepsilon \quad (3.8)$$

Depending on the values of β_1 and β_2, there can be zero, one or two unit roots (Patterson, 2000). Finding unit roots gets more complicated when there are more lags, more variables and more than one equation.

A series with *d* unit roots must be differenced *d* times to be stationary. A series that must be differenced *d* times is said to be integrated of order *d* or I(d) (Patterson, 2000).

3.3.4 Difference-stationary variables lead to spurious regressions, unless they are cointegrated

Variables that must be differenced once to be stationary are said to be difference stationary. In a simulation, Patterson (2000) finds that there is a 10% chance of two unrelated non-stationary variables yielding an R^2 greater than 0.6. The t-statistic associated with ordinary least squares regression (OLS) no longer follows the t-distribution, so the critical values used for inference are unreliable (Johnston & DiNardo, 1997). This often results in regressions appearing to be more significant than they are.

However, it is possible for a linear combination of I(1) variables to create a stationary series ς. Assuming export demand and foreign GDP are I(1) and using coefficients in Patterson (2000):

$$0.9Q^d-0.45y^f=\varsigma' \quad (3.9)$$

The values of 0.9 and 0.45 are examples of weights attached to the two variables to yield a stationary series, and are by no means unique. They are a combination that results in a stationary residual. The two I(1) variables are therefore cointegrated. This means that regressions with these variables are reliable (Patterson, 2000).

Normalising on exports and using vector notation gives

$$\begin{pmatrix}1 & -0.5\end{pmatrix}\begin{pmatrix}Q^d \\ y^f\end{pmatrix}=\varsigma \quad (3.10)$$

Vector notation presents what is called a cointegrating vector. When there are more than two variables in the system of equations, it is possible to have more than one co-integrating vector. The Johansen technique determines how many cointegrating vectors there are out of a set of I(1) variables (Patterson, 2000).

The weights are the coefficients in a co-integrating regression. Specifying exports as the dependent variable and moving world income to the right hand side yields:

$$Q_t^d = 0.5y_t^f + \varsigma_t \tag{3.11}$$

Time subscripts are added. The standard single equation cointegration test estimates equation 3.11 using OLS and examines the residuals for stationarity. It is called the OLSEG or OLS-Engle-Granger regression. Cointegration requires the residuals to be I(0) (Patterson, 2000).

3.3.5 Cointegrating relationships are a component of Error Correction Models

A cointegrating relationship is a long-run relationship between the variables that results in stationary residuals. Given this is a long run relationship, there are two influences on changes in quantity in a given period. The first is a change in foreign income in that period. However, only part of the adjustment takes place in the same period, so the second influence is an adjustment to correct for disequilibrium in prior period(s) (Patterson, 2000).

Specifying equation 3.11 in error correction format makes changes in export demand a function of changes in foreign income and of disequilibrium in the previous period(s).

$$\Delta Q_t^d = \theta_1 \Delta y_t^f + \theta_2 (Q_{t-1}^d - 0.5y_{t-1}^f) + \varsigma_t \tag{3.12}$$

θ_1 captures how quickly exports change for a given change in income. If all the adjustment happens in the same period, θ_1 would be 0.5 in this case. θ_2 captures to what extent exports adjust to restore the long run relationship. θ_2 should be negative, because exports will fall if they are higher than equation 3.12 determines (Patterson, 2000).

3.4 ECONOMETRIC APPROACHES TO NON-STATIONARY SIMULTANEOUS EQUATIONS

Provided the equation is (just- or over-) identified, simultaneous equation methods can be used. Of these, systems methods are the most appropriate because they estimate the coefficients in all the structural equations in the system at the same time, imposing all existing restrictions. They can be estimated using full information maximum likelihood techniques (Goldstein & Khan, 1978).

Systems methods presented some computation difficulties (until recently), so limited information methods were used instead. Of the limited information methods, two-stage-least squares (2SLS) is the most appropriate. This removes the problem of simultaneity bias, but it is again often impossible to derive supply-elasticities (Goldstein & Khan, 1985).

Riedel (1988) uses 2SLS. In addition to OLS, Tsikata (1999) applies 2SLS using a substitute for the export price variable. Riedel can extract price elasticities but Tsikata does not. These methods are not explicitly designed to address non-stationarity.

The Autoregressive Integrated Moving Average (ARIMA) or Box-Jenkins approach combines autoregressive (AR) and moving average (MA) processes (Gujarati, 1995). Like AR processes, MA processes are used when there is serial correlation in the error term, but the serial correlation is specified slightly differently in each case (Patterson, 2000).

In this approach, non-stationary variables are differenced until they are stationary. The main drawback is that the information contained in long-run relationships – the very elasticities this study intends estimating – are lost (Gujarati, 1995). The number of lags for each variable and error term is determined using autocorrelation measures and diagnostic checks, but can be highly subjective.

Finally, the ARIMA approach is designed for single equations. While the process can be repeated in each equation, simultaneity problems arise. Finding unit roots and determining the order of integration becomes cumbersome for systems of equations.

The alternative method is the Engle-Granger approach introduced at the end of section 3.3. Estimating a single OLSEG regression would create simultaneity problems. The example used to illustrate such problems was the change in price in the demand equation resulting from a shock to the error term (section 3.2.1). This is closely related to the broader concept of endogeneity (price was endogenous). Endogeneity makes the OLSEG estimator biased and makes the relevant t-statistics unreliable (Patterson, 2000).

For this reason, Phillips and Hansen developed the fully modified OLS (FMOLS) estimator, although it is not guaranteed to be beneficial (Patterson, 2000). In their estimate of a single demand equation Senhadji & Montenegro (1999) use FMOLS.

Muscatelli, Srinivasan & Vines (1992) use ECMs in their two-step procedure. They first use FMOLS on equations for demand and supply to get long run cointegrating relations. Then they use ECMs to explain short run changes in the variables. Using Riedel's data, they get very different results to Riedel (1988) himself. Their article is an example of how confusing the terminology can be. They refer to their approach as a systems method when it appears to be a two-stage approach modified for non-stationarity.

ECMs can be readily applied in genuine systems estimation methods, as is discussed in chapter 4.

CHAPTER 4. THE COINTEGRATING VECTOR AUTOREGRESSION METHOD

This study uses the CVAR approach, which provides a systematic way of dealing with non-stationary variables in a simultaneous equation system.

This chapter describes how the specified model can be estimated as a system of vector autoregressions (VAR). VARs must be reparameterised into vector error correction models (VECMs) to establish the cointegrating relationships among I(1) variables when there are multiple equations. The process for finding these cointegrating vectors and estimating them is complicated by the possible existence of trends and intercepts in the model. Identification issues are crucial for deriving economically meaningful estimates and are discussed along with the concept exogeneity.

In summary, the CVAR approach follows the following procedure. The variables in the model are examined to see what their order of integration is. All variables are estimated as a system in an ECM corresponding to a VAR with a chosen number of lags and subject to restrictions imposed on intercept and trend variables in the model. Amongst the I(1) variables, tests for the number of unit roots and cointegrating relations are performed. These are identified on theoretical grounds to yield interpretable estimates of long run relationships and short run dynamics.

4.1 VECTOR AUTOREGRESSIONS

A vector autoregression specifies each variable in the model as a function of lagged values of itself and lagged values of all the other variables in the system. In other words, each variable is described by an autoregressive model (Patterson, 2000). A VAR is characteristic of the top down approach to econometrics. It is based on the premise that all variables are related to all other variables without a priori assumptions or restrictions (Gujarati, 1995).

Assume a VAR with two lags and three variables. This can be represented as:

$$\begin{aligned}
y_{1,t} &= \mu_1 + \Pi_{11} y_{1,t-1} + \Pi_{12} y_{1,t-2} + \Pi_{13} y_{2,t-1} + \Pi_{14} y_{2,t-2} + \Pi_{15} y_{3,t-1} + \Pi_{16} y_{3,t-2} + \varepsilon_{1,t} \\
y_{2,t} &= \mu_2 + \Pi_{21} y_{1,t-1} + \Pi_{22} y_{1,t-2} + \Pi_{23} y_{2,t-1} + \Pi_{24} y_{2,t-2} + \Pi_{25} y_{3,t-1} + \Pi_{26} y_{3,t-2} + \varepsilon_{2,t} \qquad (4.1) \\
y_{3,t} &= \mu_3 + \Pi_{31} y_{1,t-1} + \Pi_{32} y_{1,t-2} + \Pi_{33} y_{2,t-1} + \Pi_{34} y_{2,t-2} + \Pi_{35} y_{3,t-1} + \Pi_{36} y_{3,t-2} + \varepsilon_{3,t}
\end{aligned}$$

Where $\Pi_{i,j}$ [i=(1...3); j=(1...6)] refers to the estimated coefficient and $y_{k,t-l}$ [k=(1...3); l=(1;2)] refers to the selected variable and the lag length. This can be expressed in matrix form:

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Pi}_1 y_{t-1} + \boldsymbol{\Pi}_2 y_{t-2} + \boldsymbol{\varepsilon}_t \qquad (4.2)$$

where

$\mathbf{y}_t$ is a vector of the m variables , $\boldsymbol{\mu}$ is a vector of m constants, $\boldsymbol{\Pi}_i$ is a matrix of dimension $m \times m$ for time period $t-i$ (there is one for each lag) - and $\boldsymbol{\varepsilon}_t$ is a vector of errors.

4.2 TRANSFORMING A VAR INTO A VECTOR ERROR CORRECTION MODEL

Equation 3.12 is an ECM for a single equation. A VAR can be transformed into a vector error correction model (VECM). The VECM attributes changes in each endogenous I(1) variable to adjustment to a long term cointegrating relationship and to changes in the other variables.

The simple ECM is based on the notion of adjusting to a long run cointegrating relationship, and is therefore intimately tied to the concepts of non-stationarity and cointegration. This will be discussed in a VAR context in section 4.3. Section 4.2 extends the simple ECM to allow for more variables, more lags, and eventually more cointegrating relationships. It builds up to equation 4.11, the starting point in many articles that use VECMs.

In the same way that equation 3.11 was transformed to equation 3.12, systems like 4.1 can be transformed into a VECM. Patterson (2000) gives the following example; the numbers are values for θ_2 as defined in equation 3.12[1]:

$$\Delta y_{1,t} = -1/2\xi_{1,t-1} + 1/4\xi_{2,t-1} + \varepsilon_{1,t} \qquad (4.3a)$$

$$\Delta y_{2,t} = 1/8\xi_{1,t-1} - 5/8\xi_{2,t-1} + \varepsilon_{2,t} \qquad (4.3b)$$

$$\Delta y_{3,t} = 1/4\xi_{1,t-1} + 3/8\xi_{2,t-1} + \varepsilon_{3,t} \qquad (4.3c)$$

Two cointegrating relations are given:

$$\xi_{1,t-1} = y_{1,t} - 1/8 y_{2,t} \qquad (4.3d)$$

$$\xi_{2,t-1} = y_{2,t} - 1/4 y_{3,t} \qquad (4.3e)$$

[1] While it would be nice to use variables from the trade model, the signs on the coefficients used by Patterson for illustration don't fit the economic theory discussed in chapter 2.

There is no $\xi_{3,t-1}$ – with three variables, it is possible to have up to two cointegrating relationships (Patterson, 2000). How these cointegrating relationships are established is the subject of sections 4.3-4.6. Patterson (2000) continues:

$$\begin{pmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \\ \Delta y_{3,t} \end{pmatrix} = \begin{pmatrix} -1/2 & 1/4 \\ 1/8 & -5/8 \\ 1/4 & 3/8 \end{pmatrix} \begin{pmatrix} \xi_{1,t-1} \\ \xi_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix} \tag{4.4}$$

$$= \begin{pmatrix} -1/2 & 1/4 \\ 1/8 & -5/8 \\ 1/4 & 3/8 \end{pmatrix} \begin{pmatrix} 1 & -1/8 & 0 \\ 0 & 1 & -1/4 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix} \tag{4.5}$$

In matrix notation:

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha\beta}'\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{4.6}$$

Matrix α contains the coefficients that show the speed of adjustment to long run equilibrium. β' is the matrix of long run coefficients (Patterson, 2000); each row represents a cointegrating relation. Thus, the VECM captures adjustment to equilibrium and the cointegrating relations governing that equilibrium.

Determining the number of cointegrating relations, which is discussed later, requires defining

$$\boldsymbol{\pi} = \boldsymbol{\alpha\beta}' \tag{4.7}$$

It is important to adjust the VECM for cases of more than one lag. Just as the change in y_i at time t is defined in terms of changes in all the variables in t-1, changes in y_i in t-1 are defined in terms of changes in all the variables in t-2. Therefore, a second lag is captured in the following functional form (Patterson, 2000):

$$\Delta \mathbf{y}_t = \boldsymbol{\Pi}\mathbf{y}_{t-1} + \boldsymbol{\Gamma}\Delta\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{4.8}$$

Equation 4.8 is a VECM based on VAR equation 4.2. A ρ^{th} order VAR

$$\mathbf{y}_t = \boldsymbol{\Pi}_1\mathbf{y}_{t-1} + \boldsymbol{\Pi}_2\mathbf{y}_{t-2} + \ldots\boldsymbol{\Pi}_\rho\mathbf{y}_{t-\rho} + \boldsymbol{\varepsilon}_t \tag{4.9}$$

corresponds to VECM

$$\Delta \mathbf{y}_t = \mathbf{\Pi y}_{t-1} + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{y}_{t-2} + \ldots \mathbf{\Gamma}_{p-1} \Delta \mathbf{y}_{t-(p-1)} + \boldsymbol{\varepsilon}_t \tag{4.10}$$

Some variables might be I(0) exogenous variables. For this reason, ψD_t is added to the VAR, resulting in the complete VECM (Pesaran & Pesaran, 1997):

$$\Delta \mathbf{y}_t = \mathbf{\Pi y}_{t-1} + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{y}_{t-2} + \ldots \mathbf{\Gamma}_{p-1} \Delta \mathbf{y}_{t-(p-1)} + \psi D_t + \boldsymbol{\varepsilon}_t \tag{4.11}$$

Equation 4.11 is the canonical starting point for many articles that use VECMs.

Finally, Patterson (2000) introduces a companion matrix, which is necessary for determining the number of cointegrating relationships when there is more than one lag.

$$\mathbf{C} = \begin{pmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 & . & \mathbf{\Pi}_{p-1} & \mathbf{\Pi}_p \\ \mathbf{I} & 0 & . & 0 & 0 \\ 0 & \mathbf{I} & . & 0 & 0 \\ 0 & 0 & . & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & . & \mathbf{I} & 0 \end{pmatrix} \tag{4.12}$$

4.3 NON-STATIONARITY, UNIT ROOTS AND COINTEGRATING VECTORS

In a single equation, a unit root in the series implies non-stationarity and a spurious regression, unless the series are cointegrated. This applies to a system of equations, but the process of finding unit roots and establishing cointegration is more complex.

Patterson (2000) extends the lag operator of section 3.3.3 to multiple variables:

$$\mathbf{A(L)} = \mathbf{I} - \mathbf{\Pi}_1 \mathbf{L} \tag{4.13}$$

The eigenvalues are the values for ω that solve the reverse characteristic polynominal

$$\left| \mathbf{I} - \omega \mathbf{\Pi}_1 \right| = 0 \tag{4.14}$$

where $| \ |$ denotes determinant (ibid.). While the link between 4.13 and 4.14 may be clearer, it is customary in multivariate models to define the eigenvalues as the roots υ that solve (ibid.)

$$|\mathbf{\Pi}_1 - \upsilon \mathbf{I}| = 0 \tag{4.15}$$

$\omega = 1/\upsilon$. A system is stable if all values of υ have modulus less than 1. The number of unit roots is the number of eigenvalues equal to one (ibid.). Extending equation 4.15 to multiple lags employs the companion matrix.

$$|\mathbf{C} - \upsilon \mathbf{I}| = 0 \tag{4.16}$$

The number of non-zero eigenvalues in the companion matrix is the number of unit roots in the system (ibid.).

Patterson (2000) shows how unit roots reduce the rank of $\mathbf{\Pi}$. The rank (r) of a matrix is the number of linearly independent rows (equations) it contains (Chiang, 1984). Analogously, there is a link between r and the number of linearly independent cointegrating vectors (Johnston & DiNardo, 1997).

If $1 \le r \le k$, where k is the number of variables and hence equations in the system, then r equals the number of cointegrating vectors and is the cointegrating rank. The best way to find r in this context is to determine the number of non-zero eigenvalues (Patterson, 2000).

4.4 TRENDS, CONSTANTS AND EXOGENEITY

The nature of trends and constants/intercepts in the model affects the critical values for selecting the number of cointegrating relationships (Pesaran & Pesaran, 1997). A constant or trend in the model can come from the cointegrating relations and/or trends in the data. Most sets of critical values allow for five combinations of trend and constant (ibid.); this section addresses the most important choices for this study.

A constant in the VAR can arise from two sources. The first is a constant in the cointegrating regression. Patterson (2000) agrees that, as for almost all linear regressions, most applications should allow for a constant in the cointegrating regression. The second source is a linear trend in the data, as explained using equation 4.17.

$$X_t = a + bX_{t-1} \tag{4.17}$$

If a and b are positive, X will rise over time. A visual inspection of the series in this export model is likely to reveal such trends. This justifies specifying a separate constant in the VECM (Patterson, 2000).

Like the constant, a time variable in the VAR has two possible origins. The first is a time variable in the cointegrating relationship. This is justified when the relationship between the variables changes over time (Patterson, 2000). It is quite possible, for example, for the relationship involving export demand and foreign income to have changed as world trade relative to world GDP has increased consistently over time. Patterson recommends plotting the cointegrating relations (for example, ξ_1 and ξ_2 from system 4.3) over time and seeing whether they are constant or need a time variable.

The second source is a quadratic trend in the data if the model contains unit roots (Abbott & De Vita, 2002). Visual inspections show this is unlikely. Therefore, a time variable will not be specified separately in the VECM.

LR tests for the exclusion of variables can also be applied to trends and constants (Pesaran & Pesaran, 1997).

The general form of the VECM, which allows for all 5 trend/intercept combinations is[2]:

$$\Delta \mathbf{y}_t = \mathbf{a}_0 + \mathbf{a}_1 \mathbf{t} + \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{y}_{t-2} + \dots \mathbf{\Gamma}_{p-1} \Delta \mathbf{y}_{t-(p-1)} + \boldsymbol{\psi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t \quad (4.18)$$

The additions are the intercept ($\mathbf{a}_0$) and trend variable ($\mathbf{a}_1$). The 5 cases differ according to whether each of these (Pesaran & Pesaran, 1997)

- is restricted to zero
- is restricted to the values given by the cointegrating relations
- is also allowed to consist of linear or quadratic trends in the data.

This study will leave $\mathbf{a}_0$ unrestricted. $\mathbf{a}_1$ will either be restricted to zero or to the value of the time coefficients in the cointegrating vectors.

Some variables may possibly be treated as exogenous. Exogeneity is anathema to a VAR, but can save many degrees of freedom (Patterson, 2000). In some forms, saying x is exogenous is equivalent to saying that the other variables do not Granger-cause x (Pesaran & Pesaran, 1997). The VECM only estimates equations for the endogenous variables, but the equations are expressed in terms of both the endogenous and exogenous variables.

---

[2] This equation combines those formulated by Pesaran & Pesaran (1997) and Patterson (2000).

4.6 IDENTIFICATION

This section applies identification issues to VECMs.

The two cointegrating relations from system 4.3 are:

$$\xi_{1,t-1} = y_{1,t} - 1/8 y_{2,t} \tag{4.19a}$$

$$\xi_{2,t-1} = y_{2,t} - 1/4 y_{3,t} \tag{4.19b}$$

The coefficients are restricted to zero on $y_{3,t}$ in 4.19a and on $y_{1,t}$ in 4.19b. It is necessary to have some variables in the system that are not in that equation for that equation to be identified. Alternatively, restrictions can set coefficients on two variables equal to each other (Patterson, 2000). For example, 4.19a could have been restricted as (ibid.):

$$\xi_{1,t-1} = y_{1,t} - 1/8(y_{2,t} - y_{3,t}) \tag{4.20}$$

In addition to the common normalisations, there must be at least r-1 restrictions applied separately to each of the r cointegrating vectors (Pesaran & Pesaran, 1997). The generic form is $\mathbf{R}_i\boldsymbol{\beta}_i = 0$. Each column in $\boldsymbol{\beta}$ represents a cointegrating vector. Each row in $\mathbf{R}$ represents a restriction (Patterson, 2000).

Theory guides the identification process, which usually results in there being more than the required number of restrictions. Linear homogeneity assumptions might motivate an equality restriction. Alternatively, two cointegrating vectors could be assumed to be demand and supply equations (Johnston & DiNardo, 1997).

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \beta_Q \\ \beta_{Pe} \\ \beta_{Pd} \\ \beta_{Pf} \\ \beta_{Pc} \\ \beta_{yd} \\ \beta_{yf} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{4.21a}$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_Q \\ \beta_{Pe} \\ \beta_{Pd} \\ \beta_{Pf} \\ \beta_{Pc} \\ \beta_{yd} \\ \beta_{yf} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (4.21b)$$

4.21a sets domestic prices and domestic income equal to zero, as these are not demand factors. 4.21b sets foreign prices, competitors' prices and foreign income equal to zero, because they are not supply factors. In addition to the restrictions in system 4.21, export quantity or export price is usually set equal to one to normalise each vector (Patterson, 2000; Abbott & De Vita, 2002).

Muscatelli, Srinivasan & Vines (1995) cite Riedel's assertion that estimated elasticities depend critically on normalisation. Normalising the demand equation on price and the supply equation on quantity yields infinite price elasticities while normalising the demand equation on quantity and the supply equation on price yields low price elasticities and high income elasticities, as found in single equation estimates of demand.

Muscatelli et al (ibid.) assert this problem is alleviated if full systems techniques including ECMs are used. Athukorala & Riedel (1994) rebuke this assertion, claiming the problem arises from the use of the OLS-based Phillips-Hansen estimator. They state ML techniques do solve the normalisation problem. The Johansen approach is an ML estimator of a full system, so normalising on price or quantity will not affect the results.

CHAPTER 5. DATA ISSUES

Chapter 2 introduced the variables to be used in the model. Using trade data is like eating a Vienna sausage: the less you know about its contents, the better. This is probably why so few papers (not only brief journal articles) discuss data problems and even fewer do anything about them (Dezhbakhsh, 2002). Too many authors are less transparent than they should be.

This chapter goes into more detail than chapter 2, explaining the variables used, their sources and their construction. It also shows some of the actual series. One of the aims of this study is to investigate the sensitivity of the results to different data series. The data is therefore analysed for consistency between the various constructions. The visual inspections are also used to look for possible relationships between the variables that are consistent with the theoretical hypotheses. As a framework, equation 2.1 is reproduced.

$$\begin{aligned} X^s &= f(p^e; p^d; y^d) \\ X^d &= g(p^e; p^f; p^c; y^f) \end{aligned} \qquad (5.1)$$

X^s is the volume of exports supplied
X^d is the volume of exports demanded
p^e is the export price
p^d is the domestic price
y^d represents measures of income, production capacity or capacity utilization in the exporting country
p^f is the foreign price
p^c is competitors' prices
y^f is foreign income

5.1 LENGTH OF DATA SERIES

Vector autoregressions require many observations (Patterson, 2000). They also require long time spans to allow sufficient opportunity for enough shocks to take place and for adjustment to those shocks to occur. Scarce observations might force researchers to drop variables and/or use shorter lags than otherwise.

Muscatelli, Srinivasan & Vines (1992) have quarterly data from 1972 to 1984. Hanninen & Toppinen (1999) use 60 quarterly observations. They include many price variables and use the homogeneity assumption to save degrees of freedom. Bhorat (1998) estimates monthly data from 1995-2000. While this may be a similar number of observations, the time span is insufficient. Six years is nowhere near long enough for the CVAR to estimate long-run relationships.

Current databases do not go back far enough, so some data were captured manually from various printed sources. Long time series are prone to definitional adjustments and inconsistencies. In some cases, data from various sources was merged. In others, proxies were necessary. This is a serious drawback. It is no good developing ever-sophisticated econometric techniques if the data quality is poor (Dezhbakhsh, 2002) – the garbage-in-garbage-out principle still applies. Data quality and consistency motivated Bhorat's (1998) decision to use a short data series.

This study uses quarterly data from 1975-2000 and annual data from 1961-2000. Given the CVAR's need for observations and the relatively large number of variables this study employs, the advantages of longer time series outweigh the disadvantages. 100 quarterly observations provide enough freedom to use many variables and lags if necessary and 25 years is a long-enough time span. The 40 annual observations are the bare minimum, but the time span is nice and long.

In the data capturing and merging process, all indices are set equal to 100 in the first year (1975 or 1961). The data were thoroughly inspected for nonsensical observations and obvious cases of the source's being wrong were corrected. Sometimes, later versions of a given source have a different number for a given observation than the earlier versions. The later version was used.

5.2 EXPORT VOLUME

Export volume is the variable used for export quantity demanded and for export quantity supplied.

5.2.1 There are many reasons to be concerned about South African trade data

Wood (1995) does an in-depth investigation into South African trade data: Much trade data covers the South African Customs Union (SACU), so exports from South Africa to the rest of SACU are excluded and exports by the rest of SACU to the world are included. This tends to understate exports. According to various issues of Statistics South Africa's Quarterly

Review, export volume statistics are fortunately a rare exception; they are from South Africa only.

Wood (1995) continues: Exporters may have the incentive to mis-invoice their exports, either to qualify for government subsidies or for exchange control reasons. Authorities have also played their part by listing many goods as unclassified. Wood suspects these may be arms sales. There is no way of knowing what portion of unclassified materials is manufactures. Furthermore, Rustomjee (1992, cited in Wood, 1995) argues many products are incorrectly classified as manufactures because they are simple mineral beneficiation.

5.2.2 Export volumes combine data from the TIPS database and from various Statistics South Africa publications

The trade and industrial policy secretariat (TIPS) has Standard Industrial Classification (SIC) data for manufacturing sub-sectors from 1988 in quarterly format and from 1970 in annual format. TIPS's source is the Department of Customs and Excise. The data is in real 1995 Rand values, derived using producer prices, consumer prices and national accounting deflators. The 1995 Rand values are added to arrive at total manufacturing and then converted into an index. The only source of imperfection is that the sub-sectoral data were already deseasonalised.

Quarterly data from 1975 to 1995 was taken from various issues of the *Quarterly Bulletin of Statistics*, published by Statistics South Africa. Annual data from 1961 to 1995 was taken from *Statistics Yearbooks*, also published by Statistics South Africa. Statistics South Africa also used Department of Customs and Excise data. The series were inexplicably discontinued in 1996. There were base and presumably definitional changes in 1985 and 1988. Statistics South Africa also changed from the International Standard Industrial Classification to the Standard Industrial Classification in 1988. Given the aggregate nature of the data, this is unlikely to be important.

There is an eight-year period of overlap for the quarterly data and a 25-year overlap for the annual data. The values are inexplicably greater in the data sourced from Statistics South Africa over this period. Therefore, two series are used. Both use combinations of TIPS and Statistics South Africa data. One uses TIPS data (henceforth TIPS series) from 1988 onwards for quarterly data and from 1970 onwards for annual data, using Statistics South Africa data for the rest, while the other uses TIPS data from 1996 onwards only (henceforth SSA data).

Attempts were made at extrapolating the TIPS data backwards, or, alternatively, the SSA data forwards. The period of overlap was carefully studied for a ratio or linear relationship between the data series. No such relationship was apparent, ruling out an adjustment to one of the series. Using two series is the best option. The annual series are compared in figure 5.1.

5.2.2 Export volumes have grown significantly since 1961 – almost three-fold since 1990.

The blue and red series overlap before 1970 and after 1996 because the pre-1970 data is Statistics South Africa data in both cases and the post-1996 data is TIPS data in both cases. The intermediate years are when both sources offered alternative data series. The blue line shows the SSA observations are higher than the TIPS observations.

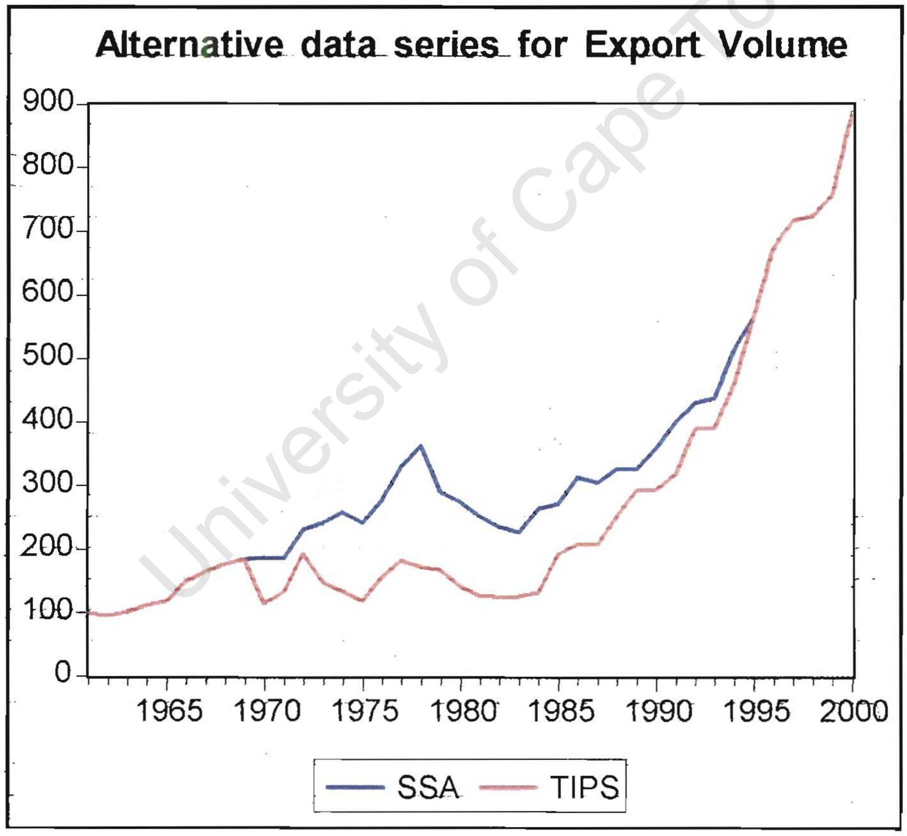

Figure 5.1: The observations based on Statistics South Africa data are higher than those in the TIPS data over the period where comparisons are available. The blue line is covered by the red line before 1970 and after 1996 as the two series are identical over these periods.

Exports have grown more or less consistently since 1961, almost doubling since 1994 and tripling since 1990. Table 5.1 presents TIPS values for manufactured exports in billions of 1995 Rands, showing much of the rise occurred from 1994 to 1996.

Year	Rbn 1995
1988	29
1994	45
1995	58
1996	68
2000	87

Table 5.1: Real Manufactured Exports

The absolute rise in the late 1990s in particular will be compared to changes in other variables. A graph in natural logarithms reveals a perhaps more surprising picture. Figure 5.2 shows that exports have been growing at more or less the same rate since the mid 1980s; export growth did not only start accelerating in the late 1990s.

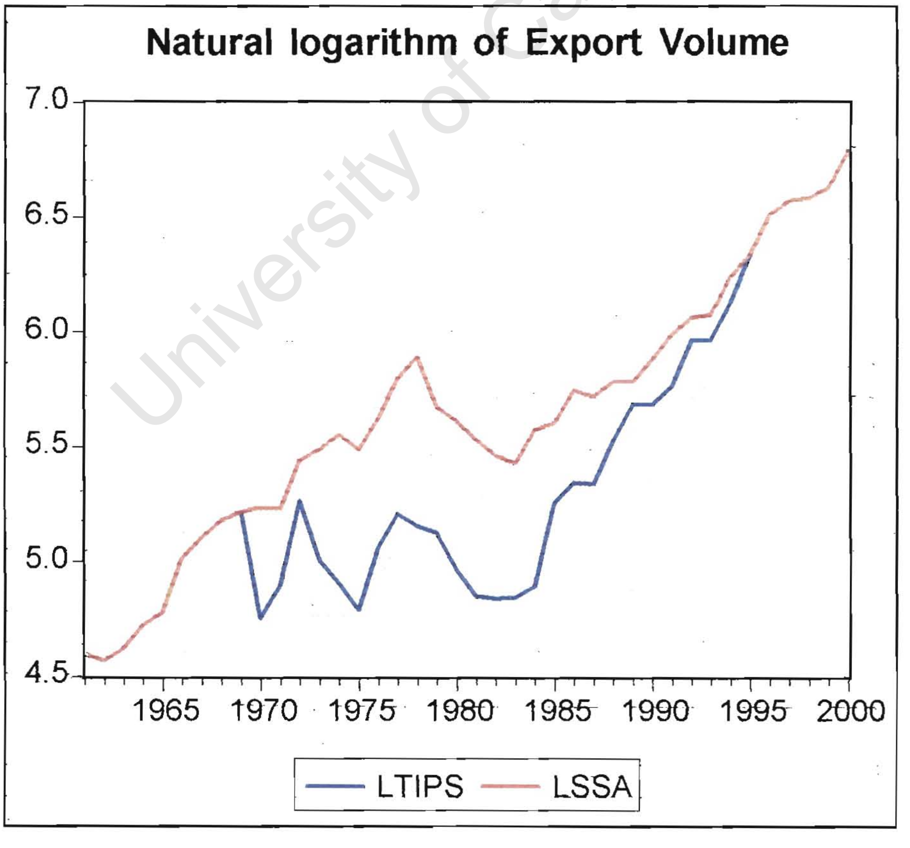

Figure 5.2: Exports have been growing at a consistent rate since before 1985.

5.3 EXPORT PRICES

Export prices are the prices South African producers get for their exports, and should be positively correlated with export supply. They are also the prices foreigners pay for South African products, and should be negatively correlated with export demand. There are two variables – export unit values and a producer price index for exports. Each of these in turn has alternative constructions.

5.3.1 All prices are converted to US Dollars

For consistency between all forms of export prices, foreign prices and competitors' prices, these must be stated in a common currency, chosen to be US Dollars. For consistency between domestic prices and export prices, the former are also converted to US Dollars. Exchange rates were sourced at the International Monetary Fund's (IMF) International Financial Statistics database.

5.3.2 Export unit values combine SIC data from the TIPS database and from various Statistics South Africa publications

Chapter 2 introduced the difference between export unit values and export price indices. The former are derived by dividing export value by export volume while the latter are based on direct measures of prices of exported goods. The estimation section refers to both forms of data as export price.

Export unit value data was taken from exactly the same sources as export volumes. As for volumes, there are disparities between the sources. The source of the disparities in both volumes and unit values appears to be a different deflator. While volumes are higher in the SSA data, prices are lower in the SSA data. Any regressions that use SSA volumes will use SSA unit values and the same applies for TIPS data.

5.3.3 Two export price indices are derived using price indices for total production and for South African consumption

Goldstein & Khan (1985) assert export price indices are superior to unit values, but Shiells (1991) suggests the difference is not important. This study will check for systematic differences in the estimates.

Statistics South Africa only started producing export price indices in 1995, so they have to be derived using price data for total South African Production (henceforth all goods) and for South African production for domestic consumption (henceforth domestic goods). The data are taken from Statistics South Africa's statistical releases of price indices. The construction is based on the following formula.

$$T = \alpha D + (1-\alpha)E \tag{5.1}$$

PPI for all goods is a weighted average of price indices for domestic goods (D) and for exported goods (E). α is the share of production that is consumed domestically. As a result, the export price index is

$$E = T - \frac{\alpha D}{1-\alpha} \tag{5.2}$$

Finding values for α is not straightforward. While manufacturing's share of domestic goods and of all goods is available, it was possible to calculate manufacturing's share of exports directly only from very recently. This is why two alternative measures for export PPI were constructed.

For the first construction, data on manufacturing's share of all goods $\left(\frac{M^T}{T}\right)$ and on manufacturing's share of domestic goods $\left(\frac{M^D}{D}\right)$ are available. To derive the share of manufactures that is consumed domestically $\left(\frac{M^D}{M^T}\right)$ requires multiplication by the share of all South African output produced for domestic consumption. This is an imperfect measure, as this share is not the same for all sectors of the economy. This data was derived using the share of exports of goods and non-factor services to GDP, supplied by the Reserve Bank.

The second construction uses the recent weightings given by Statistics South Africa, but for the entire series. The drawback is this doesn't allow for changes in the weighting of exports. For the other variables, used in both calculations, Statistics South Africa only sees it fit to change the weighting occasionally, so this simpler measure is not as inferior as it may seem at first.

5.3.4 The export PPI measures are practically identical, so only one is used, but they differ from the unit values.

Figure 5.3 shows that the first construction, export PPI with the varying weightings for domestic production (PPIEXPV), is practically the same as the second construction, export PPI with a fixed weighting for domestic production (PPIEXPF). The quarterly data are slightly different, but only from the 1990s onwards, when exports increased significantly. Therefore, only PPIEXPV is used.

The export PPI and unit values differ, although they tend to converge towards the end of the time period. The TIPS measure is especially different.

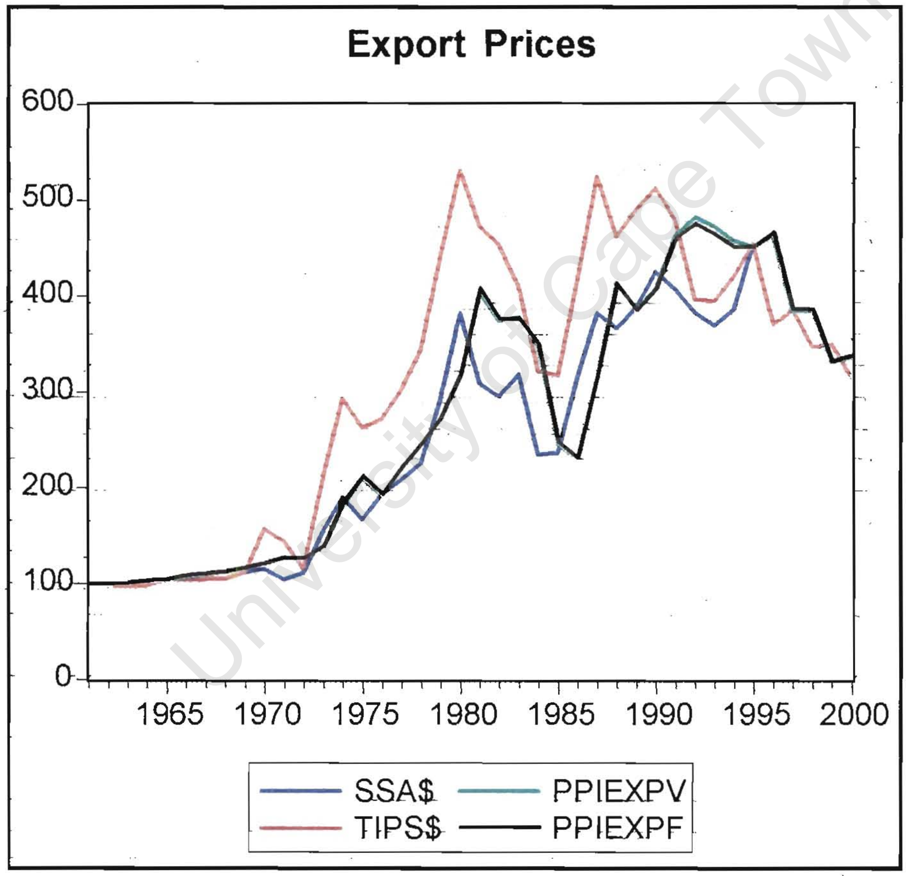

Figure 5.3: The PPI measures are practically identical and converge with the unit value series towards the end of the sample period.

5.3.5 Export prices rose sharply in the 1970s and fell in the 1990s

The key features of the data are the three-fold rise in export prices (four-fold according to TIPS data) in the 1970s, two sharp spikes in TIPS prices in 1980 and 1987, and a fall in prices from around 1990, depending on the data series. Prices have fallen by up to one third since 1995. Importantly, this latest fall shows the depreciation of the Rand has not been completely offset by rising prices. Falling export prices may have led to higher demand for exports over the last decade or so, but other prices must also be considered.

5.4 OTHER PRICES

Export prices must be compared to other prices to determine their relative attractiveness for domestic producers and foreign consumers.

5.4.1 Domestic prices are highly correlated with export prices

The domestic price index is the series used in the construction of the export price indices, as explained in section 5.3.3. Some of the correlation visible in figure 5.4 is attributable to this. Apart from the early 1990s, variable weight export prices move almost identically to domestic prices. The starting values are artificially set equal to each other, so the prices are not necessarily at the same level. The correlation between the price indices is closer than between domestic prices and unit values.

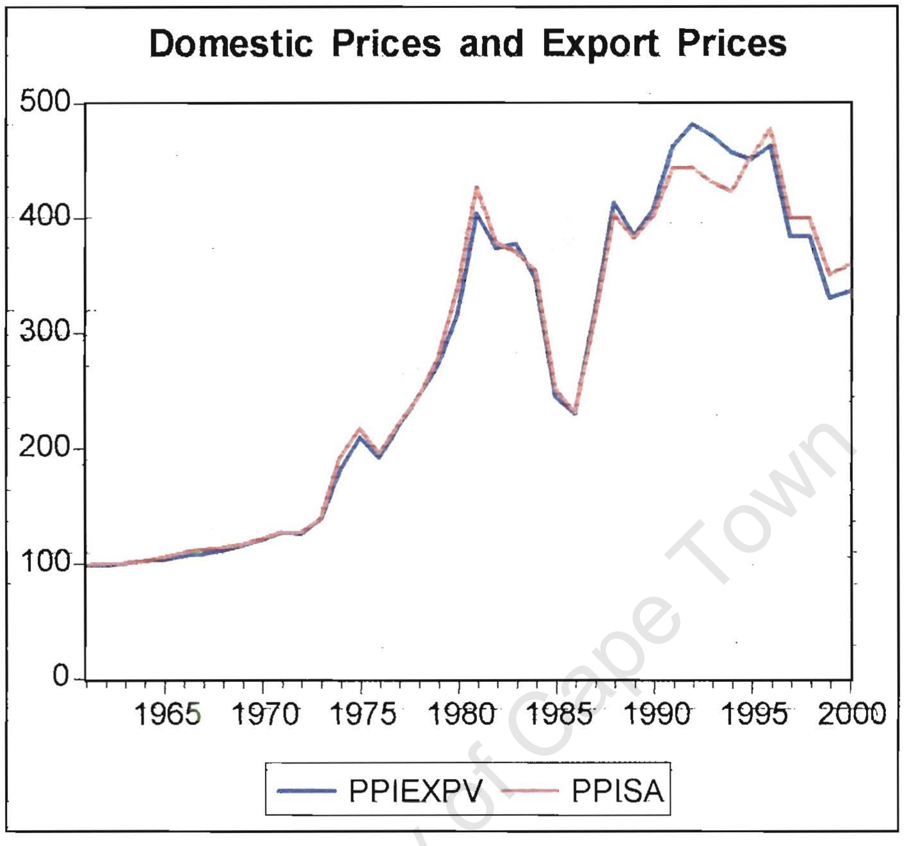

Figure 5.4: The PPI for domestic goods and export prices are almost perfectly correlated

5.4.2 Foreign prices are import-weighted averages of major trading partners' manufacturing import prices for quarterly data and manufacturing wholesale prices for annual data.

For quarterly data, manufacturing import price indices for the Untied States, United Kingdom, Germany and Japan are used. These countries were South Africa's four largest total export destinations throughout the 1990s, representing R 89 million out of R 226 billion in 2000 (Absa, 2001). The price sourced from the OECD.

There are several aggregation problems. The OECD data is in index form with a common base of 1995=100. This means any average is not a true average. In summary, aggregating such data means that absolute price changes in countries with low actual prices are overweighted.

The data are weighted by each country's import volumes. The OECD also provides volumes, but they were hard to interpet. This is because all prices indices have a common base, yet OECD export values are still the product of the price and volume data. After failing to find alternative volume measures, current values (converted to US Dollars) were used instead. The data are however still weighted by real exports, as explained next.

With satisfactory volume data, the weighting formula for each period in a four-country case would be:

$$P^f = \frac{P^1 I^1 + P^2 I^2 + P^3 I^3 + P^4 I^4}{I^1 + I^2 + I^3 + I^4} \qquad (5.3)$$

P^f is the foreign price index, P^i is the price level in country i and I^i is the value of real imports in country i. If only current values are available, P^i can be used to deflate the series:

$$I^i = \frac{N^i}{P^i} \qquad (5.4)$$

N^i is the current value of exports in country i. This means equation 5.3 can be expressed in terms of current values.

$$P^f = \left(\frac{N^1 + N^2 + N^3 + N^4}{N^1P^2P^3P^4 + N^2P^1P^3P^4 + N^3P^1P^2P^4 + N^4P^1P^2P^3} \right) \times \left(P^1P^2P^3P^4 \right) \qquad (5.5)$$

Another drawback is that the weighting is by total imports, not manufacturing imports.

For annual series, the OECD data is not available. The United Kingdom and United States have data on manufacturing prices. Germany and Japan only provide total wholesale prices, and do so from a bit later than 1961. Series with all four countries and with only the UK and US had a correlation coefficient of 0.99, which is high even if the series are not cointegrated. This motivates only using US and UK data for the index. It was constructed using the same aggregation process as the quarterly data.

5.4.3 Competitors' prices are export-weighted averages of three competitors' export prices for quarterly data and two competitors' producer prices for annual data.

The competitors chosen for quarterly data are Mexico, Hungary and South Korea. The main reason was data availability, but the three countries conveniently represent competitors that are close to export markets in North America, Eastern Europe and South East Asia. The data come from the same sources as foreign price data and are subject to the same aggregation procedure.

Producer prices are used for annual data, but Hungarian data is not available from 1961. The correlation coefficient between series including and excluding Hungary is 0.98, so only South Korean and Mexican data are used.

5.4.4 After being persistently higher for almost 20 years, export prices have moved into line with competitors' prices.

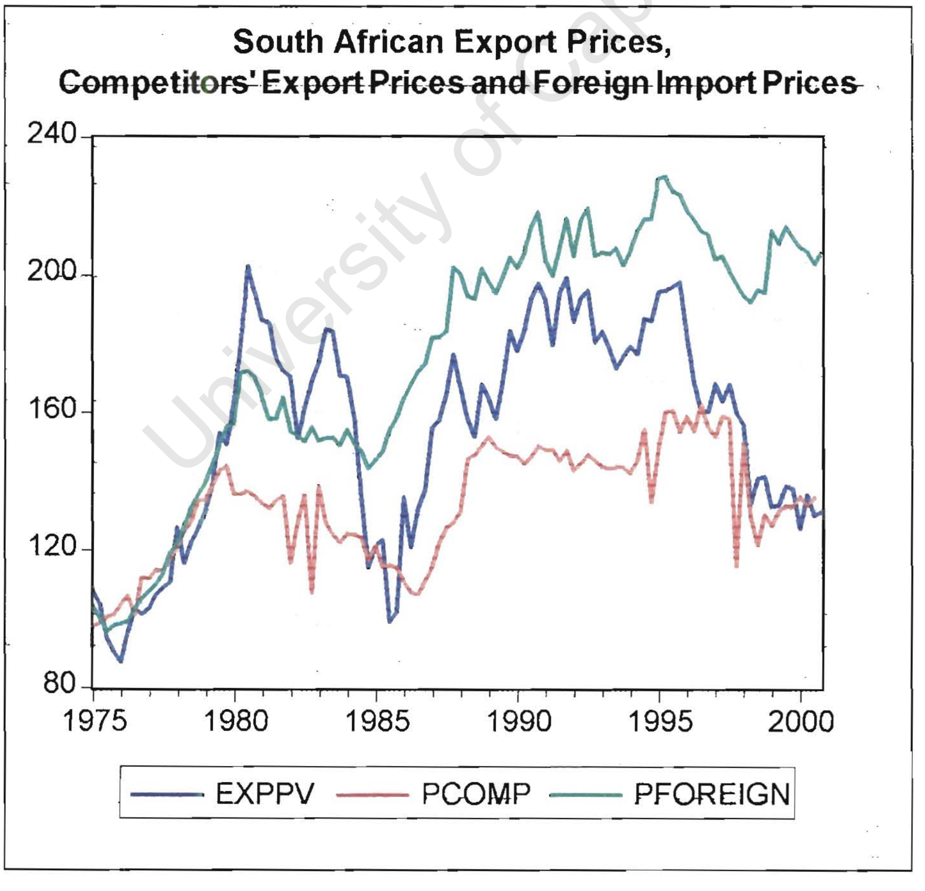

Figure 5.5: The South African export PPI has moved into line with competitors in recent times, but seems more closely correlated with foreign prices.

Figure 5.5 graphs quarterly data because the quarterly construction is better and because the effects of currency crises in South Africa's competitors are visible. The Mexican crises in the early 1980s and mid 1990s and the Korean crisis in 1997 saw sharp price falls. There were sharp rises soon thereafter, which are only partly attributable to currency recoveries. This means competitors have not held onto price advantages after a depreciation. The same can be said of South Africa in the mid 1980s, but not in the 1990s.

Using 1975 as a base, South Africa was relatively uncompetitive from the late 1970s to the late 1990s. Prices were in line with competitors' prices since then. Quarterly unit values have not fallen enough to restore the 1975 price ratio.

5.4.5 South African prices only became meaningfully lower than foreign prices after 1995

South African and foreign prices moved almost identically from 1961 until about 1980, when South Africa became relatively uncompetitive. The comparisons between South African and foreign prices in the 1980s and early 1990s depend on the choice of variable. All variables show meaningful relative cost advantage gains for South African in the late 1990s.

5.4.6 There is a close relationship between foreign prices and exports from 1975 to 1988

Quarterly data show TIPS export volumes and foreign prices moved together until about 1993. Figure 5.6 presents the closest obvious relationship between a price variable and exports. Given the complex nature of the relationship between exports and the other price variables, this is not surprising. The suggestion in figure 5.6 is that higher foreign prices entice producers to produce for export. This relationship breaks down fundamentally in 1988 for the SSA export data. The rise in exports since 1988 cannot necessarily be attributed to rising foreign prices.

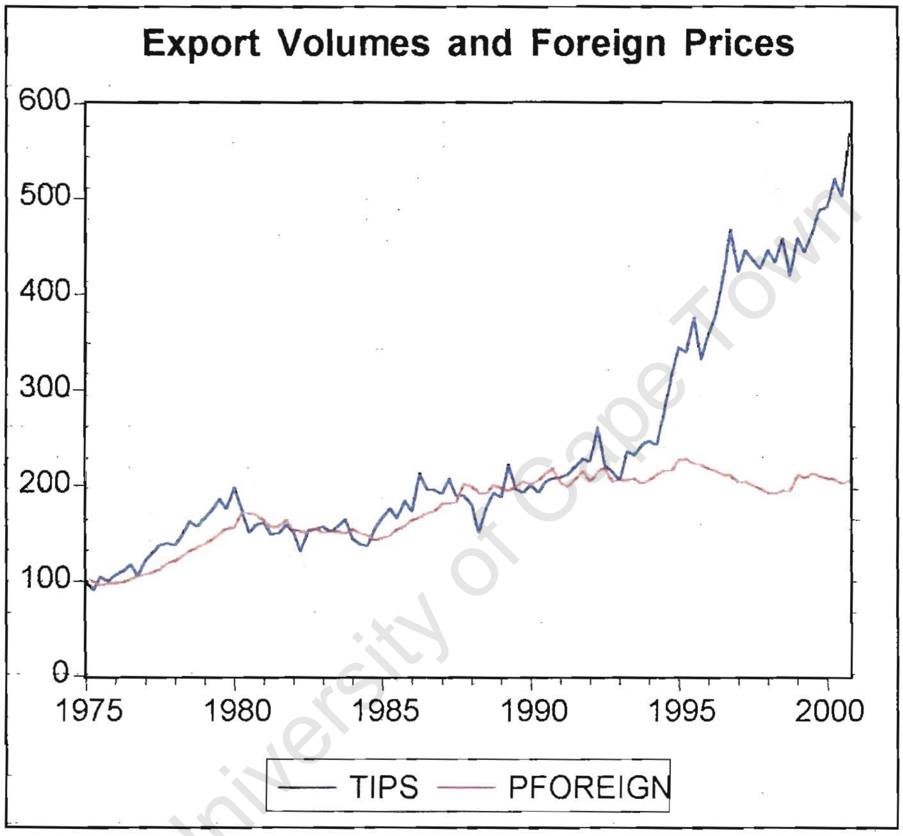

Figure 5.6: Export volumes and foreign prices were closely correlated until the early 1990s.

5.4.7 Relative prices do not reveal any clear relationships

Graphs of ratios of export prices to domestic prices, South African prices to foreign prices and South African prices to competitors' prices were inspected. The ratio patterns differed according to the variable used to represent each price. No clear relationships between relative prices and export volumes were visible.

This has two possible implications. The first possibility is that prices have a relatively limited role to play, and that the income variables to be discussed next are more important. The second possible implication is that the relative prices are acting on both the demand and supply sides of exports. Separate relationships posited for export supply and export demand are not evident in the graphs.

5.5 INCOME AND CAPACITY VARIABLES

Higher foreign incomes are generally expected to result in higher exports. The main issue is how to aggregate foreign GDPs to represent world income. While South Africa's production potential should be positively correlated with exports, section 2.4.2 explained why the relationship involving capacity utilization is subject to debate. The main data issue is the construction of potential GDP.

5.5.1 Foreign Income is represented by the United States, Untied Kingdom, Germany and Japan, aggregated using exchange rates or purchasing power parities.

As for foreign prices, the United States, United Kingdom, Germany and Japan were chosen to represent foreign income because of their importance in total South African exports. Real GDPs were obtained from the IMF's International Financial Statistics Database. For Japan, real values were not available before 1980 for quarterly data and before 1970 for annual data. Indices of production were available and were used to infer real values for these early observations.

GDP is regrettably not separated into trend and cyclical income because of data availability and further aggregation complications. Two methods are used to standardize the GDPs. The first converts each country's GDP into US Dollars at the nominal exchange rate. Exchange rates are seldom at their "equilibrium" level, so Schreyer & Koechlin (2002) recommend using purchasing power parities (PPPs) instead. These are important when one is trying to standardize volumes of production rather than values.

Both methods are used and compared. Given that real values by definition adjust for price rises and hence purchasing power over time, it is correct to take a PPP measure from one year only and apply it to the entire time series. PPPs for 1995 were sourced from the OECD. The quarterly data were deseasonalised after aggregation.

Figure 5.7 compares the two series. GDP in PPPs (GDPPPP) is virtually a straight line, but GDP in US Dollars (GDPUSD) is cyclical. While the series produce similar values from 1961 to 1985, GDPUSD is consistently higher than GDPPPP thereafter.

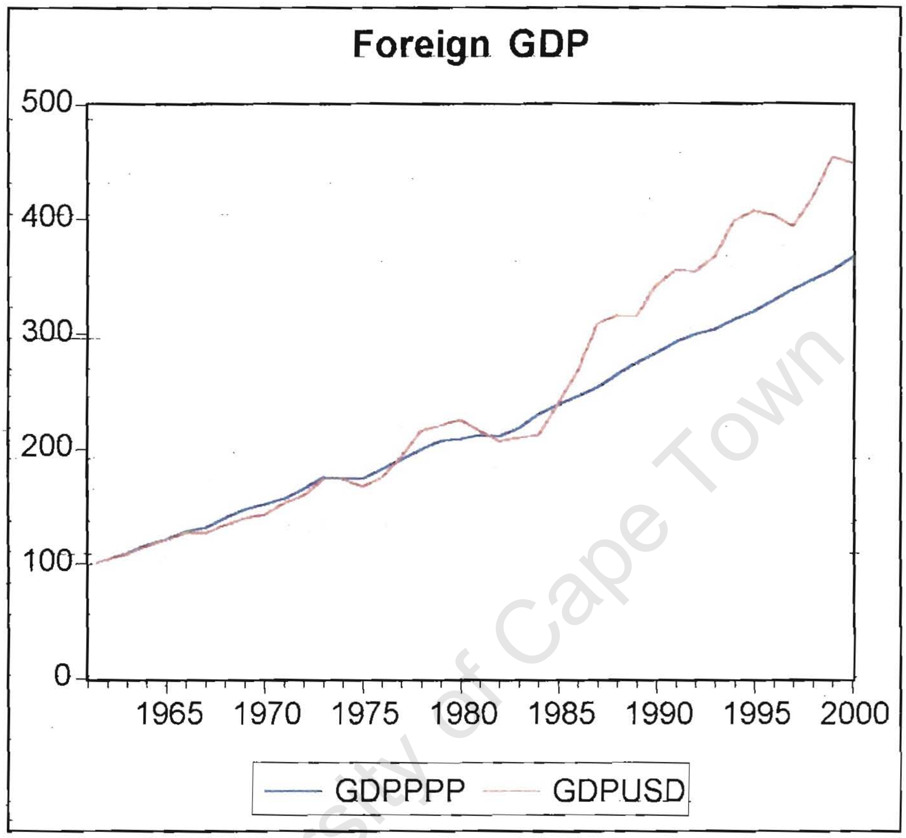

Figure 5.7: GDP in US Dollars is more cyclical than PPP GDP, which is virtually a straight line. GDP in US Dollars has been consistently higher than PPP GDP since 1985.

5.5.2 There is evidence of a positive relationship between foreign income and export volumes.

Figure 5.8 compares foreign GDP in US Dollars with export volumes. The quarterly data shows a fairly good positive relationship between the two variables (so does the annual data). They suggest higher world income leads to higher exports. Like foreign prices, world GDP cannot fully explain the rise in exports in the late 1990s. It seems that the fall in export prices in this period was a major contributor to the rise in exports.

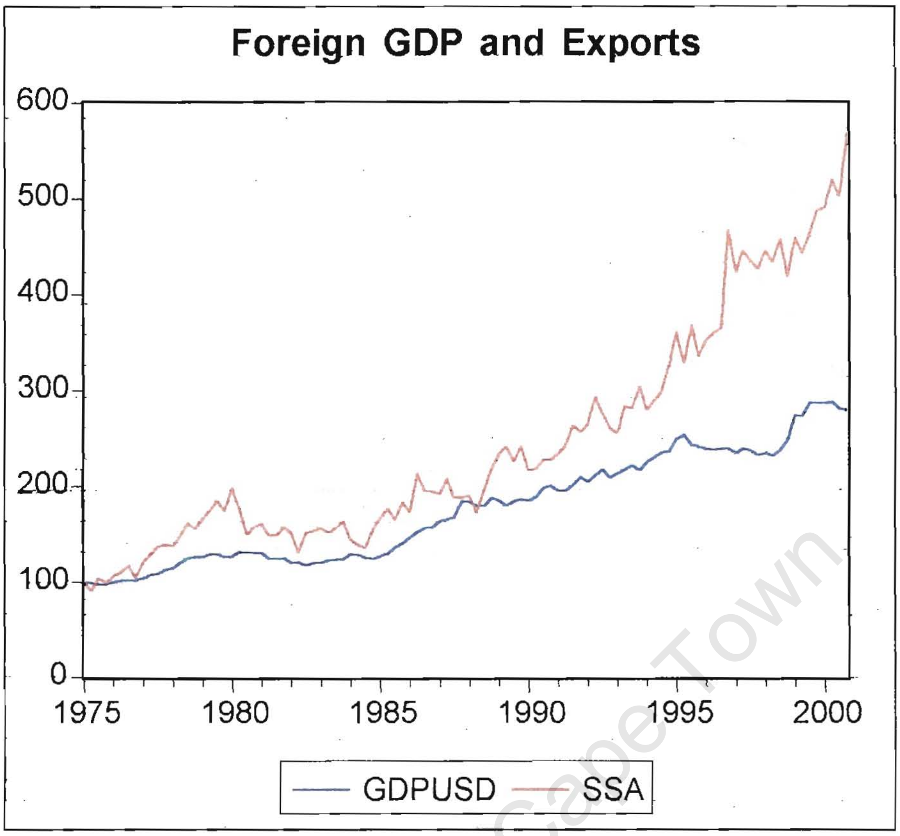

Figure 5.8: Foreign GDP in US Dollars appears to be positively correlated with exports, but cannot account for the sustained rise in exports in the late 1990s.

5.5.3 Trend GDP is derived using actual GDP and capacity utilization and is included with capacity utilization in the regressions.

Quarterly real GDP for South Africa and manufacturing capacity utilization figures were sourced from the South African Reserve Bank. Capacity utilization is not available prior to 1970, so real GDP is the only variable in the annual estimates. This is unfortunate, as the separate effects of production potential and exhaustion of that potential cannot be seen.

Capacity utilization is manufacturing output divided by potential manufacturing output. Therefore, potential manufacturing output is constructed by dividing actual output by the capacity utilization percentage. The flaw is that actual and potential GDP apply to the entire economy while capacity utilization only applies to manufacturing. In log terms:

$$\ln CU = \ln(\frac{Y}{Y^*}) = \ln Y - \ln(Y^*) \qquad (5.6)$$

$$\ln Y = \ln Y^* + \ln CU \qquad (5.7)$$

Y is actual GDP, Y^* is potential GDP and CU is percentage capacity utilization. Equation 5.7 shows that only using income imposes a restriction that sets the coefficients of potential income and capacity utilization equal to each other.

The log of capacity utilization is negative and its absolute value gets bigger as capacity utilization gets bigger. Therefore, a positive coefficient on income effectively forces capacity utilization to be negatively related to exports. This robs one of the opportunity to test the debatable relationship between capacity utilization and exports. For this reason potential GDP and actual GDP are both included in the quarterly regressions, although actual output will also be used for comparison.

5.5.4 Correlation coefficients suggest a negative coefficient on capacity utilization

Apart from the fact that exports grew faster than GDP, no graphs revealed interesting relationships. However, the correlation coefficient between exports (SSA data) and potential GDP was 0.93 – marginally bigger than the 0.91correlation with actual GDP – while the correlation with capacity utilization was –0.46. The coefficients for TIPS data were very slightly lower in each case. This provides tentative support for the vent-for-surplus argument.

Visual inspections have limited use in models like these. Only those where relatively simple one-way relationships are predicted reveal graphical relationships. The best example is foreign income and exports, although this relationship breaks down. The relationship between export price and export quantity is not a simple one, which exposes the danger of using single equation estimates instead of accepting that export prices should be positively related with export supply and negatively related with export demand. Nonetheless, it appears that the rise in exports in the late 1990s is most closely related to a fall in export price.

CHAPTER 6: ESTIMATIONS

After re-emphasising the aims of the estimations, this chapter describes the CVAR process in more detail, using one of the variable combinations as an example. There are many ways to adjust the estimations to gauge robustness. One particular variation was the choice of data set; many different combinations were tried.

Annual data estimations attempted almost every possible combination in a generally unsuccessful search for meaningful estimates. In contrast, the falsification doctrine influenced the quarterly data estimations. If a given selection of variables yielded satisfactory estimates, adjustments were made until they changed materially.

While no estimates are ideally robust to variable choice and other adjustments, the wide range of estimation combinations yields some broadly consistent results. These are presented along with the general problems encountered. The variables are estimated in natural-log form unless stated otherwise.

6.1 AIMS OF THE ESTIMATIONS

The primary aim is to extract price elasticities of demand and supply for manufactured exports. The estimates also gauge the relative importance of world income and export prices for export demand and the relative importance of domestic output conditions and export prices for export supply.

The estimates should reveal whether any of the perfect price elasticity assumptions are valid or not and whether the capacity/income variables are positively or negatively related to exports. The estimations also test the relevance of some of the new variables this study introduces into the export equations. This chapter also probes the robustness of the estimates.

6.2 ESTIMATION PROCESS

To warrant the need for the CVAR technique, the variables are tested for non-stationarity. Unrestricted VARs that make no adjustment for non-stationarity are estimated to choose the order of the CVAR. Various selection criteria help one select the number of cointegrating relations. The cointegrating relations are identified by imposing restrictions to yield long run elasticity estimates.

The long run estimates are part of the full VECM. If the estimates are satisfactory, the full VECM is analysed for significance and further insights. Although the VECM approach has ample scope for studying short run dynamics, the analysis is brief and restricted to the ECM for export volumes. Figure 6.1 summarises the procedure.

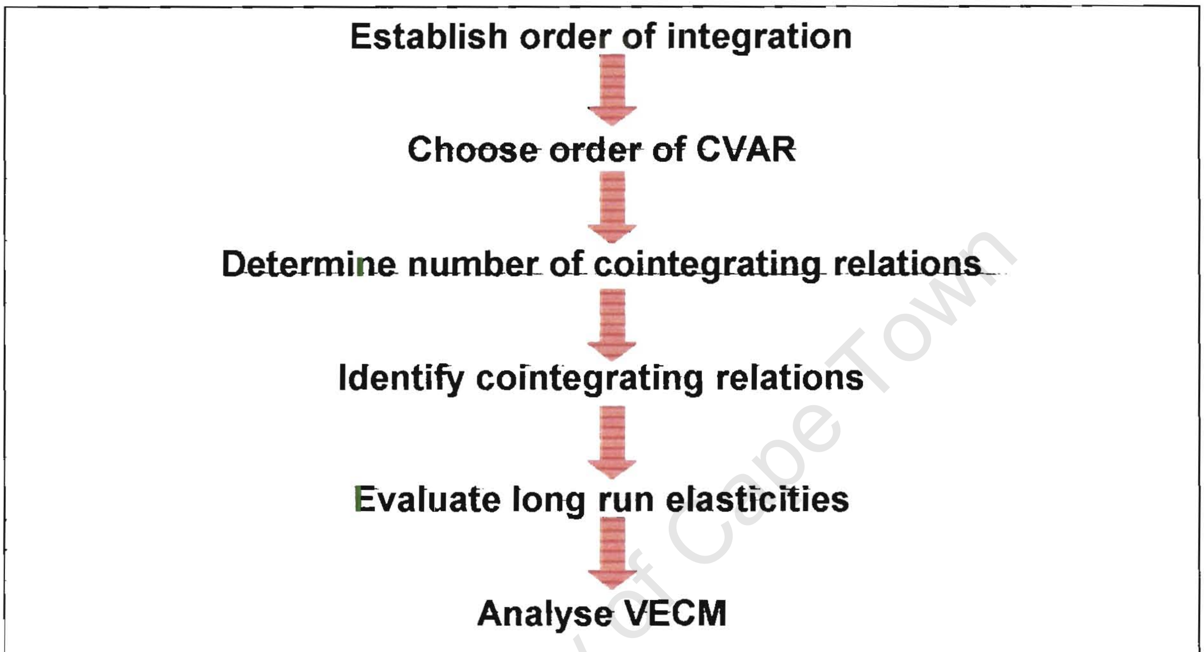

Figure 6.1 The estimation process

6.3 DETERMINING WHETHER THE VARIABLES ARE STATIONARY

Series that appear to be non-stationary, like those with an upward trend such as figure 3.1, could have been caused because they are a function of time, not because they exhibit a unit root. Such series are trend-stationary, and a time variable would prevent spurious regressions (Gujarati, 1995). It is hard to distinguish between a trend stationary process and a difference stationary process (Patterson, 2000), making tests necessary.

The augmented Dickey-Fuller (ADF) unit root test tests the hypothesis that $\delta = 1$ in the following equation (Pesaran & Pesaran, 1997):

$$y_y = \alpha + (1-\delta)\beta t + \delta y_{t-1} + \sum_{i=1}^{p+1} \delta_i \Delta y_{t-i} + \varepsilon_t \qquad (6.1)$$

They hypothesis is therefore that the series does have a unit root. The lagged differences correct for serial correlation and p is the number of lags in the model. The critical values differ if a linear trend is included in equation 6.1 (Pesaran & Pesaran, 1997). If the hypothesis is not rejected, the variable can be taken to be I(1), meriting the use of the CVAR approach. This test makes it too easy to interpret variables as I(1) when they are I(0) (Gujarati, 1995).

Microfit presents the test with up to four lags, depending on the periodicity of the series. Akaike information (AIC) and Schwarz Bayesian (SBC) criteria should be used to decide at which order to use the test statistic (Pesaran & Pesaran, 1997).

In the annual data, the ADF test with linear trend failed to reject the hypothesis that that the series is I(1) for all variables except world GDP in US Dollars (GDPUSD). The ADF statistics had the same conclusion at all lags, so it was not necessary to choose the appropriate lag length. GDPUSD cannot be included in the cointegrating relation. It can however be included in the full VECM.

In the quarterly data, the ADF tests with linear trend do not reject the I(1) hypothesis at all lags. The case of capacity utilization is ambiguous. Being a percentage, capacity utilization is unlikely to have an obvious linear trend over time, unlike the other series in the data set. However, figure 6.2 suggests there has been a fall in the mean of capacity utilization (not in log form) over time.

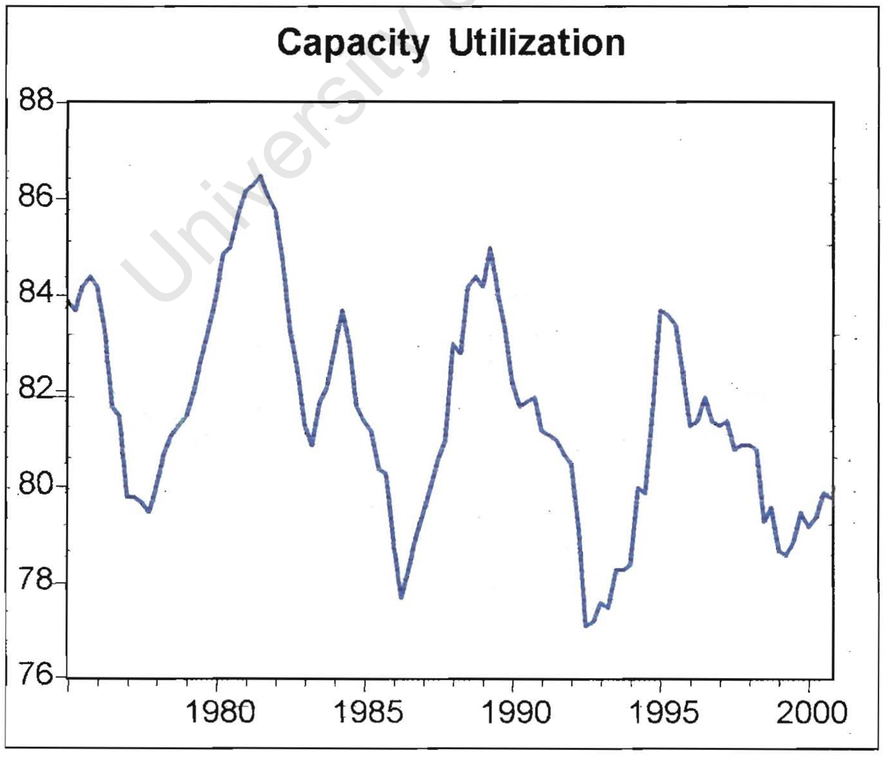

Figure 6.2: Mean capacity utilization seems to have fallen since 1975

Furthermore, a series must also have a constant variance and constant autocorrelations to be stationary (Gujarati, 1995). The ADF tests nonetheless reject the hypothesis the log of capacity utilization is I(1) at all lags, so it cannot be included in the cointegrating regression.

Capacity utilization is sometimes included in non-logarithmic format (CAPUTNL). This allows slightly easier interpretation, as the coefficient would represent the percentage change in exports for a one percentage-point change in capacity utilization rather than for a one-percent change. However, the log-form is more theoretically sound, as explained in section 5.5.3.

The ADF test with linear trend fails to reject the I(1) hypothesis for CAPUTNL at the correct lag length in terms of the AIC and the SBC, but not for the other lags. The linear trend ADF test is used because of the slight downward trend in the mean of capacity utilization.

It is important to include capacity utilization in the cointegrating vector as its sign is of particular interest. However, doing so would be justified by only the flimsiest evidence, as there is a strong case for using the ADF test with no trend, and the ADF test with a linear trend rejects the I(1) hypothesis at all lags except for one, even if it is the most appropriate one. Therefore, the alternative Phillips-Perron test is used (details are available in Pesaran & Pesaran, 1997). This test unambiguously does not reject the I(1) hypothesis, so CAPUTNL is included as an I(1) series.

Testing whether the variables are I(2) or not entails performing the same test on the differences of the data. At all lags, the ADF test with no linear trend convincingly rejects the hypothesis that the variables are I(2).

6.4 SELECTING THE ORDER OF THE VAR

The choice of lag length is usually crucial (Pesaran & Pesaran, 1997). A combination of other researchers' lags, theoretically reasonable delays, statistical selection criteria and serial correlation considerations were used to determine the best lag length.

Bhorat (1998) has a lag of eight months and international studies generally have lags of four or five quarters. The AIC and SBC can also aid in order selection. Log-likelihood ratio (LR) tests test the hypothesis that the order is ρ and not $\rho+1$.

The statistical criteria are characteristically (Patterson, 2000) ambiguous in this study, but the theory is not helpful. A hunch of five or six quarters does not help one decide between a lag of one or two years in the annual data, and the statistical criteria consistently signal lags of two or three quarters. This is broadly consistent with Bhorat's (1998) choice.

Pesaran & Pesaran (1997) advise making order selection subject to the absence of serial correlation in the individual equations. This advice was followed, and usually determines the final choice of lag-length. Output like table 6.1 was used for lag order selection.

Test Statistics and Choice Criteria for Selecting the Order of the VAR Model

Based on 99 observations from 1976Q2 to 2000Q4. Order of VAR = 5
List of variables included in the unrestricted VAR:
SSA PEXPSSA PPISA CAPUTNL POTENTIAL PCOMP
List of deterministic and/or exogenous variables:
PFOREIGN GDPUSD TIME CONSTANT

Order	LL	AIC	SBC	LR test	Adjusted LR test
5	1015.4	811.3924	546.6901	------	------
4	982.7624	814.7624	596.7724	CHSQ(36)= 65.2598[.002]	42.8474[.201]
3	946.4351	814.4351	643.1572	CHSQ(72)= 137.9145[.000]	90.5499[.069]
2	919.8954	823.8954	699.3297	CHSQ(108)= 190.9939[.000]	125.4001[.121]
1	870.6991	810.6991	732.8455	CHSQ(144)= 289.3865[.000]	190.0012[.006]
0	528.5691	504.5691	473.4276	CHSQ(180)= 973.6466[.000]	639.2629[.000]

AIC=Akaike Information Criterion SBC=Schwarz Bayesian Criterion

Table 6.1: The AIC and Adjusted LR test suggest and order of 2 lags.

The maximum order of the VAR (5 in this case) should be chosen to have a high probability of including the optimum order (Pesaran & Pesaran, 1997). In some instances, changing this maximum affects the choice of the optimum. There is a disagreement between the AIC and SBC; the former recommends 2 and the latter recommends 1. In this study, the SBC regularly signals an order of 1, no matter what the other criteria suggest.

The adjusted LR test does not reject the hypothesis that the order is 2, suggesting the order should be 2. However, it does reject the hypothesis that the order is 3 at 90% significance, suggesting the best order could be 4 or higher. This pattern is also common in many of the regressions using alternative data sets. The evidence is stronger for 2 lags in this example.

Once the lag length was selected, the diagnostics of the unrestricted VAR were analysed. In particular, the export equations were inspected for serial correlation. Better serial correlation characteristics were usually the decider when the selection criteria were evenly split, but in this case, it is just a diagnostic check for serious serial correlation.

Table 6.2 suggests it is not a problem. Serious serial correlation would have forced the choice of another order. While longer lags usually address this (Pesaran & Pesaran, 1997; Patterson, 2000), they tend to produce more serial correlation in this study. The general diagnostics are also good, suggesting the order of the VAR is suitable.

```
          OLS estimation of a single equation in the Unrestricted VAR
*******************************************************************************
 Dependent variable is SSA
 102 observations used for estimation from 1975Q3 to 2000Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 SSA(-1)                     .36732             .10346            3.5505[.001]
 SSA(-2)                     .34353            .095238            3.6070[.001]
 PEXPSSA(-1)                -.15826             .18421           -.85909[.393]
 PEXPSSA(-2)                .042207             .15957            .26450[.792]
 PPISA(-1)                   .11544             .20890            .55262[.582]
 PPISA(-2)                  -.28940             .20925           -1.3830[.170]
 CAPUTNL(-1)               .0033474            .010858            .30829[.759]
 CAPUTNL(-2)              -.0047124            .011406           -.41314[.681]
 POTENTIAL(-1)              -.20764             .82457           -.25182[.802]
 POTENTIAL(-2)               .50467             .83312            .60576[.546]
 PCOMP(-1)                   .33694             .12167            2.7692[.007]
 PCOMP(-2)                   .23938             .11627            2.0588[.043]
 PFOREIGN                   -.26840             .14960           -1.7941[.076]
 GDPUSD                      .24323             .19338            1.2578[.212]
 TIME                      .0022808           .0030847            .73939[.462]
 CONSTANT                   -1.0840             2.5128           -.43141[.667]
*******************************************************************************
 R-Squared                     .98211   R-Bar-Squared                   .97899
 S.E. of Regression           .064370   F-stat.    F( 15,  86)  314.8227[.000]
 Mean of Dependent Variable    5.4140   S.D. of Dependent Variable      .44414
 Residual Sum of Squares       .35634   Equation Log-likelihood       143.7676
 Akaike Info. Criterion      127.7676   Schwarz Bayesian Criterion    106.7679
 DW-statistic                  2.0878   System Log-likelihood         945.8731
*******************************************************************************


                              Diagnostic Tests
*******************************************************************************
*    Test Statistics  *        LM Version          *           F Version
*******************************************************************************
*                     *                            *
* A:Serial Correlation*CHSQ(   4)=   5.3684[.252]*F(   4,  82)=   1.1389[.344]
*                     *                            *
* B:Functional Form   *CHSQ(   1)=   2.6002[.107]*F(   1,  85)=   2.2235[.140]
*                     *                            *
* C:Normality         *CHSQ(   2)=   .48009[.787]*       Not applicable
*                     *                            *
* D:Heteroscedasticity*CHSQ(   1)=  .034890[.852]*F(   1, 100)=  .034218[.854]
*******************************************************************************
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
   D:Based on the regression of squared residuals on squared fitted values
```

Table 6.2: The equation for exports in the VAR does not have serial correlation.

All the variable and data combinations motivated similar orders, and occasional deviations from the chosen order had no material effect on the satisfactory estimates. So, the order choice was not crucial in this study, although this is often the case (Pesaran & Pesaran, 1997).

6.5 TESTING FOR COINTEGRATING RANK

Once the unrestricted VAR is chosen, the number of cointegrating relations to be estimated in the CVAR is established. Because the critical values depend on the nature of trends and intercepts in the cointegrating vectors (Johnston & DiNardo, 1997), this choice is made first. This study runs the estimations assuming trends in the cointegrating vectors and then runs the estimations assuming no trends in the vectors. The latter produce most of the satisfactory results.

The number of cointegrating relations depends on the eigenvalues calculated using the contents of the companion matrix. These coefficients are econometric estimates, so r is not certain. Hypothesis tests based on estimated eigenvalues are necessary. This section shows the Johansen ML technique for estimating the coefficients in the companion matrix (Johnston & DiNardo, 1997) and testing for r.

The hypothesis testing algorithm for the rank of the companion matrix is (Patterson, 2000)

- Test the hypothesis that r=0, or H(0). If not rejected, the system has no cointegrating relations and the VAR must be reformulated in first differences.
- If H(0) is rejected, test the hypothesis that $r \leq 1$, or H(1). If H(1) is not rejected, and given H(0) was rejected, there is one cointegrating vector.
- If H(1) is rejected, test H(2) and repeat the process until H(k-1). If H(k-1) is not rejected, there are k-1 cointegrating vectors.
- If H(k-1) is rejected, r=k. The matrix is of full rank and there are no unit roots. The VAR is stationary.

There are two eigenvalue based tests. The first test uses the trace statistic. It tests the hypothesis that the rank is r against the alternative that it is k by using the following log-likelihood statistic (Pesaran & Pesaran, 1997):

$$\text{trace}_i = -n\sum_{i}^{k}\ln(1-\hat{\lambda}_i) \qquad (6.2)$$

The estimated eigenvalues are ordered from largest ($\hat{\lambda}_1$) to smallest ($\hat{\lambda}_k$). i = r+1. For example, to test H(1), eigenvalues 2 to k are included in equation 6.2. An eigenvalue of zero contributes zero to the statistic but an eigenvalue close to one adds a large amount. If the statistic exceeds a certain critical value H(1) is rejected and H(2) is tested. As i rises, the statistic falls (by progressively larger amounts). Eventually, the statistic is not large enough to reject H(r) and r is determined to be the number of cointegrating vectors (Patterson, 2000).

The maximum eigenvalue statistic tests the hypothesis that rank = r against the alternative that rank = r+1 (Pesaran & Pesaran, 1997). Instead of summing eigenvalues, progressive eigenvalues are tested individually:

$$\lambda_{\max} = -n\log(1-\hat{\lambda}_{r+1}) \tag{6.3}$$

If the statistic given by 6.3 exceeds the given critical value, the hypothesis that the rank is r is rejected. If so, the hypothesis that the rank is r+1 is tested against the hypothesis that it is r+2 and so on (Patterson, 2000).

Table 6.3 is an example of the test statistics used to select the number of cointegrating relations. This formed part of an estimation that assumed no trends in the cointegrating vectors.

The maximal eigenvalue statistic rejects the hypothesis that the rank is less than 2 but fails to reject the hypothesis that the rank is less than 3, suggesting 2 cointegrating relations. The trace statistic rejects the hypothesis that the rank is 2 or less at the 90% level, suggesting the rank is 3 or more. It does not reject this hypothesis at the 95% significance level. Selecting 3 cointegrating relations is as appropriate as selecting 2. Similar ambiguities often arise in deciding between 1 and 2 cointegrating relations.

The hypothesised functional form exerts a strong bias in favour of choosing 2 vectors; one for export demand and one for export supply. While economic theory can decide in ambiguous cases in this particular study, this does not apply to all investigations. The bias is in favour of 2 vectors, even when the statistical evidence tends to support another number. This is motivated and discussed in the context of exogeneity.

Cointegration with unrestricted intercepts and no trends in the VAR

Cointegration LR Test Based on Trace of the Stochastic Matrix

102 observations from 1975Q3 to 2000Q4. Order of VAR = 2.

List of variables included in the cointegrating vector:

SSA	PEXPSSA	PPISA	CAPUTNL	POTENTIAL
PCOMP	PFOREIGN	GDPUSD		

List of I(1) exogenous variables included in the VAR:

PFOREIGN GDPUSD

List of eigenvalues in descending order:

.40719 .36941 .22217 .15272 .14069 .066834 .0000

Null	Alternative	Statistic	95% Critical Value	90% Critical Valu
r = 0	r>= 1	165.4167	122.7800	117.2600
r<= 1	r>= 2	112.0836	92.4200	87.9300
r<= 2	r>= 3	65.0522	68.0600	63.5700
r<= 3	r>= 4	39.4249	46.4400	42.6700
r<= 4	r>= 5	22.5208	28.4200	25.6300
r<= 5	r = 6	7.0556	14.3500	12.2700

Use the above table to determine r (the number of cointegrating vectors).

Cointegration with unrestricted intercepts and no trends in the VAR

Cointegration LR Test Based on Maximal Eigenvalue of the Stochastic Matrix

102 observations from 1975Q3 to 2000Q4. Order of VAR = 2.

List of variables included in the cointegrating vector:

SSA	PEXPSSA	PPISA	CAPUTNL	POTENTIAL
PCOMP	PFOREIGN	GDPUSD		

List of I(1) exogenous variables included in the VAR:

PFOREIGN GDPUSD

List of eigenvalues in descending order:

.40719 .36941 .22217 .15272 .14069 .066834 .0000

Null	Alternative	Statistic	95% Critical Value	90% Critical Valu
r = 0	r = 1	53.3331	46.0900	43.2500
r<= 1	r = 2	47.0314	39.8500	37.1500
r<= 2	r = 3	25.6273	33.8700	31.3000
r<= 3	r = 4	16.9041	27.7500	25.2100
r<= 4	r = 5	15.4652	21.0700	18.7800
r<= 5	r = 6	7.0556	14.3500	12.2700

Use the above table to determine r (the number of cointegrating vectors).

Table 6.3 The maximal eigenvalue suggests a rank of 2, while the trace statistic could be used to justify a rank of 2 or 3.

6.6 EXOGENEITY AND THE NUMBER OF COINTEGRATING RELATIONS

It is theoretically appropriate to test world income and world prices for exogeneity. Assume **x** is a vector of variables being tested for exogeneity and **n** is a vector of all the other variables. The test of block non-causality estimates equations in **x** and determines whether the coefficients of the lagged values of the variables in **n** differ significantly from zero.

LR tests are used (Johnston & DiNardo, 1997), but others are available (Patterson, 2000). The *Microfit* LR test only seems to be available after the unrestricted VAR, which means the test may be misleading (Pesaran & Pesaran, 1997). This imperfect test was used in the annual data set and world income was deemed exogenous in two variable combinations.

However, the early procedures often pointed to there being three or four cointegrating vectors when only two are expected. The additional relations could involve a relationship between world income and world prices, for example. The problem is that this relationship would be woefully incomplete. Exposing one of the drawbacks of the general-to-specific methodology, the immediate problem for the early estimations was whether to select two cointegrating relations anyway or to select more cointegrating relations and try to restrict them partially.

Both options present specification error; the former ignores possible cointegrating relations while the latter misspecifies them by omitting many variables, biasing estimates in all the vectors (Banerjee, Dolado, Galbraith & Hendry, 1993). The cointegrating relation between world GDP and its influences is especially complex, so the former option is better. Misspecification is likely to be mitigated by assuming world GDP and foreign prices are exogenous.

This, together with theoretical support and the unreliability of the test motivated using the assumption in the subsequent quarterly estimates. Further reflection suggests the exogeniety assumption should reduce the number of cointegrating vectors, as the opportunity for relations is restricted. The first few quarterly estimates supported this hypothesis. In general, the supporting evidence, although there, is not strong.

Although the quarterly estimates assume exogenous world GDP and foreign prices, all satisfactory results are tested for robustness to changing the assumption. This had an important effect in some cases. In contrast, imposing exogeneity assumptions on selected annual estimates had little or no effect.

When the data forcefully recommends a high number of cointegrating relations, selecting two cointegrating relations almost never produces reasonable results in this study.

6.7 IDENTIFICATION AND ESTIMATION

When there are two cointegrating vectors, there must be at least two restrictions per vector for each vector to be identified (Patterson, 2000). After specifying the four initial restrictions, overidentifying restrictions can be applied so that the vectors have an economic interpretation.

There can be many identification options. Luckily, theory guides them firmly in this study, so the portfolio of restrictions is limited. The restrictions fall into three categories. The first category normalises each vector, setting the quantity of exports demanded/supplied equal to one. The second category restricts coefficients in each vector to zero. In this case, all variables that determine demand are set to zero in the vector chosen to describe the export supply relationship and vice versa.

The third category is price restrictions. Price homogeneity has some theoretical justification, but the extent to which it applies in an aggregated trade context is not established. For this reason estimations are attempted with and without price homogeneity restrictions. Estimations are also run with and without the restriction that the coefficients on world and foreign prices are equal. This is often necessary for convergence and regularly produces better results. Theory is neutral in this regard, but the restriction is similar to having one international price index instead of two.

Identification is iterative. Pesaran & Pesaran (1997) recommend introducing restrictions one at a time – imposing the most likely restrictions first – in order to avoid convergence problems. I found that imposing all restrictions at once and subsequently removing the price restrictions was usually more successful. The order in which the restrictions are imposed often determines whether the algorithm converges on a coefficient estimate or not.

All the restrictions are represented by system 6.4.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{SSA} \\ \beta_{PEXPSSA} \\ \beta_{PPISA} \\ \beta_{CAPUTNL} \\ \beta_{POTENTIAL} \\ \beta_{PCOMP} \\ \beta_{PFOREIGN} \\ \beta_{GDPUSD} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (6.4a)$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{SSA} \\ \beta_{PEXPSSA} \\ \beta_{PPISA} \\ \beta_{CAPUTNL} \\ \beta_{POTENTIAL} \\ \beta_{PCOMP} \\ \beta_{PFOREIGN} \\ \beta_{GDPUSD} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (6.4b)$$

6.4a shows the restrictions placed on the vector representing the relationship involving export supply. The first row of the restriction matrix sets exports equal to 1. Rows 2,3 and 4 set demand side variables – competitors' prices, foreign prices and world GDP - equal to zero. Row 5 imposes price homogeneity restrictions. 6.4b places similar restrictions on the demand vector. The last row is the equality restriction imposed on foreign and competitors' prices.

In table 6.4, the restrictions staring with the letter *a* refer to the supply restrictions and the coefficients starting with a *b* refer to the demand side restrictions. The table also shows the coefficient estimates generated by ML techniques.

Vector 1 represents the supply relationship and vector 2 represents the demand relationship. The two vectors comprise the β matrix of long run cointegrating relationships. The way in which the restrictions are specified means the sign of the coefficients on the variables (other than export quantity) are opposite to those produced by the regression output. For example, export price has a negative coefficient and domestic price has a positive coefficient.

```
          ML estimates subject to over identifying restriction(s)
     Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                    Converged after 17 iterations
     Cointegration with unrestricted intercepts and no trends in the VAR
*******************************************************************************
 102 observations from 1975Q3 to 2000Q4. Order of VAR = 2, chosen r =2.
 List of variables included in the cointegrating vector:
 SSA             PEXPSSA         PPISA           CAPUTNL         POTENTIAL
 PCOMP           PFOREIGN        GDPUSD
 List of I(1) exogenous variables included in the VAR:
 PFOREIGN        GDPUSD
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
 a1=1; a6=0; b1=1; b3=0; b5=0; a7=0; a8=0; b4=0 ; a2+a3=0; b2+b6+b7=0; b6=b7
*******************************************************************************
                  Vector  1      Vector  2
 SSA                 1.0000         1.0000
                  (  *NONE*)     (  *NONE*)

 PEXPSSA             .17920         7.4046
                  (  .71206)     (  3.9689)

 PPISA              -.17920          .0000
                  (  .71206)     (  *NONE*)

 CAPUTNL           -.016647          .0000
                  (  *NONE*)     (  *NONE*)

 POTENTIAL          -3.6569          .0000
                  (  .63141)     (  *NONE*)

 PCOMP                .0000        -3.7023
                  (  *NONE*)     (  1.9845)

 PFOREIGN             .0000        -3.7023
                  (  *NONE*)     (  1.9845)

 GDPUSD               .0000        -2.2135
                  (  *NONE*)     (  .50254)

*******************************************************************************
 LR Test of Restrictions          CHSQ( 7)=  34.5726[.000]
 DF=Total no of restrictions(11) - no of just-identifying restrictions(4)
 LL subject to exactly identifying restrictions= 903.3383
 LL subject to over-identifying restrictions= 886.0520
*******************************************************************************
```

Table 6.4: All coefficients except the price variables in the supply equation have satisfactory signs.

Adding or removing price homogeneity restrictions often affects the signs of these two price coefficients. Table 6.5 presents an example of when removing the price homogeneity restrictions reverses the signs.

```
            ML estimates subject to over identifying restriction(s)
      Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                          Converged after 63 iterations
      Cointegration with unrestricted intercepts and no trends in the VAR
*******************************************************************************
 102 observations from 1975Q3 to 2000Q4. Order of VAR = 2, chosen r =2.
 List of variables included in the cointegrating vector:
 SSA             PEXPSSA         PPISA           CAPUTNL         POTENTIAL
 PCOMP           PFOREIGN        GDPUSD
 List of I(1) exogenous variables included in the VAR:
 PFOREIGN        GDPUSD
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
 a1=1; a6=0; b1=1; b3=0; b5=0; a7=0; a8=0; b4=0
*******************************************************************************
                  Vector  1      Vector  2
 SSA                  1.0000          1.0000
                 (    *NONE*)    (    *NONE*)

 PEXPSSA             -1.3781         12.6358
                 (    *NONE*)    (   14.5540)

 PPISA                1.1483           .0000
                 (    *NONE*)    (    *NONE*)

 CAPUTNL             .014101          -.0000
                 (    *NONE*)    (    *NONE*)

 POTENTIAL           -2.5826           .0000
                 (    *NONE*)    (    *NONE*)

 PCOMP                -.0000         -7.5484
                 (    *NONE*)    (   10.0578)

 PFOREIGN              .0000         -6.5994
                 (    *NONE*)    (    7.6788)

 GDPUSD                 0.00         -2.4043
                 (    *NONE*)    (    1.5212)

*******************************************************************************
 LR Test of Restrictions          CHSQ( 4)=   34.6992[.000]
 DF=Total no of restrictions(8) - no of just-identifying restrictions(4)
 LL subject to exactly identifying restrictions= 903.3383
 LL subject to over-identifying restrictions= 885.9887
*******************************************************************************
```

Table 6.5: This estimation produces coefficients that have the correct signs and that are in line with those achieved in other successful estimations.

To summarise the procedure so far, export volumes and unit values consisting mainly of Statistics South Africa data, GDP measured in US Dollars and the I(1) measure of capacity utilization suggested a VAR of order 2 when foreign prices and income were assumed exogenous. In the CVAR with no trends, 2 cointegrating relations were chosen, although there is some statistical support for 3. Imposing all the identification restrictions produces price coefficients with the wrong sign on the demand side. Removing the price homogeneity restrictions (and the price equality restriction, which makes no difference here) reverses the signs.

However, the coefficients in the demand vector are large. This illustrates the problem that the Johansen technique searches for *a* cointegrating relationship, not necessarily *the* cointegrating relationship. This provides another reason for performing a wide variety of estimations.

There are more restrictions than the four which identification requires. The LR test tests whether the extra restrictions do not lead to model misspecification (Patterson, 2000). If the null hypothesis of over-identification is not rejected, the structural coefficients in the cointegrating vector can be safely interpreted (Abbott & De Vita, 2002).

The LR statistic in tables 6.4 and 6.5 reject the restrictions imposed. This unfortunately happens almost always in this study, even if there are only one or two extra restrictions. Proceeding despite the rejection is a serious shortcoming, but this study would not have been able to produce a single estimate otherwise. In mitigation, the restrictions imposed have a firm grounding in economic theory. The rejection of the restrictions could well be related to the exogeneity issue and the need for more cointegrating relations.

6.8 ESTIMATING THE FULL VECM

The β coefficients of long run relations just estimated are only part of the full VECM, which is given by equation 6.5.

$$\Delta \mathbf{y}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \boldsymbol{\psi}\mathbf{D}_t + \boldsymbol{\varepsilon}_t \qquad (6.5)$$

The $\mathbf{a}_1$ term is zero because there are no trends in this particular estimation. Estimations with trends in the VAR would restrict $\mathbf{a}_1$ to the values estimated in the cointegrating vectors. The $\boldsymbol{\Pi}$ matrix is again broken up into $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. The former represents how much of the change in a variable is the result of adjustment back to the cointegrating relations given by $\boldsymbol{\beta}$, and must still be estimated.

There is only one matrix of lagged coefficients because the order of the CVAR is 2. The contents of Γ must also be estimated. The matrix of deterministic terms **D** is empty in this estimation, but would have contained the I(0) capacity utilization variable in the quarterly estimations and world GDP in US Dollars in the annual estimations.

There is an ECM for each of the endogenous variables in the VAR. Table 6.6 presents OLS estimates of the ECM for export quantity. The dependent variable is no longer conceptually export demand or export supply.

ECM for variable SSA estimated by OLS based on cointegrating VAR(2)

Dependent variable is dSSA; 102 observations used for estimation from 1975Q3 to 2000Q4

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
Intercept	-1.0894	.34746	-3.1353[.002]
dSSA1	-.34885	.098392	-3.5455[.001]
dPEXPSSA1	.13593	.16626	.81754[.416]
dPPISA1	.095540	.20748	.46048[.646]
dCAPUTNL1	-.2662E-3	.010629	-.025047[.980]
dPOTENTIAL1	-.19052	.84095	-.22656[.821]
dPCOMP1	-.042727	.10988	-.38887[.698]
dPFOREIGN1	-.58264	.37562	-1.5512[.124]
dGDPUSD1	.18777	.31388	.59822[.551]
ecm1(-1)	-.094852	.040500	-2.3420[.021]
ecm2(-1)	-.030606	.0097789	-3.1298[.002]

List of additional temporary variables created:
dSSA = SSA-SSA(-1)
dSSA1 = SSA(-1)-SSA(-2)
dPEXPSSA1 = PEXPSSA(-1)-PEXPSSA(-2)
dPPISA1 = PPISA(-1)-PPISA(-2)
dCAPUTNL1 = CAPUTNL(-1)-CAPUTNL(-2)
dPOTENTIAL1 = POTENTIAL(-1)-POTENTIAL(-2)
dPCOMP1 = PCOMP(-1)-PCOMP(-2)
dPFOREIGN1 = PFOREIGN(-1)-PFOREIGN(-2)
dGDPUSD1 = GDPUSD(-1)-GDPUSD(-2)
ecm1 = 1.0000*SSA -1.3781*PEXPSSA + 1.1483*PPISA + .014101*CAPUTNL -2.5826*POTENTIAL -.0000*PCOMP + .0000*PFOREIGN 0.00*GDPUSD;ecm2 = 1.0000*SSA + 12.6358*PEXPSSA + .0000*PPISA -.0000*CAPUTNL + .0000 *POTENTIAL -7.5484*PCOMP -6.5994*PFOREIGN -2.4043*GDPUSD

R-Squared	.25125	R-Bar-Squared	.16897
S.E. of Regression	.069410	F-stat. F(10, 91)	3.0536[.002]
Mean of Dependent Variable	.017814	S.D. of Dependent Variable	.076141
Residual Sum of Squares	.43842	Equation Log-likelihood	133.1952
Akaike Info. Criterion	122.1952	Schwarz Bayesian Criterion	107.7579
DW-statistic	2.0427	System Log-likelihood	885.9887

Diagnostic Tests

Test Statistics	LM Version	F Version
A:Serial Correlation	CHSQ(4)= 2.6189[.623]	F(4, 87)= .57317[.683]
B:Functional Form	CHSQ(1)= 1.4769[.224]	F(1, 90)= 1.3223[.253]
C:Normality	CHSQ(2)= .72250[.697]	Not applicable
D:Heteroscedasticity	CHSQ(1)= .012845[.910]	F(1, 100)= .012595[.911]

A:Lagrange multiplier test of residual serial correlation
B:Ramsey's RESET test using the square of the fitted values
C:Based on a test of skewness and kurtosis of residuals
D:Based on the regression of squared residuals on squared fitted values

Table 6.6: The adjustment coefficients on ecm1(-1 and ecm2(-1) are small but have the correct sign and are significant. The coefficients on the differenced variables are generally not significant.

Before analysing the ECM for exports, the results so far are presented in system 6.6. It follows the format of equation 6.5. In this more general format representing the whole VECM, changes in each of the endogenous variables are expressed in terms of adjustment coefficients multiplied by the error in the value of the variable, which is determined by the matrix of cointegrating relations. Changes in each of the endogenous variables are also a function of lagged changes in all the variables in the system and of an intercept term.

$$\begin{pmatrix} \Delta SSA \\ \Delta PEXPSSA \\ \Delta PPISA \\ \Delta CAPUTNL \\ \Delta POTENTIAL \\ \Delta PCOMP \end{pmatrix} = \begin{pmatrix} -1.09 \\ a_{02} \\ a_{03} \\ a_{04} \\ a_{05} \\ a_{06} \end{pmatrix} + \begin{pmatrix} -0.1 & -0.03 \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \\ \alpha_{41} & \alpha_{42} \\ \alpha_{51} & \alpha_{52} \\ \alpha_{61} & \alpha_{62} \end{pmatrix} \begin{pmatrix} 1 & -1.38 & 1.15 & 0.01 & -2.58 & 0 & 0 & 0 \\ 1 & 12.64 & 0 & 0 & 0 & -7.55 & -6.6 & -2.4 \end{pmatrix} \begin{pmatrix} SSA \\ PEXPSSA \\ PPISA \\ CAPUTNL \\ POTENTIAL \\ PCOMP \\ PFOREIGN \\ GDPUSD \end{pmatrix}$$

$$+ \begin{pmatrix} -0.35 & 0.14 & 0.1 & 0 & -0.19 & -0.04 & -0.58 & 0.19 \\ \Gamma_{21} & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & \Gamma_{68} \end{pmatrix} \begin{pmatrix} \Delta SSA1 \\ \Delta PEXPSSA1 \\ \Delta PPISA1 \\ \Delta CAPUTNL1 \\ \Delta POTENTIAL1 \\ \Delta PCOMP1 \\ \Delta PFOREIGN1 \\ \Delta GDPUSD1 \end{pmatrix} \qquad (6.6)$$

In table 6.6, ECM1(-1) is the error in the supply relationship in the previous period and ECM2(-1) applies to the demand relationship. They represent the first row in α, where there is a column for each cointegrating relation. Each subsequent row corresponds to adjustment coefficients for the other endogenous variables.

The coefficients should be negative. If export demand is higher than suggested by the other variables in the cointegrating relationship, it should fall and vice versa. The same applies to export supply. While the coefficients are negative and significant, they are small.

A coefficient of –0.03 means the change in exports will increase by –0.03 multiplied by the previous period's percentage-point change in the disequilibrium, denoted by Z. If Z is –10%, the change in exports will be 0.03% percentage-points higher than it would have been. For example, a change of 1% caused by lagged changes in the other variables rises to a change of 1.03% because of adjustment to equilibrium. A mathematical translation is available in the footnote[3].

This suggests very slow adjustment to equilibrium. Assume exports are at their equilibrium value of 95, and a shock makes the equilibrium value 105. $Z = (-10 - 0)/100 = -10\%$. The percentage change in X is therefore only 0.03 percentage points higher, suggesting a very slow journey towards its equilibrium value. Even the coefficient of -0.1 is low, although the 1% change would match the 1% change caused by lagged changes in other variables.

The results in table 6.6 suggest lagged changes in other variables also have a very small effect, if any, on export volumes. Except for lagged exports itself, the lagged difference terms are not significant. A Wald test finds all the lagged difference terms except for exports jointly insignificant. In contrast, the error correction terms are jointly significant at a significance level of less than 1%.

One could argue that both sets of coefficients are jointly inadequate at explaining changes in exports, but the f-statistic for the ECM is highly significant. All the diagnostics for this ECM are good. Sometimes, serial correlation was statistically significantly present. If so, Newey-West adjusted variances at different lags were used. These improved significance levels in all cases.

$$d \ln X = \alpha(\ln X - \ln X^*) + \phi d \ln Y$$

$$\frac{dX}{X} = \alpha\left(\frac{X - X^*}{(X + X^*)/2}\right) + \phi\frac{dY}{Y}$$

$$\frac{dX}{X} \Big/ \frac{dY}{Y} = \phi$$

$$\frac{d\left(\frac{dX}{X}\right)}{d\left(\frac{X - X^*}{X}\right)} = \alpha$$

X = value of exports

3 X^* = equilibrium value of exports

Y = world income

α and ϕ are estimated coefficients

$$d\left(\frac{dX}{X}\right) = \frac{dX^{\bullet} - dX^{\mu}}{X}$$

$dX^{\bullet}$ = change in X because of adjustment to equilibrium and lagged changes in Y

dX^{μ} = change in X because of changes in Y only; no disequilibrium to adjust to

6.9 GENERAL FINDINGS

One of the aims of the study is to measure the robustness of the estimates to changes in the CVAR, especially to using data obtained from various sources and constructed using different criteria. While there are no completely robust results, there are some consistencies and reliable ranges of elasticity estimates.

They are summarised in table 6.7. For estimates that have coefficients with the correct sign in the entire cointegrating vector, there are separate columns for quarterly data and for annual data. To gauge robustness, a third column presents results based on cointegrating VARs that had one or two incorrect coefficients.

Results or ranges separated by a semi-colon are reported when there is an obvious break in the range of estimates. If so, the values are listed in descending order of importance.

Variable	Quarterly	Annual	Imperfect	Equation
Price elasticity of Demand	–3 to –6	–3 to –6	–3.5 to –6.5	D
Price elasticity of Supply	0.7 to 1.3; 0.35	0.7 to 1	0.8 to 1.2 ; 0	S
Domestic Prices	–0.7 to –1.3; –0.35	0.7 to 1	–0.8 to –1.2 ; 0	S
SA Potential GDP	2.6 to 3.9	-	3; 2 to 4	S
SA Capacity Utilization*	–0.02 to 0.12	-	–0.127 to 0.12	S
SA GDP	2.7 to 3.7	2 to 5	5.5 to 7.3	S
Competitors' Prices	2.5 to 4	1 to 4	1.8 to 2.3; 4	D
Foreign Prices	1.5 to 2.5	1.5 to 2.5	1.8 to 2.3;4; 5.5	D
GDPPPP	3.5	1 to 1.7	3; 4	D
GDPUSD	2 to 2.5	1 to 1.7	0.9 to 2.2	D
ECM 1	–0.02 to –0.09	–0.04 to –0.08	-	S
ECM 2	–0.03 to –0.17	+0.02 to –0.12	-	D

Table 6.7: Summary of coefficient estimates. A semi-colon represents a break in the range of elasticity estimates; values/ranges are listed in descending order of importance. For example –0.8 to –1.2 ; 0 means elasticities generally ranged from –0.8 to –1.2, but there were occasional estimates close to zero. * denotes not in log format, so the coefficient must be interpreted differently to the others.

Most of the imperfect estimates considered had the correct coefficients except for the two price variables on the supply side, namely export price and domestic price. A negative export price coefficient and a positive domestic price coefficient were common. However, their absolute values were generally close to the coefficients with correct signs.

Because the two variables are closely correlated (see figure 5.4), it is possible that the estimation procedure somehow "confuses" the two variables. Price homogeneity restrictions regularly switch the signs of the price coefficients around.

6.9.1 Export demand is highly elastic

Studies that assume infinitely price elastic demand for exports do so incorrectly and hence bias their supply estimates. Both annual and quarterly data produce price elasticity estimates ranging from –3 to –6, and the imperfect cointegrating vectors raise these values slightly.

While the range is wide, these estimates are unequivocally large. The significant fall in price seems to offer a credible explanation for rising exports in the late 1990s. Demand for South African manufactured exports is highly elastic, so lower export prices would raise export revenues, not only volumes.

The distribution about this range is random; the values do not systematically vary according to the criteria chosen for the estimation.

6.9.2 The price elasticity of supply is close to1, but could be far less

The mode of the quarterly elasticity estimates is about 1.2, but this value is mainly generated by the SSA series. TIPS series have lower quarterly estimates. Annual estimates using both series are also generally lower. Using quarterly South African GDP instead of potential GDP and capacity utilization lowers estimates to about 0.35 for both series. The imperfect cointegrating vectors in general produced a similar range of estimates, but there were a number of estimates close to zero.

A researcher wanting to support a hypothesis of perfectly price-elastic supply could have done so with a carefully chosen combination of variables, but this broader study shows such an assumption should not be made.

6.9.3 The coefficient on domestic prices was close to –1

Price homogeneity restrictions obviously made the coefficients on domestic prices and export prices equal in absolute value in all estimations. Domestic price coefficients in the absence of such restrictions were a bit higher in some cases and a bit lower in others, so the restriction itself seems justified. While the Chi test should be used to test this and other restrictions more formally, this option was not meaningfully available, because almost all the restriction were always rejected.

If price homogeneity does hold, equal percentage changes in domestic and export prices will not influence export supply. Given that the two series have moved together so closely, relative prices do not seem to have had an affect on export supply. In addition, price homogeneity justifies using a price ratio instead of separate variables.

6.9.4 GDP matters; capacity utilization probably does not

The coefficient on potential GDP is as expected positive. The quarterly estimates range fairly uniformly from 2.6-3.9. There is no systematic tendency for some estimations to produce higher coefficients than others. There are no annual estimates for this variable. The imperfect results consisted of a slightly wider range of coefficients but most hovered at around 3.

Capacity utilization, when legitimately included in the VAR in non-logarithmic format, was generally positive. The small coefficients should not be interpreted as having no effect. A coefficient of 0.1 means a 1 percentage-point rise in capacity utilization leads to a 10% rise in export supply. The less satisfactory estimates produced ambiguous results.

The I(0) version of capacity utilization was included in the VECM. The coefficients range from –0.047 to 0.034. These values are close to zero and come from ECMs based on good long run estimates. They were never close to being statistically significant. Almost all terms in the ECM are insignificant, so the I(0) version's insignificance does not necessarily mean the I(1) measure is not important. However, the evidence is not strong enough to assert capacity utilization and export supply are positively or negatively related.

This is very disappointing, as the estimations cannot say much about the relationship between exports and capacity. The relative strengths of the "vent-for-surplus" and the "exports-generate-demand" arguments cannot really be evaluated. The two effects are possibly equally potent.

It might be too optimistic to suggest the effect of capacity utilization is fully captured by price adjustments. However, when price homogeneity is imposed, there is support for this in some ECMs for export prices. They show a significantly negative relationship between domestic prices and capacity utilization.

Using actual GDP yields similar results to potential GDP. Quarterly estimates ranged from 2.7 to 3.7, but annual estimates ranged from 2 to 5. The imperfect cointegrating vectors gave estimates as high as 7.3. The similarity between actual and potential GDP and the inconclusive capacity utilization results suggest the relationship between GDP and exports operates solely through the production potential influence and not through aggregate demand or capacity utilization.

6.9.5 Foreign prices and competitors' prices are important

The coefficients on foreign prices and competitors' prices are consistent when exogenous foreign prices are assumed, the former ranging from 1.5 to 2.5 and the latter ranging from 1 to 4 in the satisfactory estimates. In the unsatisfactory estimations, foreign prices occasionally had higher coefficients, especially when actual GDP was used instead of potential GDP and capacity utilization.

Imposing this equality restriction and/or price homogeneity often generated more significant VECMs and were often useful when the coefficient on one of the two price variables was incorrectly negative. Incorrect signs often resulted when exogenous foreign prices and GDP were not assumed. Price homogeneity restrictions applied to foreign prices, competitors' prices and export prices had no noticeable effect on the actual coefficient size. Again, it would have been beneficial to test this formally.

Competitors' prices, which have been ignored in other studies, are certainly an important determinant of demand for South African exports. The fact that the coefficients tended to be higher than for foreign prices vindicates the decision to include competitors' prices. The results also assert that absolute competitiveness improvements may not be sufficient to increase export demand. If our competing exporters continue to offer cheaper and cheaper goods, lower South African export prices will be required merely to preserve market share.

6.9.6 World GDP is not as important as export price

The coefficients on world income in the demand equation are positive and robust. They are similar in annual estimations. Quarterly estimations and the imperfect estimations produce far lower coefficients for GDPUSD than for GDPPPP. While the role of world income in export demand is substantial, there is strong evidence that export price is more important.

6.9.7 The speed of adjustment is slow, but lagged difference terms are not as important as the error correction terms

Satisfactory long run estimates motivate estimating error correction models for exports. About half the ECM1 terms were significant; all the significant ones were –0.07 or higher. There was a greater number of significant ECM2 terms; the range of coefficients is slightly wider, yet even low coefficients were significant. No values are reported for the imperfect estimations because ECMs were not routinely analysed. Occasional inspection produced similar coefficients.

Even the highest satisfactory coefficient of –0.17 suggests a 10% positive shock to the error term, which means export supply should be 10% lower than it currently is, will only cause an additional downward adjustment of 1.7%. Assuming no other disturbances, approximately one third of the existing error (3.27%) would still be there after six quarters.

The annual coefficients should be higher than the quarterly coefficients, as each period represents more time. The coefficients are similar to the quarterly values, suggesting the adjustment takes roughly four times as long. Despite the low coefficients, tests of the joint significance of the error correction terms were highly significant.

Individual inspection of the t-statistic probabilities on the difference terms suggested they are not important. The only exception is the lagged version of differenced exports itself. Wald tests of the joint importance of the lagged difference terms were not significant.

6.9.8 Is the model significant?

The insignificant difference terms are disappointing as there is no satisfactory way to test the significance of individual variables. It seems bizarre that, apart from an occasional mention of standard errors, I found no reference to significance tests for individual variables in each cointegrating vector. The restrictions imposed often meant even standard errors were not available. One method could be the LR test of the restrictions, but this method was not instructive as it rejected all combinations of linear restrictions. The next best alternative was to use the lagged difference terms in the full ECM, but their general insignificance makes this option unhelpful in this case.

The low adjustment coefficient and the insignificant lagged differences could mean two things. On the one hand, it could mean exports are very slow to adjust to changes in the lagged variables in the previous period, which creates an error. Exports are also slow to adjust to this error and restore long run equilibrium.

One the other hand, it could mean the VECM is wrong as a whole. There could be important variables missing and the included variables could be insignificant. However, all ECMs based on satisfactory long-run estimates were highly significant, with F-statistics ranging from 0.000 to 0.005. This suggests the model offers a good explanation for changes in exports.

6.10 REVIEW OF OTHER FINDINGS

Menhadji & Montenegro (1999) study total exports in 75 countries. The long run price elasticities of demand range from –0.02 to –4.72, with an average of about -1. Their study was a single-equation estimate. Given that this biases estimates downwards, it is no surprise that the coefficients are lower than those found in this study. World income elasticities average 1.48, ranging from 0.17 to 4.34.

The nature of the other South African studies limits direct comparability. Bhorat (1998) estimates the price elasticity of supply for total exports to be 1.3, which is close to this study's estimates despite assuming perfectly elastic demand. His coefficient on domestic prices is a high –4.7. Fallon & Pereira de Silva (1994) find their relative price has the incorrect sign and is insignificant. Their other regressions find the real exchange rate significant, but with elasticities of less than – 0.5. Tsikata (1999) also finds the exchange rate deviation from the purchasing power parity level significantly negative.

Fallon & Pereira de Silva (1994) find capacity utilization is statistically significant with elasticities exceeding –1. In contrast, Tsikata (1999) and Wood (1995) advance capacity utilization is not important. Bhorat (1998) estimates the coefficient on domestic output to have an elasticity of –1.8, suggesting the "vent-for-surplus" argument prevails and is strong enough to outweigh the likely positive effects of potential income. Both Fallon & Pereira de Silva and Wood (ibid.) find world income to be insignificant.

6.11 CONVERGENCE PROBLEMS AND INCONSISTENCIES

While section 6.9 summarised a variety of estimates that were satisfactory, the majority of estimates did *not* work. This is especially so for the annual data. In many cases, the maximum likelihood algorithms did not solve. The two methods in *Microfit* are the modified Newton-

Raphson method and the back-substitution method. There are numerous ways to adjust the solving procedure, but they in general had no effect.

One of the adjustments is the initial values the algorithm uses. The default starting point is the most recent set of estimates. In cases where the coefficients are satisfactory except for the signs on export price and domestic price, I tried switching the signs on the coefficients around and re-running the estimations. Fortunately (or unfortunately, depending on your point of view), this did not work.

I often had a converged set of estimates based on certain restrictions, tested a restriction, removed the restriction, and failed to get results for the original restrictions. It is even more serious when the algorithm does converge using the original restrictions, but at different coefficients! Tables 6.8 and 6.9 present two estimations that have exactly the same variables and restrictions yet completely different coefficients.

Finally, the number of variable and restriction combinations is high to say the least. The fact that imposing them in different orders and that the same combinations can produce different results is very serious. This means that it is virtually impossible to be completely thorough, and a good set of estimates can be missed. More importantly, it seems that, if one looks hard and long enough, one can find a set of estimates to back almost any hypothesis.

While ambiguity can be exploited in almost any empirical technique, it seems that the CVAR approach is particularly susceptible to such manipulation. Any study that produces a single set of estimates should be treated with extreme caution.

```
              ML estimates subject to over identifying restriction(s)
          Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                           Converged after 16 iterations
          Cointegration with unrestricted intercepts and no trends in the VAR
***************************************************************************
 102 observations from 1975Q3 to 2000Q4. Order of VAR = 2, chosen r =2.
 List of variables included in the cointegrating vector:
 SSA              PEXPSSA          PPISA            POTENTIAL        PCOMP
 PFOREIGN         GDPPPP
 List of I(1) exogenous variables included in the VAR:
 PFOREIGN         GDPPPP
 List of I(0) variables included in the VAR:
 CAPUT
***************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
 a1=1; a6=0; b1=1; b3=0; a5=0; a7=0; b4=0; a2+a3=0; b2+b5+b6=0; b5=b6
***************************************************************************
                  Vector  1      Vector  2
 SSA                  1.0000         1.0000
                 (    *NONE*)   (    *NONE*)

 PEXPSSA             -1.2438         6.0577
                 (    *NONE*)   (    2.9559)

 PPISA                1.2438         -.0000
                 (    *NONE*)   (    *NONE*)

 POTENTIAL           -2.9365         -.0000
                 (    *NONE*)   (    *NONE*)

 PCOMP                -.0000        -3.0289
                 (    *NONE*)   (    1.4780)

 PFOREIGN              .0000        -3.0289
                 (    *NONE*)   (    1.4780)

 GDPPPP                .0000        -3.3053
                 (    *NONE*)   (    .72523)
```

Table 6.8: This estimation of long-run cointegrating relations produce common and theoretically consistent estimates. Table 6.9 shows completely different estimates with the same inputs.

```
              ML estimates subject to over identifying restriction(s)
     Estimates of Restricted Cointegrating Relations (SE's in Brackets)
                        Converged after 21 iterations
     Cointegration with unrestricted intercepts and no trends in the VAR
*******************************************************************************
 102 observations from 1975Q3 to 2000Q4. Order of VAR = 2, chosen r =2.
 List of variables included in the cointegrating vector:
 SSA             PEXPSSA         PPISA           POTENTIAL       PCOMP
 PFOREIGN        GDPPPP
 List of I(1) exogenous variables included in the VAR:
 PFOREIGN        GDPPPP
 List of I(0) variables included in the VAR:
 CAPUT
*******************************************************************************
 List of imposed restriction(s) on cointegrating vectors:
 a1=1; a6=0; b1=1; b3=0;a5=0;a7=0;b4=0;b2+b5+b6=0;a2+a3=0;b5=b6
*******************************************************************************
                  Vector  1      Vector  2
 SSA                  1.0000         1.0000
                 (    *NONE*)   (    *NONE*)

 PEXPSSA              1.8220       -.067957
                 (    2.0927)   (    *NONE*)

 PPISA               -1.8220          .0000
                 (    2.0927)   (    *NONE*)

 POTENTIAL           -1.1275         -.0000
                 (    2.0671)   (    *NONE*)

 PCOMP                -.0000        .033978
                 (    *NONE*)   (    *NONE*)

 PFOREIGN              .0000        .033978
                 (    *NONE*)   (    *NONE*)

 GDPPPP               -.0000        -2.1754
                 (    *NONE*)   (   .098725)
```

Table 6.90. This set of results is far from satisfactory, yet it uses exactly the same inputs as table 6.8.

CHAPTER 7. CONCLUSION

The basic framework for elasticity estimates is grounded on the traditional laws of demand and supply. Export supply is a function the price of exports, the prices of domestic production substitutes and inputs, production capacity and domestic demand conditions. Export demand is a function of the price of exports, the price of substitute products in the export market, the price of substitute products produced by competitors and world income.

Choosing the data series is not as straightforward. Proxies are often necessary, and even the first-choice measures have problems. This study required the substantial sourcing, capturing and merging of series from different sources and the construction of other series.

The cointegrating vector autoregression (CVAR) technique is used because it provides an integrated way to estimate systems of simultaneous equations using non-stationary data. This is required in order to avoid biased estimates and spurious regressions. The CVAR is a reparameterisation of a vector error correction model (VECM), which describes changes in the endogenous I(1) variables in terms of lagged differences in all the system's variables and in terms of adjustment to a possible long run equilibrium relationship. The long run relationship is cointegrated, so the regressions are no longer spurious, and estimation in VAR format avoids simultaneity bias.

The Johansen Technique determines the number of cointegrating relations that exist between non-stationary data. Statistical criteria are used to validate the theoretical assertion that there are two cointegrating relations in this study. Estimating the coefficients entails finding combinations of the variables that are cointegrated. By imposing theoretically motivated restrictions, separate demand and supply equations are specified. This allows the coefficients to be interpreted as elasticities.

This paper concentrates on finding the coefficients of the long run cointegrating relationships, improving on existing South African studies by using the CVAR technique, by using a relatively large data set, and by introducing some new variables to the standard specification.

The estimation process reveals that it can sometimes be very difficult to find economically meaningful results. Paradoxically, it can be easy to find an estimate to support a given hypothesis. While this is true for most empirical methods, any studies that only show a single set of results based on the CVAR technique should be treated with particular caution.

While many estimation combinations did not converge, and many yielded nonsensical results, some broad patterns emerge. With a few exceptions, changing the data series did not alter the results systematically, but the wide range of some coefficient estimates shows the selection and construction of the data series can affect the results.

The price elasticity of demand is –3 to -6, dismissing any studies assuming South Africa is a price taker. The income elasticity of demand depends on whether the series are annual or quarterly and whether income is measured in nominal US Dollars or in terms of purchasing power parities, rangings from 1 to 3.5. The conclusion is that while both world income and export prices are important, competitiveness measures can materially increase exports volumes and revenues.

This conclusion is affirmed by the positive coefficients on foreign prices and competitors' prices. Cross elasticities ranging from 1 to 4.5 also suggest absolute competitiveness improvements may not be enough. Competitiveness improvements may be necessary merely for preserving export shares. Greater market share is likely to require improvements in competitiveness relative to other producers.

The price elasticity of supply is about 1, and the coefficient on domestic prices is about –1. Production potential is positively related to exports, having a coefficient of 2.6 to 3.9, but capacity utilization does not seem to be important. The effect of domestic demand on exports is therefore unclear, although it may work through domestic prices. The results suggest the relationship between domestic output and exports is expressed fully through potential GDP.

The adjustment to the long run equilibrium relationship is a significant determinant of changes in export quantities. The size of the adjustment coefficient suggests slow adjustment to equilibrium. While the adjustment process is statistically significant, it is hard to tell which variables in the VECM are statistically significant.

While the specification is fairly standard, there is plenty of opportunity for alternative variable constructions, which perhaps study the role of real exchange rates more directly. In a few years, research should also have the luxury of using long time series from a single source, avoiding some of the data inconsistencies encountered by this study.

The focus of this study has been on long-run elasticities, but the CVAR approach provides many opportunities for in depth analysis of dynamics. Furthermore, the CVAR approach could perhaps be used to estimate a more general model consisting of cointegrating relations, involving South African GDP and international prices for example.

BIBLIOGRAPHY AND DATA SOURCES

Abbott, A & G de Vita (2002); "Long-run price and income elasticities of demand for Hong Kong exports: a structural cointegrating VAR approach" in *Applied Economics*, 34

ABSA Group Ltd (2001); *South Africa's Foreign Trade, 2001 edition*; South African Financial Sector Forum

Athukorala, P & J Riedel (1994); "Demand and Supply Factors in the Determination of NIE Exports: A Simultaneous Error-correction Model for Hong Kong" in *The Economic Journal*, 104

Banerjee, A, J Dolado, J Galbraith & D Hendry (1993); *Co-integration, Error-correction, and the econometric analysis of non-stationary data: advanced texts in econometrics*; Oxford University Press, Oxford

Bhorat H, (1998); "Income and Price Elasticities in Manufacturing Exports" in *Trade & Industry Monitor*, Volume 5, March 1998; Trade and Industrial Policy Secretariat and Development Policy Research Unit

Chiang, A (1984); *Fundamental Methods of Mathematical Economics*; McGraw-Hill, Singapore

Dezhbakhsh, H (2002); *The Political Economy of Statistical Evidence: Economic Data Problems in Developing Countries and Their Impact on Empirical Research*; Conference on Development Issues in the New Economy, University of Cape Town

Fallon, P & L Pereira de Silva (1994); *South Africa: Economic Performance and Policies*, Discussion Paper 7; World Bank Southern Africa Department

Goldstein, M & M Khan (1978); "The supply and demand for exports: a simultaneous approach" in *The Review of Economics and Statistics*, vol. 60

Goldstein, M & M Khan (1985); "Income and Price Effects in Foreign Trade" in R Jones & P Kenen; *Handbook of International Economics, vol II*; Elsevier Science Publishers

Golub, S (2000); *South Africa's International Cost Competitiveness*, TIPS Working Paper 14_2000

Gujarati, D (1995); *Basic Econometrics 3rd edition*; McGraw-Hill; Singapore

Hanninen, Riita & A Topinen (1999); "Long-run Price Effects of Exchange rate Changes in Finish Pulp and Paper Exports" in *Applied Economics*, vol. 31 issue 8

IMF; International Financial Statistics electronic database via the University of Cape Town database

Johnston, J & J DiNardo (1997); *Econometric Methods, fourth edition*; McGraw-Hill, Singapore

Mahdavi, S (2000); "Do German, Japanese and US Export Prices Asymmetrically Respond to Exchange Rate Changes? Evidence from Aggregate Data" in *Contemporary Economic Policy*, vol. 18 issue 1

Muscatelli (1994); "Demand and Supply Factors in the Determination of NIE Exports: A Reply" in *The Economic Journal*, 104

Muscatelli, V, A Stevenson & C Montagna (1995); "Modeling [sic] Aggregate Manufactured Exports for some Asian Newly Industrialized Economies" in *Review of Economics and Statistics*, volume 77 issue 1

Muscatelli, V, T Srinivasan & D Vines (1992); "Demand and Supply Factors in the Determination of NIE Exports: A Simultaneous Error-correction Model for Hong Kong" in *The Economic Journal*, 102

Organisation for Economic Co-operation and Development; SourceOECD database found at www.oecd.org
Patterson, K (2000); *An Introduction to Applied Econometrics, a time series approach*; Palgrave, New York
Pesaran H & B Pesaran (1997); *Microfit 4.0 Windows Version*; Oxford University Press
Riedel, J (1988); "The Demand for LDC Exports of Manufactures: Estimates from Hong Kong" in *The Economic Journal*, 98
RSA (1996); *Growth, Employment and Redistribution: A Macroeconomic Strategy*; Ministry of Finance
Schreyer, P & F Koechlin (2002); *Purchasing Power Parities – Measurement and Uses*; OECD Statistics Brief
Senhadji, A & C Montenegro (1999); *Time series analysis of export demand equations: a cross country analysis*, *International Monetary Fund Staff Papers*, vol. 46 no. 3
Shiells, C (1991); "Errors in Import-demand Estimates Based Upon Unit-value Indexes" in *The Review of Economics and Statistics*
South African Reserve Bank; time series database found at www.resbank.co.za and via the University of Cape Town
Statistics South Africa (1995); *PPI Base: June 1995=100, January 1971 – June 1995, Statistical* Release P0142; Pretoria
Statistics South Africa; *Quarterly Bulletin of Statistics*, various issues; Pretoria
Statistics South Africa; *Statistical Release P0142*, various issues; Pretoria
Statistics South Africa; *Statistical Yearbook*, various issues; Pretoria
Trade and Industrial Policy Secretariat; database found at www.tips.org.za
Tsikata, Y (1999); *Liberalization and Trade Performance in South Africa, Discussion Paper 13*; World Bank Southern Africa Department
Varian, H (1992); *Microeconomic Analysis, third edition*; WW Norton & Company, New York
Wood, E (1995); *Essays on Trade and Manufacturing in South Africa*; University of Cambridge PhD thesis