CINTRAFOR

Working Paper 76

Time Series Methods for Commodity Price Forecasting: An Application to Market Pulp

Gerard Alexander Malcolm

CENTER FOR INTERNATIONAL TRADE IN FOREST PRODUCTS
UNIVERSITY OF WASHINGTON
COLLEGE OF FOREST RESOURCES
BOX 352100
SEATTLE, WASHINGTON 98195-2100

CINTRAFOR

Working Paper 76

Time Series Methods for Commodity Price Forecasting: An Application to Market Pulp

Gerard Alexander Malcolm

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science
University of Washington

1999

CENTER FOR INTERNATIONAL TRADE IN FOREST PRODUCTS
UNIVERSITY OF WASHINGTON
COLLEGE OF FOREST RESOURCES
BOX 352100
SEATTLE, WASHINGTON 98195-2100

Technical Editor: Edward L. Jenkinson, Kendall Carson, Nicole Stevens

ABSTRACT

Time series methods for commodity price forecasting:
An application to market pulp

The goal of this research is to assess the usefulness of cointegration analysis and related time series techniques for forecasting commodity prices. The analysis focuses on market pulp, a typical commodity. Important short-term factors in determining pulp prices include capacity utilization, the shipments rate, and inventories. Important long-term factors include investment behavior and costs of production. Autoregressive, moving average (ARMA), vector autoregressive (VAR) and error correction models of price and these variables are developed. Market scope is defined to include all North American and Scandinavian (Norscan) chemical paper grade pulp, and the sample period is 1976-1991. Out of sample forecasting performance of the error correction models is no better than that of the VAR model, according to the RMSE criterion. However, the forecasts generated by the error correction models have the property that sensible long-term relationships between variables are maintained. In addition the error correction models are more amenable to incorporating expert knowledge into the model-based forecasts. It is thus concluded that error correction models are useful for forecasting commodity prices.

Program authorized to offer degree: College of Forest Resources

Table of Contents

LIST OF FIGURES

List of tables

ACKNOWLEDGEMENTS

I wish to offer particular thanks to Timo Teräs for his invaluable insights into the workings of the industry, and to John Perez-Garcia for his patience with my frequent changes of topic. Thanks are also due to Antti Rytkönen, Rick Gustafson, Bruce Lippke, Barry Malcolm and Gary Helik for helpful comments and suggestions, and to Eero Ylatalo, Mikko Tahvanainen, Nina Gullstén and Marya-Leena Kilpeläinen for supplying data. Any errors in the results are my own. Finally, thanks to Alice for her understanding and support.

DEDICATION

To my parents.

1 INTRODUCTION

PULP (pulp) n. 1. A soft, moist, shapeless mass of matter. 2. A magazine or book containing lurid subject matter and being characteristically printed on rough, unfinished paper.

American Heritage Dictionary, New College Edition (1993)

Many commodity products have similar characteristics. They are relatively homogenous products that can be produced using readily available technology, and are traded internationally on competitive markets. They are based on natural resources, the availability of which is subject to shocks. They are intermediate inputs with few short-term potential substitutes. As a result of these characteristics, commodity prices tend to be volatile.

Market pulp is in many ways a typical commodity. It also has some particular characteristics that contribute further to price volatility. These include high capital intensity, long-lived capital equipment, and speculative inventory management behavior on the part of consumers. Because of this volatility, accurate forecasting of market conditions is a difficult but potentially very useful exercise. Investment decisions are made on the basis of price forecasts, and improving their accuracy can lead to better decision making on the part of producers. This may in turn reduce the volatility of prices.

Considerable energies are devoted to the task of forecasting pulp prices. A number of recognized short-term leading indicators for price movements exist, and these are widely studied by industry participants. However, less attention is paid to forecasts of the more distant future. This is surprising, given that the relevant time horizon for some of the most important decisions made by the industry (namely, investments in new capacity) is at least five years. This may perhaps be explained by a relative dearth of useful long-term leading indicators on which to base forecasts. The purpose of this research is to contribute to our understanding of how the market functions and to our ability to forecast market conditions, in both the short term and the long term.

- The first step in achieving this is to identify the most important causal factors in determining market conditions. This is done by means of a survey of existing literature and discussion with industry experts. The focus of this survey is on prices and the variables that interact with prices.

- The second step is to assess how best to model these factors in such a way as to be useful for forecasting purposes. Initially, it was intended that this be done using structural methods, focusing particularly on forecasting investment levels (and then prices). In the course of developing the research proposal, however, it became apparent that recently developed time series methods provided a superior means to this end.

Since the late 1980s, time series econometrics has undergone something of a revolution. Two key concepts are non-stationarity and cointegration. Broadly speaking, the first of these refers to a variable which does not have a fixed trend, and the second refers to two or more such variables moving 'in parallel' in the long term. Non-stationarity occurs when a variable is subject to stochastic shocks that have permanent effects on it. It has been recognized that many economic variables (including commodity prices) do exhibit non-stationarity, and it has been found that previously used methods based on the assumption of stationarity are not valid for modeling such variables. Methods have been developed to test for non-stationarity, and to model systems of non-stationary variables.

Cointegration methods provide a means for the modeler to take advantage of the effective treatment of short-term dynamics, which time series models provide, while ensuring that long-term forecasts have sensible properties. The concept of cointegration is an intuitively sensible one. A pair of variables is said to be cointegrated if they have a tendency to maintain a fixed 'equilibrium' relation to one another over the long term. In a stable cointegrating relationship, the variables will adjust to eliminate any divergence from this equilibrium. Such relationships make a great deal of sense for several variables of relevance to the present case, and forecasts for these variables which show significant and long lasting violations of these relationships are not plausible. Cointegration methods allow us to test for and model the existence of long-term equilibrium relationships between the variables in a system, within

the framework of a dynamic time series model. A focus of this research is to assess whether these methods prove to be of practical use for building forecasting models.

One difficulty encountered in this research is that little previous time series research on pulp markets has been published.[1] We are therefore obliged to adopt an incremental approach to model building. First, the time series characteristics of pulp prices themselves are explored. We then develop a series of multivariate time series models. These models will attempt to capture some of the short-term and long-term processes that are thought to determine prices. The success or failure of each model will be judged according to three criteria:

- Its statistical acceptability as a representation of the data generation process;
- The plausibility of its estimated coefficients;
- And its out of sample forecasting performance.

Forecasting performance is assessed both on the basis of a formal measure of forecast error, and on an informal assessment of the desirability of the properties of the forecasts.

While these models are designed for forecasting purposes, they may also be used to an extent for structural analysis. This research is not designed primarily to test any specific behavioral hypotheses, but where convenient, structural analysis will be used to provide insights into how the market functions.

The structure of this thesis is as follows. A description of the pulp market is provided in chapter 1. This outlines the short-term and long-term dynamics of the market. It also addresses the issue of how the 'market' should be defined in terms of its product and geographical scope. Chapter 2 provides a brief overview of the econometric concepts mentioned above, along with a description of relevant modeling methods. Data availability for the variables found to be of importance is addressed in chapter 3. The core of the research, which consists of the estimation and testing of several models of pulp prices and other variables, is described in chapter 4. The forecasting performances of the different models are compared in chapter 5. Chapter 6 concludes the study.

---

[1] A considerable amount of econometric analysis has been done, but this has almost exclusively not focused on forecasting. Exceptions include Young (1987) and Brannlund *et al* (1999), but neither of these uses the methods described herein. An important structural model is NAPAP (see Zhang *et al*, 1996).

2 THE PULP MARKET

"As paperbacks and comics took over each end of the spectrum that the pulps served, the industry saw a decline. Yet if you could define the 'heyday' of the pulps, you would have to proclaim the 30's as the 'pulp 30's.' From depression through the start of World War II, the pulps helped millions escape from their troubled lives."

J.P. Gunnison.

This chapter reviews available evidence on several topics. As a precursor to a discussion of the industry, the product of interest itself is described. The purpose of doing this is to determine an appropriate scope for the market to be analyzed. The structure of the pulp industry is then examined, along with the implications this structure has for interactions between different variables. The focus of the models to be developed is price. Price is affected by many factors, and evidence (from both anecdotal and published sources) will be collected on which of these are most important, how they are determined, and the linkages between them. Much evidence is available only for the pulp and paper industry in general (rather than for pulp producers specifically), and this is reported unless it is apparent that it will not be valid for the pulp industry. Short-term and long-term behaviors are examined separately. This information will form the basis for the model construction process that follows.

2.1 MARKET DEFINITION

As will become clear in chapter 3, it is necessary to define the product (and corresponding market) to be modeled. This definition should be made such that all products that can be reasonably considered part of a 'unified' market are included within its scope, and all products that are not part of that same market are excluded. Pulp is not a perfectly homogenous commodity, but some types of pulp are sufficiently close substitutes that it makes sense to treat them as homogenous for modeling purposes. Scope must be defined with reference to pulp types and geographical regions of production and consumption.

This problem can be restated in more concrete terms. The central point of the market to be analyzed is chosen to be the *cif* price for NBSK delivered to Northern European ports.[2] This is a benchmark price that is widely reported and analyzed. If we aim to forecast this price, what is the appropriate scope for other variables that we might use to do this? Is it wider than just Northern Europe, and if so, how much wider? And does it also extend to other types of pulp beyond NBSK?

There are no unique answers to these questions. Choice of an appropriate scope is a matter of judgement that is usefully informed by consideration of a number of factors. It is necessary to consider the physical properties and uses of the different types of pulp (which will determine whether they are potentially close substitutes) and the openness of markets to inter-regional trade (which will determine whether any such potential can be realized).[3]

2.1.1 Physical Properties of Pulp

Pulp is normally distinguished by the species used, the pulping process, and whether or not it has been bleached.[4] Each of these will affect its physical properties. The raw material used in its production is most commonly wood, but other materials may also be used. Wood pulp consists of fibers made up primarily of cellulose and

---

[2] The *cif* price includes production cost, insurance and freight costs. It is the delivered price to the dock of an importing country.

[3] An alternative approach is to consider whether market outcomes (specifically, prices) are similar between countries/products. Statistical methods to test this are available. These are discussed in chapter 4.

[4] Thus, for example, NBSK is bleached pulp produced from Northern softwood species using the Kraft process.

hemicellulose molecules. The fibers also contain residual amounts of lignin. Important physical properties of pulp (from the perspective of the user) include strength (resistance to tearing and piercing) and brightness.

Strength is determined primarily by the length of the fibers present in the pulp, and the extent to which they have been damaged in the pulpmaking process. Fiber length differs by tree species.[5] Softwoods generally produce longer fibers than hardwoods. Fiber lengths are also reduced during mechanical or semi-mechanical pulping processes. Recycled (de-inked) pulp contains shorter fibers than the equivalent virgin pulp.

Brightness depends primarily on the amount of lignin in the pulp. Because chemical pulping is based on removal of lignin from the wood, while mechanical pulping is not, the former usually have better brightness characteristics. Brightness also depends upon the bleaching process that is used, and the tree species used.

The most popular chemical pulping process is the Kraft or sulfate process. Kraft pulp can be manufactured and bleached under a range of different chemical conditions. These will affect its physical properties, as will the differences in fiber characteristics noted above. Kraft pulp is usually identified according to the fiber type used (softwood or hardwood) and whether or not it is bleached. In addition, for softwood pulp, 'northern' and 'southern' softwoods are differentiated. Within each pulp type, there is not complete homogeneity, although any differences are generally less significant than the differences between pulp 'types'.

The product we consider first is NBSK pulp. This is produced mainly in Scandinavia and North America. It is made from a range of species, including spruce, fir and Douglas fir. It is used primarily for producing both printing and writing papers (which are differentiated according to whether or not they contain mechanical pulp, and whether they are coated or not) and also tissue paper. Table 1.1 shows common furnish mixes for some of these papers.

Table 1.1: Furnish Composition of Paper Grades

	Softwood Chemical Pulp	Hardwood Chemical Pulp	Mechanical Pulp	De-inked Recycled Pulp
Newsprint (waste fiber)			0-50	50-100
Newsprint (virgin fiber)	0-10		90-100	
Uncoated mechanical	10-25		75-90	
Coated mechanical	30-40		60-70	
Uncoated woodfree	20-50	50-80		
Coated woodfree	30-60	40-70		

Source: Diesen (1998). Chapter 5, Table 2.

The closest substitute for NBSK is southern bleached softwood Kraft pulp (SBSK). This is produced using pine species from the Southern US and the Southern Hemisphere. These species have somewhat shorter fibers than the species used for NBSK, but overall the resulting difference in strength properties between the two is not great.

As Table 1.1 shows, another substitute for NBSK is bleached hardwood Kraft (BHK) pulp. Species used for producing BHK vary in different parts of the world. In Norscan countries, species used include birch and aspen, while in Brazil and Indonesia plantation-grown eucalyptus and acacia are used as well as native hardwoods. BHK is generally weaker than NBSK but it provides better printing qualities.[6] Advances in papermaking technology

[5] It also differs intra-species according to geographic location; within the tree according to whether it is juvenile or mature wood; and whether it is heartwood or sapwood. These differences are less important than inter-species differences, however.

[6] See Rudie (1997).

(specifically, the use of vertical drying rolls which place less stress on the drying paper) have allowed increasing proportions of BHK pulp to be used in the furnish mix.

Unbleached Kraft pulp is used for production of board products and other applications where brightness is not important. It is not a substitute for bleached pulp in such products as printing and writing papers.

Sulfite pulp is another type of chemical pulp. Different chemicals are used in the pulping process, and it can only be produced from a limited range of species. It is generally more expensive to produce than Kraft pulp, and is used mainly for fine papers.

Mechanical pulps are produced predominantly from softwood species. In the pulping process the wood fibers are broken and thus shortened. In addition, mechanical pulping does not remove lignin from the wood as chemical pulping does, and this makes bleaching more difficult. Various refinements of the mechanical pulping process have been developed to utilize high-pressure refiners and/or chemical pre-treatment of the pulpwood. These hybrid processes (notably BCTMP) produce stronger pulp than mechanical pulping. BCTMP is a substitute for NBSK, since both can be used in the furnishing of newsprint and mechanical printing and writing papers.[7] Mechanical pulp is cheaper than Kraft pulp (by a margin of approximately $150 per ton for BCTMP) so the amount of NBSK used in applications where mechanical pulp is potentially a substitute is generally limited.

Several other types of pulp exist, including fluff pulp and dissolving pulp.[8] However, these are not close substitutes for bleached Kraft pulp, as they are used in quite different products.

In summary, all chemical grades of bleached pulp are potential substitutes for NBSK in most products in which NBSK is used. SBSK is the closest substitute for NBSK. Anecdotal evidence on price setting behavior suggests that BHK be also unified with the NBSK market. This is not true of non-chemical pulp types (although semi-mechanical pulps are substitutes in some products). The most suitable market definition seems to be 'bleached chemical pulp', i.e. encompassing NBSK, BHK of all hardwood species, SBSK, and sulfite pulp.

2.1.2 International Tradeability And Market Access

In addition to examining potential substitutability between products, it is also necessary to consider whether markets for these products are unified across national boundaries (we are assuming that markets within countries are unified). Both barriers to exports in producing countries and barriers to imports in consuming countries are relevant. As far as is known, no export barriers exist in major pulp producing countries.

Barriers to imports may be either 'natural' factors that cause a product to be intrinsically non-tradable, or may be constructed intentionally for the purposes of limiting or preventing trade. Pulp faces few natural barriers, and can be regarded as a tradable good.[9] Transport costs usually amount to no more than $40-50.00/per metric ton. Intentional barriers to trade generally take the forms of formal tariff or quota systems imposed by importing countries, and other less formal (and less visible) barriers. Most developed countries do not impose high tariffs on pulp. Often, the fact that pulp is a relatively low value added product means that it is subject to lower tariffs than paper. We briefly consider whether non-tariff barriers exist for each of the major consuming countries.

---

[7] See Gundersby and Diesen (1985).

[8] Smook (1992) gives a detailed description of the properties of different pulp types.

[9] Christensen and Caves (1997) based their market definition of pulp and paper on the tradability of the products. They examined data on average distance shipped, percent of sales exported, and transport costs. They concluded that an appropriate market definition for all pulp and paper commodities in aggregate is the United States and Canada. Interestingly, they also noted that not much trade occurs between regions of North America, but that arbitrage nevertheless occurs when price differentials arise.

The most important pulp producing and consuming regions are Europe and North America. Free trade areas exist within each of these regions, and there are no significant barriers to trade between them.

Japan is another important region. It is a net importer but also produces a significant amount of pulp domestically (much of it from imported wood fiber). In general, Japan is well known for its trade policies and related institutions.[10] Japanese paper markets are not generally regarded as being fully integrated with global markets. Trade volumes (both exports and imports) are very low relative to the size of the Japanese industry and consumer market. This can be explained by various institutional factors, including the existence of cross-shareholding links between paper and distribution companies, and the resulting reluctance of the latter to handle foreign-produced paper. In the case of pulp, however, these barriers are not relevant, as pulp is sold directly to paper mills rather than to paper users, without the need to go through these distribution channels. In addition, a large number of Japanese-owned pulp mills in other countries exist, whose production is dedicated to supplying their parent companies in Japan, and who do not face these informal barriers to entry. Japan also purchases market pulp from non-affiliated producers. Therefore we conclude that the Japanese pulp market is unified with the international market.

China, like Japan, has a large economy and limited forest resource. Thus it is also a natural importer of forest products, including pulp.[11] Prior to 1978, China had a closed economy. Since then, it has gradually become more open, and is likely to continue to do so in the future. Nevertheless, many barriers to imports remain, both formal and informal. The industrial structure that exists in China (including a limited number of approved import companies, state ownership of most paper mills, very small scale of most mills, and the existence of non-market prices domestically) makes it unlikely that China is fully integrated with the international market. Nevertheless, China does purchase market pulp, and for this reason is best regarded as a part of the global market.

Similar conclusions may be drawn with regard to other Asian markets.[12] Some of these economies are producers of market pulp, while others are consumers only. Markets are not fully open, but this less affects pulp than is paper. Idiosyncrasies in production systems and demand are less of a barrier to commodity products such as pulp than to more specialized products.

In summary, all of the regions examined appear open to imports of pulp (if not necessarily of paper). For this reason a market definition with global scope is most appropriate.[13]

2.1.3 Market and Integrated Pulp

Another issue related to market definition is that of market vs. integrated pulp – should they be treated separately? Most pulp is produced and used at the same site, within integrated mills.[14] It is not traded on any market, and hence no price for this pulp exists. The issue here is similar to that addressed above: In developing a model of pulp prices,

---

[10] See Tilton (1996), for example.

[11] A caveat: China is a large paper consumer, but not as large a pulp consumer. This is because much of China's demand is met from non-wood or recycled fiber sources.

[12] For example, Johansson (1996) observes that "[small Asian mills] have been able to survive because of specific features in the domestic distributions systems. Fundamental differences in the technology used by the printing and converting industries in various countries, as well as heavy protection against foreign competition in the form of duties or tariffs, have both played an important role in this respect."

[13] In fact, because the models constructed here focus primarily on the supply side, the issue of the openness of consuming regions becomes moot. Nevertheless, for a comprehensive model of the market containing demand and supply side variables, this discussion is relevant.

[14] In the US, for example, only 14% of total 1997 production was market pulp, while the remainder was used internally, either at the same (integrated) mill or elsewhere.

should we include the non-market part of the industry? This decision should be based on an assessment as to whether or not events occurring in the non-market part of the industry will have an effect on the market price.

Physically, market and non-market pulps do have some differences. Market pulp is dried into sheets in preparation for shipping, whereas non-market pulp is piped in liquid form directly to the paper mill. This drying does have some effect on the paper produced, being bulkier and having greater tear strength than non-dried pulp, but being weaker in burst and tensile strengths. Market pulp is sold on the basis of viscosity and brightness, and its production therefore focuses on meeting targets for these characteristics, whereas for non-market pulp, these characteristics are only important insofar as they affect the final properties of the paper produced. Overall, however, these differences are not great.

The definition of market pulp does not include all market participants. Not all integrated mills are completely isolated from the market. They may purchase some pulp on-market, or may sell some of the pulp they produce. This may be done on a regular basis, or only occasionally (if market conditions or imbalances between their pulp and paper production levels make it attractive to do so). Conversely, not all producers of what is defined as 'market pulp' are in fact operating within the market. Some producers sell to affiliated companies in other countries.[15] These producers may well be detached from the market. Thus, the boundary between producers who are affected by (and affect) market conditions and those who do not is blurred, and does not exactly match the boundary that separates market and integrated pulp. Nevertheless, using the latter division seems preferable to including all integrated pulp producers in the market scope.

2.2 SHORT-TERM MARKET BEHAVIOUR

In this section, we focus on the price-setting process, and attempt to identify the variables that affect this process. A number of different price concepts exist. Market pulp is traded both on a spot and contract basis. Contracts are written on the basis of list prices announced by producers. However, actual transaction prices are often somewhat lower than the publicly announced list prices (depending on market conditions).

Bulk grades of pulp are typical 'commodity' products, with few quality factors to differentiate between the products of competing producers (as noted in the previous section, some quality differences do exist within each grade, however). In addition there is little loyalty to particular buyer-seller relationships. The only factor preserving relationships appears to be the existence of transaction specific assets such as dedicated warehouses.[16] Thus there is little to prevent buyers from switching suppliers if prices differ, and as a result prices for a particular grade/location of pulp are usually similar across producers.

Price change announcements are initiated by 'price leaders', who are usually large producers.[17] The set of companies who act as price leader is not necessarily stable over time.[18] A producer will not necessarily adhere to an announced price increase because of the risk of losing market share if other producers do not follow suit. Often,

---

[15] All pulp transported across national boundaries is included within the definition of market pulp, regardless of any ownership or other links between buyer and seller.

[16] Christensen and Caves (1997).

[17] For example, three Norscan major producers announced NBSK price increases in August 1999 almost simultaneously: Weyerhaeuser, Sodra Cell, and Metsa-Botnia. Aracruz, a major Brazilian producer, announced an increase in the BHK price at the same time.

[18] Steele (1995), for example, provides a fascinating description of pricing behavior in the newsprint industry, and his empirical work supports the existence of oligopolistic pricing. Certain large producers act as price leaders, and other producers tend to follow any price changes that they make. Non-leader firms do not have the same ability to drive a market-wide price increase, and will generally not move above industry prices themselves for fear of losing customers.

price increases will be only partially successful: Prices will increase slightly but not to the full extent of the increases initially announced.

Christensen and Caves (1997) provide a useful description of how pulp and paper markets operate in North America (and this description seems to apply equally well to European markets):

"Sellers are few enough that changes in list prices (apart from reflecting the prevalence of discounts) involve interacting public initiatives, responses, and revisions. An individual firm announces a change (with up to several months lead-time); rivals either do or do not match; and buyers give varying signals about their determination to resist. Initiated price increases are not uncommonly canceled or at least postponed, and firms have reported temporary losses of profit due to unsuccessful efforts to make price increases stick. Price adjustments sometimes begin and spread regionally, and some announced changes apply only to limited regions."

As markets go, that for market pulp is quite close to the concept of a perfectly competitive market. Neither buyers nor sellers are particularly concentrated.[19] Different pulp types exist but within each type there is a high degree of product homogeneity. Price is the primary competitive factor. All agents in the market have reasonable (although not perfect) information about prevailing market conditions.

Exchange rate movements will result in pressure for price changes. NBSK prices are usually denominated in US dollars, both in North America and Europe. BHK prices are denominated in Euros, in Europe. Because these are substitutes, a strengthening of the US dollar against the Euro will lead to downward pressure on the US dollar price of NBSK. Also, both paper prices and prices of inputs into pulp production are commonly denominated in local currencies. This means that a strengthening of the US dollar exchange rate will be beneficial for non-US pulp producers, and detrimental to non-US pulp users. As a result, there will be further downward pressure on the US dollar price. In reality, the US dollar will often move differently in relation to individual currencies. However, we can make the general conclusion that a strengthening of the US dollar will cause downward pressure on the dollar-denominated pulp price.[20]

Price changes are implemented at least partly on the basis of expectations about future market conditions. This makes it difficult to pin down exactly which factors determine prices, as a wide range of variables may affect expectations. However, market participants commonly focus on three leading indicators for how prices may change in the immediate future: Capacity utilization, inventory levels, and shipments. Capacity utilization can be regarded as an indicator of the supply side conditions of the market, and shipments as an indicator of the demand side conditions. Inventories are an indicator of the balance between supply and demand. Each of these is discussed in turn in the following sections.

2.2.1 Inventories

Both producers and users hold inventories. Producers will generally desire to hold a low inventory level. Users will desire to hold enough inventories to meet their paper production needs for a certain period of time. If there is a mismatch between production and consumption of pulp, this will be reflected in a change (either voluntary or involuntary) in inventory levels.

---

[19] For example, Diesen (1998, 77) shows 'forest products' as being the industry with the lowest top-five concentration ratio among 13 major industry groups. Separate data for the pulp industry itself was not located.

[20] The effect of the exchange rate is not central to this analysis, and it is not addressed in the modeling section. A simple regression of the US exchange rate (proxied by a trade-weighted real exchange rate for the US dollar) on price did reveal some explanatory power, however.

If a pulp mill is receiving a price above its marginal cost of production, it will generally be unwilling to curtail production in response to involuntary inventory buildup. Instead, its response will be to offer a lower price, either formally by reducing the list price, or informally by offering increased discounts.[21]

If inventories of a particular grade of pulp or paper fall below what is perceived to be the 'normal' level, this provides a signal of market strength. The converse is also true. Inventories data are reported regularly and are closely watched. The normal level for Norscan producer inventories of NBSK is considered to be 1.5m tons. Teras (1997) discusses the market pulp market specifically:

"Whenever inventories are well above 3-week production potential, prices tend to fall. And, whenever the inventories fall close to or below the 3-week production potential level, prices move up."

As well as involuntary inventory changes, user demand for inventories can vary quite substantially due to speculative motives. In general, pulp is held by producers and users, and not by traders whose goal is to sell these inventories at a profit. Nevertheless, users do engage in speculative behavior, by increasing or reducing their inventory levels in response to expectations about future prices.[22] In general, expectations appear to be formed adaptively; i.e. current price trends are expected to continue. This can cause expectations to be self-fulfilling: "When prices start going down no-one will buy in order not to have overvalued pulp or paper in their stock. Thus the prices drop even faster." (Gundersby 1996). This process does not continue indefinitely, but it does exacerbate the cyclicality of the market. This speculative behavior has been identified as one of the two main causes of cyclicality in the market.[23]

When market conditions are good, market psychology is also affected by the recognition that production is effectively limited by capacity:

"The papermaker – seeing apparent orders mounting, seeking pulp price hedging, and nervous about the fixed capacity of the pulp pipeline – moves fast to secure pulp stocks, and equally fast to destock, while doggedly maintaining paper production. Merchants – also believing the [supply] chain to be inflexible – exaggerate order levels into waves, each merchant seeking apparent supply security and price hedging, but in actuality stretching the chain to its limits." (Wilson, 1997)

Seasonal patterns in both demand and supply also tend to have an effect on inventories. On the demand side, paper production is subject to seasonal fluctuations. Particularly in Europe, paper production is low during the northern summer (July and August) as paper mills take holidays. Demand for paper is also low during this time, and is high prior to the Christmas period and also early in the calendar year as government organizations make purchases. Usually, pulp producer inventories increase during July and August, and market participants regard this as a period of weak prices.[24]

---

[21] For example, after signs of weakness in the market in mid-1995, papermakers reduced purchases in order to reduce inventory levels. Norscan pulp producers, however, did not respond by cutting production (at least not immediately), and instead saw their inventories rise dramatically (from 1.5m tons in June 1995 to 4m tons in February 1996). See Fromson (1997).

[22] Speculative inventory behavior also occurs on the part of some producers. Notable are Japanese trading houses (some of whom own market pulp mills).

[23] See Diesen (1998). The other main cause is bunching of investments at particular periods in time.

[24] Another timing issue is that shipments do not occur until agreement is reached between producer and buyer. When agreements are made on a quarterly basis (as has usually been the case) this may cause shipments during the first month of the quarter to be lower than the following two months. This effect will not cause seasonality in quarterly data, however.

On the supply side, there are a number of seasonal influences. Norscan producers are not willing to shut down during the northern winter because of the physical problems caused by doing so in cold temperatures. Maintenance downtime is often scheduled for Easter. Although the industry is based ultimately upon a natural resource, this does not play as important a role as it does for annual crops. Trees may be harvested at any time of year, and while severe weather may prevent harvesting, mills are usually able to accumulate sufficient chip piles to avoid any supply shortages.

In summary, an 'equilibrium' level of inventories is widely perceived to exist, and *ceteris paribus*, departure from this level will lead to price changes. However, pulp users may be willing to make short-term departures from equilibrium levels of inventory for speculative or other reasons.[25] Non-economic factors may sometimes also impinge upon this equilibrium.[26]

2.2.2 Capacity Utilization

A definition of capacity is:

"The tonnage of pulp and paper of commercial quality that could be produced per year with the full use of equipment and adequate supplies of raw materials and manpower, and assuming full demand. No allowance is made for losses due to unscheduled shutdown, strikes, temporary lack of power, etc., which cause decreases in actual production, but not in productive capacity."[27]

The number of operating days achievable per annum varies according to the specific machine type, and the pulp and paper industry, as a whole, averages 348 days.[28] Capacity utilization is the ratio of actual production to capacity. Historically, wide variations in capacity utilization have been observed. In the US (in the 1970s), "experience suggests 96% is the practical maximum and 92% is the level at which normal service and product range can be provided."[29] For the US pulp and paper industry as a whole, capacity utilization has generally fluctuated between 80% and 95% during the post-war period. However, lower rates have been experienced in other countries. For example the Japanese pulp industry, for example, capacity utilization fell from 83% in 1989-90 to 67% in 1993.[30] This was due to substantial capacity growth during the bubble period of 1986-1990, followed by a slump in demand.

As mentioned above, producers generally wish to maintain as high a utilization rate as possible. This is because of the capital-intensive nature of the industry, and is especially true of companies with high debt levels. Mills will shut down for maintenance, for extraordinary circumstances, or when it becomes unprofitable to run them. Maintenance is usually done on a regular basis, but may be postponed if prices are high. Extraordinary circumstances include machinery failures, labor stoppages and environmental issues.[31]

---

[25] An example of the latter we can look at the 'millennium effect', where pulp users maintained higher-than-usual levels of inventory to reduce the risk of delays in the delivery system affecting their production.

[26] Between 1975 and 1977, for example, Swedish producers decided for political reasons (namely, the government implemented a financing scheme to assist full production) to continue producing at full capacity despite weak market conditions. This led to a huge buildup of inventory in Sweden (of over 1 million tons).

[27] American Paper Institute, quoted in Lee, 1986.

[28] Lee (1986, 6).

[29] URS (1975) quoted in Lee (1986).

[30] Whitham (1994).

[31] For example, a pulp mill in Indonesia was shut down for four months during 1998 and again in 1999 because of

Capacity utilization is also affected by market conditions.[32] A mill becomes unprofitable to run when prices fall below the variable production cost of that mill. Theoretically, in the short term, a producer will only produce pulp if the anticipated selling price exceeds the variable costs of production and transportation to the point of sale.[33] If it falls below this level, the producer will shut down temporarily (or potentially permanently). Once the price rises above the level of costs, the producer will restart production. Thus, at an aggregate level, marginal producers will enter or exit the market as price changes, causing changes in aggregate production. The least efficient mills tend to shut down for economic reasons most frequently, and efficient mills do so less often if ever. Prices may fall to a level equal to the variable cost of efficient producers, but generally not below that.[34]

In practice, a desire to preserve customer goodwill or market share, or political or other non-economic considerations may all interfere with this pattern of behavior. In addition, producers who sell pulp to overseas affiliates will not be motivated solely by the profitability of the pulp mill. Thus, in many cases production decisions will not be based solely on the balance between price and marginal costs. For this reason, it does not seem appropriate to define any structural relationship between the balance between price and marginal costs on the one hand, and capacity utilization on the other.[35]

However, it is clear that prices will have some effect on capacity utilization, and conversely will also affected by capacity utilization. From their survey of trade journals, Christensen and Caves (1997) concluded that

"A strong influence on price movements is exerted by current and expected capacity utilization, with 93 percent widely recognized as the threshold at which some mills face excess demand, and price increases become quite likely."[36]

Because of the impracticality of operating at above 100% capacity when demand is strong, and the desire to maintain high capacity utilization rates when demand is weak, modest movements in capacity utilization can be associated with large movements in prices (P.J. Ince, 1999, pers. comm.). Thus, as for inventories, an 'equilibrium' level is recognized, and departures from this will result in pressure for price changes.

2.2.3 Shipments

Shipments are defined as the volume of pulp shipped from producers to consumers of pulp.[37] This is an indicator of the level of demand for pulp (both fundamental and speculative) and as such its importance for a price is obvious.

---

popular protests over its environmental impacts. Strikes are reasonably common events, notably in British Columbia.

[32] Constantino and Townsend (1986) present a model of producer behavior within which capacity utilization is the primary means by which companies respond to market conditions.

[33] Within North America, prices are quoted as *fob* at the paper mill. Transport costs are therefore borne by producers.

[34] Bernstein's (1992) study of the Canadian pulp and paper industry provides supporting evidence for this statement.

[35] Potentially, data is available to do this. See section 4.5.

[36] They also carried out a brief empirical test that accepted their hypothesis that price varied with marginal cost at capacity utilization levels below 93%, and increased sharply once capacity utilization exceeded that level.

[37] A caveat: Producers who are located far from their customers commonly ship their pulp initially to a consignment warehouse nearer to the consumers, prior to any actual sale being agreed upon. This practice has become increasingly common over time, as it allows for more rapid delivery to the consumer, and storage usually costs the producer little if anything. In some instances pulp shipped in this way is included in the reported shipments

Shipments also affect capacity utilization, as a second prerequisite for production (other than that it is profitable to do so) is that the producer be able to sell the pulp produced. Although this research is not intended to address the demand side of the market in any detail, it is useful to explore how demand is determined, in order to assess the extent to which shipments will be affected by prices and the other variables discussed above.

Fundamental demand for bleached chemical pulp comes from producers of certain grades of paper and paperboard, predominantly printing and writing papers. The total demand for pulp (i.e. for both market and integrated pulp) is related directly to the volume of paper production, which is determined by the level of installed papermaking capacity and paper mill capacity utilization. In the long term, paper production is driven by demand, which is related primarily to levels of overall economic activity, and to secular trends in usage of paper of various sorts. Gundersby and Diesen (1985) identify several factors important in determining paper consumption: population, per capita GDP, consumption patterns, usage of electronic media, and usage of plastics. Changes in the furnish mix used in paper production also affect the demand for pulp.

In addition to these 'fundamental factors', short-term factors also influence actual demand for pulp. Teras (1997) lists several:

- The net demand for market pulp from integrated papermakers;
- The inventory situation
- Exchange rates;
- And other factors that affect market sentiment, including production curtailments, strikes or threats to strike, political or military conflicts, wood supply and weather.

In summary, the demand for market pulp is the result of many different influences, some affecting the production of paper and some affecting the degree to which market pulp is needed for its production. In the short term, there are very limited opportunities for substitution between pulp and other inputs, and consequently the price elasticity of eliminate spacing demand is low.[38] Speculative demand for pulp appears to be directly influenced by price movements, since the latter are important in affecting expectations of future price movements. Overall, then, we can expect price to have some effect on shipments, although the effect may be limited in magnitude. There is no reason to expect inventories or capacity utilization to have a direct effect on shipments, except insofar as these affect expectations about prices.

2.2.4 Summary

For the purposes of model building, we wish to identify the most important influences on prices, and also how other variables are themselves influenced. Price is affected by inventory levels, capacity utilization, and shipments. A speculative feedback effect also exists under which price changes themselves may cause further price changes.

Capacity utilization is determined by the balance between price and marginal cost of production (as well as by non-economic factors). The marginal cost of production is also determined by capacity utilization: a fall in capacity utilization occurs when the least efficient mill shuts down, and this reduces the marginal cost of production for the industry as a whole. Shipments also affect capacity utilization. Shipments themselves are affected by a host of demand side factors (including price expectations), but not by supply side factors such as capacity utilization. The balance between demand and supply determines inventories (in a strictly accounting sense). Hence, both capacity

---

although it has not actually been sold.

[38] See Gundersby (1996). Quicke *et al* (1990) show that, for the US paper industry, demand for pulp is less elastic than demand for other inputs.

utilization and shipments will affect inventory levels. These linkages comprise the essence of the short-term dynamics of the pulp market. They are summarized graphically in Figure 1.1.

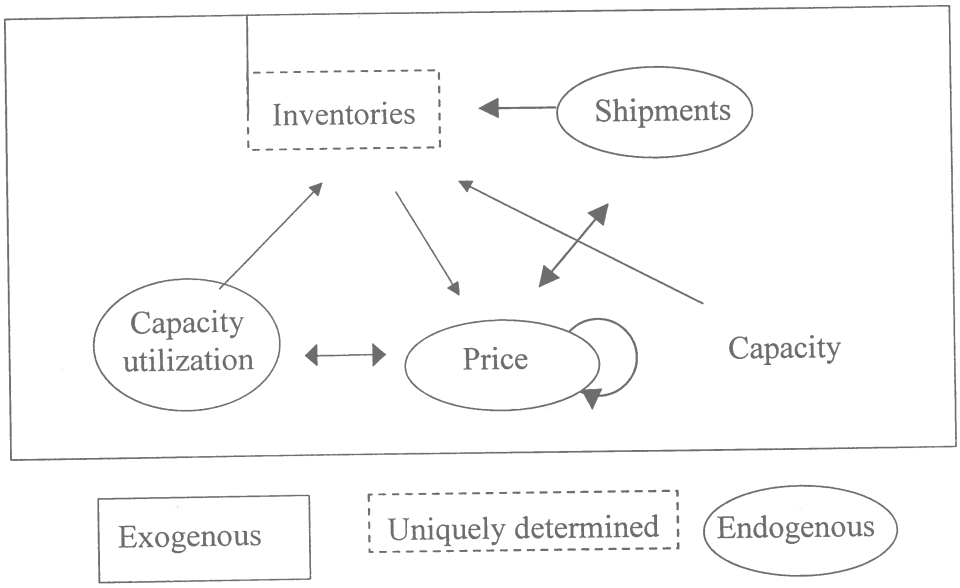

Figure 1.1: Summary of Short-Term Market Interactions

2.3 LONG-TERM MARKET BEHAVIOR

An economic definition of the 'long term' is a period sufficient for companies to make any desired changes to their capital stock and to other inputs to production that are fixed in the short term. It is evident both from anecdotal evidence and from a consideration of the nature of the industry that production capacity is a key factor affecting market conditions. The level of installed capacity at any point in time is the cumulative result of investment and shutdown decisions made by companies. In this section, the process of investment in pulp producing capacity is described, and various theories of investment behavior are presented. Investment decisions in turn are also affected by market conditions. In particular, investments will usually only be made if future prices are expected to exceed total costs (including the costs of capital employed). The importance of total costs is also explored in this section.

2.3.1 Investment

There are several ways in which pulp production capacity changes can occur. These include Greenfield construction; expansion upgrades (upgrading of existing equipment, installation of new equipment in existing facility, or purchase and installation of existing equipment from another site); and shutdowns (permanent closure of existing operations). Old mills may be idled, with the option to reopen if market conditions make it attractive to do so. This is best treated as affecting capacity utilization rather than capacity itself, given that this idling is not irreversible. Pulp lines may be switched from one grade to another. However, this also is often a reversible decision (as least for many modern mills with the capacity to process different types of fiber), and best excluded from consideration of capacity changes.

The required capital expenditures will vary according to the type of investment. For new pulp lines, it will include the cost of purchasing and installing digesters, recovery boilers and other machinery, and any new buildings and infrastructure required. The bulk of this spending will generally be incurred in the establishment phase of the project, i.e. prior to startup. For decommissioning it will include the costs of dismantling the mill, site cleanup, labor severance payments, etc.

The criteria which firms use to assess whether to undertake each of these actions will differ.[39] The effects on capacity of each decision type, and the time lag before the effects occur, will also be different. Firms use various criteria to make investment decisions. These include profitability, wood-paying capacity, cost competitiveness and the firm's ability to take risks.[40]

Construction of new mills is of particular interest and is focused on here. This is because most volatility in investment spending is due to new mills, whereas upgrades occur steadily over time.[41] Relative to market size, investments in new mills are large.[42] This is because investment in new mills is subject to increasing returns to scale up to quite a large scale, and construction of a mill below optimal size results in a significant competitive disadvantage. Capital costs are also procyclical, since a disproportionate number of mills are constructed at cyclical peaks.[43] Time lags involved in building new capacity are substantial.[44] As a result of these lags, new mills constructed at cyclical peaks often commence operations in a less favorable climate than was envisaged by their planners. This in itself contributes to the cyclicality of the industry, in a fashion reminiscent of the classic hog cycle. Industry commentators have advanced various opinions as to why investments are made in this way, and as to whether this pattern will continue in future. In order to make long-term forecasts for market conditions, the ability to forecast investment seems quite important.

'Traditional' Theories of Investment

This section briefly presents some of the most popular economic theories of investment behavior, and summarizes relevant empirical evidence on their validity. For a more complete description of these theories, the reader is referred to Jorgenson and Siebert (1968) or Chirinko (1993).

The accelerator theory is the oldest theory of investment, and simply posits a fixed relationship between production volume and capital stock. In order to reconcile the theory with observed departures from a fixed ratio, it is assumed that adjustment to a desired capital stock occurs slowly. Empirical tests have generally found the accelerator model to be inadequate in explaining investment.

---

[39] For example, the required rate of return for upgrades is normally higher than for greenfield construction, because of less certainty that the benefits will be realized (Diesen, 1998).

[40] See Diesen (1998).

[41] See Christensen and Caves (1997).

[42] Christensen and Caves (1997) calculated the average new market pulp mill scale to be 2.6% of the total market size. This of course depends on how the 'market' is defined.

[43] High investment activity results in bidding up of the prices of various inputs to construction (e.g. machinery components, contractor services). Conversely, when prices are low and investment activity is weak, investments can be made at a significantly lower cost per unit of capacity. Additionally, if the pulp price cycle coincides with the general cycle in economic activity (as shown, for example, by Gan and Kolison, 1997), then funding costs may also be higher in peak periods.

[44] Lee (1986) describes three lags: The decision lag from when conditions for investment are good to the starting of appropriations; the lag between making the decision to invest and the starting of construction; and the time required for construction. Overall, he estimated the total lag time to be 2 – 2 ½ years. Diesen (1998) provides a more recent estimate of the time lags involved: He describes pre-feasibility and basic engineering studies taking 18 months, and construction then taking two years. He also notes that the basic engineering phase for a paper mill project takes 4-8 months, but can take up to 18 months due to application procedures for the required building and environmental permits. Thus the time lag from the initial decision to investigate a project is close to four years, but the lag from the actual decision to invest (and the outlay of most of the investment) is only two years.

The neoclassical theory states that companies choose their capital levels in order to maximize net worth.[45] This is an intuitively appealing starting point because of its consistency with the decision-making tools actually used by companies. Given standard neoclassical assumptions regarding the form of the production function, the optimal capital stock can be shown to be a function of output price, production, and cost of capital. Investment is made so as to reach this optimal level. Capital is normally assumed to be malleable both *ex ante* and *ex post*, although the theory has been modified to recognize that capital is not malleable *ex post*.[46] Gradual adjustment towards equilibrium is again assumed to be due to the existence of adjustment costs associated with changing the level of capital employed. This is an unattractive assumption in the case of the pulp and paper industry: Generally it is attractive to complete any investment project as rapidly as is physically possible.[47]

This theory provides a fully specified investment function at the level of the individual company. However, with idiosyncratic companies this does not correspond to a fully specified aggregate investment function. Also, some of the standard neoclassical assumptions used to derive this function are somewhat unattractive in the present case (notably the assumption of constant returns to scale).[48] Therefore, it is attractive to make use of the general finding that these variables (price, production and cost of capital) are important, but it is not attractive to use the exact functional form that the theory yields.

Overall, empirical evaluation of the neoclassical theory has had mixed results, as summarized by Jorgenson (1996). Specifically in relation to the pulp and paper industry, Lee (1986) and Farimani et al (1988) have tested the theory. Lee evaluated the neoclassical and accelerator models for the US paper industry.[49] He found that the accelerator model did not perform as well as the neoclassical model. Capacity expansion was found to be dependent on output (lagged), capacity utilization, and the interacting effects of the relative prices on output levels. Input prices, cash flow, and profits were found not to be important. Lee also reached the counter-intuitive result that a rate of capacity utilization above 90% had a negative effect on investment.[50] Farimani *et al* (1988) applied the neoclassical framework to the US paper industry. They examined the relative importance of demand and cost of capital in influencing investment, and found the latter to be relatively unimportant.

Q theory was first proposed by Tobin (1969), and focuses on the role of the profitability of capital relative to its cost. The theory states that investment is a function of the ratio (q) of the market value of investment goods to their replacement cost. Because q is a marginal concept, it is impossible to measure, and studies done have generally used an average value instead. This theory has received a great deal of attention, but empirical testing of average-q models has met with mixed results.[51]

---

[45] See e.g. Jorgenson (1996, chapter 1).

[46] See Bischoff (1971).

[47] See Diesen (1998).

[48] Bernstein (1992) found that, for the Canadian pulp and paper industry overall, increasing returns to scale existed (with a factor of 1.15). Stier (1985) and Orr and Lee (1990) also find evidence of increasing returns to scale in the US paper industry.

[49] He also examined two other models that are not described here: the relative price model, and a putty-clay vintage capital model.

[50] Interestingly, the present study finds a negative impact of capacity utilization on price (see section 4.2.4). These two findings may be related.

[51] See for example Devereux and Schiantarelli (1990), Cuthbertson and Gasparro (1995) and references listed therein.

Meyer and Strong (1990) included q in their study of the US paper industry (discussed further below). Its significance varied depending on the method of estimation used. Under fixed effects estimation (the most appropriate of the alternative estimation methods they used) they found it to be significant, and consistent with estimates from studies outside the paper industry. Zhang and Buongiorno (1992) developed an average-q model for US paper and paperboard industries, and found that it performed well in explaining capacity investment. From this it appears that q does have some explanatory role in investment. However, they did not consider other explanatory variables, which may explain their finding.[52]

The cash flow theory is another to have attracted considerable attention. This theory posits that firms are effectively restricted in their investment spending to the amount of cash that they can generate internally, i.e. the amount of their free operating cash flows. A number of reasons have been proposed for why this might be the case, including differing costs of internal and external finance, and asymmetric information causing external funds to be unavailable at any cost.[53] External finance for companies in the pulp and paper industry may be high cost or unavailable because of the generally poor historical returns on investment in the industry. There are, however, means by which funds may be obtained in other ways. Many pulp-producing companies also own timberlands, which generate substantial positive net cash flows. They are thus not obliged to obtain all capital financing from external capital markets. Another possibility is for project financing to be supported by equipment suppliers and other interested parties (anonymous, 1998).

At a macroeconomic level, business investment and cash flows are closely correlated.[54] However, empirical investigation for industry overall has, as usual, given mixed results. Early work did not support the cash flow model. Kuh (1963) concluded that it was inferior to the accelerator model. Jorgenson and Siebert (1968) found it to be inferior to the neoclassical and accelerator models. More recent work has been more supportive of the theory. Devereux and Schiantarelli (1990), using a variety of statistical techniques, found that cash flow was a significant factor in investment decisions in the UK. Samuel (1998) used a range of econometric techniques (including fixed effects estimation) to compare each of the four theories described above (i.e. the accelerator, neoclassical, q, and cashflow models). He investigated how well each theory explained observed investment behavior in the US manufacturing sector over the period 1972-1990. He constructed *ad hoc* criteria to rank the performances of the competing models, and found that models based on the cash flow and neoclassical theories performed better than those based on accelerator and q theories.

These results were obtained from observations across all manufacturing industries. Casual observation suggests that investment behavior is far from identical in different industries. For the US pulp and paper industry since 1970, cycles in cash flow and capital expenditure match quite closely, whereas for the chemical industry, for example, there is no obvious correlation between the two variables.[55] Chapman et al (1996) offers one potential explanation for this. They found that cash flow was a much more important determinant of investment for Australian companies (between 1974 and 1990) that were 'financially constrained' than for companies that were not constrained.[56] In

---

[52] Cuthbertson and Gasparro (1995) noted that "aggregate time series studies that use only average-Q, not surprisingly perhaps, find that it is statistically significant, but studies that introduce other variables generally find that average-Q is either statistically insignificant or that although average-Q remains statistically significant, these 'other variables' (most notably output) tend to dominate the equation both statistically and in terms of their quantitative impact on investment."

[53] See Samuel (1998) or Meyer and Strong (1990).

[54] Caballero (1997).

[55] Butner and Stapley (1996).

[56] They define 'financial constraint' as occurring when net profits are insufficient to cover interest and dividend payments.

contrast q was more important for companies that were unconstrained. The majority of paper companies probably fall into the 'constrained' camp, at least since the mid-1970s.

The cash flow theory has been tested for the pulp and paper industry specifically. Lee (1986) found cash flow not to be an important determinant of capacity growth. Meyer and Strong (1990) presented a variant of the cash flow hypothesis.[57] They analyzed both 'sustaining investment' and 'discretionary investment', and found that determinants of the two types of investment differed. Using data on US firms from 1971 to 1986, fixed effects analysis showed that cash flow was the most important determinant of discretionary investment. Profitability was also found to be significant, while q was found to be insignificant.

In summary, then, both the neoclassical model and the cash flow model have sufficient evidence supporting their empirical validity that it seems appropriate to include the variables indicated by these theories in our treatment of investment. Cash flow should be included. The neoclassical model indicates that price, production volume and cost of capital are important. Production level is idiosyncratic to the individual firm, and therefore does not seem a suitable variable to use in an industry-level model. Capacity utilization seems to be an attractive proxy to use at an aggregate level. Price is generic to the industry as a whole, and is therefore well suited for inclusion in an industry-level model. Although some authors have found the cost of capital to be unimportant in determining investment, it seems appropriate to at least consider it for inclusion. The actual cost of capital will differ from company to company, depending on capital structure and perceived risk, but at least trends in the cost of capital will be common across all companies. Because the models will be estimated at an aggregate level it is not possible to incorporate q into the model framework in anything other than a very crude way. We exclude it, therefore, but are reassured in doing so that the empirical research generally suggests a minor role for this variable in determining investment.[58]

New Developments

In addition to these 'traditional' theories of investment behavior, a number of more recent theoretical areas of research pertinent to investment behavior have been developed. These address the issues of strategic behavior, expectation formation, and the problems of uncertainty and irreversibility.

Because of its capital intensity and the irreversibility of any investments made, the possibility of strategic behavior in the industry is quite plausible. Because scope for new investments is limited, and each investment is of a substantial size relative to the market, there are interactions between firms in making investment decisions. Steele (1995) found that large North American newsprint companies engaged in strategic behavior towards one another, particularly by making pre-emptive capacity investments. Christensen and Caves (1997) found that, for the North American pulp and paper industry over the 1970-1991 period, 38% of announced capacity expansions were later abandoned, and that these abandonments were most likely to occur in situations where other investments had also been announced. This finding is suggestive of investment being a form of competition between companies, aimed at capturing market share.

Gundersby (1996) portrays the pulp and paper industry as being in a consolidation phase on its way towards being a mature and more-concentrated industry. In this current phase, each company is trying to preserve or expand its market share. This can be done either by acquisition or by making new investments. Because of the cost advantage conferred by new investments, he argues that the latter is the most effective means of gaining market share:

"Investments are unavoidable strategic weapons in the struggle for long-term survival among the few large global companies which we shall see dominating the global pulp and paper business in the future."

---

[57] In fact, their 'residual funds' model hypothesized that residual cash flow (after paying interest, taxes, an 'established' level of dividends, and making 'sustaining investments') determines 'discretionary' investment spending. This is both intuitively plausible and useful for our purposes (where we are interested in causes of *changes in* capacity).

[58] Furthermore, the variables which we have included (price, cost of capital) are important elements of q in any case. In practice the inclusion of these variables may be justified with reference to either the neoclassical or q theories.

This cost advantage arises because of technical improvements. By making use of these, new paper mills have significantly lower short-term costs of production than older mills. In particular, improvements in process technology have allowed for mills of greater size (in terms of their production capacity) to be built, which are then able to realize economies of scale that older, smaller mills cannot. As is also the case for papermaking, there have been some technological improvements to pulping processes over recent decades. These include the use of additives such as polysulfide and anthraquinone to improve yields; computerization in control rooms to reduce labor requirements; and improvements in liquor recycling to improve energy efficiency. However, unlike the improvements in papermaking technology, these are not embodied in the capital structure of mills, and can be introduced in pre-existing mills with relatively little capital expenditure. The only cost advantage to new pulp mills which existing mills cannot also obtain stems from their location flexibility, which enables them to take advantage of differences in wood costs and availability between regions.

A second development attempts to model expectation formation. Investment has been modeled using Euler equations. The assumptions underlying this approach are similar to those of the neoclassical theory, but this approach allows expectations of the costs and benefits of investment to be modeled. This approach is quite new, and has not yet been tested thoroughly. Oliner *et al* (1995) compared the forecasting performances of two alternative specifications of a Rational Expectations model (one including time-to-build lags) against the 'traditional' neoclassical, accelerator and Q theory models. They found that the forecasting performance of the RE models was considerably worse than that of all the 'traditional' models.

This approach has been applied to the pulp and paper industry. Lundgren and Sjostrom (1999) estimate a model of factor demand for the Swedish pulp industry using Euler equations. They find that the user cost of capital is important in determining investment, but output levels are not. They also find weak evidence of adjustment costs in the industry.

Thirdly, another recent development in the investment literature is the explicit modeling of irreversibility and uncertainty. See for example Abel and Eberly (1997), Caballero (1997), and Caballero and Engel (1998). These factors certainly do play a role in investment in pulp mills. However, this literature is highly technical and it is difficult to extract any practical insights from it.

In summary, all of the more recent developments in investment theory (rational expectations, uncertainty, irreversibility and strategic factors) *prima facie* seem pertinent to the case of the pulp and paper industry. However, for the purposes of selecting variables to explain aggregate investment patterns, they do not seem to add much to what the traditional theories tell us.

Finally, a very important influence on investment behavior in the pulp and paper industry (which does not arise in any of the above-mentioned general theories) is environmental regulation. Since the 1970s, and especially since the 1985 discovery of dioxin in bleaching plant effluent, the industry in most regions of the world has been faced with tightening regulatory requirements. As a result environmental upgrades have been an important focus of investment expenditure. Although environmental investments do not necessarily have implications for capacity, in many cases the two are linked.[59] Changes in regulations will affect any new mills that are built. In addition, many existing mills must make environmental investments to comply with new regulations, and this is often done by installing new equipment that may have a different capacity to the equipment it replaces. The effect on capacity is idiosyncratic to an individual mill: Each mill will have a different part of the process which acts as the limiting factor in the capacity of the mill as a whole.[60] Similarly, if existing mills choose to shut down rather than comply with new regulations, this will obviously have a negative effect on capacity. In the future changes in environmental regulations can be

---

[59] This is shown by Gray and Shadbegian (1998).

[60] Notably, environmental investments have had to be made in bleaching operations. Potential to enhance capacity simultaneous with environmental improvements obviously only exists for a mill whose bleaching operations act as the bottleneck to capacity.

expected to continue. Thus, consideration of the quantitative impacts of such regulation on capacity is a useful task (although not one addressed herein).

2.3.2 Costs

Economic theory suggests that in a competitive industry, entry of new producers (or expansion of production capacity by existing producers) or exit of existing producers will occur in response to the existence of super- or subnormal profits. Consequently, the effect on price of any entry or exit will tend to drive prices towards some equilibrium level at which 'normal' profits are made, i.e. a level at which profits are just sufficient to recover the cost of capital.

Profits depend upon the margin between prices and total costs. Here we define total costs to include not only variable costs of production but also fixed costs of production and the cost of capital employed. If price is on average above total costs (which allows firms to make super-normal profits) then new firms will enter the industry, and conversely if price is on average lower (which causes firms to make sub-normal profits), then some will exit. Both exit and entry will have implications for price that will cause it to converge towards the level of total costs. Thus this theory suggests that in the long term prices will tend to converge towards the level of total costs.[61]

The pulp industry is reasonably competitive, and so we can expect this relationship to hold in the present case. A number of studies have attempted to explain prices as a function of input costs and technological change (e.g. Buongiorno and Gilless, 1980; Buongiorno and Lu, 1989; Chas-Amil and Buongiorno, 1999). These studies have made the same assumption as is used here, namely that in the long-term supernormal profits are zero. There is some evidence that subnormal profits have been maintained for long periods, but it is not clear whether this is true on a global scale. For example, Swann (1998) notes that the US industry "returned only 83% of its required cost of capital between 1960 and 1996, averaging an 8.6% return on total capital, as opposed to a 10.3% weighted average cost of capital."

It is likely that adjustment will be quite slow, for two reasons. Firstly, capital is very long-lived and immobile, and adjustments to the capital stock may therefore take a long time. Secondly, the price tends to be quite volatile, and this may cause uncertainty about the level of future profits, and restrict the willingness of market participants to take action for the sake of small-expected gains.

2.3.3 Summary

We hypothesize that investment is determined by cash flow, price, capacity utilization and cost of capital. We can combine these linkages with the short-term structure discussed above. In this overall system the level of installed capacity is no longer exogenous, but is determined by investment behavior. Cash flow is affected both by prices and by throughput. It is assumed that the cost of capital is mainly determined by capital market conditions, which are exogenous to the pulp industry.

We wish to model prices as being linked in some way (in the long term) to costs. The issue of how to treat costs is not clear-cut. Many influences on cost components (such as technical progress, and input coefficients) seem effectively exogenous. However, the prices of inputs (especially those specific to the pulp industry) will be affected by market conditions, particularly production levels. Overall it seems most appropriate to treat costs as exogenous, however, because the role of total costs in the models to be developed is that of a long-term 'attractor' for prices.

The speed at which prices will converge to this is likely to be sufficiently slow that cyclical variations in it are irrelevant in any case (this issue is explored further in section 3.5). All of the long-term linkages, together with those hypothesized in section 1.2, are shown in Figure 1.2.

---

[61] The theory is based on an assumption that companies are homogenous, and thus have identical cost levels. This is not true of the pulp industry, as is discussed further in section 3.5.

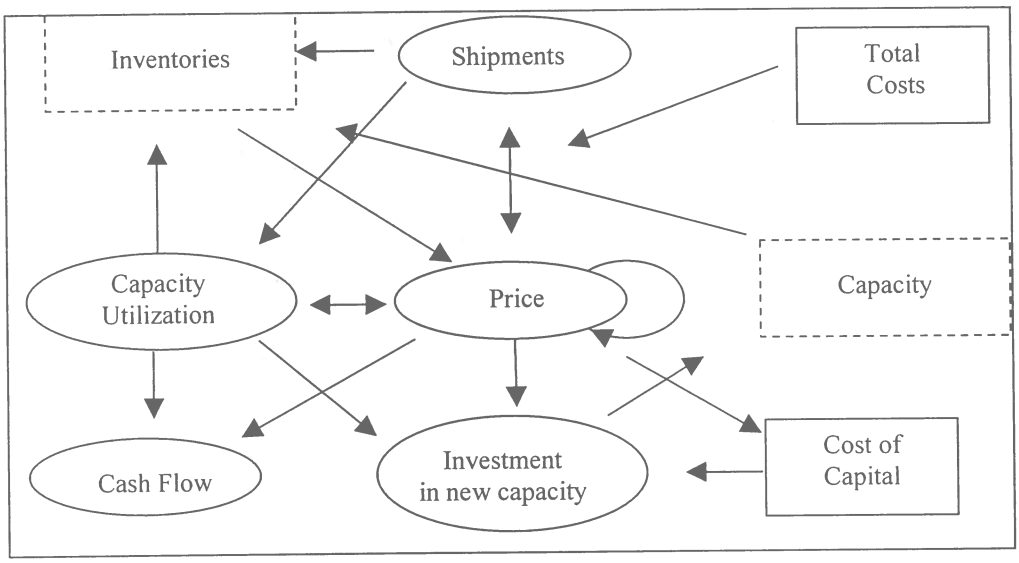

Figure 1.2: Summary of Short-Term and Long-Term Market Interactions

Of course, modeling of each of these variables is conditional on our ability to obtain data on acceptable empirical proxies for them. This is addressed in chapter 3.

3 TIME SERIES METHODS

"All models are wrong, but some are useful."

George Box

The decade since the late 1980s has seen considerable advances both in the theory of time series modeling and in its application to forecasting. In this chapter, we first attempt to summarize these theoretical developments. In light of these concepts, we then describe the different classes of time series models that are available, and present some of the model development methods that will be applied in chapter 4. Finally, the models are compared in terms of their suitability for the purposes of commodity price forecasting.

3.1 OVERVIEW OF TIME SERIES CONCEPTS

The statistical basis of econometric modeling prior to the 1980s was the assumption of stationarity. More recently, the existence of non-stationarity in economic time series has been found to be widespread, and methods have been developed to deal with this. In particular, the concept of cointegration provides a basis for methods that allow us to characterize series as having stochastic but common trends. These two concepts are outlined below.[62]

3.1.1 Non-Stationarity

Stationarity can be defined in various ways. Herein, we take it to mean covariance trend stationarity. This implies the following three properties: A potentially trending but deterministic mean over time ($E[Y_t] = \mu + \delta t$ at time t); a constant variance over time ($E[(Y_t - (\mu + \delta t))(Y_t - (\mu + \delta t))] = \sigma^2$ for all t); and a constant covariance over time ($E[(Y_t - (\mu + \delta t))(Y_{t-j} - (\mu + \delta(t-j)))] = \gamma_j$ for all t).

A stationary series is said to be integrated of order zero, or I (0). A series that is non-stationary has a higher order of integration. A series which has to be differenced n times before it becomes stationary is said to be integrated of order n, or I (n). Most economic time series are either I (0) or I (1).[63] An I (1) series is said to contain a unit root, or a stochastic trend.

A stationary series, after being subject to a shock, will have a tendency to return towards its mean level. Conceptually, whether this tendency exists for a particular series will depend on the nature of the shocks to which it is subject. If all shocks are transitory in their effects, then they will have no long-term influence on the series, and it will be stationary. If some or all of the shocks are permanent in their effects, then the series will have no tendency to return to its (trending) mean, and thus will not be stationary.

The potential existence of non-stationarity has important implications for modeling. OLS is not valid for non-stationary series. In order to model non-stationary series, it is necessary to transform them in some way to a stationary form. One common way of doing this is to take first differences (although as will be shown doing this is not always an attractive option). In any case, when we are deciding how to go about modeling a series or group of series, it is necessary to first assess whether they are stationary, before we proceed to the modeling stage. This assessment should be done both with reference to our conceptual ideas about the data-generating processes (DGPs) of the series and to empirical evidence.

---

[62] See Banerjee *et al* (1993) for a more detailed treatment.

[63] Some financial series are I (2), and some may be fractionally integrated. Neither of these possibilities is of central interest here.

Empirically, non-stationarity can be detected by testing for the presence of a unit root. Two tests are commonly used for this purpose: The Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. The ADF test is based on the assumption that the true DGP is an autoregressive model with p lags. The test is based on the following auxiliary regression (generated by subtracting y_{t-1} from both sides of a regular AR (p) model equation):

$$\Delta_1 y_t = \rho y_{t-1} + \alpha_1^* \Delta_1 y_{t-1} + \ldots + \alpha_{p-1}^* \Delta_1 y_{t-(p-1)} + \varepsilon_t$$

Depending on the hypothesized DGP, a constant and/or deterministic trend may also be included in this equation. Under the null hypothesis of a unit root, $\rho = 0$, and under the alternative, $\rho < 0$.[64] Critical values for this t-test are non-standard, and Mackinnon (1991) gives asymptotic values. Before carrying out the test, it is necessary to select a value for p, which is done by selecting a maximum lag length on the basis of theory, and then iteratively estimating the equation with successively lower maximum lags until a statistically significant coefficient on the maximum-lagged value is found.

Phillips and Perron (1988) proposed the PP test. It is a non-parametric test; i.e. it is not based on any specific assumption about the form of the DGP. It is based on a regression of an AR (1) process:

$$\Delta_1 y_t = \alpha + \beta y_{t-1} + \varepsilon_t$$

which may also be adjusted to allow for the inclusion of a constant or deterministic trend. The t-statistic is adjusted to account for possible heteroscedasticity and autocorrelation, as follows:

$$t_{PP} = \frac{\gamma_0^{1/2} t_b}{\omega} - \frac{\left(\omega^2 - \gamma_0\right) T s_b}{2\omega\hat{\sigma}}$$

where t_b and s_b are the t-statistic and standard error associated with β, σ is the standard error of the test regression,

$$\text{and } \omega^2 = \gamma_0 + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q+1}\right)\gamma_j ,$$

$$\text{where } \gamma_j = \frac{1}{T}\sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}$$

and q is the truncation lag chosen in accordance with the Newey-West procedure.[65] The asymptotic distribution of this statistic is identical to that of the ADF test statistic.

In developing our models of price and related variables, we will make use of these tests to check our *a priori* expectations about the stationarity characteristics of the series.

3.1.2 Cointegration

Engle and Granger (1987) formalized the concept of cointegration. It builds upon the concept of non-stationarity. Cointegration is a useful concept because it allows us to formally express ideas about equilibrium relationships in the context of time series models, and to model the consequences of any deviations from them.

Cointegration between two or more *non-stationary* time series is said to occur when the series are subject to common shocks. An implication of this (which allows us to test for the existence of cointegration) is that a *stationary* linear combination of the series exists. Thus, a set of non-stationary variables $y_{1t}, y_{2t} .. y_{nt}$ are cointegrated if there exists a vector a such that

---

[64] See Hamilton (1994, section 17.7) for the derivation of this.

[65] See Hamilton (1994, section 17.6) or QMS (1998, 331) for further details.

$$ay_t = [a_1 \quad a_2 \quad \dots \quad a_n][y_{1t} \quad y_{2t} \quad \dots \quad y_{nt}]' \sim I(0)$$

The vector of coefficients a is known as a cointegrating vector, and is not unique.[66] A group of n variables may have up to n-1 linearly independent cointegrating relationships. These relationships may involve only a subset of the variables in the group. The three cointegrating relationships identified in the course of this research are all relationships between only two variables.

The existence of cointegrating relationships (and their accompanying cointegrating vectors) can be established either with reference to relevant theory or by empirical testing of the data. Here, theory provides us with strong indications of the cointegrating relationships that are likely to exist, and we also examine the data to see whether it provides any supporting evidence. Johansen (1988) proposed the most widely used method of testing for the existence of cointegration. For a multivariate system, construction of an error correction model requires that we know how many cointegrating relationships between the variables in the system exist, and that we know the cointegrating vectors associated with each relationship. Johansen's test provides a useful empirical method of testing how many relationships appear to be present in the data. The method for carrying out this test is shown in Appendix B.

Once identified, a cointegrating relationship can be expressed in the context of an error correction model. This model type is described in section 2.2.3 below.

3.2 TIME SERIES MODELLING METHODS

In this section we first present three commonly used alternative classes of time series model. The ARMA model is the simplest, representing a single variable as a function of its own lagged values and shocks to it. The VAR model is the multivariate equivalent of the ARMA model (although usually without lagged shocks). The error correction model is a VAR model expressed in a way that allows us to impose cointegrating relationships upon the model structure.

The structure of the models is presented, along with various techniques that are used in making each model operational. After describing the models, previous empirical tests of their relative forecasting success are reviewed, and some issues relating to their usefulness for forecasting in the present context are discussed.

3.2.1 Auto Regressive Moving Average Models

Auto-regressive moving average models became popular due to the work of Box and Jenkins (1970). These models are not based on economic theory, but are a useful and relatively simple way to model the dynamic patterns of individual time series. In general, an ARMA (p, q) model takes the form:

$$y_t = a + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

where y is the variable of interest, ε is an independent and identically distributed error term, a is a constant, ϕ and θ are coefficients on the autoregressive and moving average terms respectively, and subscript t represents time.[67]

A requirement of this class of models is that the series y_t be stationary. The first step in constructing an ARMA model is to test for evidence of non-stationarity (using the tests described above). If it is found to be present we consider means by which the series can be transformed into a stationary form.

---

[66] If a is a cointegrating vector, then for any scalar k, ka is also a cointegrating vector.

[67] Hamilton (1994) provides a detailed discussion of each of these models.

To make the ARMA model operational, it is necessary to choose lag lengths for the autoregressive and moving average components, i.e. values for p and q. Box and Jenkins advocated the use of visual analysis of the series' autocorrelation function (ACF) and partial autocorrelation function (PACF).

The ACF measures the correlation coefficient of the current value of the series with the value of the series over each time lag. Its values are calculated as

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \text{ where } \gamma_k = E[(p_t - \mu)(p_{t-k} - \mu)]$$

is the kth order autocovariance of the series.

Each value of the PACF (ψ_k) measures the correlation between the same two variables after taking into account the predictive power of all of the smaller-lagged values of the series. Values for each ψ_k are estimated using separate regressions of the following form:

$$p_t = c + \eta_1 p_{t-1} + \eta_2 p_{t-2} + \ldots + \eta_{k-1} p_{t-k+1} + \psi_k p_{t-k} + \varepsilon_t$$

The shape of the ACF and PACF give clues about the most appropriate values for p and q. For an AR (p) process, the PACF will decline abruptly at lags greater than p, while the ACF will decline asymptotically (at a rate depending on the AR coefficients). For an MA (q) process, the ACF will decline abruptly at lags greater than q, and the PACF will decline gradually.

Formal statistical tests can also be used to select an appropriate lag structure. Two commonly used statistics are the Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC) statistics. Both of these are conceptually similar to the adjusted R^2 criterion of OLS regression in that they evaluate the fit of the model against the number of parameters included. They are calculated for a given model as follows:

$$AIC = n\log\left(\frac{RSS}{n}\right) + 2k, \quad SIC = n\log\left(\frac{RSS}{n}\right) + k\log n$$

where RSS is the sum of squared residuals from the estimated model, n is the number of effective observations used and k is the number of ARMA parameters to be estimated. The model that gives the smallest value for these statistics is preferred. The SIC penalizes the inclusion of additional regressors more than the AIC (whenever $n > 7$).

Once these values have been selected, an ARMA model can be estimated using OLS. The estimated model can be examined for stability, cyclicality, and other features. It is a useful starting point for analysis, but does not allow us to model the interactions between multiple variables.

3.2.2 Vector Auto Regression Models

In constructing a dynamic multivariate model, it is necessary to make a number of decisions, including the variables that are important parts of the system; which of the variables affect which other variables; the dynamic structure of these relationships, and the quantitative size of the interrelationships. Usually, available theory and prior knowledge help us to address the first and perhaps the second of these, but not the latter two. Economic theory usually provides little guidance as to the specific dynamic behavior of economic systems. As a result, decisions on the appropriate dynamic structure of the system are often made on the basis of the available data. For structural modeling, it is necessary to impose restrictions upon the set of relationships that exist in order to estimate coefficients. This creates a concern that the restrictions imposed may not be valid, which leads to spurious results. If we begin with an

inappropriately restrictive structure, we may fail to include all true relationships, and instead detect non-existent relationships.[68]

For this reason, a general and unrestricted specification is an attractive starting point for modeling. The VAR approach provides this, as it does not require any limitations to be imposed on interrelationships between variables. The only assumptions required are that the variables in the model (and an appropriate transformation for them if necessary) be known.

This type of model was popularized by Sims (1980), and is now widely used.[69] It is a multivariate extension of the ARMA model. In principle it is possible to include both autoregressive and moving average components in a multivariate model, but this is cumbersome. Commonly only autoregressive components are included.

The first step in constructing a VAR model is to decide upon the set of variables for inclusion. Usually theory will give an indication of the central variables related to the problem of interest. The classical VAR model requires that these variables be stationary, as above. The second step in model construction is to determine an appropriate number of lags to include. Here we rely solely on the multivariate analogues of the AIC and SIC criteria described above.

Once we have decided upon a set of variables $y_{1t}, y_{2t}, \ldots y_{nt}$ to include, and a lag length p, the VAR model can be specified as follows.

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} \Pi_{10} \\ \Pi_{20} \\ \vdots \\ \Pi_{n0} \end{bmatrix} + \begin{bmatrix} \Pi_{111} & \Pi_{211} & \cdots & \Pi_{n11} \\ \Pi_{121} & \Pi_{221} & \cdots & \Pi_{n21} \\ \vdots & \vdots & & \vdots \\ \Pi_{1n1} & \Pi_{2n1} & \cdots & \Pi_{nn1} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ \vdots \\ y_{nt-1} \end{bmatrix} + \ldots + \begin{bmatrix} \Pi_{11p} & \Pi_{21p} & \cdots & \Pi_{n1p} \\ \Pi_{12p} & \Pi_{22p} & \cdots & \Pi_{n2p} \\ \vdots & \vdots & & \vdots \\ \Pi_{1np} & \Pi_{2np} & \cdots & \Pi_{nnp} \end{bmatrix} \begin{bmatrix} y_{1t-p} \\ y_{2t-p} \\ \vdots \\ y_{nt-p} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{bmatrix}$$

where Π_{abi} shows the effect of variable y_a on variable y_b at i lags. Here for convenience the trend term has been omitted, but if desired it can be included in the same way as for the ARMA model. This can be expressed in matrix notation as $y_t = \Pi_0 + \Pi_1 y_{t-1} + \ldots + \Pi_p y_{t-p} + e_t$. The VAR method is attractive in that it does not require us to impose any restrictions on these coefficients.[70]

Because of the multitude of effects contained within the model, it is difficult to gain a good understanding of the dynamics of the models from an examination of the coefficients alone. One commonly used interpretive tool is the Impulse Response Function (IRF). This shows the net effect of a shock to one variable on another at different lags, and is a good way to gain an understanding of the dynamics of the models (if not the causal mechanisms at work). The IRF for the effect of y_a on y_b is defined as the set of derivatives

$$\frac{\partial y_{bt+s}}{\partial e_{at}} \text{ for all } s = 1, 2, 3\ldots$$

Initially, Sims' model formulation was in reduced form, i.e. not representing any structural relationships between the variables in the model. However, it has subsequently been shown that structural systems can be represented as VAR

---

[68] See Franses (1998, 192) for an example of this.

[69] Useful treatments of various aspects of the topic include Canova (1995), Holden (1995), Mills (1998) and as usual Hamilton (1994).

[70] One cost of this (as Hamilton (1994) observes) is that in practice the standard errors associated with the impulse response functions estimated using VAR models are often large, and few precise inferences can be drawn.

models, and structural coefficients can under certain restrictions be extracted from the estimated reduced form model.[71] This process is described in Appendix A.

3.2.3 Error Correction Models

The VAR model discussed above is not able to model non-stationary variables, except by first transforming them into stationary form. Thus, it is unable to exploit the existence of cointegration for forecasting. In a situation where cointegration exists, use of an unrestricted VAR model of differences will result in inefficient forecasts.[72] To explicitly model cointegration, error correction models are commonly used.[73] The popularity of this type of model has occurred because of the finding of Granger (1986) that any group of cointegrated variables can be modeled as having been generated by an error correction model.[74] This type of model enables us to combine short-term dynamics and long-term theoretical ideas about equilibrium into a single framework.[75]

The error correction model framework is based upon a VAR structure, with cointegrating relationships imposed upon this. Our starting point is the general expression for a reduced form VAR (p) system:

$$y_t = \Pi_0 + \Pi_1 y_{t-1} + \ldots + \Pi_p y_{t-p} + e_t$$

$$y_t = \alpha + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \varepsilon_t$$

This can be expressed in first differences:[76]

$$\Delta y_t = \Pi_0 + \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{p-1} \Delta y_{t-p+1} + \zeta_0 y_{t-1} + e_t$$

Where $\zeta_i \equiv -\left[\Pi_{i+1} + \ldots + \Pi_p\right]$ for all $i = 1,2,\ldots, p\text{-}1$, and $\zeta_0 \equiv -\left[I_n - \Pi_1 - \ldots - \Pi_p\right]$.

Apart from the term $\zeta_0 y_{t-1}$, this model is identical to a VAR (p-1) model of Δy. We can use this term to impose cointegrating relationships upon the model. If there are h cointegrating relationships between the components of y with cointegrating vectors represented in the rows of the hxn matrix A, then it follows that there exists an nxh matrix B such that $\zeta_0 = BA$.[77] Substituting for ζ_0 in the equation above gives us

$$\Delta y_t = \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{p-1} \Delta y_{t-p+1} + \Pi_0 + Bz_{t-1} + e_t$$

---

[71] See Hamilton (1994, 324-36).

[72] See e.g. Holden (1995).

[73] Examples of such work include Sarker (1993), Tegene and Kuchler (1994), Kuiper and Thijssen (1996), Kuiper and Meulenberg (1997), and Ghali (1998), among others.

[74] This is known as the Granger Representation Theorem.

[75] Schoonbeck and Sterken (1995) discuss recent trends in dynamic macroeconomic modeling. They comment that "one of the very attractive features of this type of model is that from the outset a consistent and clear distinction is made between the long-run equilibrium and short-run adjustment dynamics."

[76] The fact that these two representations are equivalent is demonstrated by Hamilton (1994, 549).

[77] See Hamilton (1994, 579) for the proof of this.

Where $z_t \equiv Ay_t$.

This is the form in which the error correction model is estimated. By defining and including z_t on the RHS of this equation, we are effectively imposing both the number of cointegrating relationships in the model and the cointegrating vectors associated with them. We can interpret z_t as the 'errors' and B as the matrix of 'correction factors'. The size of the coefficients in B will determine the speed at which the variables adjust to eliminate any 'disequilibrium' in the cointegrating relationships.

The error correction model is expressed in differences because the variables in y_t are non-stationary. If a system contains a subset of variables that are stationary, then it is not necessary to express these in differences, and the model can be formulated as a partially differenced system.[78]

3.3 PREVIOUS COMPARISONS OF FORECAST ACCURACY

A number of researchers have tried to evaluate the usefulness of the different classes of time series models for forecasting purposes. Comparisons between forecasts are based on criteria which measure the divergence of the forecasts from realized outcomes (a commonly used criterion is described in chapter 5).[79] This section summarizes some of the comparisons that have been made.

Much evaluation has focused on forecasts of macroeconomic variables. Engle and Yoo (1987) compared the forecasting performance of an unrestricted VAR model with that of an error correction model (with cointegrating relationships estimated from the data), and found that the error correction model gave better forecasts in the long term, but not in the short term.[80] LeSage (1990) carried out an empirical comparison of an error correction model and various types of classical and Bayesian VAR models, using labor market data. He found that the error correction model produced superior forecasts when the variables in the model did in fact exhibit cointegration. Duy and Thoma (1998) examined the issue of whether imposing cointegrating relationships via an error correction model improved model forecasting performance for macroeconomic variables. They found that it often did, especially in cases where strong evidence for cointegration between variables existed. They also found that error correction models with theory-based cointegrating relationships were more successful than those with databased cointegrating relationships. Hoffman and Rasche (1996) carried out empirical tests, and found that imposing cointegration did not greatly improve forecast accuracy.

Assessment of forecast accuracy for models of commodity prices has also given mixed results. Fanchon and Wendel (1992) carried out a comparison of a VAR model in levels, a Bayesian VAR and an error correction model. The subject of their analysis was cattle prices. They found the VAR model to produce the best forecasts. Lord (1991) developed an error correction model that he found to outperform ARMA models in forecasting most commodity prices. This model contained both supply and demand side aspects.

Christofferson and Diebold (1998) argued that the reason for earlier findings of better forecasting by error correction models than VAR models was due to the formers recognition of integration of variables, rather than recognition of cointegration *per se*. They tested this hypothesis empirically, and found that error correction models were no more accurate than univariate ARMA models (in differences) in long-term forecasting of cointegrated variables. This finding is at odds with that of LeSage (1990), who controlled for the effect of imposing integration by comparing an

---

[78] Hamilton (1994, 652) provides details.

[79] A caveat upon comparing forecast accuracy is that the results of doing so depend upon how the data is transformed, and the data selected for forecast evaluation. There is no data-invariant comparison that can be made. See Clements and Hendry (1993).

[80] Engle and Yoo also explored the theoretical implications of cointegration. They showed that in a cointegrated system, Granger causality (described in section 3.1) in at least one direction must exist.

error correction model to a VAR model in differences, and found the error correction model to generate better forecasts. Christofferson and Diebold also found that error correction models performed better at short horizons than at long horizons. Again, this is at odds with LeSage (1990), who found exactly the opposite.

In summary, neither the VAR nor the error correction model seems to be consistently superior to the other. Presumably, the relative performance of the two depends greatly on the variables that are being forecast: In particular, whether cointegration is present; and the speed at which convergence to equilibrium occurs relative to the forecast horizon. Granger (1986) argued that "error correction models should produce better short run forecasts and will certainly produce long-term forecasts that hold together in economically meaningful ways." Even if the first part of this statement does not hold true in every case, the second part is sufficient to justify use of the error correction model.

3.4 MODELLING STRATEGIES

Preliminary to any model building, it is necessary to determine the characteristics of the variables of interest, to assess whether they are non-stationary, and if so, whether they are cointegrated. This can be done both with reference to theoretical ideas about the nature of the shocks to which they are subject, and by the empirical means of graphical analysis and statistical testing.

In a situation where at least some of the variables of interest are non-stationary (and possibly cointegrated), Hamilton (1994, 652) presents three alternative ways in which modeling may proceed. The first is to estimate a VAR model in levels and allow the data to impose its own restrictions, without any explicit cointegrating relationships. This method is used in section 4.2. This is considered a useful first step, and should be followed by an assessment of which series are likely to be nonstationary. If no evidence of non-stationarity is found, then the level VAR model will be adequate.

The second method is to estimate a VAR model with any variables suspected to be non-stationary modeled in differences. Hamilton observes that

"The drawback to this approach is that the true process may not be a VAR in differences. Some of the series may in fact have been stationary, or perhaps some linear combinations of the variables are stationary, as in a cointegrated VAR. In such circumstances a VAR in differenced form is misspecified."

A particular problem with this approach in the present context is that an unrestricted model with price expressed in differences has the property that forecasts for prices are highly dependent on the price level at the time the forecasts are made (since the historical mean price change is close to zero). While we do not necessarily regard price as stationary in the sense that it will be fully mean reverting, we do expect that some mean reverting tendency will exist. For this reason, we do not use this approach.

The third method involves estimating error correction models (in differences, and with the inclusion of error-correction terms). Here, two error-correction models are constructed, each focusing on different cointegrating relationships. Cointegrating relationships may be determined either on a theoretical or empirical basis. Hamilton comments

"The disadvantage of the third approach is that, despite the care one exercises, the restrictions imposed may still be invalid – the investigator may have accepted a null hypothesis even though it is false, or rejected a null hypothesis that is actually true. Moreover, alternative tests for unit roots and cointegration can produce conflicting results, and the investigator may be unsure as to which should be followed."

The approach taken here (in both sections 4.3 and 4.4) is to base the imposed cointegrating relationships on theory, but to use statistical methods as a check to see whether or not empirical evidence contradicts the existence of these relationships. We are fortunate that theory does give a reasonably strong indication of the cointegrating relationships (and corresponding cointegrating vectors), which should exist. This approach will reduce the likelihood that the cointegrating relationships imposed are invalid.

4 DATA

"The pricing of pulps has been a controversial issue in the field for many years. There are always pulps for sale - at Pulpcon, from dealers, bookstores, antique dealers, paper shows, etc - but the current value of specific titles and issues is often open for debate. This is because most pulps are relatively scarce and sell infrequently, and when they do sell price is usually a private matter between buyer and seller. This makes publication of an accurate price guide impossible. There simply are not well known prices for all but the most commonly traded titles."

J.P. Gunnison.

This section discusses the data that is available to proxy the variables of interest. For cash flow, no suitable proxy has been located. For some of the other variables, only an imperfect proxy is available. Unless specifically indicated otherwise, all data were obtained from Jaakko Poyry Consulting Oy. Data were provided on a confidential basis, and for this reason are not reported herein. The following sections discuss the proxies for each variable in turn.

4.1 PRICE

The benchmark pulp price series is the NBSK *cif* price in Northern Europe, denominated in US dollars per ton. This price series is not only a benchmark price, but has also been reported on a consistent basis for a long period of time. Monthly prices are available for January 1950 to December 1959, and quarterly figures for March 1960 onwards. Since the models will be developed on a quarterly basis, an average of the monthly prices is taken to convert them to this basis. The resulting series is shown in Figure 3.1.

Data on a number of other pulp prices are also available, but these are not used in the modeling process. As was discussed in chapter 1, an important issue in choosing suitable proxies for non-price data is the scope of the market to be considered. In that section, we considered the degree of substitutability of different pulp types and tradability across regions. An alternative approach to examining the conditions that may result in products being part of a unified market is to examine outcomes. A necessary condition for products to be part of a unified market is that movements in their prices are similar.

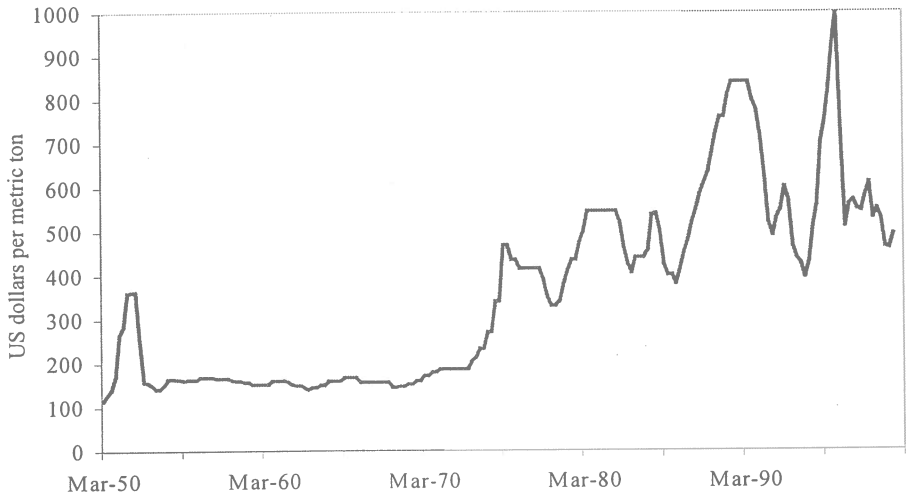

Figure 3.1: Nominal NBSK Price

We can use these price data for this purpose. To assess similarity we need to consider interactions between price and other variables. The issue of interest is whether markets for different pulp types sold in different regions are substantively unified or not. 'Unified' is taken to mean that the prices are subject to common shocks, and will move together.[81]

Price data on different pulps is not available on a consistent basis. Time coverage and frequency both vary by product. The regions of most interest are the largest ones, as events in these regions will have the most influence on the market as a whole. The different price series available for Kraft pulps are summarized in Table 3.1.[82]

---

[81] Yet another method that could be used to assess whether markets are unified is to examine the responses of different prices to non-price shocks. If changes in non-price variables pertaining to a product have substantively similar impacts on the Northern Europe NBSK price as do changes in those same non-price variables pertaining to Northern Europe NBSK, then that product should be included in the market scope. For example, if demand growth in Europe has the same impacts on the Northern Europe NBSK price as does demand growth in the US, then the US and European NBSK markets should be considered unified. If new investments in BHK capacity have as significant an impact on NBSK prices as do investments in NBSK capacity, it is attractive to include BHK in the market scope. This method is not applied here.

[82] Another source of information on prices is trade unit values. However, trade unit values are an inferior indicator of market prices because not all trade occurs at the contemporaneous spot price.

Table 3.1: Summary of Kraft Pulp Price Data Availability

Pulp type	Location	Time span	Freq.
NBSK	Northern Europe	1950-1999	Qtrly
BHK(birch)	Northern Europe	1970-1999	Qtrly
BHK(eucalyptus)	Northern Europe	1987-1999	Qtrly
NBSK	USA	1950-1979 1980-1999	Ann. Qtrly
BHK (northern and southern), SBSK	USA	1952-1979 1980-1999	Ann. Qtrly
NBSK, BHK (birch)	Germany	1980-1999	Qtrly
SBSK, BHK (eucalyptus), BHK (southern mixed)	Germany	1986-1999	Qtrly
NBSK, BHK (birch), SBSK, BHK (eucalyptus), BHK(southern mixed)	France, Italy, UK	1986-1999	Qtrly

The Northern European prices are on a *cif* basis, while the other prices are for delivery within particular countries. Some data is also available for other pulp types. Visually, we can examine the different series. Figure 3.2 compares NBSK prices in Northern Europe and the US since 1950. Interestingly, the price spike in Northern Europe in the early 1950s is completely absent from the US market, but after this, the series appear to have moved in concert. Exact comparison is made difficult by the different frequency of the data before 1980, but certainly since 1980 the two series have been very similar.

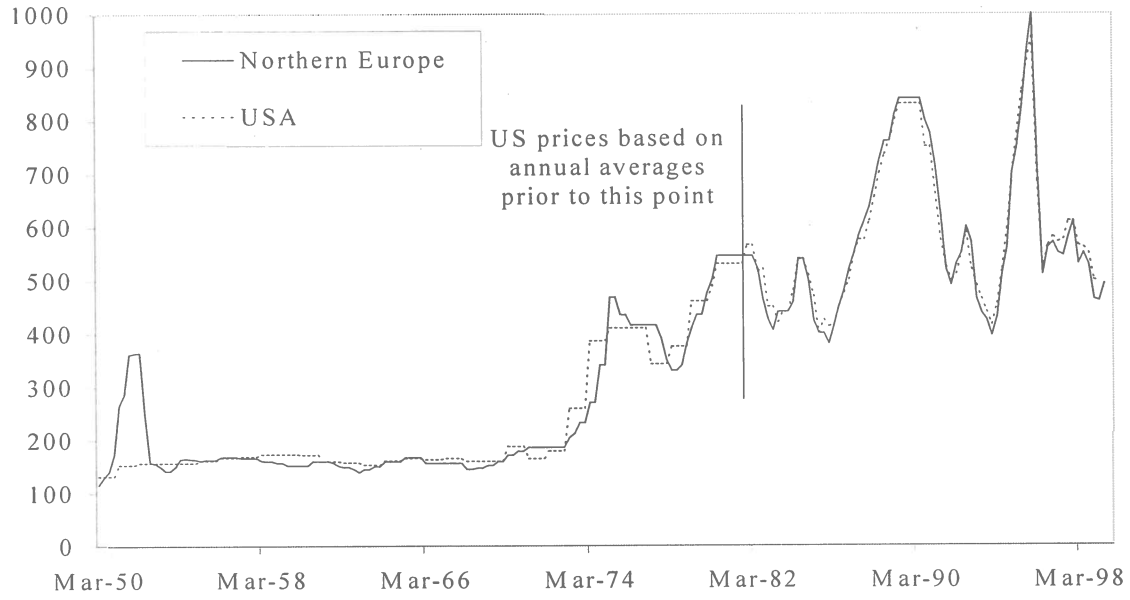

Figure 3.2: NBSK Prices in Northern Europe and USA

For a shorter time period, we can also compare these to prices in other locations (Figure 3.3). All of the NBSK prices in Europe and the US seem very similar.

We next examine prices of other types of pulp in the USA (Figure 3.4). Absolute levels of the prices of different pulps vary: For example, the price of eucalyptus BHK is continuously lower than that of NBSK. However, price changes in the series are closely matched. Prices of the various types of bleached Kraft pulp appear to follow similar patterns.

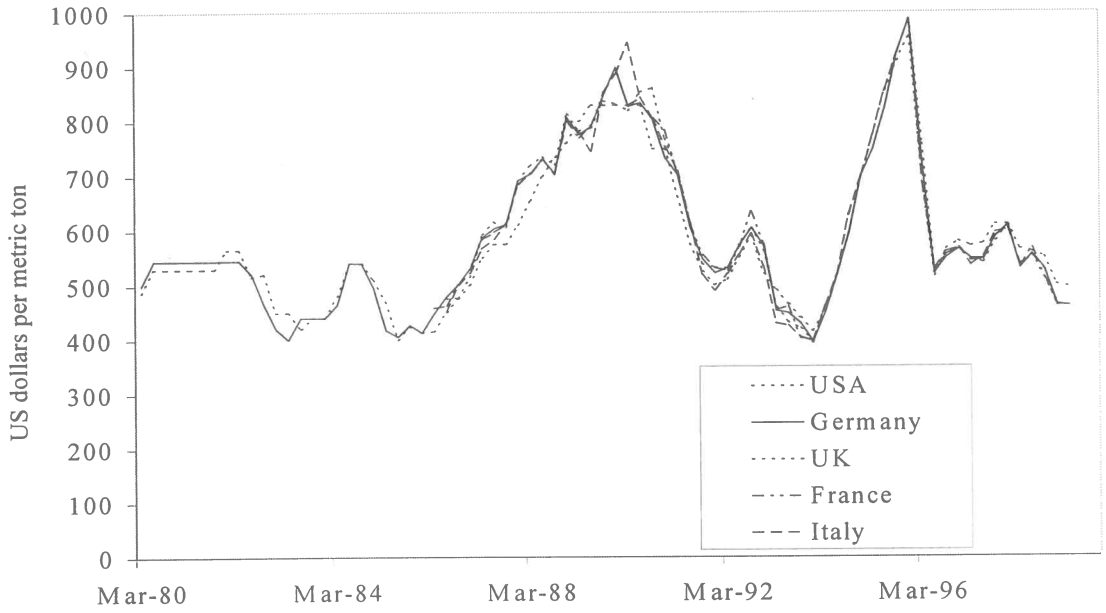

Figure 3.3: NBSK Prices in Different Locations

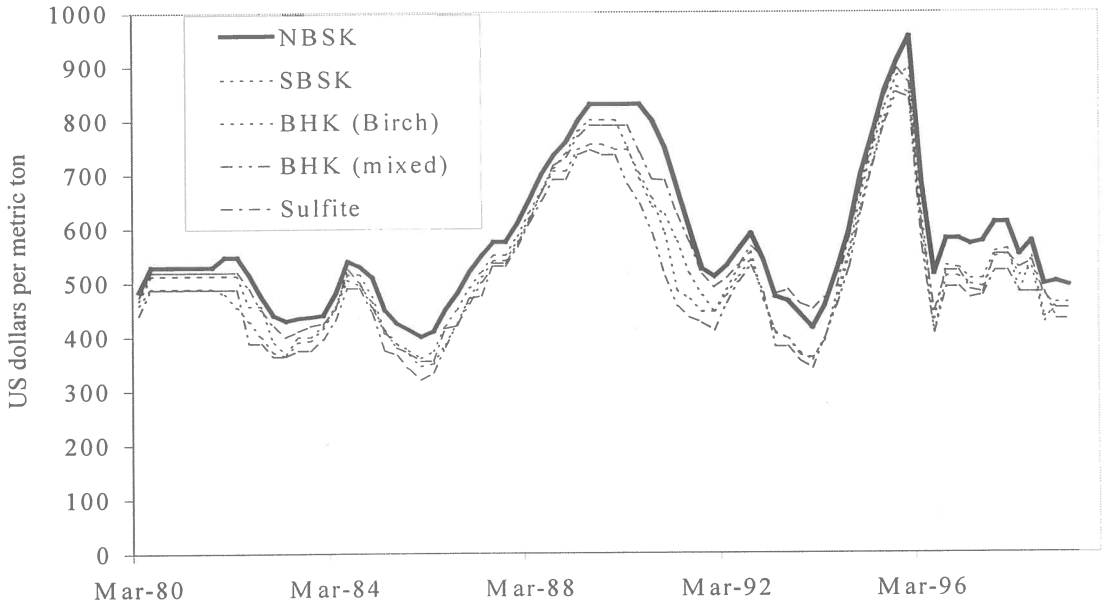

Figure 3.4: US Prices for Bleached Kraft Pulps

Figure 3.5 shows prices of other pulps in the US. Here the picture is less clear. Most of these pulps are not substitutes for Kraft pulp, as discussed above. Despite this, the price series show a surprising degree of commonality, at least in terms of price changes. For some products (unbleached softwood Kraft pulp, dissolving pulp and fluff pulp), periods of divergence from trends in the NBSK price can be observed.

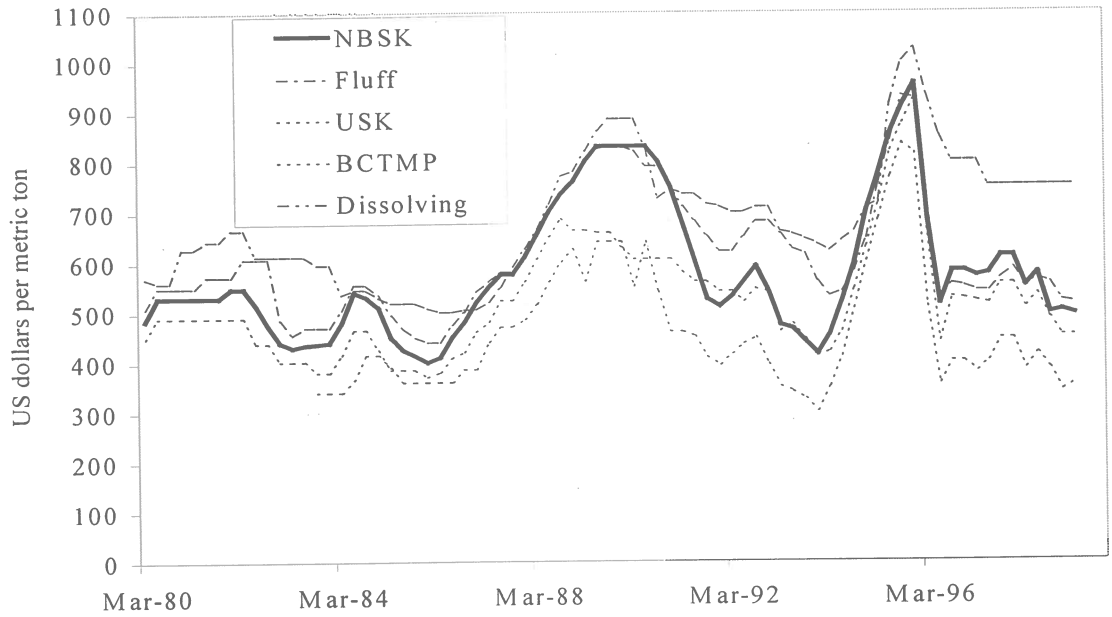

Figure 3.5: US Prices for Other Pulps

There are two possible reasons why prices of other pulps are so closely correlated with the NBSK price. The first is that the markets are in fact unified. This is plausible in some cases. According to anecdotal evidence, one product that has had an active influence on the NBSK price is BHK. In recent years, as production of Kraft pulp in tropical and sub-tropical regions has increased, the share of BHK in the overall market for Kraft pulp has increased, and increasing production in these regions has had a downward influence on the NBSK price.

The second possibility is that the NBSK price is a standard against which other prices are passively set. Historically, NBSK has been a dominant pulp type in terms of production volume, and anecdotal evidence suggests that it has been a benchmark grade against which other prices are set. If this is the case, then conditions in the markets for these other pulps will not affect the NBSK price. In a sense there is only a one-way connection between the two. In this case it is not desirable to treat markets as being unified. Also, some shocks will affect these other markets that do not affect the NBSK market, and so we would not necessarily expect prices to maintain a fixed relationship in the long term (although this may not be apparent over a series covering only a relatively short period of time).

One potential means of distinguishing these two types of relationships is by means of Granger-causality testing. We examine whether lagged values of one variable *a* significantly explain another variable *b*. If we find that they do, we can state that variable *a* 'Granger causes' variable *b*.[83] In this present case, such a finding would indicate that shocks to the market for pulp type *a* do affect pulp market *b*.

For example, it has been suggested that the BHK price has 'led the market down' in recent years, in which case we would expect to see a Granger-causal link from this price to NBSK. On the other hand, if other prices simply react to the NBSK price, then no Granger-causality would be detected. In order for this type of testing to detect a linkage, data must be of a sufficiently high frequency. From visual inspection of the price data, it appears that most price changes occur contemporaneously in all pulp types. This suggests that the (hypothesized) response of other prices to changes in the NBSK price occur within the same quarter. It therefore seems unlikely that Granger-causality

[83] The presence of Granger causality does not necessarily indicate true causality. Both variables may be influenced by some other variable, with different lags. True causality may even run in the reverse direction, if for example agents form expectations about variable *b* in advance, which then influence variable *a*.

testing using quarterly data will shed light on the issue, and it is not carried out. The availability of higher-frequency data would make this testing more useful.

Another statistical approach that may be used is to test whether the Law of One Price holds between different pulp types. A large literature on this Law exists, and recent tests of it have used cointegration analysis to determine whether different prices are subject to common shocks.[84] No application of these tests to pulp prices has been located.[85] A necessary requirement for this test to be valid is that the original series be non-stationary themselves (if they are stationary, then any linear combination of them will also be stationary). Empirically, there is no evidence of non-stationarity in these series for the period over which most of them are available.[86] Therefore, as an empirical matter, it would be very difficult to reject the hypothesis of cointegration based on this data, and so it does not seem appropriate to use cointegration tests in this instance.

In summary, statistical evidence on which pulp products form a unified market is inconclusive. We continue to rely on the assessment made in section 1.1.1 that global bleached chemical pulp is a suitable market scope.

4.2 PRODUCTION, SHIPMENTS, INVENTORIES

The data on these variables is described in terms of product scope and geographic scope. As regards product definition, data is available for 'total chemical paper grade pulp' as well as for more disaggregated categories (bleached and unbleached sulfite and Kraft pulps). Given the degree of substitutability between pulp types, the aggregate data seems most suitable for our use.

As regards geographical definition, the situation is less ideal. Data availability is different for three groups of countries. The first is the Norscan countries (the USA, Canada, Sweden, Norway and Finland) for whom data on the variables required here are available over the entire period of interest.[87] The second is a group of non-Norscan countries for whom data is currently available, but has not always been so over the period of interest. Actual data on inventories is available for all pulp for a number of countries (Austria, Belgium, France, Spain, Argentina, Japan, South Korea, Swaziland and New Zealand), and for Kraft pulp only for several others (Portugal, Brazil and Chile). The third group is made up of the remaining non-Norscan pulp producers, for whom actual data is not available. These are Mexico, Morocco, South Africa, Indonesia, Taiwan, Thailand, Germany, Switzerland, and China. Some of these non-Norscan countries have quite stable pulp industries that have existed for a long time, whereas others have recently developed and/or growing industries. Estimated data for these countries (again only for part of the period of interest) is available.

In cases where non-Norscan data was available, it was examined to see how similar the Norscan series was to the overall total (Norscan plus non-Norscan) series. The series were different in terms of their absolute levels, and occasionally different in terms of the quarter-to-quarter changes in them. The overall upward trends are somewhat higher for the total market than for Norscan, for all variables. Overall, however, a high degree of similarity between

---

[84] A useful description of the Law and summary of empirical work is provided by Officer (1989). Cointegration methods have been used by, for example, Ardeni (1989), Vogelvang (1992), Jung and Doroodian (1994), Gordon and Hannesson (1996), Bose and McIlgorm (1996) and Hanninen (1996). Ravallion (1986) and Goodwin, Grennes and Wohlgenant (1990) have also used other methods, for example.

[85] Buongiorno and Uusivori (1992) tested whether the LOP held between US exports of NBSK, BHK and chemical pulp to different destinations, and found evidence that it did. However, they did not test the LOP between products, nor did they examine the market prices in those different destinations.

[86] Conceptually, however, there are strong reasons to believe that prices are non-stationary (as will be discussed in chapter 4).

[87] The period chosen for analysis is 1976:3 onwards. The reasons for this are discussed in section 4.1.2.

the series was observed. The total calculated by using actual data only and the total calculated using the estimated data as well were also compared, and found to be rather different in terms of their quarter-to-quarter changes. This implies that the improvement gained from including actual non-Norscan data may in fact not be substantial, and may not be worth the introduction of uncertainty which doing so entails. Although its scope is more limited than is ideal, the Norscan data is at least of relatively good quality.

It is therefore concluded that use of Norscan statistics is acceptable, at least given the alternatives available. Although this data is obviously not fully inclusive, it is attractive for a number of reasons. Firstly, market commentators often focus on Norscan data, and thus it can be regarded as affecting market conditions in its own right. Secondly, the Norscan share of total market pulp capacity is quite large.[88] Finally, the data is available in a regular and timely fashion, which makes it useful for maintaining up-to-date forecasts.

From a forecasting perspective, differences in the absolute levels of the series are not of concern, as long as we can be reasonably confident that historical differences will be maintained in future. Differences in trends, and differences in patterns of the series (for example, if changes in the total series are not reflected in the Norscan series) are of more concern. This creates a concern that forecasting models estimated on the basis of historical Norscan data will not be valid for the future. A number of the shocks to the market have originated outside of Norscan, at least in recent years. Exclusive focus on Norscan data may not allow us to capture the effects of these shocks. A useful avenue for future work is to expand the geographical coverage of data on these variables.

Production, inventories and shipments data are collected by producer associations in each country, and are announced on a coordinated basis each month, usually about two weeks after the end of the month in question. The source of these data is Metsateollisuus Ry.

Data on inventories, production and shipments are reported separately. Data on inventories covers only those held by producers.[89] Conceptually, the change in inventories from one quarter to the next should be exactly equal to the difference between production and shipments, if all of these are defined on a consistent product and geographic basis. For a number of reasons, however, there may be discrepancies in these statistics. These include the fact that some pulp produced for sale on market may consequently be used internally; the fact that some pulp is shipped to a consignment warehouse overseas without a final buyer, but has to be invoiced for customs purposes; post sale rejection of shipments due to quality problems with the pulp; and discrepancies between reported and actual inventories which are not discovered until year-end. Each of these events leads to inconsistencies between the different statistics. Furthermore, the effects from company to company may not be the same. Overall, however, these discrepancies are slight.

4.3 CAPACITY

Capacity data is less readily available than data on other quantity variables. This data is also collected by producer associations in the various countries, and is summarized by the FAO on an annual basis (e.g. FAO, 1998). *Ad hoc* procedures are used to convert this to a quarterly basis. Quarterly capacity data is available from 1984 onwards, and pre-1984 values can be inferred from available data on the shipment rate (shipments/capacity) and shipments levels.

4.4 COST OF CAPITAL

As a proxy for cost of capital, we develop an estimate of the real interest rate. This is based on a US lending rate deflated by year-on-year US consumer price inflation. Both of these are obtained from IFS (1999). The real US rate

---

[88] At the beginning of the period of interest (1976), the Norscan share of total market pulp capacity was 77%. However, this share has declined over time, particularly due to the expansion of capacity in Latin America and Asia. In 1997, the Norscan share was 62%.

[89] Only limited data (from Europe) is available on user inventories.

provides the most suitable proxy because a large proportion of borrowing is denominated in US dollars.[90] US wholesale price inflation was also considered as a deflator. However, this gave a more volatile series, and was not used for this reason. The series is shown in Figure 3.6.

[90] Initially, an average real interest rate was calculated, with weights based on capacity shares by country. However, the resulting series seemed unappealing for a number of reasons, both conceptual and practical.

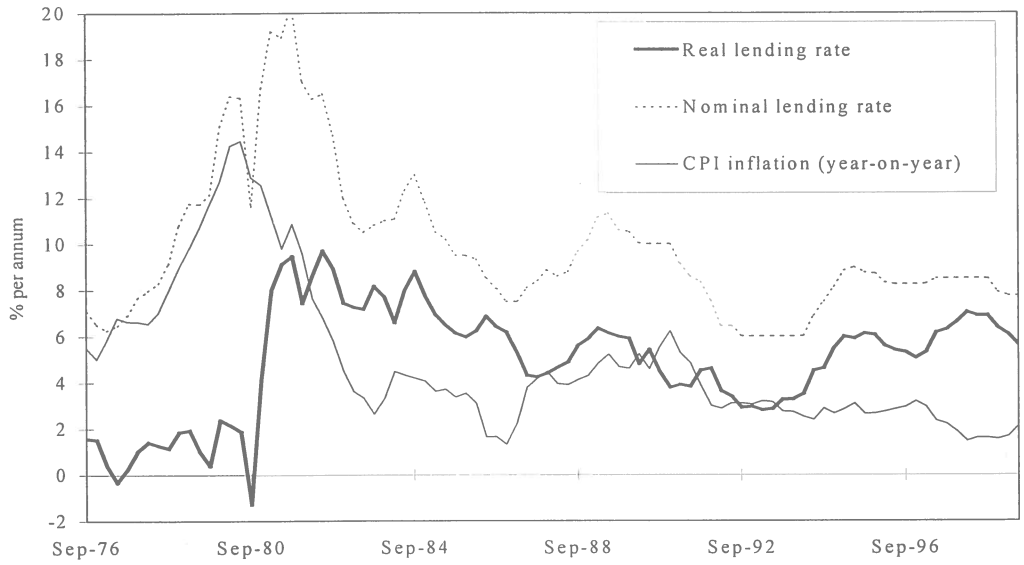

Figure 3.6: US Real Lending Rate

4.5 TOTAL COSTS

Cost data is much less readily available than for the other variables. Two alternative measures of total costs exist, but neither is available on a consistent time series basis. We describe the two, and consider which is more appropriate for purposes of estimating the error correction model.

Jaakko Poyry Oy Consulting has a unique database on the costs of every large and medium sized mill in the industry. Cost components available include costs of production at the mill, costs of transportation to various points, and capital costs. These costs are calculated on the basis of input requirements and unit costs. Total costs are built up from a very detailed level. Costs are available on a mill-gate basis, *cif* at a particular location (such as Northern Europe) or on a 'full cost' basis defined as the *cif* price plus capital cost. We can use this capital cost as a proxy for the 'normal profit' that companies in a competitive industry should earn in the long term. From this database, 'cost curves' for the entire industry at a given point in time can be derived. Historically, this has not been done very often, but curves are available for a number of points in time.

At any point in time, there is a considerable difference in costs between the most efficient producers and the least. For example, for NBSK production in March 1999, the lowest full cost was estimated at approximately $380 per ton and the highest at $730. The costs of most mills are much more similar: Excluding the cheapest and most expensive deciles of capacity gives a full cost range of $500 - $610.

The second source of cost data is estimated full costs of specific but hypothetical mills (some of which were in fact subsequently built) at various locations. These are calculated in a similar way to the actual costs, namely by means of engineering knowledge on input requirements in conjunction with unit costs.

These costs exhibit some variation, both over time and by location. Because only limited numbers of estimates are available, it is difficult to assess what proportion of these differences is due to idiosyncratic (mill-specific) factors as opposed to universal differences in cost levels across time and space. Only the latter are of interest from our aggregate perspective.

Conceptually, what is the cost level to which prices should converge in the long term? The simple economic theory on which the price/cost relationship is based does not provide much insight into this issue, based as it is on an assumption of identical companies. To address this we first need to consider the reasons for the differences in costs across the spectrum of existing mills, and in relation to hypothetical mills. It is difficult to pinpoint a single reason for the differences in costs. Mill age and location are both relevant, as well as idiosyncrasies of each individual mill.

Mills at the low end of the cost spectrum are, on average, somewhat younger than mills at the high end, but there is also a considerable amount of variation within the spectrum, and mill age is not the primary factor determining costs. Mill location seems more important, with mills in a certain area generally clustered at a certain part of the cost spectrum.[91] One important reason why location is important is that exchange rates vary over time, causing mills in each location to gain or lose competitiveness accordingly. There are exceptions to this pattern, however, indicating that idiosyncratic factors are also important.

Concentration of mills in a particular country is limited by fiber availability, and increased competition for fiber in a location will cause costs to rise. This implies that migration of capacity from high cost to low cost regions will tend to have the effect of equalizing costs across regions. If we compare the Pacific Northwest region of the US to Chile, for example, wood costs are currently cheaper in the latter area. This means that capacity will tend to move (over a long time period) from the Pacific Northwest to Chile. This in turn will affect the relative costs of fiber in the two regions, though, and at some point there will no longer be a tendency for such movement to occur. The overall effect of this process is to flatten the cost spectrum. The process will of course never be complete, as new shocks to cost levels in particular regions will occur from time to time. Nevertheless, we can envisage the following equilibrating situation: If price is at the level of average total cost, then producers in low cost regions will make supernormal profits, and producers in high cost regions will make subnormal profits. Over time, new capacity will be developed in low cost regions, which will tend to drive up the cost of production in those regions; and conversely capacity will decline in high cost regions, reducing the costs of production of the remaining producers. Conversely, if price is above or below the average cost level, then capacity change will be unbalanced, and equilibrium will not result.[92] Thus the average level of total costs seems the most likely point to which price will converge. In particular, convergence to the average is more plausible than convergence to either end of the spectrum.

Another reason to prefer actual costs to hypothetical costs is that the latter are more sensitive to cyclical patterns. Because we are hoping to develop some measure of costs towards which prices converge over the long term, we would like this measure of costs to be unaffected by short or medium-term changes in market conditions. As noted in section 1.3.1, costs of investment are positively related to levels of investment activity, which are positively related to prices. Variable costs of production are also cyclical to a degree. For example, high production levels may bid up wood input costs.

This problem affects hypothetical costs and actual costs differently. For existing mills that have already incurred construction costs, capital costs are insensitive to cyclical patterns occurring after their construction. For hypothetical mills that are assumed to incur construction costs at the time the observation is made, capital costs are sensitive to cyclical patterns. Both actual and hypothetical mills are affected by changes in variable costs. The fact that we wish to avoid this cyclical pattern of costs suggests that actual costs are more appropriate for our purposes.

Hypothetical cost estimates have the advantage of being more 'forward-looking' than the actual costs. In particular, any technological advances that have only recently occurred will be more fully reflected in the former. This is attractive, given that we can expect such advances to be gradually taken up by the industry as a whole.

Overall, we prefer the actual average full cost, because we can be reasonably confident that it represents a plausible point of convergence. As an average it is less subject to idiosyncratic factors than the hypothetical estimates. It also has the advantage of being less affected by short-term cycles than the hypothetical costs, but has the disadvantage that it does not capture any existing technological improvements to the extent that they have not yet been fully taken up by the industry.

---

[91] In fact, newer mills are generally in lower-cost locations, and it seems that location is more important than their younger age *per se* in providing a cost advantage.

[92] For purposes of this discussion we assume the overall level of demand is constant over time. This is a simplifying assumption not critical to the conclusions reached.

The next problem we face is one of data availability. Actual cost spectra has been estimated only occasionally. Data is available on an annual basis for the 1990s, but not prior to that (at least not in easily accessible, electronic form). Where actual cost spectra are available, these are used, and where they are not available, hypothetical costs are used. In order to construct a full time series, simple linear extrapolation between the available observations is carried out. This is acceptable because costs are not expected to exhibit high levels of short-term volatility. We find that in some cases where both actual and hypothetical costs can be observed for the same point in time, they are similar, but at other times not similar. Where there is conflict we prefer the actual costs. The total cost series, along with the observations used to construct it, is shown in Figure 3.7.[93]

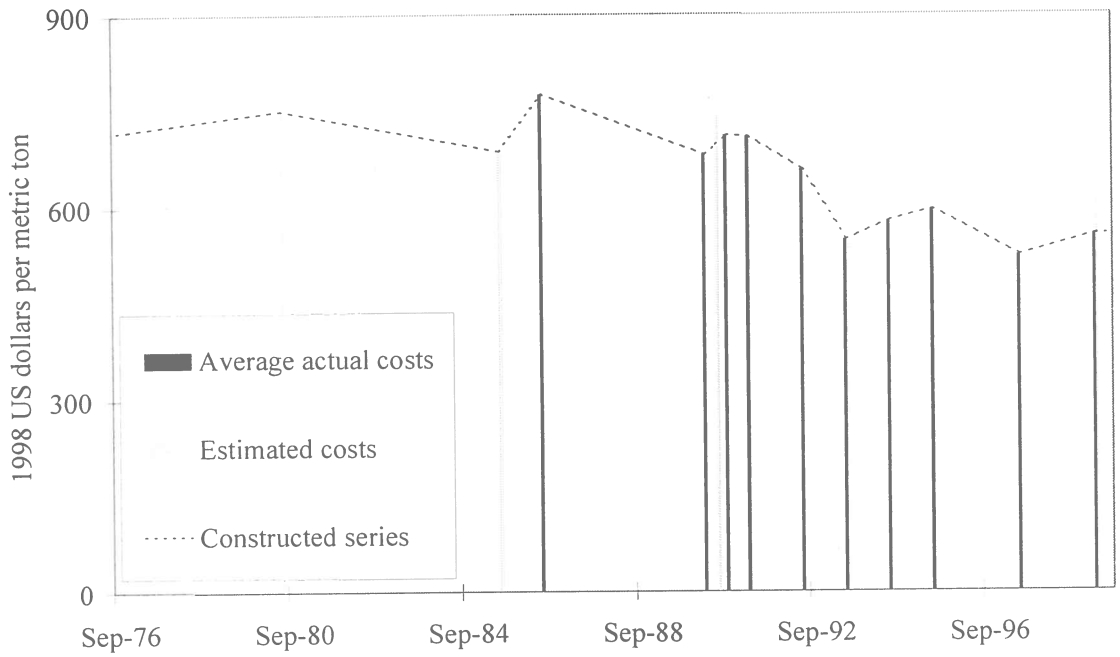

Figure 3.7: Total Costs of NBSK

As can be seen, few observations are available for the earlier part of the period of interest. Nevertheless, no superior data is available.

[93] As will be discussed in section 4.1.5, prices are deflated into real terms. The costs shown in Figure 3.7 are deflated in order to be consistent with this treatment.

5 CHAPTER 4: MODELS

"There are two things you are better off not watching in the making: sausages and econometric estimates."

Edward Leamer

This section explores the possibilities for using the techniques outlined in chapter 2 to construct time series models of pulp prices. The approach used is an incremental one: Four models are estimated, in increasing order of complexity. The first step is to analyze prices in isolation from consideration of other variables. For this purpose, the characteristics of the price series itself are first investigated, and an ARMA model is constructed. The second model is an attempt to represent the short-term structure of the market (as described in section 1.2) using a VAR model. The third and fourth steps analyze different aspects of the long-term structure of the market. The third includes investment as well as price, and uses an error correction framework to model the links between production, shipments and capacity (as described in section 1.3.1). The final step also uses the error correction technique, this time to model the long-term link between prices and total costs (as described in section 1.3.2). For each model, the following steps are required: Selection of appropriate variables for inclusion in the model; selection of an appropriate lag structure; specification of the model's reduced form; estimation of the coefficients; assessment of the economic plausibility and statistical significance of the estimated coefficients; and assessment of the statistical adequacy of the model.[94]

5.1 UNIVARIATE ANALYSIS OF PRICES

Here we examine how prices behave, in isolation from other variables. There are a number of attributes which time series may potentially have that are relevant from a modeling perspective. These include deterministic and stochastic trend patterns, seasonality, structural breaks, non-stationarity, and heteroscedasticity. The presence of all of these is tested for. The second part of this section attempts to develop a simple univariate time series model of prices. This will serve as a basis for comparison for the multivariate models to be developed in the following sections.

It is apparent that prices exhibit a considerable degree of cyclicality, at least in the post-1974 period (see Figure 3.1). An interesting feature of the series is that major peaks (with the exception of the most recent) are plateaus. One possible explanation for this is that producers chose not to raise prices as much as market fundamentals would have allowed, and as a result some time passed before prices began to fall again. Christensen and Caves (1997) suggested the reason for this price restraint was to avoid encouraging new investments. This price restraint may well have benefited producers, by enabling high prices to be maintained for longer periods of time. This did not occur during the 1995 peak. Its absence may possibly be explained by the fact that the nature of the market has expanded from an exclusive 'club' of producers mainly within the Norscan countries, to now include producers in other countries who may have different time horizons. In particular, hardwood Kraft producers with low production costs but high debt levels may be more aggressive in adjusting prices both upwards and downwards.

Another interesting feature of the series is that a 'minor peak' of short duration has followed each of the last three major peaks. A possible explanation for this is that producers are unwilling to accept that the higher prices achieved earlier are no longer justified, and they attempt to regain these price levels immediately, before market conditions are able to sustain the higher prices again.

The original series is transformed into logarithmic form. The original series is represented hereafter as P_t and the log series as p_t. This is a commonly applied transformation. The effect of taking logs is to transform an exponentially growing series into a linearly growing one. Thus, for example, steady inflation will appear as a linear trend, which is convenient for modeling purposes.

---

[94] In some cases other steps are also required. These are discussed on an *ad hoc* basis.

5.1.1 Heteroscedasticity

A series that exhibits differing levels of volatility during different time periods is said to be heteroscedastic. Visual examination of the series suggests immediately that prices have become more volatile since the early 1970s. It seems likely that this is related to the changes in the overall economic environment that occurred at that time, as a result of the breaking down of the fixed exchange rate system, and also to the first oil shock of 1974.

One simple way to test for conditional heteroscedasticity that does not first require the development of a model for prices is given by Franses (1998, 24). We regress the variance of the series on its lagged value:

$$(p_t - p_{t-1})^2 = \alpha + \rho(p_{t-1} - p_{t-2})^2 + u_t, \qquad t = 1950{:}3,\ 1950{:}4, \ldots\ 1999{:}1.$$ [95]

A positive coefficient on the lagged value indicates that volatility is serially correlated. For the sample of 1950:3 – 1999:1, we obtain an estimated value for ρ of 0.308 (with an estimated standard error of 0.068). This suggests that the series is heteroscedastic, consistent with visual examination of the series.

One means of dealing with this heteroscedasticity is to construct a model that is specifically designed to account for it, e.g. some variant of the autoregressive conditional heteroscedasticity (ARCH) model. Another approach is to focus attention only on a subsample of the data within which heteroscedasticity is less evident. In this case we can proceed to estimate an ARMA model without the need to take account of potential heteroscedasticity. Given that we are primarily interested in forecasting the level of prices rather than their volatility, the latter method is chosen.

5.1.2 Structural Breaks

The purpose of this section is to examine the price series for evidence of 'structural' changes in the nature of the underlying data-generating process. The reason for doing this is that any such change threatens the validity of using a model estimated on the basis of the entire series for forecasting purposes. An econometric model estimated using data that pre-dates a fundamental change of this nature cannot be regarded as very useful for forecasting values of the series that post-date the change. It is more useful to use only post-change observations for estimation, despite the loss of degrees of freedom that this involves.

Prices rose dramatically in 1950 and 1951, and then fell back to their original levels in 1952. For a long period following that, prices were extremely stable. Between 1953 and 1972, the quarterly change in the nominal price level never exceeded $14, and was quite often zero (for 47 quarters of a total of 80). The price ranged between $139 and $186 (in nominal terms). This pattern clearly changed over the following years: by 1975:1 the price had risen to $468. Quarterly price changes between $10 and $50 became common (with occasional larger changes), and static prices less common. This pattern has continued up until the present. Clearly the period between the early 1950s and early 1970s was quite different from the period following. Which is the most appropriate time to treat as the point at which this 'structural break' occurred? 1973:1 was the first quarter of what turned out to be a sustained price rise to previously unattained levels, and we may therefore regard that point in time as being a break in the series.

We also observe that between 1973:3 and 1976:2, price changes (both positive and negative) were made on a half-yearly basis, so that a quarter of no change was followed by a quarter of non-zero change. This was a result of the institutional practices of that time period.[96] Since this pattern has not been observed since 1975, we may consider this to be an 'interim' period partway between the earlier period of stable prices and the later period of prices changing on a quarterly (or more frequent) basis. Given that this pattern will have a significant impact on the results of estimation (as regards the short-term dynamics of the market), it is appropriate to exclude this interim period from

---

[95] All the analysis is done on a quarterly basis. The figures following the colon in a date represent the quarter. For example, 1976:3 represents the September quarter of 1976.

[96] In particular, it was due to the semiannual timing of a price-setting meeting which was held between major Scandinavian producers and UK buyers, and which formed the basis for overall market prices.

the sample for purposes of estimation.[97] We therefore choose 1976:3 as the start of our sample for model estimation purposes. There is no visual evidence of structural breaks having occurred after this date, and neither have the institutional structures of the industry changed greatly.[98]

5.1.3 Seasonality

As noted in section 1.2.1, various seasonal effects occur. Both demand and supply side factors are subject to seasonal patterns, and we may expect to see this seasonality transmitted to prices.

Seasonality can be tested for by the inclusion in a model framework of seasonal dummies for each except one of the seasons. The significance or otherwise of the estimated coefficients on these dummy variables will indicate whether seasonality is important. To get a preliminary indication of the overall importance of seasonality, a regression of the real price (in logarithmic form) on only these dummies is carried out (estimated over 1976:3 – 1999:2):

$$\underset{}{p_t} = \underset{(0.05)}{6.48} - \underset{(0.06)}{0.03}\, Q1 - \underset{(0.06)}{0.02}\, Q2 + \underset{(0.06)}{0.01}\, Q3$$

Where *Q1*, *Q2* and *Q3* are the dummy variables for the first three-quarters and the figures in parentheses are estimated standard errors. None of the dummy variable coefficients are statistically significant, and the R^2 statistic for this equation is only 0.004. This suggests that we can ignore seasonality in the following analysis. This is a somewhat surprising finding, given the number of seasonal factors that exist. It suggests that the primary adjustment to these factors occur in quantities (in particular, via changes in inventories) rather than prices.

5.1.4 Deterministic Trend

In broad terms, a trending series is one that does not exhibit fluctuations of uniform variance around some fixed mean. Two types of trends can be defined within the context of ARMA modeling, each having fundamentally different implications. These two types are deterministic and stochastic trends. A deterministic trend in a series is represented by means of a time-varying regressor in the model. A stochastic trend is identified by the presence of a unit root in a series (which was described in section 2.1.1).

One simple way to test for the presence of a deterministic trend is to evaluate the statistical significance of the time-varying regressor in isolation. The following equation is estimated using the sample period 1976:3 – 1999:2:

$$p_t = c + \delta t + \varepsilon_t$$

Where c is a constant and t takes the values 1,2,3… for 1976:3, 1976:4, 1977:1…. The estimated equation is

$$p_t = \underset{(0.047)}{6.060} + \underset{(0.001)}{0.005}t + \varepsilon_t$$

---

[97] Originally, 1973:1 was taken as the sample starting point. With this starting point, both AIC and SIC suggested use of an ARMA (2,1) model. The estimated model over this sample had a negative AR (1) coefficient, and, upon further investigation, it was found that changing the sample starting point to 1976:3 had the effect of changing the sign of *all* of the estimated coefficients! This gave an indication both that the choice of sample starting point is quite important, and also that the ARMA (2,1) specification was not very robust.

[98] One recent institutional change that may cause price volatility to decline is the introduction, since 1996, of pulp futures markets. However, futures trading has not yet gained widespread acceptance, and opinion is divided on whether or not this will in fact lead to any reduction in spot price volatility (see e.g. Fromson (1997) and Finchem (1998) for opposing arguments).

The estimated coefficient δ is significant at a 95% confidence level, and indicates annual price growth of approximately 2%. This is in nominal terms. To remove the effects of general inflation, we can obtain a real price series by deflating this nominal series by some measure of general inflation. As the price series is denominated in US dollars, an appropriate deflator should relate to US prices. The prices are deflated using the US Wholesale Price Index (with the 1998 annual average WPI level taken as the base period) to give a real series.

When the same analysis as above is carried out for this real price (again in logarithms), a negative trend is found, and is also significant. However, given the magnitude of fluctuations in the series, this result is quite sensitive to the sample period chosen. The sample period used here begins at a peak and ends in a trough. Estimation is redone over the period 1978:3 – 1999:1 as these two points appear similar in terms of their position in the cycle (i.e. both are just after troughs). This results in a finding of a non-significant trend (at a 95% confidence level). We tentatively conclude that real prices do not exhibit a deterministic trend. This issue will be revisited once a model of the series has been constructed.

5.1.5 Stochastic Trend

A non-stationary series is said to contain a stochastic trend. Conceptually, whether a given series contains a stochastic trend or not will depend upon the types of shocks to which it is subject. If it is subject to permanent shocks, then it will contain a stochastic trend. Shocks that may affect pulp prices include demand shocks, input cost shocks, technology shocks, and speculative shocks (due to changes in expectations).

Some of these shocks will be transitory. As described in section 1.3.2, economic theory suggests that in the long term companies in a competitive industry will make normal profits. The margin between prices and costs will be just sufficient for them to earn a 'normal' profit on capital employed.[99] Adjustment will occur by way of changes in capacity of the industry as a whole. This implies that shocks to demand will be reflected by changes in capacity, and will not have lasting effects on prices (even though the adjustment time may be quite long).

Technology shocks will be permanent. Broadly speaking, two types of shocks to input possibilities are important for the pulp and paper industry. One is application of advances in scientific knowledge that allow the input/output ratio in production to be reduced, while the other is changes in environmental regulations preventing the industry from using certain inputs or processes. In general, these two will have opposite effects on costs. Since prices and costs are linked in the long term, they will permanently affect the 'equilibrium' price level, if the change in costs is itself permanent. Changes in input costs may be either transitory or permanent, depending on the nature of the market for the particular input involved. Therefore we can conclude that, at a theoretical level, pulp prices are subject to at least some permanent shocks, and therefore the series will be non-stationary.

Before proceeding to test for non-stationarity, we note that, as an empirical matter, it will not necessarily be detectable in a finite sample. A non-stationary series will appear to be stationary if the permanent shocks do not occur over the period the series is observed; are not statistically significant or distinguishable from a deterministic trend; or have offsetting effects on the series, so that the net effect is small.

We use the unit root tests described in chapter 2 to test for the presence of a stochastic trend. These tests are run on the log of the nominal price series from 1976:3 to 1999:1. In doing these tests, we allow for the possible presence of a deterministic trend in the series, as well as a non-zero intercept term.

Before carrying out the ADF test it is necessary to select a maximum lag length. This represents the longest period over which lagged values of prices have an influence on current prices. A maximum lag length of 12 quarters (3 years) is selected, as this is consistent with the discussion of investment lags (2 – 2 ½ years) earlier. Testing shows that only a 1-quarter lag is significant, so this is the specification used in the ADF test. The ADF test statistic is –

---

[99] An implication of this is that the margin between costs and prices will be stationary. This suggests that forecasting margins may be more successful than forecasting prices *per se,* if no acceptable method to deal with non-stationarity of prices is available.

3.17. This causes us to reject the hypothesis of a unit root at a 90% significance level, but not at a 95% significance level. The PP test statistic is –2.37. This causes us to accept the hypothesis of a unit root at a 90% significance level. Overall, this is sufficient evidence for the presence of a unit root in the data.

This presents us with the problem of how best to transform the series to a stationary form (which is necessary for the statistical analysis to follow). What is the reason for this non-stationarity? This is a difficult question to answer conclusively, but may provide useful clues as to how best to transform the series.

Empirically, most non-stationarity in other commodity prices has been found to be the result of inflation. Ardeni (1989) investigated nominal prices of a number of different commodities (not including pulp or any paper products), and found consistent evidence that they had unit roots, using the tests outlined above. Deaton and Laroque (1992) did the same for real prices, and found no evidence of non-stationarity. Thus, one way to remove at least a part of the non-stationarity from the series is to deflate it by some measure of general inflation.

Another possible cause of non-stationarity is permanent technology or regulatory shocks, as discussed above. There is no obvious method that we can use to remove the effects of these shocks from the data. The only approach available to us in this case is to assume that such shocks, while permanently affecting the level of the series, will not affect future *changes* in the level of the series. We can then examine these changes.

Thus, two methods for transforming the series present themselves: to deflate it into real terms, or to take first differences. Neither of these is guaranteed to provide us with a stationary series, but the same tests as above can be done on the transformed series. The two series are shown in Figure 4.1, and are quite different. The deflated series has persistent cycles (as does the nominal price series) while the differenced series exhibits less serial correlation.

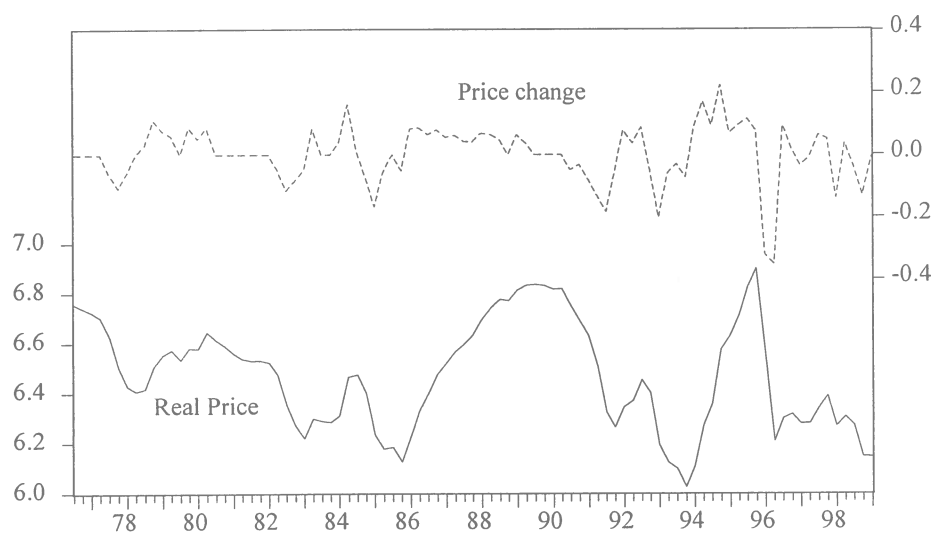

Figure 4.1: Transformed NBSK price

The choice of which method to use will have important implications for forecasting. A characteristic of ARMA forecasts is that shocks up until the time of forecasting will gradually disappear from the forecasts (if the ARMA specification is stable), and the series will be forecast to converge to its mean value. If forecasts of price change are made, they will gradually converge to a value equal to the deterministic trend growth (which in the present case is close to zero). This implies that the price level will be forecast to stay at or near its present level. In contrast, a model of the price level will forecast the price level to return towards its mean value. Given the considerable cyclicality evident in the original series, the latter is more attractive, and for this reason deflation is preferred over differencing. Deflating into real terms to remove the stochastic trend is acceptable in the present context. General inflation is not the only cause of non-stationarity in prices, but empirically has been found to be a major cause of it.

Here we are trading off conceptual correctness (which supports differencing) against likely plausibility of the results (which supports deflating). This tradeoff is made palatable by the expectation that the degree of divergence from the deterministic trend over the forecast horizon will be limited.[100]

Visually, the deflated series appears stationary, although its mean-reverting behavior is slow. We use the same tests as above to test formally for stationarity. A deterministic trend is allowed for in both of these tests as well as a constant. For the ADF test, we again choose a maximum lag length of 12, and find that only a 1-quarter lag is significant. The ADF test statistic is -4.57, and the PP test statistic is –4.02. Both of these allow us to reject the hypothesis of a unit root at the 99% confidence level. We conclude that the real price series appears stationary.

5.1.6 ARMA Model

Having decided upon an appropriate variable to model, we proceed to construct a simple ARMA model of it. Background to this family of models was discussed in section 2.2.1. We use the methods described in that section to select an appropriate lag structure. In this case, the ACF declines asymptotically, while only the first two values of the PACF are significant at the 95% confidence level (Figure 4.2). This suggests that an AR (2) model is appropriate.

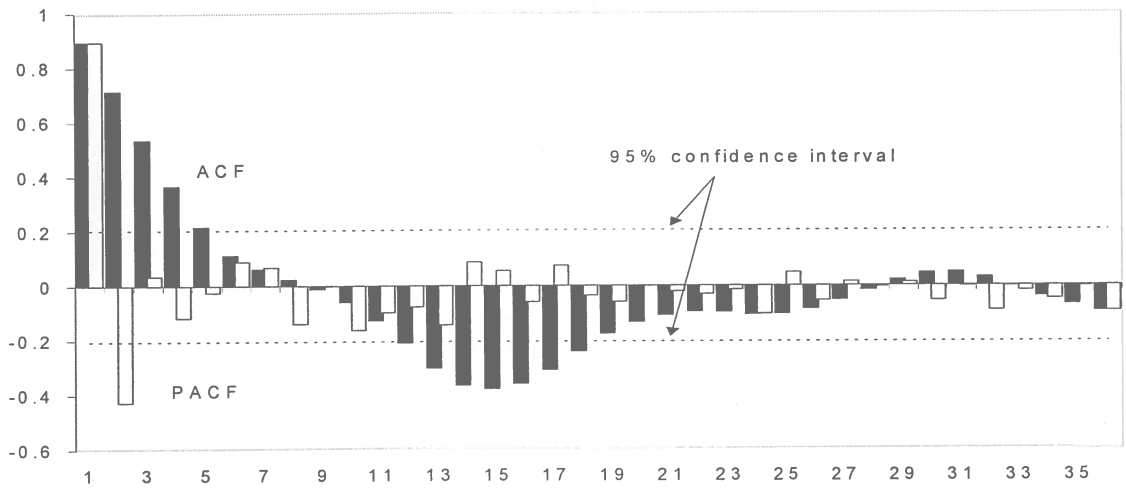

Figure 4.2: ACF and PACF of Real NBSK Price

The AIC and SIC statistics (also described in section 2.2.1) are calculated for all permutations of $p = 0,1,2,3$ and $q = 0,1,2$ based on data from 1976:3 to 1991:4. Post-1991 data is reserved for out-of-sample testing and is therefore not used either in the estimation process for any of the models or in the preliminary analysis relating to their specification.[101] Results are shown in Table 4.1.

---

[100] A more suitable approach than both of these will be discussed in section 4.4.

[101] One exception is made: In sections 4.3 and 4.4 cointegration testing uses the full sample. This is because, in these cases, we intend to impose these cointegrating relationships regardless of the empirical support for them. The cointegration tests are therefore not part of the model development process, but can be regarded as a separate test for the 'reasonableness' of the assumed cointegrating relationships.

Table 4.1: AIC and SIC Tests of ARMA Models

ARMA(p,q)	q = 0	q = 1	q = 2
p = 0	-0.422 -0.353	-1.468 -1.365	-2.060 -1.923
p = 1	-2.570 -2.466	-2.849 -2.711	-3.005 -2.833
p = 2	-2.993 -2.853	-2.960 -2.785	-2.957 -2.747
p = 3	-2.943 -2.767	-2.995 -2.784	-2.963 -2.717

The AIC statistic is shown above the SIC statistic for each case.

According to the AIC, an ARMA (1,2) model is preferred, and according to the SIC, an AR (2) model is preferred. Because it allows for a more intuitive interpretation, and because the margin of preference is slightly larger, the AR (2) specification is selected.

One representation of the general form of an AR (2) model with a constant and deterministic trend is: [102]

$$p_t = c + \delta t + u_t, \text{ Where } u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t .$$

The model is estimated using OLS, and is as follows:

$$\underset{(0.17)}{p_t = 6.45} + \underset{(0.004)}{0.001}\, t + u_t \quad \text{, where} \quad u_t = \underset{(0.11)}{1.53}\, u_{t-1} - \underset{(0.11)}{0.62}\, u_{t-2} + \varepsilon_t$$

Estimated standard errors of the coefficients are given in parentheses. The coefficient on the time trend is insignificantly different from zero, while all other coefficients are significant.

We first examine the stability properties of the estimated equation. The eigenvalues for this equation are

$$\lambda = \frac{1.53 \pm \sqrt{(1.53)^2 - 4(0.62)}}{2} = 0.76 \pm 0.19i$$

The fact that these eigenvalues are complex means that forecasts made with this model will have sinusoidal components, and thus will follow a cyclical pattern.[103] This is an attractive result in the present case, where the series clearly does follow a cyclical path. To examine whether the model is stable or not, we calculate the moduli of these eigenvalues:

$$R = \sqrt{(0.76)^2 + (0.19)^2} = 0.78$$

---

[102] This representation is adopted for convenience, as it is used by the Eviews software package that is used for this section. An alternative is $p_t = c^* + \delta^* t + \phi_1 p_{t-1} + \phi_2 p_{t-2} + \varepsilon_t$, where c and c^*, and δ and δ^*, are not identical.

[103] See Hamilton (1994, 16).

Since $R<1$, we conclude that the amplitude of the sinusoidal patterns will decay over time, and that the system is stable. In other words, the impact of any shock will gradually disappear over time, at a quarterly rate of R. We can also estimate the length of these cycles:

$$L = \frac{2\pi}{\cos^{-1}(a/R)} = \frac{2\pi}{\cos^{-1}(0.763/0.786)} = 26 \text{ Quarters}$$

This implies that the cycles in projections will be about 6½ years from peak to peak. This is reasonably close to what historically has been the case.[104] In practice, the rate of decay of the oscillations will mean that the cycle will be very small well before the end of the first cycle, and so this cyclical behavior will not be quantitatively significant.[105]

It is informative to rewrite the estimated model as an infinite-order MA model. This can be done for any covariance-stationary process, according to the Wold decomposition. Hamilton (1994) presents a method to calculate the MA coefficients as follows:

$$\varphi_j = f_{11}^j \ ,$$

Which is the [1,1] element of the matrix $F = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}$ raised to the power of j.

Calculating these coefficients by means of matrix multiplication is somewhat tedious, but fortunately a recursive solution is possible:

$$f_{11}^j = \phi_1 f_{11}^{j-1} + \phi_2 f_{11}^{j-2}$$

Given $f_{11}^0=1$ and $f_{11}^1=\phi_1$, we can calculate the remaining coefficients. These are plotted in Figure 4.3. This shows that any shock to prices will initially be amplified, and then gradually disappear. The cyclicality of the model is swamped by its decay.

---

[104] Anecdotal evidence suggests that the cycle is shortening. The current work sheds no light on this issue, however.

[105] This can be seen from the MA coefficients shown in Figure 4.3.

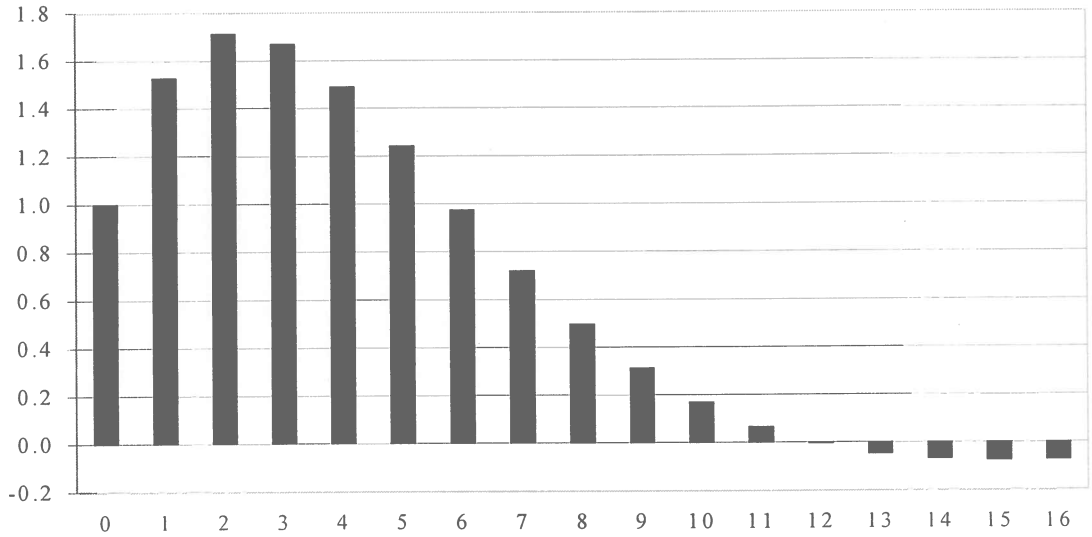

Figure 4.3: Wold representation MA Coefficients of AR (2) Model

We next assess the statistical validity of this model as an adequate description of the data generating process. Ideally, the residuals from the estimated model should be Gaussian white noise, i.e. independent and identically normally distributed. They are graphed in Figure 4.4.

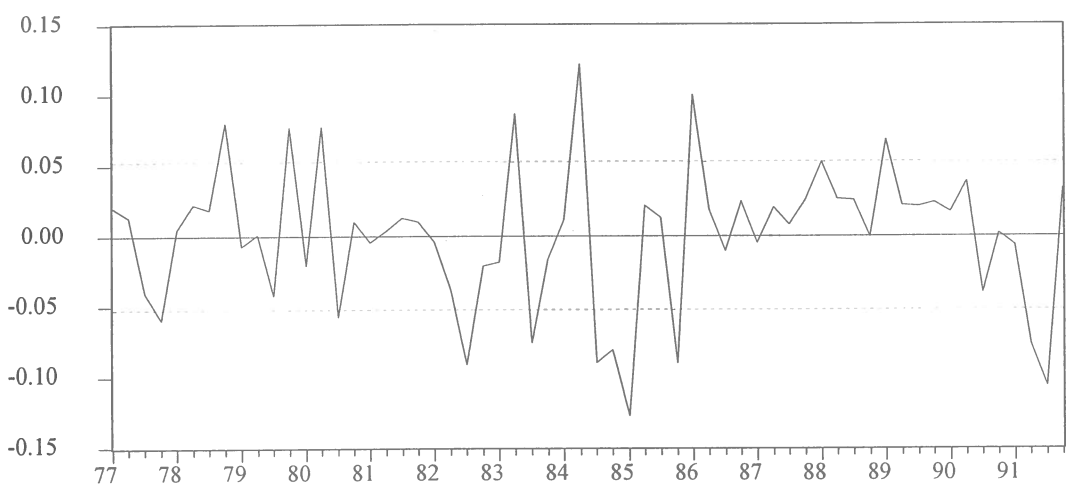

Figure 4.4: Estimated Residuals of Ar (2) Model

No violations of the assumptions made are visually evident. Both the skewness (-0.2) and kurtosis (3.2) of the distribution of the residuals are quite close to the values of a normal distribution (0 and 4 respectively). A test statistic often used is the Jarque-Bera statistic, which considers both of these:

$$JB = \left(\sqrt{\frac{n}{6}} \cdot s\right)^2 + \left(\sqrt{\frac{n}{24}} \cdot (k-3)\right)^2 \sim \chi^2(2)$$

Where s and k are skewness and kurtosis respectively. In this case the JB statistic is 0.72 (against a critical value at 95% confidence level of 5.99) that leads us to accept the hypothesis that the residuals are normally distributed.

Independence of the residuals requires that there be no serial correlation between them. There is little visual evidence of serial correlation except over the period from 1987 to 1990, when the residual is consistently above zero. The autocorrelation function of the residuals is calculated, and none of the autocorrelation coefficients are found to be significantly different from zero at the 95% confidence level (Figure 4.5).

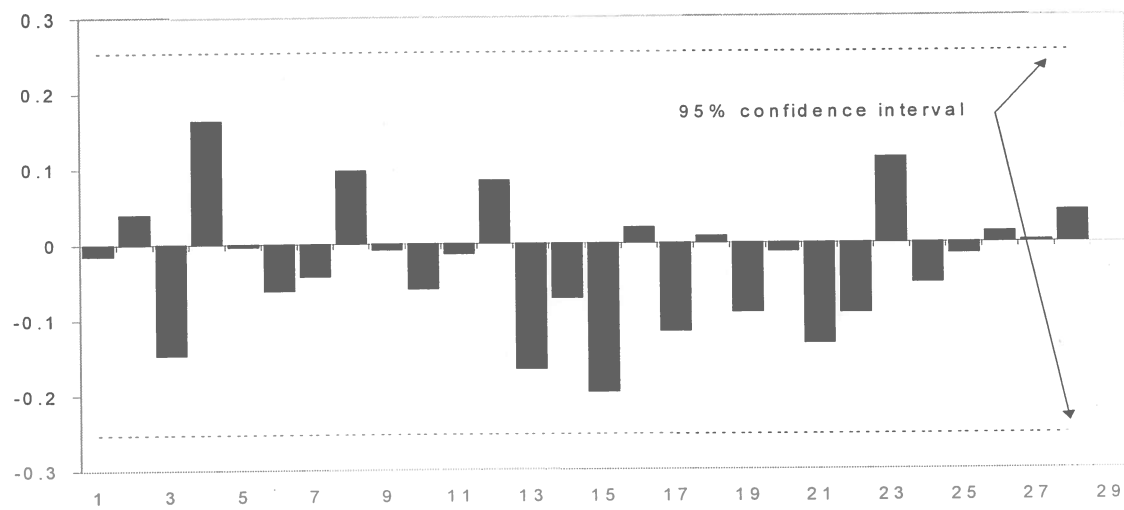

Figure 4.5: Residual ACF of AR (2) Model

The Durbin-Watson statistic is not appropriate as a test for serial correlation in this case, due to the presence of a lagged dependent variable on the RHS of the model. More suitable tests are the Ljung-Box (1978) Q test and the Breusch-Godfrey LM test. The Q test is a joint test for the significance of the first k residual autocorrelations. The test statistic is

$$Q = n(n+2)\sum_{j=1}^{k}\frac{\rho_k^{\,2}}{(n-j)}$$, where ρ_k is the estimated residual ACF parameter.

Asymptotically, these are distributed as χ^2 with k - p - q degrees of freedom (where p and q are the AR and MA parameters). The Q statistic allowing for up to 3 orders of autocorrelation is 1.51, against a 95% critical value of 3.84. Statistics allowing for higher orders of autocorrelation also indicate that no serial correlation is present.

The Breusch-Godfrey test is based on the LaGrange Multiplier principle for comparing different models. To assess the adequacy of an AR (2) model, for example, as compared to an AR (2+r) or ARMA (2,r) model, the following estimation on the estimated residuals of the AR (2) model is carried out:

$$\hat{\varepsilon}_t = \alpha_1 p_{t-1} + \alpha_2 p_{t-2} + \alpha_3 \hat{\varepsilon}_{t-1} + \alpha_4 \hat{\varepsilon}_{t-2} + \ldots + \alpha_{2+r}\hat{\varepsilon}_{t-r} + \upsilon_t$$

The statistic is calculated as nR^2, and is asymptotically distributed as χ^2 with r degrees of freedom. In this case we choose $r = 2$, and calculate a statistic of 0.15, as against a critical value at 95% confidence level of 5.99. This allows us to accept the hypothesis of no serial correlation in the residuals.

The AR (2) model appears to provide a statistically valid description of the series. We can now use this model to re-examine some of the issues looked at earlier; namely deterministic trend, seasonality, and heteroscedasticity.

A deterministic trend was included in the estimated AR (2) model above. The estimated coefficient on this variable is positive (indicating a trend real price increase of approximately 0.1% per quarter, or 0.4% per year) but not significantly different from zero. When the time trend was estimated in isolation, the result was very sensitive to the sample period chosen. Estimation of the deterministic trend in the context of an explicitly cyclical model reduces

this dependence. The sample period used is from near a peak in 1976:3 to a trough in 1991:4, but despite this an insignificant trend is found. This strengthens our earlier conclusion that the deterministic time trend is unimportant.

Seasonality is tested for in two ways. Firstly a seasonal AR component is included, so that the error process for the model estimated becomes

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \zeta u_{t-4} - \zeta\phi_1 u_{t-5} - \zeta\phi_1 u_{t-6} + \varepsilon_t$$

Where ζ is the seasonal parameter. This parameter is found to be insignificantly different from zero at a 99% confidence level. The second method tried is simply to include seasonal dummies in the estimated regression. The estimated coefficients on all of these are not significantly different from zero. We therefore continue to regard seasonality as empirically unimportant.

We test for the presence of auto-regressive conditional heteroscedasticity. Engle (1982) proposed a LaGrange Multiplier test. The test is carried by means of an auxiliary regression of the squared residuals on their lagged values (with up to q lags), with the test statistic equal to nR^2, as before. Testing for ARCH up to 2 lags gives a test statistic of 2.96, against a 95% critical value of 5.99. We therefore accept the hypothesis of no heteroscedasticity within the time period 1976:3 – 1991:4.

In summary, the AR (2) model estimated appears to be statistically acceptable. The coefficients are plausible and offer some insights into the cyclical nature of prices. The analysis does not offer any evidence of seasonality or a deterministic trend. The forecasting performance of the model, along with that of the other models, will be assessed in chapter 5.

5.2 VECTOR AUTOREGRESSION MODEL

The purpose of this section is to construct, estimate, and evaluate a multivariate model of the pulp market consistent with the discussion of the short-term market structure as discussed in section 1.2.

5.2.1 Variable Selection

Our conceptual starting point is Figure 1.1. In refining this to an operational selection and definition of variables, the issues we need to consider are whether each of the variables is stationary, and any transformation of the raw data that may be desirable. The following variables are considered for inclusion in the model: demand, production (capacity x capacity utilization), shipments and inventories. These variables are not independent. Producer inventories (*inv*) are uniquely determined by the cumulative difference between production (q) and shipments (s):

$$inv_t = inv_{t-1} + q_t - s_t \quad \Leftrightarrow \Delta inv_t = q_t - s_t$$

Thus, including all three of these variables (with inventories differenced) as explanatory variables will result in perfect multicollinearity, and parameter indeterminacy.[106] To avoid this, we can exclude one of the three parameters from consideration. Because available evidence indicates that all three of these are regarded as important indicators of market conditions, we have no strong *a priori* indication of which should be excluded. We leave this issue open for now.

The data available dictate that the 'market' concept used is truncated to exclude any consumer behavior beyond the purchase of pulp. Pulp is considered to leave the market once it is sold to a consumer, and any choices made by consumers as to their inventory levels do not affect our analysis (at least not explicitly). The proxy used for demand is shipments from producers to consumers. Consistent with this, the inventory data is limited to producer inventories. For the purposes of forecasting price, this treatment is attractive because changes in demand for

---

[106] In fact, given the data discrepancies between the three variables, multicollinearity will not be perfect, but it will result in highly uncertain parameter estimates nevertheless.

consumer inventories may often have an impact on prices, independent of any change in the underlying usage rate (i.e. paper production rate).

Variables included in a classical VAR model should be stationary, both individually and jointly. If some variables are non-stationary, then they must be transformed into stationary form (often by differencing). As discussed in section 4.1.5, real price is conceptually non-stationary, but empirically, modeling it as if it were stationary is preferable to modeling it in differences. We continue to treat real price as stationary in this section.

We next examine capacity. In the long term, the level of capacity will be directly related to the level of demand for pulp, which in turn will be related indirectly to global levels of aggregate demand. Considerable energies have been devoted to analyzing the nature of GDP (and other macroeconomic variables) and at present it is widely regarded as being non-stationary (see Blanchard and Quah, 1989). Apart from macroeconomic variables, pulp demand is affected by developments in technology and preferences, which are also likely to be non-stationary in their effects.

Because of potentially varying income elasticity's of demand for pulp, as well as changes in demand not caused by GDP (such as technological developments); demand for pulp is unlikely to be a linear function of GDP. Nevertheless, non-stationarity in GDP and in the other factors which affect pulp demand imply that the latter is also likely to be non-stationary. This in turn implies that capacity will also be non-stationary.

We next consider production and shipments. Both of these are effectively limited by capacity. Production will rarely exceed 100% of capacity even for isolated periods. While it may fall significantly below capacity for extended lengths of time, in the long-term demand growth and restriction in capacity growth will cause it to return to some level close to capacity. Historically this has averaged about 90%. Given that production is a prerequisite for any shipments, capacity also forms a constraint on the latter. Shipments may differ from production through the buildup or drawdown of inventories, and this means that capacity does not form as strong a short-term constraint as it does for production. In terms of the time-series properties of production and shipments, both will be non-stationary, given that they will track capacity.[107]

Capacity is determined solely by long-term decisions on investment, whereas production and shipments are also affected by short-term decisions related to market conditions. Therefore we can expect considerably more volatility in the latter two variables than in the former. This implies that non-stationarity will be more difficult to detect (either visually or statistically) in those series. Nevertheless this difference in volatility does not impinge upon the conceptual conclusion regarding the non-stationarity of these series. The three variables are shown in Figure 4.6, and the data appear to confirm the discussion above.

---

[107] This implies that all three series will be cointegrated (this will be discussed in detail in section 4.3).

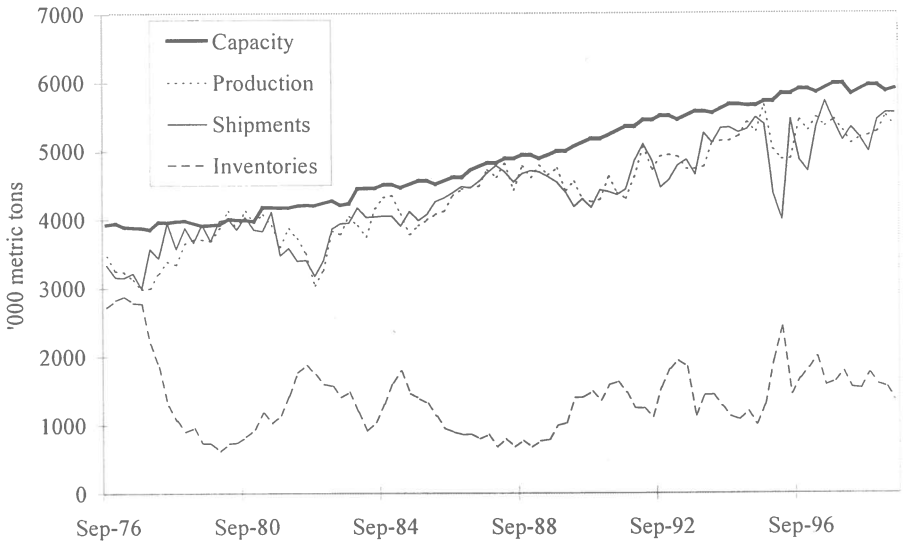

Figure 4.6: Norscan Chemical Paper Grade Pulp Volumes

We next turn to inventories. Inventories are held by suppliers as a result of any mismatch between production and shipments, and are held by users as a result of any mismatch between shipments and use. Thus, we would expect the volume of inventories to be related to the level of shipments (or production) in the industry. However, this is not a direct relationship, and the proportion of inventories to sales may change due to changing business practices (such as changes in the frequency of shipments, introduction of just-in-time inventory management, or changes in the extent of coordination between buyers and sellers. These influences will affect inventories, but not capacity. Therefore, we can expect inventory levels to be non-stationary, but not to be cointegrated with capacity (nor production or shipments). Inventories are also shown in Figure 4.6. Visually it is difficult to ascertain whether the series is non-stationary or not, because of the degree of volatility present and the short time period over which data is available.

As noted above, variables to be included in a VAR model should be stationary. Untransformed, none of the candidate variables meet this requirement, at least according to these theoretical considerations. One possible transformation that can be made is to take differences of the variables. In some situations this is acceptable: For example, Duy and Thoma (1998) found that VAR models expressed in differences performed much better in forecasting macroeconomic variables. Here, however, it is not attractive to model price in differences. The effect of forecasting price differences using time series methods is that the implied forecasts for the levels of the variables will be highly dependent on their values at the time of forecasting. This is unattractive in the present case, since prices are cyclical. Also, according to our understanding of the market, it is most appropriate to consider price *levels* to be a function of inventory and capacity utilization *levels* rather than *changes* (see section 1.2). Therefore, given that we have chosen to model prices in levels rather than differences from it seems appropriate to model the other variables in level form also.

An alternative transformation is to calculate ratios of the non-stationary variables. Where two non-stationary variables are cointegrated (with the cointegrating vector having a zero intercept) we would expect the ratio of the two variables to be stationary. From the discussion above, we can therefore conclude that any ratio calculated using production, shipments and capacity is likely to be stationary. The ratio of production to capacity is the commonly discussed capacity utilization rate. The ratio of shipments to capacity is the shipment rate.

Because inventories are not cointegrated with capacity, the ratio of the two is not expected *a priori* to be stationary, and is therefore not a candidate for inclusion in the model. The only transformation available that will allow us to convert it to a stationary form is to difference. We therefore choose to exclude inventories from the model. Although inventories are undoubtedly an important factor in determining short-term market conditions, the fact that they are collinear with shipments and production (as discussed above) means we can do this without concern for omitted variable bias.

Including both capacity utilization and shipment rates in the model has the disadvantage that both contain capacity as the denominator, so that some multicollinearity between the two variables will result. This will cause parameter estimates to be imprecise. However, because variation in capacity is relatively small compared to variation in production and shipments, it is expected that this problem will be relatively minor.

ADF and PP tests are conducted on both of the variables to be included in the model as a check on the reasoning above. No deterministic trends are allowed for. The maximum ADF lag allowed is 12, as before. The results of these tests are presented in Table 4.2. The unit root hypothesis is rejected for each variable, confirming the reasoning above.[108]

Table 4.2: Unit Root Tests for Capacity Utilization and Shipments Rate

	No. of ADF Lags	ADF Test Statistic	PP Test Statistic
Capacity Utilization	12	-3.67***	-4.00***
Shipments Rate	10	-3.36**	-4.99***

5.2.2 Seasonality

Both production and shipments exhibit seasonal patterns. There are a number of reasons for these patterns, including both demand and supply side factors, as discussed in section 1.2.1. The quarterly averages for production, shipments and capacity are shown in Table 4.3.

Table 4.3: Averages by Quarter for Norscan Volume Measures

Quarter	Capacity	Shipments	Production
January-March	4876	4370	4458
April-June	4898	4508	4359
July-September	4851	4271	4357
October-December	4862	4353	4342

Capacity shows very little seasonal variation, as expected. Production is slightly higher in the first quarter of the year than for the other three-quarters. Shipments show the most seasonal variation, with the second quarter being notably higher than the other quarters, and the third quarter lower. These differences are not large, however, with the second quarter peak being only about 5% higher than the third quarter. Seasonal dummy variables will be included in the model to be estimated, and retained if they are found to be significant.

5.2.3 Model Estimation and Statistical Evaluation

The first issue to address in construction of a multivariate model is which variables to treat as endogenous and which to treat as exogenous. *A priori* knowledge leads us to expect that each of the variables modeled in this case will potentially affect all the others. This suggests that all should be treated as endogenous. As an empirical check on this, Granger causality tests may be used (described in section 3.1). The results of Granger causality tests on the three variables are shown in Table 4.4.

---

[108] Throughout this and following sections the following notation is used to indicate the statistical significance of test statistics and estimated coefficients: * represents significance at 90% confidence level, ** represents significance at 95% confidence level, and *** represents significance at 99% confidence level.

Table 4.4: Granger Causality Tests of Price, Capacity Utilization, Shipments Rate

Null hypothesis:	Test statistic
s does not Granger cause *p*	4.19**
p does not Granger cause *s*	3.31**
u does not Granger cause *p*	0.39
p does not Granger cause *u*	4.65**
u does not Granger cause *s*	9.51***
s does not Granger cause *u*	11.40***

At the 95% confidence level, we find evidence supporting the existence of Granger-causality in each case; with the exception that capacity utilization does not appear to Granger-cause price. We therefore treat all variables as endogenous.

As with the ARMA models, we have little *a priori* indication of the number of lags to include in the model, and instead use statistical criteria to make this decision. Multivariate analogues to the statistics we examined for the ARMA models (the Akaike and Schwarz Information Criteria) can be computed for VAR models.[109] AIC and SIC statistics are shown for 1 to 5 lags in Table 4.5. A VAR model with two lags is unanimously preferred.

Table 4.5: AIC and SIC Tests of VAR Models

	1	2	3	4	5
AIC	-9.92	-10.53	-10.31	-10.37	-10.25
SIC	-9.49	-9.77	-9.23	-8.98	-8.53

Unlike the univariate case, we do not allow for any time trends in this case, since no significant trend in prices was detected earlier, and we do not expect *s* or *u* to have deterministic trends. We now have the following model:

$$\begin{bmatrix} p_t \\ s_t \\ u_t \end{bmatrix} = \begin{bmatrix} \Pi_{p0} \\ \Pi_{s0} \\ \Pi_{u0} \end{bmatrix} + \begin{bmatrix} \Pi_{pp1} & \Pi_{sp1} & \Pi_{up1} \\ \Pi_{ps1} & \Pi_{ss1} & \Pi_{us1} \\ \Pi_{pu1} & \Pi_{su1} & \Pi_{uu1} \end{bmatrix} \begin{bmatrix} p_{t-1} \\ s_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} \Pi_{pp2} & \Pi_{sp2} & \Pi_{up2} \\ \Pi_{ps2} & \Pi_{ss2} & \Pi_{us2} \\ \Pi_{pu2} & \Pi_{su2} & \Pi_{uu2} \end{bmatrix} \begin{bmatrix} p_{t-2} \\ s_{t-2} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} e_{pt} \\ e_{st} \\ e_{ut} \end{bmatrix}$$

Which is estimated using OLS on observations from 1976:3 to 1991:4, as was the case for the AR (2) model estimated above. The estimated model is

$$\begin{bmatrix} p_t \\ s_t \\ u_t \end{bmatrix} = \begin{bmatrix} 0.05 \\ 0.43** \\ 0.23 \end{bmatrix} + \begin{bmatrix} 1.33*** & 0.32* & -0.23 \\ -0.18* & 0.36*** & 0.12 \\ 0.13 & 0.25* & 0.31*** \end{bmatrix} \begin{bmatrix} p_{t-1} \\ s_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} -0.38*** & 0.49*** & -0.26 \\ 0.15 & 0.84*** & -0.53*** \\ -0.15 & 0.41*** & -0.14 \end{bmatrix} \begin{bmatrix} p_{t-2} \\ s_{t-2} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} e_{pt} \\ e_{st} \\ e_{ut} \end{bmatrix}$$

As well as minimizing information criteria, the number of lags should be chosen so as to eliminate any temporal correlation in the residuals. The potential existence of contemporaneous cross-residual correlation is not a problem, as it is consistent with the assumptions of the model. Visual examination of the residuals for each estimated equation (not shown here) does not indicate the presence of residual correlation. To assess this formally, LaGrange Multiplier tests for autocorrelation in the residuals are carried out. The results show no evidence of autocorrelation for any of the variables, as reported in the second column of Table 4.6.

---

[109] See Franses (1998, 202). Other criteria also exist, as discussed by Canova (1995, 62) but these are not used here.

Table 4.6: Residual Tests of Unrestricted VAR (2) Model

	LM Test for Residual Autocorrelation	Normality Test	ARCH Test	White's Hetero-Scedasticity Test
p	0.68	1.41	2.09	1.39
s	0.37	0.29	0.66	1.61
u	1.49	1.87	0.86	1.59
Model	0.84	3.98	n/a	1.21

None of these statistics are significant at the 90% confidence level.

Tests for other potential statistical problems, namely heteroscedasticity, autoregressive conditional heteroscedasticity, and non-normality are also carried out, and reported in Table 4.5.[110] These consistently show that the VAR (2) model does not appear to violate any of the required assumptions. We conclude that the VAR (2) model is statistically acceptable.

As noted above, there is reason to suspect that seasonal patterns are important. We include seasonal dummy variables in the system, and re-estimate the model. None of the estimated coefficients on the dummy variables are significant at the 90% confidence level in either the equation for the shipment rate or the equation for capacity utilization. As a result we do not retain these dummy variables in the model.

The next stage of model construction is to assess the significance of the estimated coefficients and eliminate any that appear unimportant. Of the 18 coefficients in Π_1 and Π_2, 9 are statistically insignificant at the 90% confidence level. We assess piecewise whether these can be restricted to be equal to zero, using a Likelihood Ratio test. The ratio is calculated by comparing the log likelihood functions of the unrestricted model already estimated and the model estimated with these coefficients restricted to zero. Where the loss in degree of fit is small relative to the degrees of freedom gained, we accept the restriction.

Where both first and second order coefficients from a certain variable to another are insignificant, we attempt to restrict both simultaneously. Where a second order coefficient is insignificant, we attempt to restrict it, but where only a first order coefficient is insignificant and the corresponding second order coefficient is not, we do not attempt to restrict the former. The results of these tests are shown in Table 4.7.

[110] See Doornik and Hendry (1994, 213) for a description of these tests.

Table 4.7: Parameter Restriction Tests of VAR (2) Model

Restricted parameters	Log-Likelihood	Number of π Parameters to be Estimated	Test Statistic
None	594.65	18	n/a
Π_{pu1}, Π_{pu2}	593.46	16	2.37
Π_{ps1}, Π_{ps2}	592.03	16	5.23*
Π_{up1}, Π_{up2}	590.96	16	7.36**
Π_{uu2}	594.04	17	1.21
Π_{pu1},Π_{pu2}, Π_{ps1}, Π_{ps2}, Π_{uu2}	586.79	13	15.71***
Π_{pu1}, Π_{pu2}, Π_{ps1}, Π_{ps2}	588.00	14	13.3***
Π_{pu1}, Π_{pu2}, Π_{uu2}	592.25	15	4.79
Π_{ps1}, Π_{ps2}, Π_{uu2}	591.46	15	6.37*

We see that restrictions on π_{pu1} and π_{pu2}, π_{ps1} and π_{ps2}, and π_{uu2} are all acceptable at the 95% confidence level, while restrictions on π_{up1} and π_{up2} are not. We next evaluate the effect of simultaneously imposing all of the individually acceptable restrictions, but find that this is not acceptable. From an examination of the unrestricted model's estimated coefficients, we note that the lagged effects of price on both shipments and utilization 'switch' in the sense that the first order lag is of the opposite sign to the second order lag. It is hypothesized that restricting both of these effects significantly worsens the model's performance, while restricting only one set of them does not. We next test the effect of combining alternative pairs of the three individually acceptable restrictions, and find that simultaneous restriction of π_{pu1}, π_{pu2}, and π_{uu2} is the most attractive of these. We re-evaluate the statistical properties of the restricted model, and this time find evidence of residual autocorrelation in the residuals of the equation for capacity utilization. To account for this, it becomes necessary to again allow the second order autocorrelation term to enter the equation in an unrestricted form. Thus, the only restrictions that appear desirable are on π_{pu1} and $\pi_{\pi u2}$. Again we evaluate the statistical properties of this model, as shown in Table 4.8.

Table 4.8: Residual Tests of Restricted VAR (2) Model

	LM Test for Residual Autocorrelation	Normality Test	ARCH Test	White's Hetero-Scedasticity Test
p	0.67	1.47	2.02	1.36
s	0.54	0.22	0.70	1.67
u	1.78	0.85	0.82	1.60
VAR model	0.90	3.71	n/a	1.27

None of these statistics are significant at the 90% confidence level.

These are acceptable, and so we accept this restricted model. The reduced form of the restricted model is as follows:

$$\begin{bmatrix} p_t \\ s_t \\ u_t \end{bmatrix} = \begin{bmatrix} 0.03 \\ 0.35 \\ 0.09 \end{bmatrix} + \begin{bmatrix} 1.31 & 0.32 & -0.22 \\ -0.26 & 0.38 & 0.14 \\ 0 & 0.29 & 0.35 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ s_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} -0.36 & 0.50 & -0.27 \\ 0.22 & 0.87 & -0.56 \\ 0 & 0.47 & -0.20 \end{bmatrix} \begin{bmatrix} p_{t-2} \\ s_{t-2} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} e_{pt} \\ e_{st} \\ e_{ut} \end{bmatrix}$$

The dynamics of this model are more difficult to understand than those of the AR (2) model. For example, at first glance the autoregressive coefficients for shipments (0.38 and 0.87) suggest instability, as their sum is greater than unity. This will not necessarily occur in the multivariate context, however. IRFs (described in section 2.2.2) can be used to gain some understanding of the dynamics of the model. These are shown in Figure 4.7.

Since all of the IRFs converge to zero, this indicates that the model is in fact stable. Because the coefficients above (and thus the IRFs also) are reduced-form rather than structural, we do not attempt to place an economic interpretation upon them. This is done in the following section.

5.2.4 Structural Interpretation

In a structural model, variables may be dependent on the contemporaneous values of other variables. This presents a problem for estimation, as these contemporaneous variables are also endogenous and therefore will be correlated with the error terms in the equations. The solution to this problem is to estimate the VAR in reduced form, with each variable being a function of lagged variables only. An implication of this approach is that the estimated coefficients are not structural coefficients upon which we can place an economic interpretation.

Structural coefficients can be derived from the estimated model if it is possible to place certain restrictions on the structural form. This process is not necessary for forecasting purposes, but it is useful in terms of evaluating the plausibility of the estimated coefficients, and for understanding the mechanisms underlying any forecasts generated.

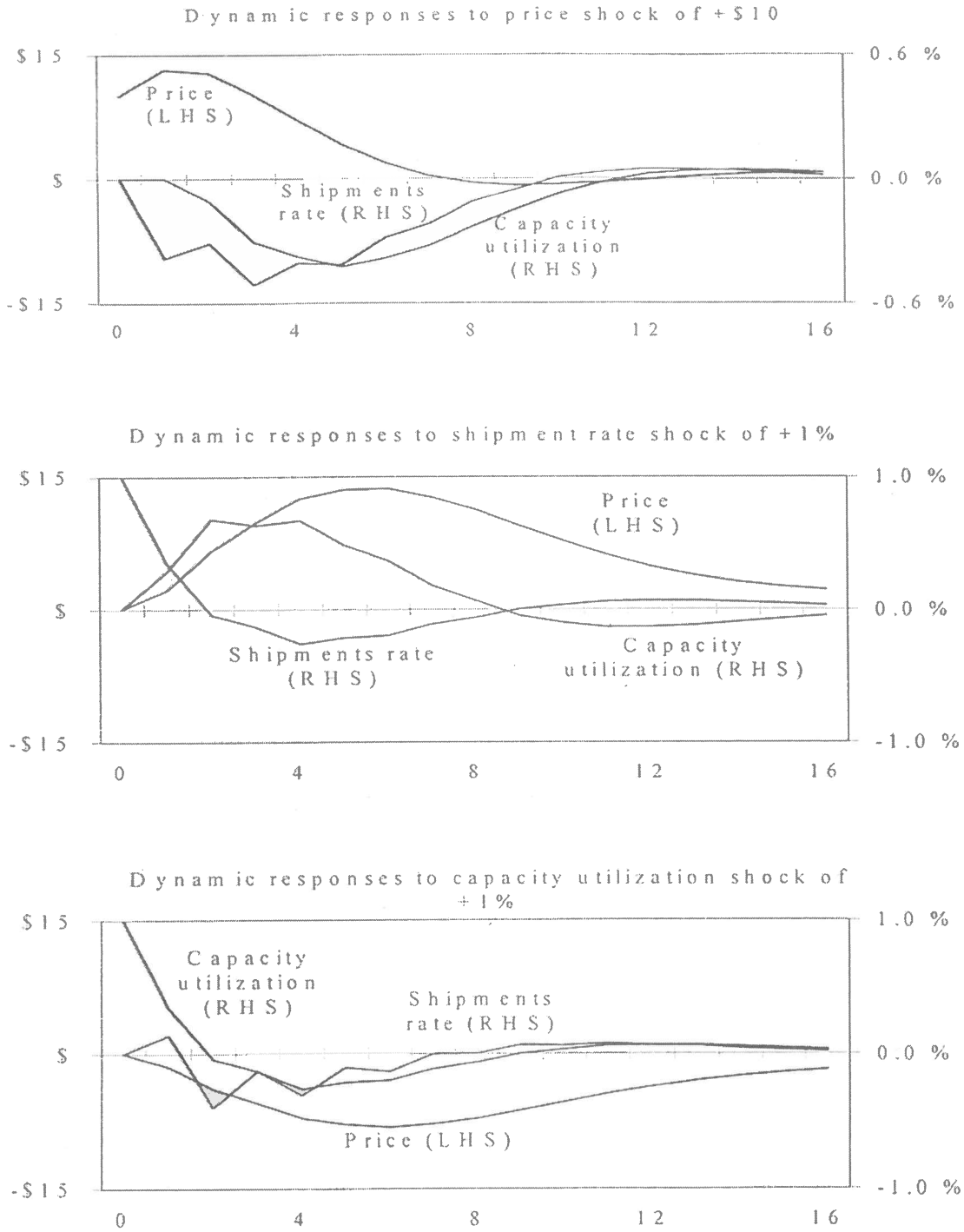

Figure 4.7: IRFs of Restricted VAR (2) Model

We use restrictions on the β parameters to extract the structural parameters. The method for doing this is described in Appendix A. There are two types of justification for restrictions of this sort: Either they can be based on economic theory, or they can be based on knowledge about the timing of availability of information on each variable.[111] In this case economic theory and *a priori* knowledge about the market do not provide strong justification for any restrictions. Thus the latter method is used. The motivation for this approach is the idea that a variable can only influence other variables once its value is known with certainty. In order to determine which contemporaneous influences we can exclude, we consider the time at which the market knows the proxy for each variable.

Data on Norscan inventories, production and shipments are usually released simultaneously, on a monthly basis, about two weeks after the end of the month. On a quarterly basis, then, the total is not known with certainty until after the quarter is finished, but data for two months of the three is known. Capacity becomes known in a less timely fashion but usually does not change greatly from quarter to quarter.

The situation regarding prices is more complicated, and has also changed over time. For most of the period under consideration, list prices have been set on a quarterly basis.[112] Producers in advance usually announce increases in list prices, with lead times of up to several months. These advance announcements do not provide a completely accurate indication of future prices, as uncertainty usually exists as to whether announced price changes will in fact occur. There is a considerable amount of published commentary on the likelihood of announced price changes being realized, and also on the general prospects for prices in the very short term (i.e. 1-3 months ahead). Price reductions happen with less warning, for obvious reasons. Therefore announcements do not provide certainty as to future prices.

Prices applying to most sales, however, are known in advance because they are done on a contract basis, with only a small proportion of sales being made on a spot basis. This creates a natural lag between announcements of price changes and their actual effect, which is estimated to average approximately two months in length. Final agreement on transaction prices (i.e. list prices less any discounts offered) is sometimes not reached until after delivery, early in the following quarter. The proxy we use for price is the list price, so this is not relevant here.

Overall, then, it appears that prices become known in advance of each quarter, while shipments and production data are gradually revealed as the quarter progresses. The most plausible assumption regarding non-causality, then, is that prices are not contemporaneously affected by either capacity utilization or shipments, i.e. $\beta_{sp} = \beta_{up} = 0$. On a monthly basis, it also seems likely that capacity utilization and shipments will not have contemporaneous effects on one another, but this is not necessarily true on a quarterly basis. We assume that capacity utilization has no contemporaneous effect on shipments, i.e. $\beta_{us} = 0$.

Using these restrictions, we follow the steps outlined in Appendix A to derive structural coefficient estimates. This gives an estimated structural model:

$$\begin{bmatrix} 1 & 0 & 0 \\ -0.12 & 1 & 0 \\ 0.05 & -0.73 & 1 \end{bmatrix} \begin{bmatrix} p_t \\ s_t \\ u_t \end{bmatrix} = \begin{bmatrix} 0.03 \\ 0.35 \\ -0.17 \end{bmatrix} + \begin{bmatrix} 1.31 & 0.32 & -0.22 \\ -0.41 & 0.34 & 0.17 \\ 0.25 & 0.02 & 0.23 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ s_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} -0.36 & 0.50 & -0.27 \\ 0.27 & 0.81 & -0.53 \\ -0.18 & -0.14 & 0.19 \end{bmatrix} \begin{bmatrix} p_{t-2} \\ s_{t-2} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{pt} \\ \varepsilon_{st} \\ \varepsilon_{ut} \end{bmatrix}$$

The price equation is unchanged from the reduced form. The autoregressive coefficients for price (1.31 at one lag and -0.36 at two lags) are reminiscent of those estimated for the AR (2) model. Lagged values of shipments have a positive impact on prices, unsurprisingly. Lagged values of utilization have a negative effect. *Prima facie* this appears counterintuitive, since a high capacity utilization is usually taken to be an indicator of strong market

---

[111] See Canova (1995) for further discussion of these methods.

[112] This has changed in recent years. Prices are now set on a more frequent basis.

conditions. One possible interpretation for this negative coefficient is as follows: A high capacity utilization *without a correspondingly high shipments rate* results in increasing inventories, which causes prices to fall.

The effect of price on shipments is difficult to understand. The coefficients are 0.12, -0.41 and 0.27 at 0, 1 and 2 lags respectively. Buyer psychology may play a role here, in attempting to time purchases so as to obtain favorable prices. The net effect is very small, indicating a low overall responsiveness of demand to price. The effect of utilization on shipments also displays a 'switching' effect, with the 1-lag coefficient being positive and the 2-lag coefficient negative.

Capacity utilization is strongly affected by contemporaneous shipments. As expected, the effect is positive (when the equation is rearranged to place the coefficient on the RHS). Quantitatively, this is the dominant effect on capacity utilization, with lagged effects being much smaller in magnitude. The finding that shipments rather than prices primarily determine capacity utilization is an interesting one. The autoregressive coefficients (0.23 and 0.19) are quite small, indicating that adjustments to capacity utilization in response to changed market conditions are quite rapid.

It is not our central purpose to test hypotheses about the structural coefficients of the model. However, an examination of their plausibility is a useful (informal) test of the validity of the model. It is difficult to conclusively assess the acceptability of the estimated model from the structural parameters, because multiple possibilities exist for several of the coefficients (for example, either positive or negative effect of capacity utilization on price would be plausible). However, some of the effects are unambiguous (for example, the effect of shipments on price) and none of these have incorrect signs. In combination with its adequate statistical performance, this leads us to accept the VAR (2) model.

5.3 ERROR CORRECTION MODEL OF INVESTMENT

The purpose of the next two sections is to construct, estimate, and evaluate two multivariate models of the pulp market with the ability to capture aspects of the long-term market structure discussed in sections 3.3.1 and 3.3.2. Variables additional to those useful in short-term modeling need to be included in the long-term models. In particular, we are interested in the interrelationships between investment and prices (addressed in this section), and between prices and costs (addressed in section 4.4). Neither of the models used above is very suitable for these purposes. Both the ARMA and VAR approaches focus only on short-term linkages between variables. They do not provide any insight into the existence of long-term relationships, nor do their forecasts ensure that any such relationships will be maintained.

The error correction approach does allow these long-term relationships to be estimated and maintained. This type of model can be regarded as a VAR model augmented with 'error correction' terms. The 'errors' measure short-term divergences from the cointegrating relationships, and these serve the purpose of 'correcting' the variables that they affect towards values consistent with the cointegrating relationships found in the system.

5.3.1 Variable Selection

In this section, we attempt to model the dynamics of investment as well as those of price. Section 3.3.1 concluded that factors important in determining investment were price, cost of capital, capacity utilization and cashflow. We also wish to include the variables already found to be important influences on price. No data is available for cashflow. We include variables representing price (p), capacity (c), production (q), shipments (s) and the real interest rate (r).

As before, price is treated as stationary, and is transformed to logarithmic form. Capacity, production and shipments are not transformed to logarithmic form because doing so affects the hypothesized cointegrating relationships between them in a way that makes testing of these relationships more difficult. The real interest rate is treated as

stationary. It is expected to be limited to a finite range.[113] This is consistent with how others have treated real interest rates.

The cost of capital is treated as exogenous, as the pulp industry is not of a sufficient size to significantly affect the macroeconomy (at least in the US). This gives us a total of four variables to be modeled (as well as one exogenous to the system) which we order as follows:

$$y_t' = [p_t \quad c_t \quad q_t \quad s_t]$$

As before, the first step in model construction is to select a maximum allowable lag length. Prior to imposing any cointegrating relationships upon the model, a VEC model is indistinguishable from a VAR model. The process of determining an appropriate lag length is therefore identical. For most of the variables, it seems appropriate to continue with the 2-lag structure used previously. Investment is an exception. Because of the physical delays between making investment decisions and the resulting capacity actually being available, time lags of up to three or four years may be relevant in this case. For the equation for investment, therefore, it is appropriate to examine the importance of longer lags. This is done in isolation from the rest of the system. In doing this, we strike a data problem: Increasing the number of lags allowed both increases the number of coefficients to be estimated and reduces the available number of observations. The maximum number of lags that can be allowed for using the 1976:3 - 1991:4 data is 8. We therefore test all lag lengths below 8. The information criteria are reported in Table 4.9.

Table 4.9: AIC and SIC Tests of Capacity Equation

	0	1	2	3	4	5	6	7	8
AIC	10.74	10.90	10.85	10.65	10.60	10.46	10.26	9.87	9.83
SIC	10.78	11.12	11.26	11.24	11.39	11.43	11.42	11.21	11.36

Here the two criteria give conflicting recommendations. The AIC indicates that it is appropriate to allow for up to 8 lags in investment (and possibly more). The SIC has a local preference for including seven lags, but overall indicates that zero lags are preferred. The latter finding suggests that this model may not be very successful in modeling the short-term dynamics of capacity investment. As a compromise we specify a model with 7 lags in the investment equation, and 2 lags in the other 3 equations. Maintaining different lags in the different equations does not present any particular problems for estimation.

5.3.2 Cointegrating Relationships

The concept of cointegration was described in section 2.1.2. In constructing the VAR model, we were able to avoid taking explicit account of cointegration by ensuring that all of the variables included in the model were stationary. In this section, we aim to use the concept of cointegration to express long-term relationships between variables, and to utilize this for forecasting. This allows us to express the variables as differences (to ensure they are stationary) and still ensure that the *relative* levels of the variables do not diverge too far from equilibrium.

Because cointegration is a long-term phenomenon, and will not necessarily be evident over the short term, we have more of a concern for sample size here than in the previous sections. We have a sample covering 92 quarters, of which 30 have usually been reserved for purposes of evaluating out-of-sample forecasting accuracy, and 2 are needed for the calculation of lagged variables. In comparison, Toppinen (1998) had data for 137 quarters, and she observed that this was "well within the range required in order for the asymptotic properties of multivariate cointegration analysis to hold (e.g. Johansen and Juselius (1990, 1994) used samples of 57 and 63 quarterly

[113] This is done on a theoretical rather than empirical basis. Empirically, over the period studied statistical tests (ADF and PP tests) do not allow us to reject the null hypothesis of a unit root in the real interest rate at the 90% confidence level.

observations, respectively)." Because the results of the cointegration tests will not have any bearing on the model specification, we use the full sample for this testing.

Capacity, shipments and production tend to move together. This means that a cointegrating relationship will exist between any two of them, giving a total of two linearly independent cointegrating vectors. For the sake of economic meaningfulness, it is convenient to postulate one relationship between capacity and production and another between production and shipments. In both cases the latter variable is effectively constrained by the former.

Under normal circumstances production will be somewhat less than capacity (as discussed in section 1.2.2) and as a result capacity utilization will be less than 100%. Evidence from the literature suggests that 93% represents an 'equilibrium' capacity utilization at least in terms of price. Examining actual outcomes gives a historical average over the period from 1976:3 to 1999:2 of 90% for Norscan producers. As a compromise, we take 91% as being the long-term 'equilibrium level', and posit a cointegrating relationship as follows:

$$0.91c - q \sim I(0)$$

The relationship between production and shipments is simpler, given that production is a physical necessity for shipments. We can posit a 1:1 relationship between the two:

$$q - s \sim I(0)$$

Because both of these relationships are expected to hold even over the short term, we may expect to find empirical evidence in support of their existence using the current data set.

We use Johansen's method (described in section 2.1.2) to test for the presence of cointegration in the data. The critical values for the test statistics depend on the nature of the data and of the cointegrating relationships allowed for. Here, we admit the possibility that the quantity variables all have an upward deterministic trend.[114] Here we test for the presence of cointegration in the system of variables $y_t' = [c_t \quad q_t \quad s_t]$. The results are shown in Table 4.10. We do find tentative evidence of two cointegrating relationships.

Table 4.10: Johansen Tests of Capacity, Production, Shipments

Rank	Eigenvalue	λ –max statistic	95% critical value	Trace Statistic	95% Critical Value
<1	0.34	37.80**	20.78	55.14**	29.51
<2	0.18	17.33*	14.04	17.34*	15.20
<3	0.00005	0.005	3.96	0.005	3.96

Having found evidence that there are two cointegrating vectors, we now look for evidence that the two vectors are those hypothesized above. A likelihood ratio test can be used to evaluate restrictions on the cointegrating vectors. We express the two hypothesized cointegrating vectors as follows:

$$A = \begin{bmatrix} k & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

[114] This implicitly has the effect also of allowing for non-zero constants in the cointegrating vectors, even though in the present case we do not expect constant terms to enter the vectors. The relevant critical values in Hamilton (1994, 767-7) are those of Case 3.

Where k is the 'equilibrium' level of capacity utilization, which is chosen to be 0.91. If we impose this on the model described in section 2.2.3, the log likelihood function and parameters for this system can be estimated similarly to the unrestricted system.[115] The likelihood ratio test statistic is

$$2\left(L_A^* - L_0^*\right) = -T\sum_{i=1}^{h}\ln\left(1-\hat{\lambda}_i\right) + T\sum_{i=1}^{h}\ln\left(1-\tilde{\lambda}_i\right)$$

Where $\tilde{\lambda}_i$ is the restricted estimate. This statistic is distributed as χ^2 with $h\ (n\text{-}q)$ degrees of freedom, where q is the number of variables involved in the cointegrating relationships. The estimated test statistic is 3.56. This is not significant at the 95% confidence level, and we therefore accept the hypothesis that the 2 cointegrating vectors in the system are those suggested by theory.

5.3.3 Model Estimation and Statistical Evaluation

Having found evidence to support the existence of these cointegrating vectors, we are now in a position to specify the following rather complicated model:[116]

$$\begin{bmatrix} p_t \\ \Delta c_t \\ \Delta q_t \\ \Delta s_t \end{bmatrix} = \begin{bmatrix} \alpha^p \\ \alpha^c \\ \alpha^q \\ \alpha^s \end{bmatrix} + \begin{bmatrix} \zeta^1_{pp} & \zeta^1_{cp} & \zeta^1_{qp} & \zeta^1_{sp} & 0 \\ \zeta^1_{pc} & \zeta^1_{cc} & \zeta^1_{qc} & \zeta^1_{sc} & \zeta^1_{rc} \\ \zeta^1_{pq} & \zeta^1_{cq} & \zeta^1_{qq} & \zeta^1_{sq} & 0 \\ \zeta^1_{ps} & \zeta^1_{cs} & \zeta^1_{qs} & \zeta^1_{ss} & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ \Delta c_{t-1} \\ \Delta q_{t-1} \\ \Delta s_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta^2_{pp} & 0 & 0 & 0 & 0 \\ \zeta^2_{pc} & \zeta^2_{cc} & \zeta^2_{qc} & \zeta^2_{sc} & \zeta^2_{rc} \\ \zeta^2_{pq} & 0 & 0 & 0 & 0 \\ \zeta^2_{ps} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{t-2} \\ \Delta c_{t-2} \\ \Delta q_{t-2} \\ \Delta s_{t-2} \\ r_{t-2} \end{bmatrix} +$$

$$\ldots + \begin{bmatrix} \zeta^3_{pc} \\ \zeta^3_{cc} \\ \zeta^3_{qc} \\ \zeta^3_{sc} \\ \zeta^3_{rc} \end{bmatrix}' \begin{bmatrix} p_{t-3} \\ \Delta c_{t-3} \\ \Delta q_{t-3} \\ \Delta s_{t-3} \\ r_{t-3} \end{bmatrix} + \ldots + \begin{bmatrix} \zeta^6_{pc} \\ \zeta^6_{cc} \\ \zeta^6_{qc} \\ \zeta^6_{sc} \\ \zeta^6_{rc} \end{bmatrix}' \begin{bmatrix} p_{t-6} \\ \Delta c_{t-6} \\ \Delta q_{t-6} \\ \Delta s_{t-6} \\ r_{t-6} \end{bmatrix} + \begin{bmatrix} \zeta^7_{pc} & \zeta^7_{rc} \end{bmatrix} \begin{bmatrix} p_{t-7} \\ r_{t-7} \end{bmatrix} + \begin{bmatrix} B^1_p & B^2_p \\ B^1_c & B^2_c \\ B^1_q & B^2_q \\ B^1_s & B^2_s \end{bmatrix} z_t + \begin{bmatrix} \varepsilon^p_t \\ \varepsilon^c_t \\ \varepsilon^q_t \\ \varepsilon^s_t \end{bmatrix}$$

$$\text{Where } z_t = \begin{bmatrix} 0.91 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} c_{t-1} \\ q_{t-1} \\ s_{t-1} \end{bmatrix}.$$

This expression is greatly complicated by the need to include different numbers of lags for the different variables, in conjunction with the fact that some variables are differenced, and some are not.[117] The presence of two lagged values for price and the real interest rate, against only one for the remaining variables, is due to the fact that the former variables have not been differenced. The coefficients on the real interest rate, except for those affecting capacities, are all restricted to zero because it is assumed that the real interest rate has no influence on pulp market conditions except through its effect on investment.

---

[115] See Hamilton (1994, 649) for details.

[116] See Hamilton (1994, 652) for a similar example (although his example mistakenly omits the term for the last lag on the undifferenced variable).

[117] A much cleaner example of an error correction formulation is given in section 5.4.

This model is estimated using full information maximum likelihood.[118] The estimated model is as follows:

$$\begin{bmatrix} p_t \\ \Delta c_t \\ \Delta q_t \\ \Delta s_t \end{bmatrix} = \begin{bmatrix} 0.2* \\ -80 \\ -236 \\ 1361 \end{bmatrix} + \begin{bmatrix} 1.22*** & 0.0001 & 0.0001** & -0.0001*** & 0 \\ 77 & -0.5*** & -0.3** & 0.4** & 26*** \\ 661 & -0.5 & 0.1 & -0.4** & 0 \\ -709 & 0.4 & 0.4** & -0.7*** & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ \Delta c_{t-1} \\ \Delta q_{t-1} \\ \Delta s_{t-1} \\ r_{t-1} \end{bmatrix}$$

$$+ \begin{bmatrix} -0.26* & 0 & 0 & 0 & 0 \\ 357** & -0.2 & -0.3** & 0.4** & -35*** \\ -619 & 0 & 0 & 0 & 0 \\ 501 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{t-2} \\ \Delta c_{t-2} \\ \Delta q_{t-2} \\ \Delta s_{t-2} \\ r_{t-2} \end{bmatrix} + \begin{bmatrix} -318* \\ -0.2* \\ -0.1 \\ 0.4*** \\ 5.2 \end{bmatrix}' \begin{bmatrix} p_{t-3} \\ \Delta c_{t-3} \\ \Delta q_{t-3} \\ \Delta s_{t-3} \\ r_{t-3} \end{bmatrix} + \begin{bmatrix} 222 \\ -0.1 \\ -0.04 \\ 0.3*** \\ 4.9 \end{bmatrix}' \begin{bmatrix} p_{t-4} \\ \Delta c_{t-4} \\ \Delta q_{t-4} \\ \Delta s_{t-4} \\ r_{t-4} \end{bmatrix}$$

$$+ \begin{bmatrix} -188 \\ -0.3*** \\ 0.02 \\ -0.03 \\ -2.8 \end{bmatrix}' \begin{bmatrix} p_{t-5} \\ \Delta c_{t-5} \\ \Delta q_{t-5} \\ \Delta s_{t-5} \\ r_{t-5} \end{bmatrix} + \begin{bmatrix} -302* \\ -0.2*** \\ 0.1*** \\ -0.1** \\ -0.1 \end{bmatrix}' \begin{bmatrix} p_{t-6} \\ \Delta c_{t-6} \\ \Delta q_{t-6} \\ \Delta s_{t-6} \\ r_{t-6} \end{bmatrix} + \begin{bmatrix} 175 & -0.5 \end{bmatrix} \begin{bmatrix} p_{t-7} \\ r_{t-7} \end{bmatrix} + \begin{bmatrix} -0.0001** & -0.0003*** \\ 0.16*** & 0.46** \\ 0.24* & -0.71*** \\ 0.14 & -0.16 \end{bmatrix} z_t + \begin{bmatrix} \varepsilon_t^p \\ \varepsilon_t^c \\ \varepsilon_t^q \\ \varepsilon_t^s \end{bmatrix}$$

We can observe that many of the estimated ζ coefficients in the investment equation are insignificantly different from zero. This may because of the variability of lag lengths from one investment project to the next (due both to project idiosyncrasies and to the fact that different forms of capacity change will usually involve different time lags). In addition, the overall explanatory power of the estimated equation is low. This indicates that a quarterly model may not be a very effective way of forecasting investment.

The statistical adequacy of the model can be evaluated using the same statistics as were calculated for the VAR (2) of section 4.2. Statistics for each equation are shown in Table 4.11.[119] No statistical problems appear.

---

[118] The models in this and the next section are implemented using the PC-FIML software package. See Doornik and Hendry (1994).

[119] Some care had to be taken in calculating these because of the different sets of RHS variables in each of the equations. In particular, this affected the results for the test for residual autocorrelation. Each equation was estimated separately, and the statistics calculated from those.

Table 4.11: Residual Tests of Investment VEC Model

	LM test for Residual Autocorrelation	Normality Test	ARCH Test (1 to 4 lags)
p	0.64	1.98	1.63
Δc	0.29	0.11	0.12
Δq	0.77	2.14	0.88
Δs	0.20	0.59	0.86
Model	n/a	5.20	n/a

None of these statistics are significant at the 90% confidence level.

It is more complicated algebraically to derive the structural model in this case, and this is not attempted.[120] However, we assess the coefficient estimates of the investment equation. If a variable is not affected by the contemporaneous values of any of the other endogenous variables then its structural form is identical to its reduced form. Because of the physical lags involved in installing capacity, it is plausible that capacity will not be so affected, and thus we can place a structural interpretation on the reduced form coefficients. Only about half of the coefficients on lagged variables in this equation are significant. All of the autoregressive coefficients are negative. Of the significant coefficients on lagged values of each of the other variables (price, change in production, change in shipments, and the real interest rate), no set of coefficients are of consistent signs at different lags. *A priori,* we would expect coefficients to be of consistent signs, with the different lags being the result simply of the different timing of different investment projects. This suggests that the estimated coefficients may be an artifact of the sample used, rather than being of true economic significance.[121] As a result, our confidence in the out-of-sample forecasting ability of this model (at least as far as forecasting investment goes) is weakened.

We conclude that using a quarterly model to forecast capacity change has a number of drawbacks. Firstly, investment is intrinsically a long-term decision, and so it may be less affected by quarter-to-quarter fluctuations in market conditions than the other variables considered here.[122] Secondly, quarterly capacity data is less reliable than annual data (in fact, the former is derived from the latter using various methods). Thirdly, the need to allow for a large number of lags weakens the statistical power of the quarterly model, and results in a large number of insignificant or difficult-to-interpret estimated coefficients. An annual model may be better suited to modeling capacity change. In terms of the most appropriate means of incorporating investment into a quarterly price model, either avoiding the need to include capacity (as was done in the VAR model above, for example) or treating it as exogenous may be more effective than the treatment used in this model.[123] As for forecasting prices, this model is not expected to be greatly different from the VAR model, given that price is still treated as a stationary variable.

---

[120] Boswijk (1992) discusses structural analysis in error correction models.

[121] One possible means of addressing this is to restrict the number of RHS variables in the investment equation. We tested the exclusion of all lagged variables whose coefficients are not found to be significant at the 90% confidence level. However, a likelihood ratio test on the re-estimated model indicated that doing this led to significant deterioration in the model's performance, so we did not persist in this approach.

[122] Jorgenson *et al* (1970) make the important point that, because of considerable continuity in investment programs, a large portion of investment spending in any given quarter is likely to be the result of conditions which occurred earlier.

[123] One alternative short-term forecasting method for capacity is to calculate the aggregate effects of announced new capacity changes. Announcements of new mills or pulp lines are often made up to three years in advance, and expansions of existing lines are also publicly announced. Jaakko Poyry Consulting Oy maintains a database of this

We next examine the estimated B matrix; i.e. the 'correction' matrix that indicates how each variable is affected by divergences from equilibrium in the cointegrating relationships. *A priori*, we may expect such divergences to affect all of the variables in the system. Price is expected to respond negatively to both excess capacity and excess production. Capacity should also respond negatively, at least to excess capacity, although it is unclear whether this effect will be detected at a one-quarter lag. Production should respond positively to excess capacity (since this is equivalent to 'inadequate production') and negatively to excess production. Shipments may be less responsive to either cointegrating relationship than production, but are expected to be affected positively (if at all) by excess production. From the estimated model above, the empirical results are:

$$B = \begin{bmatrix} -0.0001 & -0.0003 \\ 0.16 & 0.46 \\ 0.24 & -0.71 \\ 0.14 & -0.16 \end{bmatrix}$$

The first column of B indicates the responses of the variables to excess capacity, and the second column indicates the responses to excess production. The coefficients on price are both as expected, and both are statistically significant. The smaller size of the coefficients on price is to be expected given that the absolute size of this variable (in logs) is about three orders of magnitude smaller than the others. The coefficients on capacity (the second row) are the opposite of what is expected, and the reasons why capacity is positively affected by both cointegrating relationships are not clear. This reinforces the conclusion above that this model is inadequate for modeling capacity. The coefficients on production are as expected, and are statistically significant. The coefficients on shipments are not significantly different from zero.

We can evaluate the convergence properties of the model by examining how the variables that enter into the cointegrating relationships are affected by any divergence from 'equilibrium':

$$\frac{\partial CR1_t}{\partial CR1_{t-1}} = \frac{\partial(0.91c_t - q_t)}{\partial CR1_{t-1}} = 0.91\frac{\partial c_t}{\partial CR1_{t-1}} - \frac{\partial q_t}{\partial CR1_{t-1}} = 0.91(0.16) - 0.24 = -0.094$$

$$\frac{\partial CR2_t}{\partial CR2_{t-1}} = \frac{\partial(q_t - s_t)}{\partial CR2_{t-1}} = \frac{\partial q_t}{\partial CR2_{t-1}} - \frac{\partial s_t}{\partial CR2_{t-1}} = -0.71 - (-0.16) = -0.55$$

The fact that both of these are negative indicates that convergence to equilibrium will occur in both cases. For example, if excess capacity exists in a given time period, then production and capacity will change in such a way that the net effect on excess capacity in the following period is negative. Because any divergence from 'equilibrium' will also affect other variables in the system, which in turn will affect the variables entering into the cointegrating relationships, the ultimate path of reconvergence following any disequilibrium is not obvious. However, it does appear that such convergence will occur. The rate of convergence is much greater for the second cointegrating relationship than for the first. This makes intuitive sense.

In summary, this model is not fully successful. As regards the error correction part of the model, we have detected empirical evidence in support of the existence of the hypothesized cointegrating relationships, we have been able to include them in the model, and doing this has resulted in stable long-term properties and (mostly) plausible correction coefficients. However the short-term dynamic component of the model, especially as regards the capacity investment equation, is less successful. Estimated coefficients are not consistently significant. This indicates that this type of model is not useful for short-term forecasts of new capacity. Because of this, it is unlikely that the price forecasts provided by this model will be significantly better than those of the VAR model previously estimated.4.4 ERROR Correction Model of Costs

---

information.

In this section, we attempt to use the concepts applied in the previous section for another purpose: The testing and application of a different cointegrating relationship, namely between prices and costs. Up until now we have been obliged to treat the price variable as stationary, despite a conceptual understanding that it would be non-stationary. This was necessitated by our inability to transform price in any way other than by differencing, and the fact that forecasting price differences results in unattractive implied forecasts for price levels (i.e. forecasts that are highly dependent on the initial price level). Modeling costs as being cointegrated with prices allows us to model price differences that will (potentially at least) have more attractive forecast properties. In the long term, forecasts for prices should converge towards costs, and not be dependent on the initial price level. Conceptually, such an approach is superior to modeling prices in levels, and empirically it is superior to modeling price differences in an uncointegrated way.

5.3.4 Variable Selection

In this model, we continue to model the quantity variables of importance, but we revert to the treatment that was used in the VAR (2) model of section 4.2, namely modeling the stationary variables capacity utilization and the shipments rate. This allows us to focus attention on the price/cost linkage.

Should costs be treated as endogenous to this system or exogenous? As discussed in section 1.3.2, market conditions do have some effect on costs. However, this is likely to be a relatively minor effect, and may perhaps best be excluded in the interests of parsimony. We can use Granger causality tests to evaluate the extent to which costs are dependent on prices and the other variables. We test for causality at a maximum of 2 lags, and at a maximum of 8 lags (allowing for possible longer-term effects than have been detected previously). These are shown in Table 4.12, where c is total costs (transformed to logarithmic form) and the other variables are as previously defined. We also assess whether c Granger causes the other variables within 2 quarters, in order to consider whether to include it as an explanatory variable in the model (outside of the cointegrating vector).

Table 4.12: Granger Causality Tests of Total Costs and Other Variables

x	H_0: x does not Granger cause c		H_0: c does not Granger cause x
	Maximum 2 lags	Maximum 8 lags	Maximum 2 lags
p	0.971	0.787	0.914
s	0.456	0.211	1.944
u	0.010	0.694	2.000

None of these statistics are significant at the 90% confidence level.

These results show no evidence that costs are affected by any of the other variables.[124] We therefore treat costs as exogenous. The set of endogenous variables included is the same as for the VAR (2) model:

$$y_t' = [p_t \quad s_t \quad u_t]$$

Because these are the same, we do not need to repeat the process of selecting an appropriate lag structure, but rather can continue to use two lags.

[124] As shown by Engle and Yoo (1987), Granger causality between cointegrated variables will exist in at least one direction. The lack of evidence of Granger causality between price and costs here is presumably related to the fact that there is not strong empirical evidence for a cointegrating relationship between the two.

5.3.5 Cointegrating Relationship

The theory defining the relationship between price and total cost was described in section 1.3.2. Essentially, price is expected to converge to the level of total cost, although this process may take quite some considerable time to occur. The two series are shown in Figure 4.8.

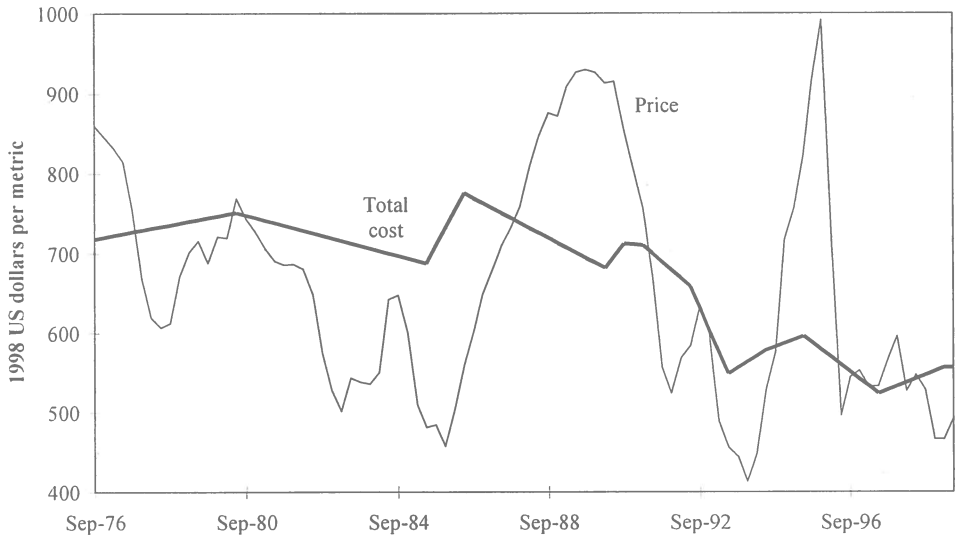

Figure 4.8: Price and Total Cost of NBSK

As can be seen, the two series do not move together nearly as closely as was the case for production and shipments, for example. Nevertheless, it is at least plausible that in the long term they do move together. Thus we posit an equality between the two variables entering the cointegrating relationship, which can be represented as:

$$p - c \sim I(0)$$

Because this relationship is expected to hold only over the long term, we may not find strong empirical evidence to support it. This will not cause us to conclude that the relationship does not exist, however.

For such a cointegrating relationship to be valid, both variables must be non-stationary. The stationarity properties of price were discussed at length in section 4.1.5. Here, we treat price as non-stationary. Here we choose to treat it as such. We test for a unit root in costs, using the entire sample. The ADF test statistic (with 8 lags, from a maximum of 12) is -1.45, and the PP test statistic is -2.07. Critical values at the 90% confidence level are -3.16 and -3.15 respectively, so we can comfortably accept the hypothesis that costs are non-stationary also.

As was the case for the previous model, we first use Johansen's test to determine how many cointegrating relationships are evident in the data. Since no evidence of a unit root was found in either capacity utilization or the shipment rate (in section 4.2) we can expect this system to contain 3 cointegrating vectors.[125] However, we are not sure that empirical support will be found for all of these. To sharpen the analysis, we consider just the two variables *p* and *c*. Carrying out Johansen's tests for these variable gives a weak indication that one cointegrating vector exists in the data (Table 4.13).

---

[125] This is because each of the two stationary variables, if included the set of variables to which Johansen's test is applied, will in isolation be found to form an $I(0)$ cointegrating vector.

Table 4.13: Johansen Tests of Price, Total Costs

Rank	Eigenvalue	λ –max statistic	95% Critical Value	Trace Statistic	95% Critical Value
<1	0.138	13.39*	14.04	14.46*	15.20
<2	0.012	1.065	3.96	1.065	3.96

As before, we use a likelihood ratio test to assess whether this single cointegrating vector is the hypothesized one, i.e. whether it is [1, -1]. The test statistic is 1.12, against a $\chi^2(1)$ 95% critical value of 3.84. Therefore we accept the hypothesis that the cointegrating vector between these two variables is [1, -1], consistent with theory.

5.3.6 Model Estimation and Statistical Evaluation

Given that price is regarded here as non-stationary, it is transformed into differenced form. We assume on the basis of the evidence from the Granger causality tests reported in Table 4.12 that costs do not affect any of the endogenous variables. Costs enter the model only through the cointegrating relationship with price. The model to be estimated is specified as follows:

$$\begin{bmatrix} \Delta p_t \\ s_t \\ u_t \end{bmatrix} = \begin{bmatrix} \alpha^p \\ \alpha^s \\ \alpha^u \end{bmatrix} + \begin{bmatrix} \zeta^1_{pp} & \zeta^1_{sp} & \zeta^1_{up} \\ \zeta^1_{ps} & \zeta^1_{ss} & \zeta^1_{us} \\ \zeta^1_{pu} & \zeta^1_{su} & \zeta^1_{uu} \end{bmatrix} \begin{bmatrix} \Delta p_{t-1} \\ s_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta^2_{sp} & \zeta^2_{up} \\ \zeta^2_{ss} & \zeta^2_{us} \\ \zeta^2_{su} & \zeta^2_{uu} \end{bmatrix} \begin{bmatrix} s_{t-2} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} B_p \\ B_s \\ B_u \end{bmatrix} z_t + \begin{bmatrix} \varepsilon^p_t \\ \varepsilon^s_t \\ \varepsilon^u_t \end{bmatrix}$$

$$\text{Where } z_t = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ c_{t-1} \end{bmatrix}.$$

At this stage, we allow for disequilibrium in the cointegrating relationship to affect all of the endogenous variables, i.e. none of the B coefficients are restricted to be zero. The model is estimated using FIML, and the results are as follows: text box

$$\begin{bmatrix} \Delta p_t \\ s_t \\ u_t \end{bmatrix} = \begin{bmatrix} -0.32*** \\ 0.19** \\ 0.15 \end{bmatrix} + \begin{bmatrix} 0.36*** & 0.33* & -0.21 \\ -0.15 & 0.35** & 0.13 \\ 0.15 & 0.25* & 0.31* \end{bmatrix} \begin{bmatrix} \Delta p_{t-1} \\ s_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} 0.50*** & -0.28 \\ 0.83*** & -0.52*** \\ 0.40** & -0.15 \end{bmatrix} \begin{bmatrix} s_{t-2} \\ u_{t-2} \end{bmatrix} + \begin{bmatrix} -0.06 \\ -0.04 \\ -0.01 \end{bmatrix} z_t + \begin{bmatrix} \varepsilon^p_t \\ \varepsilon^s_t \\ \varepsilon^u_t \end{bmatrix}$$

None of the estimated pairs of ζ coefficients are clearly insignificant, and as a result none are considered for restriction. The coefficients on the lagged values of s and u are similar to the unrestricted equivalents in the VAR model of section 4.2. The levels form autocorrelation coefficients of p can also be derived from this equation, and are identical to the estimated coefficients of the VAR model.[126] This is not surprising given the similarity of the two models.

The t-values for the three B coefficients are -1.59, -1.46 and -0.37 respectively. None are significantly different from zero at the 90% confidence level. This is perhaps not surprising given the slow adjustment times expected, and the poor quality of the cost data. Therefore, their lack of statistical significance notwithstanding, we do not attempt to place restrictions on these coefficients. We regard their statistical insignificance as an indication that adjustment is slow, rather than indicating that no adjustment occurs. What do these B parameters mean? The negative value for B_p is expected, as is its numerically small value. This indicates that any 'too-high' price will cause price to fall (slightly). The negative value for B_s suggests either that shipments fall in a situation of 'excess profits', or when

---

[126] $\gamma_{pp1} = (1 + B_p) + \zeta^1_{pp} = (1 - 0.06) + 0.36 = 1.31$, $\gamma_{pp2} = -\zeta^1_{pp} = -0.36$.

capacity rises. The latter is more plausible. If this is the case, the smaller value for B_u suggests that both capacity and production rise in such a situation.[127] This is also plausible.

We next consider the statistical performance of the model. Results of the various tests are shown in Table 4.14. Given the similarity of this model to the earlier VAR (2) model, its adequate statistical performance is not surprising.

Table 4.14: Residual Tests of Cost VEC model

	LM Test for Residual Autocorrelation	Normality Test	ARCH Test	White's Hetero-Scedasticity Test
Δp	1.54	1.33	1.85	0.91
s	0.38	0.19	0.91	1.43
u	1.74	1.62	0.80	1.09
Model	0.85	3.76	n/a	1.15

None of these statistics are significant at the 90% confidence level.

In summary, these results show that cointegration can be applied to a VAR framework without greatly affecting the latter in terms of the estimated short-term dynamic coefficients and statistical properties of the model. However, it is expected that fundamental differences between the long-term forecasts of the two will be evident. The VAR model is obliged to assume that price is stationary, and forecasts for price will therefore return to the historical mean (possibly trending) level. The error correction model treats prices as non-stationary, and forecasts for prices will not be mean reverting, but rather will tend to follow forecasts for costs. Conceptually, the latter is much more attractive, although of course it is only one part of a comprehensive model of the pulp market.

[127] A model that treated capacity, shipments and production as separate variables (such as that of section 4.3) would be able to shed more light on this.

6 FORECASTING COMPARISON

Economists have forecast 9 of the last 5 recessions.

Anonymous

In this chapter we conduct a comparison of the forecasting accuracy and characteristics of each model. We first present an informal discussion of the nature of the forecasts that the different models generate. A formal criterion for measuring forecast accuracy is then presented. This does not provide a comprehensive indication of forecast quality, however. The criterion does not measure the ability of models to correctly identify turning points, or their ability to be used in conjunction with judgmental forecasting methods. These are also discussed.

6.1 FORECASTING METHODS

Forecasts from each of the 4 estimated models are compared, along with 'naïve' forecasts and a set of actual published forecasts. Before this comparison is done, we briefly discuss the nature of the forecasts from each model.

To understand the nature of forecasts generated by these models, it is useful to consider some examples. Figure 5.1 shows the forecasts made by all four of the models in December 1992, and Figure 5.2 shows the forecasts made in September 1995. For all of the models, short-term forecasts for price depend on its recent history. For the multivariate models, short-term forecasts also depend on the recent values for other variables. In December 1992, all of the models forecast (correctly) a short-term drop in prices. The decline forecast by the multivariate models was greater than that forecast by the AR model, suggesting that inclusion of other variables may have provided some useful signal in this case.

None of the models forecast the dramatic rise in prices that occurred in 1995, and in fact model based forecasts made closer to the time of the spike also failed to do so. This demonstrates that the models are not able to successfully forecast all short-term developments. It also suggests that factors other than the 'fundamentals' of capacity utilization and shipments levels contributed to the price spike at that time.

In the long term, forecasts for price generated by the cost error correction model will be close to the (exogenous) forecasts that are made for costs. Long-term forecasts for prices generated by the other three models, on the other hand, will be determined by the estimated deterministic trends in those models. Thus, conceptually the forecasts from the former are quite different from those of the latter models, even though graphically the differences may not be apparent. For the forecasts shown, actual cost levels were used as a proxy for cost forecasts.[128] As these were somewhat lower than had historically been the case, the long-term forecasts for the cost error correction model are also somewhat lower than forecasts of the other models.

If prices at the time a forecast is made differ greatly from the long-term trend or 'equilibrium', then they will be forecast to return towards that level. Because the price in September 1995 was considerably above its trend level and also above costs, all the models forecast prices to rapidly drop.

---

[128] In generating forecasts for the investment EC model, actual outcomes for the real interest rate (which that model treats as exogenous) were also used. In a sense, this gives these models an advantage over those that do not make use of any exogenous variables. In practice, though, costs changed little over the period from 1991 onwards.

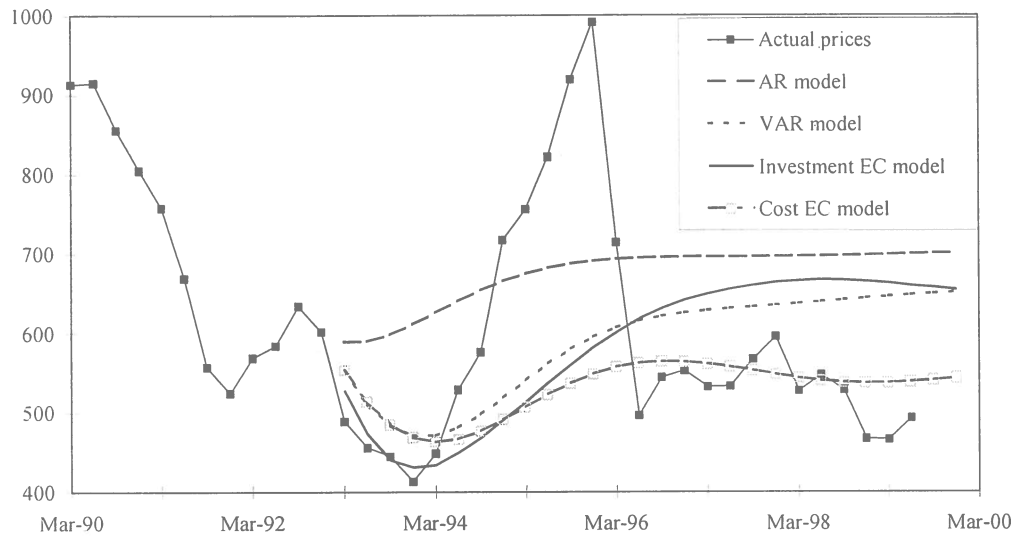

Figure 5.1: Model forecasts for real NBSK price in December 1992

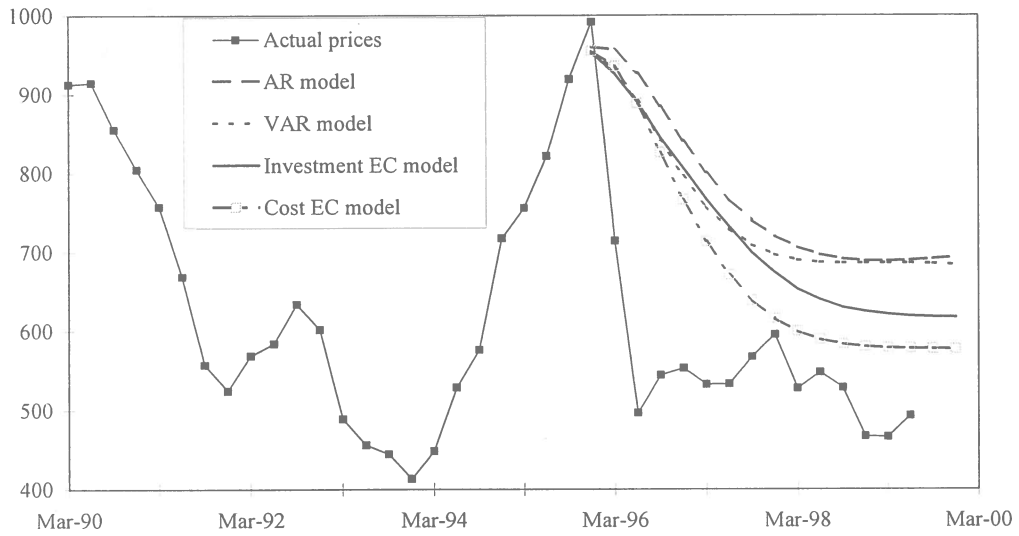

Figure 5.2: Model Forecasts for Real NBSK Price in September 1995

The forecasts for prices made by the VAR model and the investment error correction model are quite similar. This is not surprising, given that both admit a similar set of variables as influences on prices, and neither includes costs. The difference between these two models is in their forecasts for the other variables. The VAR model forecasts for capacity utilization and the shipments rate are constrained (by assumption) to have no deterministic trends, but to the extent that historical means of the two variables differ, so will long-term forecasts for them. The historical means over the period 1976:3-1991:4 were 89.9% and 90.4% respectively. Forecast differences will therefore be small, but will imply a continuous decline in inventories, which is not plausible. The investment error correction model goes some way to avoiding this problem by ensuring that the two series move together.[129]

---

[129] Even the latter model does not explicitly require any particular characteristics for inventories. In the short term,

Finally, it is apparent that the degree of price fluctuation in the forecasts is much less than that actually observed. This is because the models have no information about what shocks will occur in the future, and are thus obliged to forecast shocks of zero for every future period. This is of course unrealistic. It should be interpreted, as the result of a lack of information available to the models, rather than as a forecast that prices volatility will in fact decline. The only way to address this issue in the current context is to use expert opinion on likely future events to adjust the model-based forecasts. This topic is addressed in section 5.3 below.

6.2 FORMAL ASSESSMENT

Forecasts are generated for all endogenous variables in each model, but actual values are used for forecasts of the exogenous variables (real interest rate for the investment error correction model and total costs for the cost error correction model).

'Naïve' forecasts are also calculated. A 'naïve' economic agent will expect the future values of variables to remain the same, as they are at the time the expectation are formed. Forecasts will therefore be simply

$$E[p_{t+i} \mid I_t] = p_t \quad \text{for all } i = 1, 2, 3\ldots$$

where I_t is the information set available to the agent at time t. This expectation formation mechanism is the simplest possible. It makes no use of any information available other than the current price. It is not presented as a realistic expectations-formation mechanism, but rather as a benchmark for assessing how useful alternative methods are in improving forecast accuracy.

Occasional published forecasts are available, from a number of sources.[130] As far as can be ascertained all of these are made at least partly on a judgmental basis. Casual observation suggests that the published forecasts vary considerably in their accuracy. These forecasts do not form a consistent group that is directly comparable with the model-generated forecasts. They are not available for every quarter over the relevant time span (1991:4 to 1994:2); the forecast frequency is sometimes annual rather than quarterly; the time horizon to which they forecast is sometimes less than five years; and they generally forecast nominal rather than real prices.

Despite these problems, an attempt is made to formally calculate the accuracy of these forecasts. First the forecasts are interpolated where necessary (i.e. where forecasts are made on an annual basis). They are then deflated to real terms by use of actual price deflators (using the same US wholesale price index that was used to deflate the nominal series used in constructing the real price data series for the model estimation process).

This gives a total of six sets of forecasts to compare. Various summary statistics on forecasting accuracy can be calculated. These are based on measuring the divergence between forecasts and actual outcomes, and include the root mean squared forecast error (RMSE), the mean absolute percentage error, and the Theil inequality coefficient.[131] The most commonly used is the RMSE. This is calculated as its name suggests:

$$RMSE_{t,n} = \left[\frac{1}{n} \sum_{i=t+1}^{t+n} \left(P_i - \hat{P}_i \right)^2 \right]^{1/2}$$

---

the model has a similar problem to the VAR model in that the implied forecast for inventories may increase or decrease, with no mean-reverting tendency, but this does not continue in the long term.

[130] All of these forecasts are proprietary, and for this reason the actual forecasts are not shown here.

[131] For comparisons of these different criteria, see Armstrong and Collopy (1992), Armstrong and Fildes (1995), and Diebold and Lopez (1996).

Where t is the time at which forecasts are made, $t+n$ is the forecast horizon, and $\hat{P}_i$ is the forecast of variable P_i. The variable forecast is the real price level.[132]

Forecast accuracy at different horizons is assessed. Short term, medium term and long term are arbitrarily defined as the forecasts 1-4 quarters, 5-12 quarters and 13-20 quarters in the future respectively. Forecasts are generated from starting points of each quarter between 1991:4 and 1994:2 (a total of 11 sets of forecasts) and the short-term, medium-term, and long-term RMSE statistics for each starting point are calculated. Averages of each of these are then calculated. The models were estimated using only data up to 1991:4 so that enough observations would still be available to do this. The results are shown in Table 5.1.

Table 5.1: RMSE Statistics

	'Naïve' Forecast	AR Model	VAR Model	Investment EC Model	Cost EC Model	Published Forecasts
Short term	106	85	74	67	65	116
Medium term	224	175	189	190	208	177
Long term	123	159	120	136	120	167
Average	160	150	139	144	144	161

The absolute level of these numbers is not of interest, but in relative terms, a lower figure indicates more accurate forecasts. Firstly, we can observe that, relative to the times at which most of the forecasts were made (from December 1991 to March 1994) the price spike occurred in the 'medium-term' future, and thus led to a high forecast error for the medium term.[133] This can be seen from the results for the naïve forecasts. A period that did not contain such a dominant feature might be more useful for discriminating between the models.

All of the time series models do seem to give more accurate forecasts than the naïve model. As was exemplified in Figure 5.1, all of the models had a degree of success in forecasting the trough in 1993, and this is probably the main reason for the good short-term performances relative to the naïve forecasts. Also, the multivariate models were more accurate than the AR model, and this also seems to account for their better performance here. The published forecasts did not successfully forecast this trough.

The error correction models do not perform any better than the VAR model, and in fact slightly worse. For the investment error correction model, this is because its forecasts for prices are very similar to those of the VAR model. For the cost error correction model, it is because the actual path of the cost 'attractor' over the period of interest was not greatly different from the zero deterministic trend 'attractor' of the VAR model.

These RMSE statistics do not provide a perfect measure of forecast quality, for several reasons. They are sensitive to outlying forecast errors; they do not place any value on the identification of cyclicality and turning points; and they place no value on the consistency of forecasts for cointegrated variables, (which is achieved by error correction models).[134]

---

[132] All of the models developed forecast log prices p_t, whereas we are interested in actual prices P_t. Granger and Newbold (1976) showed that the forecast derived by taking the exponent of the forecast for the log price is inefficient in the sense that it does not minimize the expected forecast error. However, adjusting forecasts to minimize the expected forecast error is analytically complicated. Here we use this simple approach.

[133] Duy and Thoma (1998, 296) discuss this problem. They note that outliers (very poor single forecasts) can cause a model to exhibit large RMSE values even if non-outlying forecasts are good.

[134] Henriksson and Merton (1981) describe a test for how well forecasts predict turning points. See Christofferson and Diebold (1998) for a discussion of the last point.

The published forecasts have similar medium- and long-term RMSE statistics to the models. Some (but not all) of the published forecasts contain a cyclical pattern similar to that which actually occurred. However, even slight mistiming of the cycle results in a RMSE statistic that is no lower, and possibly even higher, than that of forecasts with little or no cyclical component. A forecast that captures the cyclical pattern of prices, even slightly mistimed, seems intuitively preferable to one that does not. Thus we cannot conclude directly from the RMSE statistics that the published forecasts are inferior to the model forecasts.

6.3 INFORMAL ASSESSMENT

No model can provide a perfect view of the future. In essence, any time series model characterizes the evolution of price (or any other variable) as a combination of influences from the past and of new shocks. In using time series models for forecasting, the models themselves can only provide information about the former of these two. Because the models offer no information about the latter, a purely model-based forecast is obliged to assume that no new shocks occur. As a result, such forecasts contain a low degree of volatility compared to historical experience.

The information set available to the forecaster is likely to be wider than that available to the model, since the latter consists only of historical observations on the variables in the model. If information regarding probable future shocks is available to the forecaster, this information should be provided to the model. In all likelihood, model based forecasts will be improved by the use of such information. Two examples of future shocks are provided, to demonstrate the differences between the models in terms of their abilities to be used for this purpose.

The first example is a short-term one. It is expected that two distinct effects will occur in late 1999, and these are known collectively as the 'millennium effect'. Firstly, newspaper and magazine publishers will run an atypically high number of supplements to their regular publications, which will lead to an increase in demand for paper, and in turn to an increase in demand for pulp. Secondly, because of a concern that supply logistics may be temporarily disrupted by computer failures, many papermakers will temporarily increase their inventories of pulp in order to avoid any shortages.

How can this information be imparted to the models? By default, forecasts are made assuming shocks of zero in the future, but this assumption can easily be changed if desired. For the AR model, the only variable available to be shocked is price. Thus we are obliged to assess the likely impact of these effects on prices outside of the modeling framework, which is a difficult task. In the multivariate models, we can impose shocks on shipments or the shipment rate. Concrete information about the impact of the millennium effect on shipments is much easier to obtain (by talking directly to industry participants about their intentions) than information on the price effect. This allows us much more confidence in the resulting forecast. Thus, multivariate models that contain variables such as shipments and others that are likely to be subject to shocks are superior to models that do not.

The second example is a long-term, hypothetical one. A technological advance has just occurred which will allow pulp production costs to be reduced. This technology has not yet been implemented, but is expected to be gradually adopted by the industry over the next decade or so. How can we incorporate this information into our model forecasts? In the case of the AR, VAR and investment error correction models, we are obliged to do so in an *ad hoc* fashion. It is not possible to effectively model a permanent shift in price (which is what we expect the technological advance to lead to) simply by imposing a shock on one or more of the variables in these models. In contrast, using the cost error correction model, we can do just this.[135] To allow such a shock to be effectively modeled, not only do costs have to be included in the model, but they also need to have a permanent effect on prices. Applying cointegration methods is the most convenient way to do this.

In summary, the models do appear to forecast somewhat better than other methods, although they did not successfully forecast all of the movements in prices over the testing period. Inclusion of non-price variables in the

---

[135] A further advantage of the explicit link between price and costs is that detailed cost models (such as were used to derive the cost data used herein) can be used to assess the likely impact of shocks to individual cost components on overall cost levels.

models gives an improvement in forecasting performance, and it also allows the model to incorporate judgmental information more easily. The introduction of cointegrating relationships to the models does not seem to improve price-forecasting performance. However, it does have other advantages, namely in enhancing the long-term plausibility of the forecasts, and (if the variables included would not be included in a purely short-term model) further enhancing the ability of the model to incorporate judgmental information. Therefore both of these steps are desirable.

7 CONCLUSIONS

"Everybody just be calm and cooperate with them and this will all be over soon!"

The Coffee Shop Manager, Pulp Fiction

Conclusions address three issues. Firstly, what are the important factors and interrelationships that influence pulp market conditions? Secondly, are the time series methods demonstrated herein useful for modeling these interrelationships? Thirdly, what elements should a comprehensive forecasting model for pulp prices contain?

A priori, an examination of the literature and anecdotal evidence suggests that important short-term influences on prices include capacity utilization, inventories, shipments, seasonal factors, and exchange rate movements. The empirical analysis done herein confirms the importance of capacity utilization and shipments. Because inventories are related to these two variables, we can also conclude that they are important. Seasonal factors do not seem important, and exchange rate movements are not examined.

In the long term, investment behavior is expected to play a key role in determining market conditions. A well-developed literature on the determinants of investment exists. This research addressed the issue of how to incorporate investment behavior into a model of the industry, but was not very successful in doing so. Another important long-term factor in determining prices is costs. Here, weak evidence of a cointegrating link between the two variables was found.

Four models were constructed, in increasing order of complexity. Each was designed to focus on a particular issue, and none can be regarded as being a comprehensive forecasting model. The advantage of each model over its predecessors and some of the important empirical results from estimation are summarized in Table 6.1.

Table 6.1: Summary of Model Properties

Model	Improvement Over Earlier Models	Important Insights
'Naïve' model		
Autoregressive	Allows recent price history to have an impact on price	Price has a cyclical pattern.
Vector autoregressive	Allows non-price variables to have an impact on price	Shipments positively affect price and capacity utilization; capacity utilization negatively affects price.
Error correction (investment)	Ensures long-term consistency between capacity, production, and shipments.	Production and price both respond to any 'disequilibrium' between capacity and production, and production and shipments.
Error correction (costs)	Ensures long-term consistency between prices and total costs.	Price adjusts (slowly) to the level of costs.

Are the cointegration techniques applied here useful for inclusion in forecasting models? According to the formal criterion used for evaluating forecast accuracy (the RMSE), error correction models were no better than the simpler VAR model at forecasting prices. However, other properties of the forecasts generated by the two error correction models are attractive, relative to those of the VAR model. The investment error correction model ensures long-term consistency between each of the three quantity variables in the model, which the VAR model implicitly fails to do. The cost error correction model anchors the long-term forecasts for prices to forecasts for cost levels, a much more appealing approach than having long-term price forecasts reliant on a deterministic trend with no explicit conceptual basis for its existence. It also gives us the ability to model the effects on prices of shocks to cost components. Both of these properties are highly desirable in a forecasting model.

Furthermore, these properties can be obtained at little cost (except in terms of data requirements and model complexity). The hypothesized cointegrating relationships appear to be congruent with the data available.

Therefore, imposing them does not interfere with the short-term dynamics of the system. For these reasons we conclude that the error correction models are in fact valuable.

Finally, a comprehensive forecasting model structure is proposed. None of the models developed here contains all of the elements necessary for an adequate model for forecasting pulp prices. The model development process has given a number of results that are relevant to the design of a comprehensive model. Firstly, production and shipments are useful in modeling short-term price dynamics, but are not sufficient to capture all price movements (given that models including these variables did not successfully forecast the 1995 price spike). Secondly, investment is difficult to forecast accurately in the short term (at least on a quarterly basis and using the variables included here). The most effective short-term forecasts for capacity investment can perhaps be made with reference to information based on actual announcements of investment activities by producers, rather than by econometric modeling. Thirdly, strong conceptual reasons exist (supported by empirical evidence) for ensuring that equilibrium amongst the quantity variables (capacity, production and shipments) is maintained in the long term, while weaker but still convincing conceptual and empirical reasons exist for maintaining a link between prices and costs.

The issue of how to treat 'fundamental demand' has not been addressed in detail here (although the list of factors in section 1.2.3 is a useful starting point). Further research is needed into the issue of whether the fundamentals of demand are best modeled simultaneous with the rest of the variables pertaining to the pulp market, or whether they can be treated as exogenous. In the long term, quantity variables are likely be driven by the demand side. We can envisage the following long-term chain of causality: Fundamental demand affects shipments, which in turn affects production, which in turn affects capacity. In addition, the balance between production and shipments determines inventories, and any movement away from 'equilibrium' should lead to an adjustment in production.

There are some other variables that are potentially important also. The exchange rate was not focused on here, but is expected to have some effect on prices. Implicitly, the long-term effect of exchange rate movements is captured in the cost variable, but short-term effects are also likely to be important. In addition, for forecasting nominal prices, inflation is important. In light of this discussion, the structure shown in Figure 6.1 is recommended for an effective forecasting model of pulp prices.

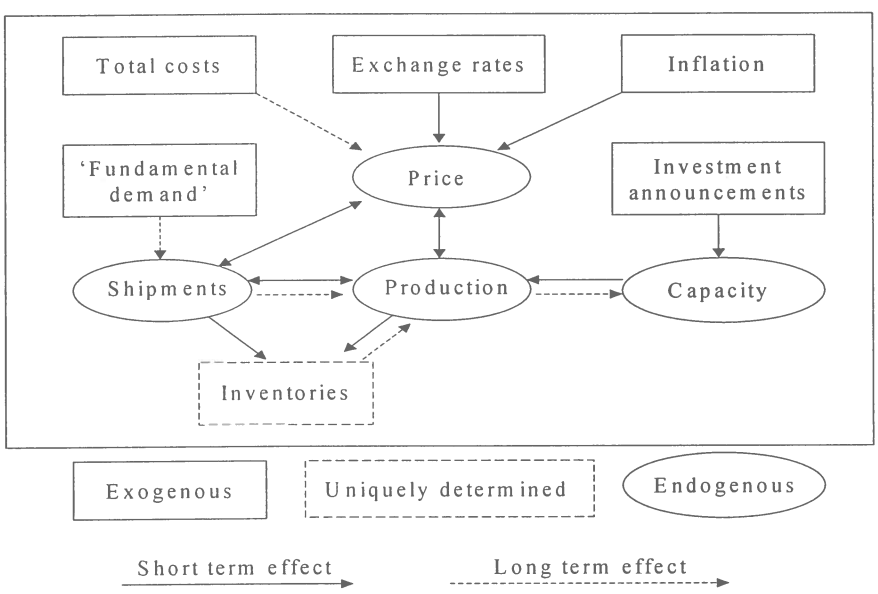

Figure 6.1: Proposed Structure for Comprehensive Forecasting Model

Not all aspects of the proposed model have been explored in depth here. However, the present research has identified the most important factors affecting market conditions, and it has also demonstrated that the methods used herein are an effective means of modeling the interactions between them.

8 BIBLIOGRAPHY

Abel, A.B. and J.C. Eberly (1997), "The Mix and Scale of Factors with Irreversibility and Fixed Costs of Investment," *Working Paper 6148*, NBER, Cambridge MA.

Anonymous (1998), "Irrational Exuberance Pulp and Paper Style," *Pulp and Paper International* 39(6), 5.

Ardeni, P.G. (1989), "Does The Law of One Price Really Hold?" *American Journal of Agricultural Economics* 71, 661-669.

Armstrong, J.S. and F. Collopy (1992), "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons," *International Journal of Forecasting* 8(1), 69-80.

Armstrong, J.S. and R. Fildes (1995), "On the Selection of Error Measures for Comparisons Among Forecasting Methods," *Journal of Forecasting* 14, 67-71.

Banerjee, A., J. Dolado, J.W. Galbraith and D.F. Hendry (1993), *Cointegration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press, Oxford.

Bernstein, J.I. (1992), "Price Margins and Capital Adjustment: Canadian Mill Products and Pulp and Paper Industries," *International Journal of Industrial Organization* 10, 491-510.

Bischoff, C.W. (1971), "Business Investment in the 1970s: A Comparison of Models," *Brookings Papers on Economic Activity* 1, 13-58.

Blanchard, O.J. and D. Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances" *American Economic Review* 79(4), 655-673.

Bose, S. and A. McIlgorm (1996), "Substitutability among Species in the Japanese Tuna Market: A Cointegration Analysis," *Marine Resource Economics* 11, 143-155.

Boswijk, H.P. (1992), "Cointegration, Identification, and Exogeneity," Ph.D. Thesis, Tinbergen Institute, Amsterdam.

Box, G.E.P. and G.M. Jenkins (1970), *Time series analysis: Forecasting and control*, Holden-Day, San Francisco.

Brannlund, R., K.G. Lofgren and S. Sjostedt (1999), "Forecasting Prices of Paper Products: Focusing on the relation between Autocorrelation Structure and Economic Theory," *Journal of Forest Economics* 5(1), 23-44.

Buongiorno, J. and J.K. Gilless (1980), "Effects of Input Costs, Economies of Scale, and Technological Change on International Pulp and Paper Prices," *Forest Science* 26(2), 261-275.

_____ And H.C. Lu (1989), "Effects of Costs, Demand, and Labor Productivity on the Prices of Forest Products in the United States," *Forest Science* 35(2), 349-363.

_____ And J. Uusivori (1992), "The Law of One Price in the Trade of Forest Products: Cointegration Tests for US Exports of Pulp and Paper," *Forest Science* 38(3), 539-553.

Butner, R.E. and C.E. Stapley (1996), "Capital Effectiveness: The Paper Industry versus Others," TAPPI Engineering Conference Proceedings, TAPPI Press, Atlanta.

Caballero, R.J. (1997), "Aggregate Investment," *MIT Dept of Economics Working Paper 97-20*, Cambridge, MA.

Caballero, R.J. and E.M.R.A. Engel (1998), "Nonlinear Aggregate Investment Dynamics: Theory and Evidence," *Nber Working Paper 6420*, Cambridge, MA.

Canova, F. (1995), "The Economics of VAR Models," in K. Hoover (ed.), *Macroeconometrics: Tensions, Developments and Prospects*, Kluwer Academic Publishers, Dordrecht.

Carter, R.L.J. (1968), *The Market for Woodpulp: A Global and Regional Study*, Praeger, New York.

Chapman, D.R., C.W. Junor and T.R. Stegman (1996), "Cash-Flow Constraints and Firms' Investment Behavior," *Applied Economics* 28, 1037-1044.

Chas-Amil, M.L. and J. Buongiorno (1999), "Determinants of Prices of Paper and Paperboard in the European Union from 1969 to 1992," *Journal of Forest Economics* 5(1), 7-21.

Chirinko, R.S. (1993), "Business Fixed Investment Spending: Modeling strategies, Empirical Results, and Policy Implications," *Journal of Economic Literature* 31, 1875-911.

Christensen, L.R. and R.E. Caves (1997), "Cheap Talk and Investment Rivalry in the Pulp and Paper Industry," *Journal of Industrial Economics* 45(1), 47-72.

Christoffersen, P.F. and F.X. Diebold (1998), "Cointegration and Long-Horizon Forecasting," *Journal of Business and Economic Statistics* 16(4), 450-58.

Clements, M.P. and D.F. Hendry (1993), "On the Limitations of Comparing Mean Square Forecast Errors," *Journal of Forecasting* 12, 617-637.

_____ And _____ (1998), *Forecasting Economic Time Series,* Cambridge University Press, Cambridge.

Constantino, L.F. and G.M. Townsend (1986), "Modeling Short-Run Producer Behavior as an Operating Rate Decision: The Canadian Sawmilling and Pulp and Paper Industries," Forest Economics and Policy Analysis Project Report 86-10, University of British Colombia, Vancouver.

Croon, I. (1995), "Pulp and Paper Industry: A Dynamic but Cyclic Affair," *European Papermaker* 3(1), 24-7.

Cuthbertson, K. and D. Gasparro (1995), "Fixed Investment Decisions in UK Manufacturing: The Importance of Tobin's Q, Output AND Debt," *European Economic Review* 39, 919-941.

Deaton, A. and G. Laroque (1992), "On the Behavior OF Commodity Prices," *Review of Economic Studies* 59(1), 1-23.

Devereux, M. and F. Schiantarelli (1990), "Investment, Financial Factors, and Cash Flow: Evidence from U.K. Panel Data," in R.G. Hubbard (ed.), *Asymmetric Information, Corporation Finance, and Investment*, University of Chicago Press, Chicago.

Diebold, F.X. and J.A Lopez (1996), "Forecast Evaluation and Combination," NBER Technical Paper 192, Cambridge, MA.

Diesen, M. (1998), *Economics of the Pulp and Paper Industry*, TAPPI Press, Atlanta.

Doornik, J.A. and D.F. Hendry (1994), *PcFiml 8.0: Interactive Econometric Modeling of Dynamic Systems*, International Thomson Publishing, London.

Duy, T.A. and M.A. Thoma (1998), "Modeling and Forecasting Cointegrated Variables: Some Practical Experience," *Journal of Economics and Business* 50, 291-307.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica* 50, 987-1008.

_____ And C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 55, 251-276.

_____ And B.S. Yoo (1987), "Forecasting and Testing in Cointegrated Systems," *Journal of Econometrics* 35(1), 143-159.

Fanchon, P. and J. Wendel (1992), "Estimating VAR Models under Non-Stationarity and Cointegration: Alternative Approaches for Forecasting Cattle Prices," *Applied Economics* 24(2), 107-217.

FAO (1998), *Pulp and Paper Capacities Survey 1997-2002*, FAO, Rome.

Farimani, M., J. Buongiorno and H.E. Thompson (1988), "A Financial Model of Investment, with an Application to the Paper Industry," *Applied Economics* 20 (6), 767-783.

Franses, P.H. (1998), *Time Series Models for Business and Economic Forecasting*, Cambridge University Press, Cambridge.

Ferguson, K.H., K.J. Finchem and K. Porter (1996), "Emerging Commodity Players Shape North America's Pulp, Paper Future," *Pulp & Paper* 70(9), 63-87.

Finchem, K. (1997), "Embattled Paper Companies Seek Wise Capital Investment Strategies," *Pulp & Paper* 71(9), 51-57.

_____ (1998), "Futures Trading Offers Potential to Reduce Pulp, Paper Volatility," *Pulp & Paper* 72(3), 59-68.

Fromson, D. (1997), "Market Pulp Volatility Still Likely Despite Efforts to Stabilize Market," *Pulp & Paper* 71(3), 101-104.

Gan, J. and S.H. Kolison (1997), "Business Cycles and Forest Products Industries," *Forest Products Journal* 47(4), 23-26.

Ghali, K.H. (1998), "Public Investment and Private Capital Formation in a Vector Error-Correction Model of Growth," *Applied Economics* 30(6), 837-44.

Goodwin, B.K., T. Grennes and M. Wohlgenant (1990), "Testing the law of one price when trade takes time," *Journal of International Money and Finance* 9, 21-40.

Gordon, D.V. and R. Hannesson (1996), "On Prices of Fresh and Frozen cod fish in European and U.S. Markets," *Marine Resource Economics* 11, 223-238.

Granger, C.W.J. (1986), "Developments in the Study of Cointegrated Economic Variables," *Oxford Bulletin of Economics and Statistics* 48, 213-227.

_____ And P. Newbold (1976), "Forecasting Transformed Series," *Journal of the Royal Statistical Society* B 38, 189-203.

Gray, W.B. and R.J. Shadbegian (1998), "Environmental Regulation, Investment Timing, and Technology Choice," *Journal of Industrial Economics* 46(2), 235-256.

Gundersby, P. (1996), "Cycles and Investments - Causes and Effects," mimeo.

Gundersby, P. and M. Diesen (1985), "Will There be any New Greenfield Mills?" in L. Haas (ed.), "Market Pulp: a Comprehensive Update on Worldwide Supply, Demand and Production Trends for Market Pulp," Proceedings of PPI's 4th International Pulp Symposium, Brussels.

Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press, Princeton.

Hanninen, R.H. (1996), "The Law of One Price in United Kingdom Soft Sawnwood Imports – A Cointegration Approach," *Forest Science* 44(1), 17-23.

Henriksson, R.D. and R.C. Merton (1981), "On Market Timing and Investment Performance II: Statistical procedures for Evaluating Forecast Skills," *Journal of Business* 54, 513-533.

Hoffman, D.L. and R.H. Rasche (1996), "Assessing Forecast Performance in a Cointegrated System," *Journal of Applied Econometrics* 11, 495-517.

Holden, K. (1995), "Vector Autoregression Modeling and Forecasting," *Journal of Forecasting* 14, 159-166.

IMF (1999), *International financial statistics June 1999*, International Monetary Fund, Washington, D.C.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12, 231-54.

_____ (1995), *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models*, Oxford University Press, Oxford.

_____ (1996), "Likelihood-Based Inference for Cointegration of Some Nonstationary Time Series," in D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (eds.), *Time series models in Econometrics, Finance and Other Fields*, Chapman and Hall, London.

Johansson, L. (1996), "Asia keeps up the pace," *Pulp and Paper International* 38(4), 25-27.

Jorgenson, D.W. (ed.) (1996), *Investment: Capital Theory and Investment Behavior* volume 1, MIT Press, Cambridge MA.

_____ And C.D. Siebert (1968), "A Comparison of Alternative Theories of Corporate Investment Behavior," *American Economic Review* 58(4), 681-712, reprinted in Jorgenson (1996).

_____, J. Hunter and M.I. Nadiri (1970), "A Comparison of Alternative Econometric Models of Quarterly Investment Behavior," *Econometrica* 38(2), 187-212, reprinted in Jorgenson (1996).

Jung, C. and K. Doroodian (1994), "The Law of One Price For U.S. Softwood lumber: A Multivariate Cointegration Test," *Forest Science* 40(4), 595-600.

Kuh, E. (1963), *Capital Stock Growth: A Micro-Econometric Approach*, North Holland, Amsterdam.

Kuiper, W.E. and M.T.G. Meulenberg (1997), "Cointegration and Prediction Analysis of Market Supply in the Dutch Pig-Farming Industry," *European Review of Agricultural Economics*, 24(2), 285-312.

_____ And G.J. Thijssen (1996), "Determinants of Private Investment in Dutch Agriculture," *European Review of Agricultural Economics* 23(4), 455-72.

Lee, K.J. (1986), "Econometric Models of Capacity Expansion for The Pulp, Paper and Paperboard Industry," Ph.D. Thesis, University of Wisconsin, Madison.

LeSage, J.P. (1990), "A Comparison of The Forecasting Ability Of ECM And VAR Models," *Review of Economics and Statistics* 72(4), 664-671.

Ljung, G.M and G.E.P. Box (1978), "On a Measure of Lack of Fit In Time Series Models," *Biometrika* 65, 297-303.

Lord, M.J. (1991), "Price Formation in Commodity Markets," *Journal of Applied Econometrics* 6(3), 239-254.

Lundgren, T. and M. Sjostrom (1999), "A dynamic factor demand model for the Swedish pulp industry – An Euler equation approach," *Journal of Forest Economics* 5(1), 45-65.

Mackinnon, J.G. (1991), "Critical Values for Cointegration Tests," in R.F. Engle and C.W.J. Granger (eds.), *Long-Run Economic Relationships: Readings in Cointegration*, Oxford University Press, Oxford.

Meyer, J.R. and J.S. Strong (1990), "Sustaining Investment, Discretionary Investment, and Valuation: A Residual Funds Study of The Paper Industry," in R.G. Hubbard (ed.), *Asymmetric information, corporation finance, and Investment*, University of Chicago Press, Chicago.

Mills, T.C. (1998), "Recent Developments in Modeling Nonstationary Vector Autoregressions," *Journal of Economic Surveys* 12(3), 279-312.

Orr, B. and K.J. Lee (1990), "A Nonparametric Analysis of Variable Cost Minimization in The United States Paper Industry," *Applied Economics* 22, 839-852.

Officer, L.H. (1989), "The Law of One Price: Two Levels of Aggregation," in T. Grennes (ed.), *International Financial Markets and Agricultural Trade*, Westview Press, Boulder.

Oliner, S., G. Rudebusch and D. Sichel (1995), "New and old models of business investment: A comparison of forecasting performance," *Journal of Money, Credit and Banking* 27(3), 806-826.

Osterwald-Lenum, M. (1992), "A Note with Quantiles of The Asymptotic Distribution of The Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Economics and Statistics* 54(3), 461-71.

Phillips, P.C.B. (1987), "Time Series Regression with A Unit Root," *Econometrica* 55, 277-301.

_____ And P. Perron (1988), "Testing for a Unit Root In-Time Series Regression," *Biometrika* 75, 335-346.

QMS (1998), *Eviews 3 Users guide*, Quantitative Micro Software, Irvine.

Quicke, H.E., J.P. Caulfield and P.A. Duffy (1990), "The Production Structure of the U.S. Paper Industry," *Forest Products Journal* 40(9), 44-48.

Ravallion, M. (1986), "Testing Market Integration," *American Journal of Agricultural Economics* 68(1), 102-109.

Reinsel, G.C. and S.K. Ahn (1992) "Vector Autoregressive Models with Unit Roots and Reduced Rank Structure: Estimation, Likelihood Ratio Tests, and Forecasting," *Journal of Time Series Analysis* 13, 353-375.

Rudie, A.W. (1997), "Wood and How It Relates to Paper Products," *TAPPI Journal* 81(5), 223-228.

Samuel, C. (1998), "The Investment Decision: A Re-examination of Competing Theories Using Panel Data," *Applied Economics* 30, 95-104.

Sarker, R. (1993), "A Maximum Likelihood Cointegration Analysis of Canadian Lumber Exports," *Canadian Journal of Agricultural Economics* 41(1) 97-110.

Schoonbeck, L. and E. Sterken (1995), "Introduction and Overview," in L. Schoonbeck, E. Sterken and S.K. Kuipers (eds.), *Methods and Applications of Economic Dynamics*, North-Holland, Amsterdam.

Sims, C.A. (1980), "Macroeconomics and reality," *Econometrica* 48, 1-48.

Smook, G.A. (1992), *Handbook for Pulp and Paper Technologists*, Angus Wilde Publications, Vancouver.

Steele, T.W. (1995), "Capacity Expansion AND Strategic Behavior in the Canadian-United States Newsprint Industry," Ph.D. Thesis, University of Wisconsin, Madison.

Stier, J.C. (1985), "Implications of Factor Substitution, Economies of Scale, and Technological Change for the Cost of Production in the United States Pulp and Paper Industry," *Forest Science* 31(4), 803-812.

Swann, C.E. (1998), "Capital Projects: Equity Builders or Money Pits? Capacity Expansions in the Paper Industry," *North American Papermaker* 9 (80).

Tegene, A. and F. Kuchler (1994), "Evaluating Forecasting Models of Farmland Prices," *International Journal of Forecasting* 10(1), 65-80.

Teras, T. (1997), "Pulp Market 1997," *Foexpress* 1, 6-7.

Tilton, M. (1996), *Restrained Trade: Cartels in Japan's Basic Materials Industries,* Cornell University Press, Cornell.

Tobin, J. (1969), "A General Equilibrium Approach to Monetary Theory," *Journal of Money, Credit and Banking* 1, 15-29.

Toppinen, A. (1998), "Incorporating Cointegration Relations in a Short-Run Model of the Finnish Sawlog Market," *Canadian Journal of Forestry Research* 28, 291-298.

Trendel, O. and D. Paun (1999), "Pulp, Paper and Packaging Industry: A Literature Review of Business-Oriented Publications," CINTRAFOR working paper 68, University of Washington, Seattle.

Uusivori, J. and J. Buongiorno (1991), "Pass-Through of Exchange Rates on Prices of Forest Product Exports from The United States to Europe and Japan," *Forest Science* 37(3), 931-948.

Vogelvang, E. (1992), "Hypotheses Testing Concerning Relationships Between Spot Prices of Various Types of Coffee," *Journal of Applied Econometrics* 7(2), 191-201.

Whitham, J. (1994), "The Foreign Expansion of The Japanese Pulp and Paper Industry," in Proceedings of the Third International Symposium on Pulp and Paper, University of Washington, Seattle.

Wilson, R. (1997), "Where Were You When the Cycle Died?" *Pulp and Paper* 71(2), 132.

Young, R. (1987), "The RISI International Pulp and Paper Model," in P.A. Cardellichio, D.M. Adams and R.W. Haynes (eds.) "Forest Sector and Trade Models: Theory and Applications," CINTRAFOR, University of Washington, Seattle.

Zhang, D. and J. Buongiorno (1992), "Capacity Changes in the United States Paper and Paperboard Industries: *Q* Theory and Empirical Models," *Canadian Journal of Forest Resources* 23, 72-80.

_____, _____, And P.J. Ince (1996), "A Recursive Linear Programming Analysis of the Future of the Pulp and Paper Industry in the United States: Changes in Supplies and Demands, and The Effects of Recycling," *Annals of Operations Research* 68, 109-139.

Appendices

APPENDIX A: STRUCTURAL VAR ANALYSIS

APPENDIX A: STRUCTURAL VAR ANALYSIS

This appendix demonstrates the relationship between reduced form and structural VAR models. We start from a structural model of variables $y_{1t} \ldots y_{nt}$ with no restrictions on the relationships between the variables, and a maximum of p lags. Each variable has both contemporaneous and lagged impacts (up to a maximum lag length of p) on all the variables, including itself:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} 0 & \beta_{21} & \cdots & \beta_{n1} \\ \beta_{12} & 0 & \cdots & \beta_{n2} \\ \vdots & \vdots & & \vdots \\ \beta_{1n} & \beta_{2n} & \cdots & 0 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} + \begin{bmatrix} \gamma_{111} & \gamma_{211} & \cdots & \beta_{n11} \\ \gamma_{121} & \gamma_{221} & \cdots & \beta_{n21} \\ \vdots & \vdots & & \vdots \\ \gamma_{1n1} & \gamma_{2n1} & \cdots & \gamma_{nn1} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ \vdots \\ y_{nt-1} \end{bmatrix} + \begin{bmatrix} \gamma_{11p} & \gamma_{21p} & \cdots & \beta_{n1p} \\ \gamma_{12p} & \gamma_{22p} & \cdots & \beta_{n2p} \\ \vdots & \vdots & & \vdots \\ \gamma_{1np} & \gamma_{2np} & \cdots & \gamma_{nnp} \end{bmatrix} \begin{bmatrix} y_{1t-p} \\ y_{2t-p} \\ \vdots \\ y_{nt-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

Where β_{ab} is the coefficient for the contemporaneous effect of y_a on y_b, and γ_{abn} is the coefficient for the n-period lagged effect of y_a on y_b. As for the ARMA model, time trends may also be included. It is necessary to assume that the error terms in these equations are distributed normally, independent from one another, and independently over time. The variance-covariance matrix D is therefore diagonal.

Because of the potential contemporaneous correlation that results from period t values of the variables entering the RHS of the equations, we cannot use OLS to estimate these equations. We can reorganize the model as follows:

$$\begin{bmatrix} 1 & -\beta_{21} & \cdots & -\beta_{n1} \\ -\beta_{12} & 1 & \cdots & -\beta_{n2} \\ \vdots & \vdots & & \vdots \\ -\beta_{1n} & -\beta_{2n} & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \gamma_{111} & \gamma_{211} & \cdots & \beta_{n11} \\ \gamma_{121} & \gamma_{221} & \cdots & \beta_{n21} \\ \vdots & \vdots & & \vdots \\ \gamma_{1n1} & \gamma_{2n1} & \cdots & \gamma_{nn1} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ \vdots \\ y_{nt-1} \end{bmatrix} + \ldots + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

Or in matrix notation,

$$\beta y_t = \alpha + \Gamma_1 y_{t-1} + \ldots + \Gamma_p y_{t-p} + \varepsilon_t$$

If β is invertible, we have

$$y_t = \beta^{-1}\alpha + \beta^{-1}\Gamma_1 y_{t-1} + \ldots + \beta^{-1}\Gamma_p y_{t-p} + \beta^{-1}\varepsilon_t$$

Which is equivalent in form to the reduced form VAR model:

$$y_t = \Pi_0 + \Pi_1 y_{t-1} + \ldots + \Pi_p y_{t-p} + e_t$$

Generally, an estimated reduced form VAR will not provide us with sufficient coefficients to derive all of the unrestricted structural form coefficients. In order to obtain a unique solution, therefore, it is necessary to impose restrictions on the structural form. There are various types of restrictions that may be imposed on the system. These include exclusion restrictions (whereby structural parameters for contemporaneous effects are set to zero), normalization restrictions (whereby variances are set to 1), and long-run restrictions (which take the form of restrictions on the IRFs).

In this research, the first type of restriction is used. Imposing the restrictions $\beta_{ij} = 0$ for all $j>i$ allows us to extract the structural coefficients from the reduced form estimates. The ordering of the y variables should be chosen so as to make these restrictions acceptable. We start from the estimated variance-covariance matrix -:

$$\Omega = E[e_t e_t'] = E\left[\beta^{-1}\varepsilon_t\left(\beta^{-1}\varepsilon_t\right)\right] = \beta^{-1}D\left(\beta^{-1}\right)$$

Triangular factorization of Ω allows us to derive unique values for β^{-1} and D.[136] The estimated β matrix can now be used to derive estimates for the structural parameters:

$$\hat{\alpha} = \beta\hat{\Pi}_0 \text{ and } \hat{\Gamma}_1 = \beta\hat{\Pi}_1 .$$

These are the structural parameters shown in section 4.2.4.

[136] See Hamilton (1994, section 4.4).

APPENDIX B: JOHANSEN'S COINTEGRATION TEST

APPENDIX B: JOHANSEN'S COINTEGRATION TEST

This appendix outlines the method proposed by Johansen (1988, 1995, 1996) to test for the number of cointegrating relationships present in a group of variables. The starting point for this test is the familiar VAR *(p)* model, which for variables $y_t' = y_{1t} \ldots y_{nt}$ can be represented as follows:

$$\Phi(L)y_t = \alpha + \varepsilon_t$$

Where $\Phi(L) \equiv I_n - \Phi_1 L - \ldots - \Phi_p L^p$ and $\varepsilon_t \sim$ i.i.d. $(0,\Omega)$.

Making use of the Wold representation, we can manipulate this expression to show that if y_t is I (1) and has h linearly independent cointegrating vectors $a_1, a_2 \ldots a_h$ then we can decompose Φ_1 as follows:

$$\Phi(1) = -BA$$

Where B is some *(nxh)* matrix and A is an *(hxn)* matrix with rows $a_1, a_2 \ldots a_h$. This implies that the rank of Φ_1 is h.

Another way of representing the VAR *(p)* system is in first differences:

$$\Delta y_t = \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + \zeta_0 y_{t-1} + \varepsilon_t$$

Where

$$\zeta_i \equiv -\left[\Phi_{i+1} + \ldots + \Phi_p\right] \text{ for all } i = 1,2,\ldots, p\text{-}1,$$

And

$$\zeta_0 \equiv -\left[I_n - \Phi_1 - \ldots - \Phi_p\right] = -\Phi(1) = BA\ .$$

We can make use of this equivalence between ζ_0 and $-\Phi_1$. We estimate ζ_0, and test whether its rank is h. If we do not reject this hypothesis we can infer that there are h cointegrating vectors for the system.

It is not acceptable to estimate ζ_0 using OLS, because y_t is non-stationary. Instead we use full-information maximum likelihood. We face the complication that the log likelihood is a function of all of the coefficients contained in the equation above, and also of Ω. Auxiliary regressions are used to concentrate the log likelihood to be a function of ζ_0 only:

$$\Delta y_t = \hat{\Pi}_1 \Delta y_{t-1} + \ldots + \hat{\Pi}_{p-1} \Delta y_{t-p+1} + \hat{\pi}_0 + \hat{u}_t$$

$$y_{t-1} = \hat{\aleph}_1 \Delta y_{t-1} + \ldots + \hat{\aleph}_{p-1} \Delta y_{t-p+1} + \hat{\theta} + \hat{v}_t$$

Where $\hat{u}_t$ and $\hat{v}_t$ are vectors of residuals from these two equations. We effectively use these estimates $\hat{\Pi}_i$ and $\hat{\aleph}_i$ to substitute out ζ_i (for i=1,2,...p-1) and use $\hat{\pi}_0$ and $\hat{\theta}$ to substitute out α from the log likelihood function. Maximizing the concentrated log likelihood then turns out to be equivalent to choosing ζ_0 so as to minimize the following expression:

$$\left| (1/T) \sum_{t=1}^{T} \left[(\hat{u}_t - \zeta_0 \hat{v}_t)(\hat{u}_t - \zeta_0 \hat{v}_t)' \right] \right|$$

(Where T is the sample size) subject to the condition that ζ_0 be of rank h. It can be shown that the minimum value for this expression is a function of the canonical correlation's between $\hat{u}_t$ and $\hat{v}_t$.[137] The squares of the canonical correlation's are equivalent to the eigenvalues λ_i ($\lambda_1 > \lambda_2 > \ldots > \lambda_n$) of

$$\Sigma = \hat{\Sigma}_{VV}^{-1}\hat{\Sigma}_{VU}\hat{\Sigma}_{UU}^{-1}\hat{\Sigma}_{UV}$$

Where sample covariance matrix $\hat{\Sigma}_{ij} = (1/T)\sum_{t=1}^{T}\hat{i}_t\hat{j}_t'$ for all $i, j = u, v$.

These eigenvalues can then be used to express the maximum value of the log likelihood function as follows:

$$L^* = -(Tn/2)(\ln 2\pi + 1) - (T/2)\ln\left|\hat{\Sigma}_{UU}\right| - (T/2)\sum_{i=1}^{h}\ln(1-\lambda_i)$$

Intuitively, these eigenvalues can be considered somewhat analogous to the R^2 of a normal OLS regression: Each eigenvalue measures the amount of residual error present, for a given linear combination of variables. If a particular vector were a cointegrating vector, we would expect the residual error to be small. Similarly, if it is not a cointegrating vector, we would expect the error to be large (indicated by an eigenvalue close to zero). Since the eigenvalues are ordered, their eigenvectors effectively represent the 'most likely candidates' to be linearly independent cointegrating vectors. The testing procedure is in essence an evaluation of how many of these eigenvalues are significantly different from zero.[138]

By evaluating the equation above for different values of h, we are able to construct a sequence of likelihood ratio tests, and thereby arrive at a preferred value for h. The null hypothesis for each test is as follows:

$H^j(r)$: The number of cointegrating vectors in the system is at most r.

Because the hypothesis admits any number of cointegrating vectors less than or equal to r, hypotheses are usually tested in sequence, beginning with $r = 0$ and potentially continuing to $r = n$-1 (if earlier null hypotheses are all rejected).[139] j refers to the structure of the model used for estimation, discussed further below. The alternative hypothesis can take either of two forms:

$H^j(r+1)$: The number of cointegrating vectors in the system is at most $r+1$; or

$H^j(n)$: The number of cointegrating vectors in the system is at most n.

The test statistics for these two are known as the λ-max and trace test statistics respectively. In both cases the statistics are calculated as likelihood ratios $2(L_A^* - L_0^*)$ (see the equation above). For the λ-max statistic this gives a value of

$$\tau_{\lambda\max} = -T\ln(1-\hat{\lambda}_{r+1})$$

and for the trace test statistic gives a value of

---

[137] Hamilton (1994, 639-41) and Banerjee *et al* (1993, section 8.2) give details.

[138] How does this relate to testing the rank of ζ_0? If each row of ζ_0 is linearly independent, then we can normalize it to a triangular representation. Its rank is equal to the number of non-zero diagonal elements, which are its eigenvalues.

[139] The greatest number of cointegrating vectors feasible is n-1. If we reject the hypothesis that there are at most n-1 cointegrating relationships, then we conclude that all of the components of y_t are in fact stationary themselves.

$$\tau_{trace} = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i)$$

These are derived from asymptotic arguments. Reinsel and Ahn (1992) suggests using the factor $(T - np)$ rather than T, as being more reliable for finite samples. In general, critical values for these tests are non-standard. Test statistics are not distributed as χ^2, but rather as functions of Brownian motion. Specific critical values have been developed by Osterwald-Lenum (1992). As well as depending on h, these critical values depend on the nature of the economic relationships underlying the hypothesized cointegrating relationships, and on whether or not we choose to allow for linear deterministic trends in the data.[140]

[140] Critical values also differ according to whether non-zero constants are allowed in the cointegrating vectors.