Modelling Non-Ferrous Metals Markets

Clinton Watkins and Michael McAleer

Department of Economics

University of Western Australia

Abstract: This paper evaluates the significance of the empirical models and the distributional properties of prices in non-ferrous metals spot and futures markets published in leading refereed economics and finance journals between 1980 and 2002. The survey focuses on econometric analyses of pricing and returns models applied to exchange-based spot and futures markets for the main industrially-used non-ferrous metals, namely aluminium, copper, lead, nickel, tin and zinc. Published empirical research is evaluated in the light of the type of contract examined, frequency of data used, choice of both dependent and explanatory variables, use of proxy variables, type of model chosen, economic hypotheses tested, methods of estimation and calculation of standard errors for inference, reported descriptive statistics, use of diagnostic tests of auxiliary assumptions, use of nested and non-nested tests, use of information criteria, and empirical implications for non-ferrous metals.

Corresponding Author:

Michael McAleer

Department of Economics	Tel:	+618 9380 3400
University of Western Australia	Facsimile:	+618 9380 1016
35 Stirling Highway,	Email:	Michael.McAleer@uwa.edu.au
Crawley, WA 6009		
AUSTRALIA		

1 Introduction

The analysis of commodity spot and futures markets is important in economics and finance. A wealth of theoretical and empirical research has seen the development of numerous empirical models to analyse the distributional properties of prices in non-ferrous metals spot and futures markets. Empirical models in the literature explain price relationships in levels or first differences, returns on commodity assets, and the variance in returns or volatility. Several models of price levels, including relationships between spot and futures prices and between prices in markets for different metals, use recent advances in the analysis of non-stationary models and data. Particular attention has been paid to efficiency and risk premia in metals markets through testing of the Efficient Market Hypothesis, the Unbiased Expectations Hypothesis (UEH) and the Speculative Efficiency Hypothesis (SEH). Implications of the theory of storage, speculation, and metals market supply and demand fundamentals for the price and return behaviour of contracts on futures exchanges have been important areas of research. Models of volatility are motivated by concerns regarding risk in the markets, the pricing of derivatives, and the development of hedging strategies for non-ferrous metals producers, consumers and stock holders.

Non-ferrous metal commodities play an important role in most national economies, and their prices have impacts on the extraction, processing and manufacturing sectors. The London Metal Exchange (LME) is the major international market for the main industrially-used non-ferrous metals, namely aluminium, aluminium alloy, copper, lead, nickel, tin, zinc and silver. Most studies surveyed in this paper analyse data from this exchange. The LME is used worldwide by producers and consumers of non-ferrous metals as a centre for spot, futures and options trading in these metals. Three primary functions are performed by the non-ferrous metals markets on the LME. First, the exchange provides a market where non-ferrous metal industry participants can hedge against risks arising from price fluctuations in world metals markets. Second, settlement prices determined on the LME are used internationally as reference prices for the valuation of activities relating to non-ferrous metals. Third, the LME also provides appropriately located storage facilities to enable market participants to take or make physical delivery of approved brands of non-ferrous metals. The LME is the most important market for the pricing of non-ferrous metals world wide. Approximately 95% of the total world trade in copper futures occurs though the LME, with the bulk of the remaining 5% in the copper market on the Commodity Exchange of New York (COMEX). Smaller regional markets typically participate only in spot trade of non-ferrous metals. One exception is the Shanghai Futures Exchange (SHFE), on which a small volume of futures for aluminium and copper are traded primarily for the Chinese domestic market. The copper settlement price determined on the LME is effectively the world copper price [Gilbert, 1996].

This paper analyses the econometric techniques applied to futures and spot price and returns data. A more accurate view of futures prices for metals is of particular interest to participants in industries reliant on the production or consumption of metals, such as miners, smelters, refiners, rolling mills, extrusion plants, metals merchants, and fabricators. Energy providers, banks, investment funds and, to some extent, speculators, are also active participants in metals futures markets. At a macroeconomic level, commodity prices play an important role in the economy of many countries, including Australia. Developing economies are particularly reliant on commodity production in the generation of national income. A greater understanding of the relationship between futures and spot prices has important policy implications for commodity-dependent nations for key indicators such as exchange rates, inflation and economic growth.

Since around 1980, non-ferrous metals markets have attracted substantial academic interest. This review of empirical papers published in leading refereed economics and finance journals between 1980 and 2002 focuses on econometric analyses of pricing and returns models applied to exchange-based spot, forward and futures markets for the main industrially-used non-ferrous metals, namely aluminium, copper, lead, nickel, tin and zinc. The research papers under review were located by a search of the ECONLIT bibliographical database.

There are two primary purposes of the paper, namely to: (i) provide an entrée to the body of empirical literature concerned with the pricing of futures contracts in markets for industrially-used non-ferrous metals; and (ii) evaluate the significance of empirical models published in leading economics and finance journals since 1980.

Detailed descriptive classifications are provided for 45 studies of industrial metals prices published in refereed journals. Appendix 1 contains a full and detailed classification for each paper considered. Published empirical research is evaluated in the light of several key criteria. The journal in which each study was published and the year of publication are tabulated for each of the 45 studies. Economic hypotheses tested are classified into nine types, and the significance of several important empirical models is discussed. Aspects of the data analysed include the exchange providing the data, the type of futures, forward or spot trades examined, the industrially-used non-ferrous metals considered, and importantly, the sampling frequency and the size of the sample used for modelling.

Several dependent variables are used in the multitude of empirical models contained in the 45 studies. The choice of dependent variables is tabulated, and the variables are classified within 21 types. Models of spot and futures prices and returns applied in non-ferrous metals markets contain several types of explanatory variables. The frequency with which each type of explanatory variable is used is discussed. Frequently, the variables specified in economic or financial models are not directly observable. In order to develop a specification of these models for empirical testing, proxy variables and generated regressors are used. The econometric consequences of including proxy variables and generated regressors in a model are reviewed, and those used in the 45 studies are identified.

Prior economic theory provides relationships to estimate and test, and several types of econometric model specification are used. The types of models and their frequency of use are discussed, followed by the methods of estimation and calculation of standard errors for inference. Most studies examined provide some means of evaluating the statistical adequacy of the empirical specifications used to model economic and financial relationships. A detailed discussion of the reported descriptive statistics and information criteria is followed by the use of diagnostic tests of auxiliary assumptions required to specify the models. In light of the fact that many of the papers have considered competing specifications, the use of nested and non-nested tests is also considered. The table in Appendix 1 classifies each empirical paper according to the descriptions given above, and each is discussed in the following sections.

2 Published Research on Non-Ferrous Metals Markets

Forty-five empirical studies published between 1980 and 2002 are evaluated in this paper. While the review is not intended to be an exhaustive study of commodity market models, the papers analysed reflect the quality of empirical work in metals markets over the last two decades. In addition to this body of work, many authors consider metals markets to be a subset of larger commodity market analyses. However, in order to provide a focus on metals market empirical analysis, commodities other than metals are not included.

This paper analyses empirical studies of the main non-ferrous industrially used metals, namely, aluminium, copper, lead, nickel, tin and zinc. Precious metals are only included to the extent that they appear in papers that are otherwise concerned with industrial metals. The papers reviewed are given in the Appendix, namely Agbeyegbe [1992], Ben Nowman and Wang [2001], Bracker and Smith [1999], Bresnahan and Suslow [1985], Brunetti and Gilbert [1995], Canarella and Pollard [1986], Chang et al. [1990], Chowdhury [1991], Fama and French [1988], Franses and Kofman [1991], Fraser and MacDonald [1992], French [1983], Gilbert [1986, 1995], Goss [1981, 1983, 1986], Gross [1988], Hall [1991], Hall and Taylor [1989], Hardouvelis and Kim [1995], Heaney [1998, 2002a, 2002b], Hill et al. [1991], Hsieh and Kulatilaka [1982], Hussey and Quiroz [1997], Kocagil [1997], Krehbiel and Adkins [1993], Labys et al. [1998], MacDonald and Taylor [1988a, 1988b, 1989], MacKinnon and Olewiler [1980]. McKenzie et al. [2001], McMillan and Speight [2001, Moore and Cullen [1995], Ng and Pirrong [1994], Sephton and Cochrane [1990a, 1990b, 1991], and Shyy and Butcher [1994], Slade [1991], Teyssiere et al. [1997], and Varela [1999].

Table 1: Journals publishing research on non-ferrous metals

Journal	Number of Papers
Applied Economics	6
Applied Economics Letters	1
Applied Financial Economics	1
Bell Journal of Economics	1
Bulletin of Economic Research	2
Economics Letters	1
European Journal of Finance	1
International Economic Review	1
International Journal of Forecasting	1
Kentucky Journal of Economics and Business	1
Journal of Applied Econometrics	1
Journal of Banking and Finance	1
Journal of Business	1
Journal of Finance	2
Journal of Financial Economics	1
Journal of Futures Markets	10
Journal of Money, Credit and Banking	1
Managing Metals Price Risk[1]	1
Oxford Bulletin of Economics and Statistics	1
Quarterly Journal of Economics	1
Resources Policy	3
Review of Financial Economics	1
Review of Futures Markets	2
Revista de Analisis Economico	1
The Manchester School	2
Total (in 25 Journals)	45

Note:
1. Refers to a chapter of an edited book rather than a journal.

2.1 Journals Containing Research on Industrial Metals Markets

The 45 published papers under review appeared in 24 economics and finance journals, and also in an edited book. Empirical analyses of industrially used metals markets appear most frequently in the *Journal of Futures Markets*, which includes a substantial proportion of the applied econometric analysis of metals futures markets. This is followed by *Applied Economics*, with six papers. *Resources Policy* contains numerous papers relating to aspects of ferrous and non-ferrous metals, three of which are relevant to this review. *Bulletin of Economic Research*, *Journal of Finance*, *Review of Futures Markets*, and *The Manchester School* each contain two of the surveyed papers. Eighteen publications each contains only one of the papers considered. In general, the journals are those of an applied nature. One paper appearing in *Managing Metals Price Risk,* although not a journal, is included in the survey as it contributes to research on volatility in metals markets where there is little published empirical research.

2.2 Year of Publication

Table 2 shows the publication year for each of the 45 empirical analyses of non-ferrous metals markets reviewed in this paper. All the publications appeared between 1980 and 2002. The greatest number of papers published in any one year is five, which occurs in 1991. In the four-year period 1988-1991, 15 papers were published, which is the greatest number of any four-year period in the last two decades. Although there are four years in which no published papers appear, there are no consecutive years of zero publications.

Table 2: Publication year for research on non-ferrous metals

Year of Publication	Number of Papers
1980	1
1981	1
1982	1
1983	2
1984	0
1985	1
1986	3
1987	0
1988	4
1989	2
1990	4
1991	5
1992	2
1993	1
1994	2
1995	4
1996	0
1997	3
1998	2
1999	2
2000	0
2001	3
2002	2
Total	45

3 Economic Hypotheses Analysed in Empirical Research

Table 3 provides nine types of economic hypothesis investigated in the literature, and the frequency with which these hypotheses are examined in the 45 empirical studies. Each paper is classified under the most relevant hypothesis.

Table 3: Economic hypotheses tested

Economic Hypothesis	Frequency
Efficient market hypothesis	13
Speculative efficiency hypothesis	8
Common (stochastic) trends	4
Theory of storage and cost-of-carry model	5
Speculation, hedging and volatility	2
Price and returns volatility processes	6
Risk premia and CAPM	1
Other futures market related	3
Other metals market fundamentals related	3
Total	45

The EMH was the most frequently analysed. Thirteen papers examine the EMH (including papers that consider the UEH as evidence for market efficiency), and develop empirical models to test the hypothesis. For a market to be efficient, prices must fully and instantaneously reflect all available relevant information, and no profit opportunities are left unexploited [Fama, 1976]. The EMH asserts that agents form rational (or model consistent) expectations and quickly arbitrage away any deviations of expected returns consistent with normal profits. Over the last two decades, much of the empirical work on metals forward and futures markets has examined issues related to market efficiency and unbiasedness. However, the evidence of market efficiency provided by empirical studies is mixed. Several definitions of market efficiency existing in the literature have led to some confusion.

Chowdhury [1991] uses two specifications of a cointegration model to evaluate whether the EMH is supported for metals traded on the London Metal Exchange (LME). Cointegration between spot (futures) prices in one metals market with spot (futures) prices in another metals market is presumed to indicate inefficiency. The presence of cointegration between two speculative markets for two different assets implies predictability [Granger, 1986]. Predictability between two different markets violates the EMH. Alternatively, cointegration is expected between spot and futures prices for the same underlying asset in an efficient market. This is because, even though the spot and futures price series may be non-stationary, they do not drift apart in an efficient market [Hakkio and Rush, 1989]. If spot and futures prices in one market do drift apart, it is likely that the futures price may not be the best predictor of the next period spot price using all publicly available information. In this respect, Hakkio and Rush [1989] show that cointegration is not necessary, but is sufficient, for the EMH to hold. Agbeyegbe [1992] argues that cointegration does not imply inefficiency, but merely that unanticipated changes dominate movements in prices.

Hill et al. [1991] test the efficiency of metals futures markets in relation to the “cold fusion” announcement of 1989. For a futures market to be generally price efficient with respect to all publicly available information, it must also prove efficient with respect to any given information set. Platinum and palladium futures markets are shown to be efficient with respect to the “cold fusion” announcement.

MacDonald and Taylor [1988b] find monthly price series for lead, tin and zinc are I(1). However, none of the metals is cointegrated with each other. The authors argue that the absence of a long-run relationship between the metals supports the EMH. Sephton and Cochrane [1990, 1991] examine efficiency of the LME aluminium, copper, lead, nickel, tin, and zinc markets in terms of forecast errors, and the joint hypothesis of risk neutrality and rational expectations. Tests involving forecast errors evaluate whether lagged forecast errors aid in predicting current forecast errors. The empirical analysis of forecast errors in single market models rejects efficiency for aluminium, lead, tin and zinc, while the copper and nickel markets are efficient. Multiple market models do not reject efficiency for aluminium and lead only. Using a methodology attributed to Fama [1984a,b], Sephton and Cochrane [1991] produce results contradicting previous published empirical papers using the same methodology. Efficiency is rejected for the copper and tin markets, where risk premia cannot be rejected. Stability tests show the results for all models, except for tin, are questionable. An alternative methodology, such as one based on cointegration, is required for valid inferences regarding market efficiency.

Futures or forward prices that are biased predictors of future spot prices are frequently found in non-ferrous metals markets. This is often argued to be evidence of inefficiency. Brenner and Kroner [1995], however, argue that the systematic difference between spot and futures prices may be due to carrying costs. Arguments for a bias in forward or futures prices in predicting future spot prices include risk premia and carrying costs, both of which may be time varying. The unbiased predictor hypothesis, or SEH, states that the futures price at time t for a contract with maturity length j is an unbiased predictor of the spot price that will prevail in the market at time t+j, given the information available to market participants at time t. That is, the futures price is an unbiased predictor of the future spot price. In the literature, speculative efficiency has been associated with the EMH and the rational expectations hypothesis (REH). However, Bilson [1981] demonstrates that the unbiased predictor hypothesis is not a necessary condition for either the EMH or the REH to hold. Markets may be efficient in that there is no opportunity for risk adjusted excess returns, but in which there is a predictable bias in the futures price forecast. Expectations may be rational even though futures prices are not equal to expected spot prices, due to the presence of transactions costs, information costs and risk aversion.

The SEH is examined in eight papers, and includes tests of futures price unbiasedness that are not considered to provide evidence for or against market efficiency. Under the joint hypothesis of risk neutrality and rational expectations, an empirically testable form of speculative efficiency can be defined so that the futures price is an unbiased predictor of the realised future spot price. The definition has the advantage of not requiring a proxy variable or generated regressor for the expected spot market price. Moore and Cullen [1995] specify their cointegration model according to this definition, and speculative efficiency cannot be rejected for the LME aluminium, copper, lead and zinc markets, while it is rejected for the nickel market. The hypotheses of unbiased expectations and absence of a risk premium in COMEX copper, gold, and silver, and NYMEX platinum futures markets, are tested by Krehbiel and Adkins [1993] using a cointegration model. Absence of a risk premium is rejected for copper only, while the unbiased expectations hypothesis is rejected for every model except platinum. A joint test for both hypotheses is rejected for all markets, except copper.

Speculative efficiency is examined by Hsieh and Kulatilaka [1982] under two informational assumptions, full information and incomplete information, and non-zero risk premia are found. Canarella and Pollard [1986] show that markets on the LME are speculatively efficient using different methods to Hsieh and Kulatilaka [1982], and attribute their conflicting results to the use of different econometric methods. The statistical procedures in Hsieh and Kulatilaka [1982] are criticised as failing to take account of the overlapping data by using, for example, the Hansen and Hodrick [1980] technique for serial correlation. However, the Hansen and Hodrick [1980] procedure has been shown to work poorly in small samples [Pesaran and Slater, 1980]. Several of the papers concerned with speculative efficiency estimate models based on risk premia. Hall [1991] analyses the risk premium in the LME copper, lead, tin and zinc markets using GARCH in mean (GARCH-M) and stochastic GARCH-M models, and provides support for the existence of time-varying risk premia.

Four papers examine metals markets for common stochastic trends, equilibrium parity relationships between markets, lead-lag relationships between markets, and common cyclical patterns. Franses and Kofman [1991] test for flow parity relationships between forward prices for aluminium, copper, lead, nickel and zinc on the LME. One cointegrating relationship exists between the five metals, so that a long-run relationship exists between the forward price series. If efficiency is defined such that a random walk is the best forecasting scheme, the LME is inefficient. Similarly, Agbeyegbe [1992] tests for common stochastic trends among copper, lead and zinc spot prices on the LME, and finds one relationship between the three metals, and a bivariate relationship between copper and lead. However, in both Franses and Kofman [1991] and Agbeyegbe [1992], the authors do not interpret cointegration as evidence of inefficiency.

International linkages between the Shanghai Metal Exchange (SHME), operating under strict Chinese Government controls, and the LME are investigated by Shyy and Butcher [1994]. Spot and forward prices for copper on the SHME are cointegrated with the respective copper spot and forward prices on the LME, and it is claimed that the SHME prices coincide with those of the world market. Although not explicitly acknowledged by the authors, the analysis of markets between exchanges involves several problems that complicate the analysis, including exchange rate conversion and trading day differences. Trading on the SHME starts well before that of the LME, and one would expect that if the SHME is important with respect to world metal prices, information from the SHME trade would be accommodated by participants in the LME exchange.

Price cycles in metals markets are thought to exist, owing to the characteristics of metal supply and demand. Supply is price inelastic in the short run, whereas metal demand responds rapidly to changes in industrial production, which is closely related to the business cycle. The apparent cyclical behaviour of metals prices is discussed in the literature (see Fama and French [1988] and Teyssiere et al. [1997]). Labys et al. [1998, 1999] explicitly examine price cycles for non-ferrous metals.

Implications of the theory of storage are tested in five papers, using models of the cost-of-carry relationship and the convenience yield on holding inventories. Fama and French [1988] propose a refinement to the Samuelson hypothesis. Futures prices are less variable than spot prices when inventory is low, so that the Samuelson hypothesis holds. When inventory is high, spot and futures prices have roughly the same variability (that is the Samuelson hypothesis does not hold) because the marginal convenience yield on inventory declines at higher inventory levels (but at a decreasing rate). Their empirical analysis supports their refinement of the Samuelson hypothesis.

Two papers consider the effects of speculation and hedging on the volatility of prices for non-ferrous metals. Slade [1991] examines the volatility of non-ferrous metal spot prices over time, with reference to periods under which producer list pricing and exchange based pricing are dominant. Producer pricing has been prevalent in many non-ferrous metals markets until the mid 1970s or 1980s. The empirical analysis by Slade [1991] asserts that producer-pricing regimes generated prices which are less volatile than those arising from exchange-based spot, futures and options trading. Brunetti and Gilbert [1995], however, find that non-ferrous metals markets have not become more volatile over the same period. Furthermore, Figuerola-Ferretti and Gilbert [2001] argue that Slade's [1991] analysis is not robust to the deletion of silver from the sample, and that silver contained an atypical influence of the Hunt Brothers' manipulation of the market, so that it biases the analysis.

Speculation is said to have a number of desirable features, including stabilisation of prices by elimination of arbitrage opportunities. Alternatively, the sophisticated speculator can exploit the naive forecasting rules of less sophisticated agents, thereby destabilising prices. Increased speculation in futures markets does not have a stabilising effect on spot prices, according to tests by Kocagil [1997] using aluminium, copper, gold, and silver contract data. Speculative trade, relative to non-speculative trade, cannot be observed, so a critical condition in terms of observable variables derived from a theoretical futures market model is tested. However, a futures market is, by its nature, a speculative market, and it is difficult to isolate part of futures trade as speculative trade. It may be more sensible to consider the effects of increased or decreased trading volume on the stability of spot prices.

Processes describing the volatility of non-ferrous metals spot and futures prices and returns are examined in six papers. Volatility refers to the variation in an asset price, or variation in the returns based on holding the asset, and may be analysed over a variety of periods, most frequently based on monthly, weekly, daily, and intra-daily data. Brunetti and Gilbert [1995] use a monthly measure of volatility, which is generated using daily data. They find that by examining averages of the monthly data over three-year horizons that: non-ferrous metals volatility has been consistent from 1972 to 1995; temporal patterns of volatility movements are similar in aluminium, copper, lead, nickel, tin and zinc; there are two periods of high volatility in 1973-1974 (except for aluminium and nickel, where there is no data), and 1988-1990, periods that are characterised by tight supply-demand balances in non-ferrous metals markets; and volatility over 1993 to 1995 was lower than its historic average level (except for tin, where it was higher). Brunetti and Gilbert [1995] note that long-term volatilities are of most interest to producing and consuming non-ferrous metals industries. However, they show that a daily measure yields the same results as their monthly measure.

Commodity market models relating prices and their volatility to market fundamentals typically hold that, in periods of excess supply, the impact of production or consumption variability is attenuated through stocks. However, in periods of excess demand, when stocks have been exhausted, the price is required to adjust sufficiently to balance the market. Gilbert [1989] suggests an asymmetry of the price response to stock level between periods of excess demand and excess supply. Prices, therefore, move more in periods of excess demand than in periods of excess supply.

The consistency of metals futures market risk premia with the systematic risk under the Sharpe-Lintner Capital Asset Pricing Model (CAPM) is examined in one paper. A risk premium commensurate with the systematic risk of each contract is identified by Chang et al. [1990] for copper, platinum and silver futures prices, using a model based on the CAPM. This model assumes that the return on a financial asset comprises a risk premium and the return on a risk-free asset. The risk-free asset return represents the time premium for committed capital and the risk premium on a financial asset is comprised of the product of the systematic risk and the risk premium on the market portfolio. Agents holding financial assets are compensated only for bearing those risks that cannot be diversified away. Chang et al. [1990] present results that support the Keynesian normal backwardation in the context of the CAPM. Under normal backwardation, the forward price will, on average, be less than the expected spot price to generate a return to speculative activity.

Hypotheses related to metals futures markets and pricing, and hypotheses related to metals market fundamentals related hypotheses are examined in six papers.

4 Analysing the Empirical Studies

In this section, the following empirical issues are discussed with respect to the 45 empirical papers reviewed: sample data; dependent variables; explanatory variables, including proxy variables and generated regressors; model specifications; method of estimation; descriptive statistics; diagnostic testing; and nested and non-nested tests. The analysis draws on the empirical classification of the papers included in Appendix 1.

4.1 Sample Data

The sample data used in each empirical analysis is evaluated in this section in light of the commodity exchange to which it relates, the non-ferrous metals markets included in the analysis, whether the focus of the modelling is on futures, forward or spot markets, the sampling frequency of the data used, and the number of observations in the sample. Several papers use more than one sample, and authors frequently consider data from multiple commodity exchanges within one analysis. Between one and nine metals are considered in the analyses (for example, Fraser and MacDonald [1992] and Labys et al. [1998] both consider nine metals), either in univariate single market models, or in a multivariate framework, relating several markets (for example, Franses and Kofman [1991]).

Table 4 indicates the reported source of the metals price data used in the 45 empirical papers. By far the most frequently analysed non-ferrous metals markets are those of the LME. The LME is the largest market for futures contracts in non-ferrous metals, and is also an exchange for spot transactions where physical delivery takes place. Of the 45 papers, 36 use data for at least one of the metals traded on the LME. The next most frequent sources of data are the New York Mercantile Exchange (NYMEX) and its subsidiary, the Commodity Exchange of New York (COMEX), which are cited as the data source in six and seven papers, respectively. In three papers, metals data from the Chicago Board of Trade (CBOT) are analysed, two papers consider spot prices for tin determined on the Kuala Lumpur Tin Exchange, and one paper uses data for the markets of the Shanghai Metal Exchange (SHME), now part of the Shanghai Futures Exchange (see Shyy and Butcher [1994]). Producer list prices are included in the models of three papers, namely, Gilbert [1995], MacKinnon and Olewiler [1980] and Slade [1991]. For a comprehensive discussion of the role of producer list prices in the international markets for non-ferrous metals markets, see also Figuerola-Ferretti and Gilbert [2001]. Four papers do not state the exchange from which their data are obtained, but in one case it would appear the authors are using data from COMEX. No papers considered in this analysis include data from the Tokyo Commodity Exchange (TOCOM), as the exchange is principally concerned with precious metals (namely, gold, platinum, palladium, and silver), although aluminium is also traded.

A number of characteristics of the different markets and exchanges are frequently cited in the non-ferrous metals literature, and in both the academic and practitioner literatures on commodity markets, in general. In this paper, a number of key points related to empirical modelling are discussed. However, a comprehensive comparison of the various world markets is beyond the scope of this review.

Table 4: Source of price data

Exchange	Frequency
CBOT	3
COMEX	7
KL Tin Exchange	2
LME	36
NCE	1
NYMEX	6
Presumably COMEX	1
Producer List Price	3
SHME	1
Not Stated	3
Total[1]	63

Note:
1. Some studies used data from more than one exchange.

Many authors cite advantages for the empirical modelling of LME metals data (see, for example, Fama and French [1988]). LME markets have high trading volume relative to metals markets on other exchanges, particularly in aluminium and copper, which reduces any non-trading effect in the data. Unlike other futures exchanges, simultaneous spot and futures prices are quoted on the LME. Contracts traded on the LME differ from traditional futures traded on the other exchanges, and there has been some debate in the literature as to whether the LME contracts are forward or futures. This issue is addressed in later sections of the paper, and LME contracts will be referred to as futures. Futures prices on the LME are quoted for fixed maturities, that is, there is a new contract for each business day. For example, there is a 3-month futures price quoted on each business day, and for each trading day there is a new contract with a maturity of 3-months (9-, 15- and 27-month contracts are also quoted for most metals on the exchange, but 3-month contracts are the most liquid in each metals market). The price for a particular vintage of contract is quoted once on the first day of the contract, and not again until maturity. In contrast, the futures contracts, exchanges such as COMEX, NYMEX, and CBOT, do not have fixed maturities, but have a three to four week delivery period at the beginning of the maturity month. Moreover, in these markets, spot prices for analysis are unavailable, and so futures prices for maturing contracts are used a proxy for spot prices. As futures contracts in most commodities do not mature monthly, using maturing futures as spot prices reduces sample sizes in these markets. Moreover, there are no limits on LME spot or forward prices, in contrast to most futures markets, in which there are daily limits. If prices reach these limits, trading stops, so prices do not represent equilibrium prices. For this reason, empirical studies often drop limit days from their sample.

Furthermore, authors have cited the lack of analysis of various issues in commodity markets, relative to the importance of the exchange to world markets in industrially-used metals, as a reason for analysing LME markets. For example, the fact that there is little research in metals markets on time-varying risk premia is noted in MacDonald and Taylor [1989]. More generally, Hsieh and Kulatilaka [1982] note that metals have no discontinuities in production, and are not subject to such seasonal patterns commonly observed in agricultural commodities. However, Slade [1991] notes the prevalence of outliers in LME spot price series during the 1970s, where logarithmic returns are greater than 15%, and that extremes are less frequent in the 1980s.

Many authors consider issues related to the storable nature of metals, and effects of inventories on the market in the context of production involving significant lags (see Bresnahan and Suslow [1985], Brunetti and Gilbert [1995], Fama and French [1988], Gilbert [1995], Heaney [1998, 2002a, b], Ng and Pirrong [1994] and Slade [1991]). Otani [1983] defines a market where inventory stock commodities are traded and prices are determined so as to equate the demand for inventory stock with the existing amount available as an inventory stock market. Canarella and Pollard [1986] observe that the LME, an example of an inventory stock market, shows market behaviour that occurs when the desired level of production is not necessarily equal to the desired level of consumption. The distributions of prices or returns are frequently evaluated with respect to dependencies on the underlying physical availability of metals.

Table 5: Metals markets analysed

Metals Markets Modelled	Frequency
Aluminium	19
Aluminium Alloy	1
Copper	41
Gold	9
Lead	30
Nickel	13
Palladium	2
Platinum	6
Silver	13
Tin	24
Tungsten	1
Zinc	28
Total [1]	187

Note:
1. Studies consider between 1 and 9 metals markets.

The non-ferrous metals have several properties that are important in relation to modelling futures and forward prices. Metals are storable commodities and are not subject to seasonal production. Fama and French [1988] find metals spot and forward prices have a strong business cycle component. Precious metals, in particular, are considered a store of wealth, and demand increases in periods of instability, anticipated high inflation and currency depreciation. Some authors differentiate their treatment of the precious and the main industrially-used metals, considering that fundamentally different forces affect the markets. However, there is no recent comprehensive analysis of this issue.

Table 5 identifies the particular non-ferrous metals markets analysed in the 45 empirical papers, and the frequency with which they are modelled. Almost all papers, specifically 41 of the 45 considered, include copper in their analysis, either using data from the LME copper spot and/or futures markets, and/or the COMEX copper futures market (in one paper, copper data from the LME and the SHME were used). Copper is widely regarded as the most competitive non-ferrous metals market, with a low level of industry concentration in both production and consumption of the metal. The metal is widely used in the manufacturing of durable goods, is an important war material, and has some interesting substitution relationships with aluminium. Copper is also one of the most liquid of the LME markets, which is reflected in many results in the literature. While copper demand is highly volatile, smelting and refining industries have capacity constraints in the very short run, and the lag involved in increasing mine capacity is several years. Producer pricing for copper was abandoned around the mid-1970s. Prior to this, LME spot prices and producer list prices were observed to diverge substantially. Now the LME copper spot price can be considered as the world price for copper (see Bresnahan and Suslow [1985] and Gilbert [1996] for further discussion). In an early study of copper prices, Labys, Rees and Elliot [1971] failed to find any departure from a random walk using time series models.

Thirty of the 45 papers analyse the lead market, exclusively considering data from the LME. Heaney [1998, 2002a] considers lead market data only, noting that the deliverable spot asset underlying futures in copper, tin and zinc were all subject to changes over the period 1976 to 1995. In contrast, there are no fundamental changes in the definition of the deliverable asset for the lead contract during that time [Heaney, 1998]. The zinc market was considered in 28 papers, also exclusively using LME spot or futures data. Sephton and Cochrane [1991] note that, up to and including September 1985, zinc was classed as either zinc standard or zinc high grade, after which only the high grade metal price was reported. Questions have been raised in the literature over the robustness of models to structural changes in the data that occur due to changes in contract specifications, particularly in the definition of the commodity traded.

Data from the tin market were modelled in 24 papers, referring to either LME spot prices and/or futures contracts, or to spot prices determined on the KL Tin Exchange. The tin market is frequently excluded from empirical models as the LME tin contract was suspended between November 1985 and July 1989 due to the collapse of the International Tin Council, the body underwriting the tin contract. Shortly thereafter, several changes were made to LME futures contracts, importantly the introduction of a clearing house to underwrite futures contracts, in line with conventional practice for futures on other exchanges. Aluminium data are modelled in 19 papers, based predominantly on LME spot and futures prices. The current aluminium contract specification was introduced in July 1987. Producer prices and LME prices for aluminium are used in a model of the aluminium market in Gilbert [1995]. Producer prices for aluminium, copper, lead, nickel, silver zinc are included in models in Slade [1991]. Figuerola-Ferretti and Gilbert [2001] extend and re-analyse Slade's [1991] data using non-econometric methods to re-evaluate the hypothesis that the move to exchange-based trading has led to higher variability in prices relative to the producer pricing regime that was predominant until the mid-1980s. LME spot and/or futures prices for nickel were included in 13 papers. Nickel and aluminium were first traded on the exchange in 1979 and 1978, respectively.

Four precious metals appear in Table 5, namely gold, palladium, platinum and silver. Precious metals are only considered to the extent that they are jointly analysed with industrially-used non-ferrous metals. Silver, however, has substantial applications as an industrial metal. Moreover, a silver contract was launched on the LME in May 1999. Silver data appear in the models of 13 papers, gold in nine papers, and platinum and palladium in six and two, respectively. Aluminium alloy, traded on the LME since October 1992, appears in one paper, as does tungsten.

Table 6 indicates the market of focus in each empirical paper. Some authors treat contracts traded on the LME as forwards, while others maintain the contracts are futures. In this context, the market of focus is reported based on the interpretation in the paper under review. Issues relating to whether the LME contracts should be considered as futures or forward commodities are discussed below. However, for simplicity, the LME contracts will be referred to as futures, unless reference is being made to the interpretation of an author describing the contracts as forwards. Of the 45 empirical papers, 17 focus on futures markets for non-ferrous metals, 11 examine models relating to LME forward markets, and one paper considers LME forward markets and futures on other exchanges. Several papers consider spot or cash markets, including 11 papers that examine hypotheses focusing on spot markets for metals, three consider spot and LME forward markets, and two examine spot and futures markets.

The theoretical difference between the pricing of futures and forward contracts is illustrated in Cox et al. [1981] using an arbitrage-based model. According to their model, the essential difference between futures and forward prices is related to the difference between holding a long-term bond and rolling over a series of one-day bonds, respectively. Thus, futures and forward price models will not be equivalent unless the interest rate is non-stochastic.

Table 6: Type of market analysed

Market Type of Focus	Frequency
Forward[1]	11
Forward[1] and futures	1
Futures	17
Spot	11
Spot and Forward[1]	3
Spot and Futures	2
Total	45

Note:
1. LME futures markets are treated as forward markets by some authors.

Goss [1981, pp133-134] defines futures contracts as "... financial instruments dealing in commodities or other financial instruments for forward delivery or settlement, on standardised terms. They are traded on organised exchanges in which a clearing house interposes itself between buyer and seller and guarantees all transactions, so that the identity of the buyer or seller is a matter of indifference to the opposite party." LME contracts are frequently treated in the empirical literature in an identical manner to futures contracts. Moore and Cullen [1985] and Goss [1986] argue that, although the LME contracts are called forward contracts, they possess many of the properties of futures contracts. The LME contracts are standardised with regard to size, metal purity, and delivery location. There are arrangements for initial margins and margin calls, and there has been an organised clearing house (guaranteeing the contract) since 1987. In fact, Goss [1986] states that LME contracts are futures contracts in the sense in which the term is usually applied in the literature. Some points where the LME contracts differ significantly from futures contracts are that LME 3-month contracts are available on a daily basis, and not quoted for a limited set of dates per year, as is the case with futures contracts. For data prior to 1987, there was no clearing house and the LME was a principals market. Delivery frequently occurs under LME contracts, which is not the case for most futures contracts. In his exchange with Goss [1986] regarding the validity of previous analyses of LME efficiency, Gilbert [1986] highlights important differences between forward and futures contracts. Forward contracts nominate a day of delivery, while futures contracts state a delivery month. LME forward contracts were not routinely marked-to-market over the samples used in Goss [1981, 1983, 1986], so that the pricing of LME contracts will differ from futures market contracts with similar delivery dates.

The sampling frequency of the data used is provided in Table 7. Much of the econometric and statistical modelling has been conducted using monthly data. Twenty-four papers reported using monthly data, and one paper did not state the sampling frequency, but presumably used monthly data. Various means of collecting monthly data were used, including the first, middle and last trading day or the month. Monthly averages of weekly or daily data were frequently used for estimation. Daily data are available for spot and futures prices on the LME and other futures markets, and were used in 12 papers. Weekly data were used in two papers, while one of the two also employed monthly data. Quarterly data were used in six papers, one of which also used 4-monthly data. Finally, annual data were used in only one paper. No models were estimated using high frequency or intra-daily data of any kind.

Table 7: Sampling frequency of data

Sampling Frequency	Frequency
Intra-daily	0
Daily	12
Weekly	2
Monthly	24
Presumed Monthly	1
Quarterly	6
4-Monthly	1
Annual	1
Total[1,2]	47

Notes:
1. One instance of both weekly and monthly
2. One instance of both 4-monthly and quarterly

Almost all data are seasonally unadjusted. The data series may contain seasonal fluctuations of a deterministic or stochastic nature, but no investigation of seasonality or seasonal unit roots has been conducted. If a futures or forward price series contain a seasonal pattern, modelling of seasonality will more accurately reflect the nature of the variable in the model. Determining the existence of a seasonal unit root in a futures price series is important for understanding how a shock will affect the series. For example, a shock to a futures price series with a seasonal unit root will have a permanent effect on the seasonal pattern of the series, rather than a permanent effect on the level of the series, as is expected if a zero frequency unit root is present. Canarella and Pollard [1986] note moderate seasonal tendencies in the copper market, in which prices increase in the first quarter and fall in the second quarter of the year.

Empirical analysis of various hypotheses in futures markets frequently encounter sampling problems related to the frequency of observations relative to the contract length (for a comprehensive discussion, see Hsieh and Kulatilaka [1982]). Overlapping data are encountered when the sampling frequency is greater than the futures or forward contract period. Consider the forecast error u_{t+k} encountered when analysing the SEH:

$$u_{t,k} = s_{t+k} - f_{t+k}, \tag{1}$$

where s_{t+k} is the natural logarithm of the spot price at time t+k, and f_{t+k} is the natural logarithm of the futures price at time t for delivery at time t+k. When observations are non-overlapping, that is, when k = 1, u_{t+k} is serially uncorrelated, so that,

$$E\left(u_{t,k}, u_{t+h,k}\right) = 0. \tag{2}$$

However, when observations are overlapping, there will be no serial correlation only when $h < k$. When $h \geq k$, the forecast error will follow a moving average process as a result of new information that becomes available within the contract interval. The regression model will have serially correlated errors if data are overlapping, which can lead to biased and inconsistent estimates, and hence invalid inferences, if OLS estimation is used.

Four approaches to the overlapping data problem are evident in the literature: select a sampling frequency that will avoid overlapping data, use a modelling procedure that accommodates serial correlation, use averaged data, or ignore the problem altogether. The last two approaches generate inadequate models. Gilbert [1986] shows that using averaged data does not avoid serial correlation. In a cointegration framework, monthly averaged data are used by Chowdhury [1991] to examine market efficiency. Weekly data aggregated from daily observations are used by Kocagil [1997] in a linear regression model to test the hypothesis that futures speculation stabilises spot prices. Krehbiel and Adkins [1993] choose to take equally-spaced non-overlapping observations of spot and futures prices to avoid problems of overlapping data. However, this choice of data set places limits on the analysis of futures and forward markets. For example, for 3-month futures contracts, the sampling frequency must be no more than quarterly to have non-overlapping data, but much higher frequency data are often available. A low sampling frequency chosen to avoid overlapping observations involves a loss of information due to the exclusion of observations from the sample. Canarella and Pollard [1986], taking into account the consequences of overlapping data for estimation of models to test the SEH, use three sets of estimation methods and models. OLS is used to estimate a linear regression model with non-overlapping data, an ARMA model incorporates the moving average structure of induced by using overlapping data, and FIML is used to estimate a system of bivariate equations. Results from the three methods are qualitatively consistent.

A further problem in using averaged data has been identified by Gilbert [1986], who proposes that the EMH cannot be tested on monthly averaged data. Using monthly average data implies, for example, that the agent could buy at the average of the January forward price and sell at the average of the April spot price. However, the average of the any month's spot price is a construct for the analysis only, and not necessarily a price available to any of the transactors in the market at the time. Thus, it would not be surprising if unexploited arbitrage opportunities existed. Goss [1986] concedes this point, but claims it does not make any difference to his results.

Table 8 provides a breakdown of the sample sizes used. Note that numerous papers used more than one sample, particularly when estimating univariate or single market models for several metals. Under these circumstances, the various metals considered in a paper often have different sample sizes. For papers using two or three samples, each sample was reported in the table. Where more than three samples were used, only the largest and smallest of the samples is reported. The table does not include the use of sub-samples (refer to Appendix 1 for information on the use of sub-samples in particular papers). Several papers broke their samples into sub-samples to model structural change, or to test hypotheses relating to different time periods in the data. The largest number of sub-samples used in any one paper is 15 (see Bracker and Smith [1999]), but two or three sub-samples are frequently considered (for example, see Fama and French [1988], Chen et al. [1990], MacDonald and Taylor [1989], Sephton and Cochrane [1990], and Slade [1991]).

In eighteen instances, papers report sample sizes between 101 and 150 observations, which is the most frequently reported sample size range for the 45 papers. On eight occasions, data are used with sample sizes between 151 and 200. Small samples were commonly used in the empirical models considered (for example, ten samples between 50 and 100 observations are reported, and surprisingly, four reported samples contain fewer than 50 observations). Samples between 201 and 400 observations are reported on 10 occasions. Several papers use large samples, reporting between 401 and 1000 observations, between 1001 and 4000 observations, and over 4001 observations on four, seven and 4 occasions, respectively. Clearly, most of the reported samples contain fewer than 151 observations. The appropriate sample size depends on the specification and estimation method used for the model in question. Many of the samples reported may be quite adequate for linear regression models estimated using ordinary least squares. However, samples of fewer than 200 observations are insufficient for cointegration models estimated using the Johansen Maximum Likelihood method (for example, Agbeyegbe [1992], Fraser and MacDonald [1992], Heaney [1998], and Krehbiel and Adkins [1993]).

Table 8: Sample sizes used

Number of Observations	Frequency
Less than 50	4
50 to 100	10
101 to 150	18
151 to 200	8
201 to 250	3
251 to 300	4
301 to 400	3
401 to 500	2
501 to 1000	2
1001 to 1500	2
1501 to 2000	1
2001 to 3000	3
3001 to 4000	1
Over 4001	4
Total[1,2]	65

Notes:
1. Some studies used more than one sample;
2. In 13 papers where more than 3 samples are used, only the smallest and largest of the samples are reported.

Care must be taken when compiling a data set from sources commonly used for empirical research in commodity markets or finance, such as newspapers, industry publications and commercial data vendors. Sephton and Cochrane [1990, 1991] show that discrepancies in the reporting of forward and spot prices have yielded inaccurate empirical analyses. In addition, Sephton and Cochrane [1990] demonstrate several analyses of LME market efficiency have not matched forward and prompt prices correctly. They also claim that previous studies by MacDonald and Taylor [1988a, b, 1989] are flawed due to not matching futures prices with appropriate spot or prompt data. The prompt price may be defined as the spot price on the delivery date associated with the sampled forward price.

4.2 Dependent Variable

The dependent variables used in the models specified in the 45 papers are listed in Table 9. Most papers specified several models and, in this context, the frequency of use indicates the number of papers in which the dependent variable appeared at least once. Different measures of the spot price figured prominently among the dependent variables, with the spot price, the natural logarithm of the spot price, and spot returns each used in eight papers, and the first difference of the spot price used in two papers. Producer price variables are analogous to spot prices in that they are quoted only for immediate delivery, and not on a forward basis. Two papers used producer price variables, one of which was a dichotomous dependent variable (see Slade [1991]). Gilbert [1995] adjusts the spot price using the US Dollar exchange rate.

The futures or forward price was used in five papers, its (natural) logarithm appeared in two papers, and the first difference in the futures or forward price was used as a dependent variable in one paper. Kocagil [1997] de-trended the futures price to remove a deterministic trend. Futures or forward returns were specified as the dependent variable in six papers. The variance of prices and the variance or covariance of returns (for both spot and futures prices) were used in three and six papers, respectively.

Several of the dependent variables used are based on the difference between spot and futures prices. Let $f_{t,n}$ represent the futures (or forward) price in time t with maturity in period n, and s_t represent the spot price at time t. Six papers reported the forecast error as the dependent variable, and three papers used the logarithm of the forecast error. The forecast error, $u_{t,n}$, is defined as:

$$u_{t,n} = s_{t+n} - f_{t,n}. \qquad (3)$$

In two papers, the realised futures or forward return ($r_{t,n}$) is used as the dependent variable, which is defined as:

$$r_{t,n} = \frac{[f_{t,n} - s_{t+n}]}{f_{t,n}}. \tag{4}$$

The futures basis ($b_{t,n}$), which is defined as:

$$b_{t,n} = f_{t,n} - s_t, \tag{5}$$

is used as the dependent variable in two papers. The basis is expressed in logarithmic form, and is adjusted by the interest rate in Heaney [2002b].

Table 9: Dependent variables

Dependent Variable	Frequency
Spot price	8
Log of spot price[1]	8
First difference in spot price	2
Futures or forward price[2]	5
Log of futures or forward price	2
First difference in futures or forward price	1
Producer price[3]	2
Spot returns	8
Futures or forward returns	6
Realised futures or forward return	2
Variance of prices	3
Variance or covariance of returns	6
Log of futures or forward basis[4]	2
Forecast error	6
Log of forecast error	3
Production / consumption / stocks	5
Futures market volume variables	3
Interest rate variables	2
Excess gain variables	2
Exchange rate variables	1
No dependent variable indicated[5]	4
Total[6]	81

Notes:
1. The spot price is adjusted using exchange rates in Gilbert [1995];
2. Kocagil [1997] uses a detrended futures price;
3. One dichotomous dependent variable for producer pricing included;
4. Includes interest adjusted basis;
5. A depenent variable was not indicated for cointegration models estimated using the Johansen ML method (see Agbeyebe [1992], Franses and Kofman [1991], Heaney [1998] and Krehbiel and Adkins [1993]);
6. Some studies used more than one dependent variable.

Variables reflecting the state of the physical market in metals, including measures of metals production, metals consumption and stocks or inventories, appear as dependent variables on five occasions, typically in models relating prices to metals market fundamentals. On three occasions, variables are used that represent volumes on metals futures markets, such as transaction volume and open interest. Interest rate and excess gain variable are used on two occasions, and an exchange rate variable is used in one paper. In four cases, no dependent variable is listed in which cointegration models are estimated using the Johansen Maximum Likelihood method (see Agbeyegbe [1992], Franses and Kofman [1991], Heaney [1998], and Krehbiel and Adkins [1993]).

4.3 Explanatory Variables

The choice of explanatory variables for the empirical models in each of the 45 papers is provided in Table 10. Explanatory variables are listed according to their frequency of use, and are classified as either current or lagged. Spot and futures prices are frequently used in various measures as explanatory variables in either current period or lagged values. Spot prices are reported as explanatory variables in 15 instances, 10 of which are current and 5 lagged. An expected future spot price is included as a current explanatory variable. The natural logarithm of spot prices is used on five occasions, two of which are lagged, and one lagged value of the first difference in spot prices is also used. Producer prices are reported as explanatory variables once in current and once in lagged form. The futures (or forward) price is used on seven occasions as a current period explanatory variable, and on four occasions as a lagged explanatory variable. Logarithmic futures prices are used seven times in the current period, and in four occasions as lagged variables. A lagged first difference of the forward or futures price is used on one occasion.

Futures and spot returns appeared as explanatory variables four and six times as a lagged variable, respectively. The spot return is also used on one occasion as a current period variable. Frequently, the first difference of logarithmic spot or futures returns are stated as explanatory (or dependent) variables, which is equivalent to (and reported in the table as) a returns variable. The autocorrelation coefficient of spot returns is used on two occasions, once current and once lagged. In six instances, the lagged variance or conditional variance of returns is cited as an explanatory variable, typically to capture returns volatility. Market portfolio returns are used in two instances, one of which refers to returns on a metals market portfolio.

Several measures of the difference between spot and futures prices are used as explanatory variables in the models analysed. The futures basis, as defined in equation (5), is used on two occasions, and the logarithmic futures basis is reported once in current terms, and four times as a lagged value. Ng and Pirrong [1994] create a basis variable that is adjusted using a risk-free interest rate and storage cost. Three different measures of the forecast error are considered (see equation (3)). The forecast error is used in levels on one and five occasions for the current and lagged periods respectively. Logarithmic forecast errors are used once as current and lagged values, and the lagged first difference of the forecast error is also used on one occasion. The realised forward return, defined in equation (4), is used as a lagged explanatory variable in one paper.

Letting I_t represent the information set at time t, the forward premium ($p_{t,n}$) is defined as:

$$p_{t,n} = f_{t,n} - E\left[s_{t+n} \middle| I_t\right]. \tag{6}$$

Two variables that may explain the forward premium, namely the risk premium and the convenience yield, appear in four models and two models, respectively. Where agents are not risk neutral, the futures or forward premium may be interpreted as a risk premium. Alternatively, convenience yield may explain a forward premium. Interest rate variables are used on four occasions, and once as a lagged explanatory variable. In some cases, the interest rate represents a risk-free rate. The logarithm of changes in futures contract margins is used as an explanatory variable in one paper, as is an exchange rate variable as a lagged value.

In eight cases, variables representing the physical metals balance, non-ferrous metal production and consumption, and inventory levels are current period explanatory variables. On four occasions, authors include lagged values of these variables in their models. In most cases, these variables are used in models to test hypotheses on the relationships between price or returns variables and the demand and supply for physical metal. Macroeconomic and metals sector variables are used on eleven occasions, often to related metals prices with macroecononomic factors, and frequently reflect variables such as industrial production.

Dummy variables are used on seven occasions. For example, Slade [1991] uses dummy variables for the Hunt Brothers' silver market bubble, as well as to indicate the period of producer price dominance, and that in which exchange and producer pricing co-exist. A deterministic trend is employed in one paper. Two papers do not state any explanatory variables. Labys et al. [1998] use a structural time series model, and McKenzie et al. [2001] specify a model in which the dependent variable is an error process.

Table 10: Choice of explanatory variable

Type of Explanatory Variable	Frequency	
	Current	Lagged
Spot price [1, 5]	10	5
Log of spot price [2]	3	2
First difference in spot price	0	1
Futures or forward price [5]	7	4
Log of futures or forward price	6	2
Log of futures to forward price ratio	1	0
First difference in futures or forward price	0	1
Producer price [5]	1	1
Spot returns	1	6
Futures or forward returns	0	4
Realised futures or forward return	0	1
Risk premium [5]	4	0
Convenience yield [5]	1	1
Variance or conditional variance of returns [5]	1	6
Futures or forward basis	2	0
Log of futures or forward basis [3]	1	4
Forecast error	1	5
Log of forecast error	1	1
First difference of forecast error	0	1
Production / consumption / stocks [5]	8	4
Returns on (metals) market portfolio [5]	2	0
Macroeconomic and metals sector variables [5]	11	0
Log of change in futures contract margins	1	0
(Risk-free) interest rate variables [5]	4	1
Exchange rate variables	0	1
Autocorrelation coefficient of spot returns	1	1
Producer price residual	1	0
Dummy variables	7	N/A
Deterministic trend	1	N/A
No explanatory variables indicated [4]	2	N/A

Notes:
1. One instance each of a deflated spot price and an expected spot price;
2. In Gilbert [1995] the spot price is adjusted using an exchange rate index;
3. The basis is adjusted for interest rate and storage in Ng and Pirrong [1994];
4. Labys et al. [1998] and McKenzie et al. [2001] do not indicate explanatory variables due to the use of structural time series models with non-stochastic regressors and a naive model, respectively;
5. Includes proxy variables and/or generated regressors.

Theoretical models for the pricing of futures and forward contracts, as is the case with many economic and financial relationships, are often specified in terms of variables that are not directly observable. In order to specify empirically estimable versions of these models, proxy variables or generated regressors are used for the unobservable variables. In Table 10, ten of the reported types of explanatory variables include proxy variables and/or generated regressors. Table 11 provides more detail on these variables, and indicates the frequency of use of proxy variables and generated regressors separately. The econometric implications of proxy variables and generated regressors are summarised briefly below.

A proxy variable in a regression necessarily implies the presence of measurement error as the correct variable is not used. Measurement error in an explanatory variable causes serious problems for OLS estimation. Including incorrectly measured variables in the linear regression model violates the assumption that the explanatory variables are exogenous, that is, the process generating the explanatory variables is uncorrelated with the process generating the error term. When one or more explanatory variables are not exogenous, OLS yields biased and inconsistent estimators. Garber and Kepper [1980] conduct a comprehensive analysis of the errors-in-variables problem, which provides two cases that are instructive for dealing with incorrectly measured variables.

Where only one proxy variable is used in a linear regression model, the OLS estimator of the proxy variable coefficient will be attenuated (that is, the absolute bias in the estimate is confined between a minimum of zero and a maximum value of the true (absolute) value of the parameter). Reducing the measurement error in the proxy variable reduces the absolute bias in the estimates of the coefficients for both the proxy variable and for the other (correctly measured) variables in the model. Furthermore, including the proxy variable in the model is better than excluding it, since the latter will maximise the absolute bias of the proxy variable coefficient estimate (that is, invoke omitted variable bias).

If two or more proxies are included in a linear regression model with correctly measured variables, one or both of the proxy variable estimates will be attenuated. Improving the measurement error in one proxy variable reduces the absolute bias in its estimate if and only if its coefficient is attenuated. Thus, reducing the measurement error in a proxy may not be beneficial, and the bias in the estimate of the coefficient of the correctly measured variable is not necessarily reduced. Omitting one or more proxy variables from the model will not necessarily increase the absolute bias in the estimated coefficient of the correctly measured variable. Excluding one or more proxies may be preferable to their inclusion. One solution is to use instrumental variable estimation, which yields consistent but inefficient estimators. An instrument correlated with the proxy variable and uncorrelated with the error term is required. However, a suitable instrument can frequently be difficult to obtain.

Unobservable expectational or surprise variables in economic models have led to the use of generated regressors in empirical work. Pagan [1984] identifies three theoretical situations in which generated regressors are produced. A predictor generated regressor occurs where the regressors are constructed as predictions values from another equation. Generated regressors may also be created using the residuals in levels from another regression, and a residual generated regressor in variance occurs where the variance, rather than the level, of the unobservable variable is of interest. Four methods of estimating models with unobservable variables, which must be generated, are defined in Oxley and McAleer [1993] (see also McAleer and McKenzie [1992], and McKenzie and McAleer [1994, 1997]). Estimating models with generated regressors using a systems method avoids the generated regressor problem. Maximum likelihood provides an efficient method of estimating models with generated regressors, but it is seldom used. Two-step OLS estimators suffer from downward bias in t-ratios and upward bias in standard errors, as the method erroneously assumes that the generated regressor is a non-stochastic regressor that may simply replace the unobserved variable in the structural equation. Although easy to implement, Pagan [1984] shows that the errors produced by this method are non-spherical, with the consequent invalid inferences. However, valid inferences may be obtained using a 2-step OLS procedure involving a correction to the conventionally programmed OLS standard errors.

As shown in Table 11, several proxy variables and generated regressors are used in the empirical literature. While some papers acknowledged the use of proxies for unobservable variables, it appears none explicitly considered the econometric implications of proxy variables or generated regressors for estimation and inference. Proxy variables are used for the risk-free interest rate and the expected risk-free interest rate on four occasions, the return on a market portfolio or a metals market portfolio on two occasions, convenience yield on two occasions, and the expected spot price of a metal and the metals price trend in one paper each. Chen et al. [1990] perform market model regressions that require a measure of excess return on the market for which a combination of two market indices are used. A frequently used proxy for the expected future spot price is the logarithm of realised spot prices.

Metals market fundamental characteristics, including the state of metals spot and futures markets, were represented by proxy variables on six occasions. For example, the state of the copper market was approximated by the lagged rate of change of the difference between the spot price and an interest-adjusted future spot price in Bresnahan and Suslow [1985].

Inventory proxies were used on two occasions, in levels and in first differences. An inventory to consumption ratio was used as an explanatory variable, where the stock level is a proxy and the consumption trend is a fitted value generated regressor (see Brunetti and Gilbert [1995]). Fama and French [1988] observe that inventory data are always problematic when analysing implications of the theory of storage using a model that includes inventory variables. The accuracy of inventory data is questionable because of substantial hidden stocks, and the difficulty in defining what stocks constitute inventory compared with working stocks, either in transit or used in production.

Other generated regressors included the futures risk premium in four instances, the detrended futures price, and the metals production shock and producer transactions price, in one case each. Unconditional variance variables for spot and futures prices, or their returns, were generated in five instances. Conditional variances for metals spot and futures returns were used in two cases. For example, Hall [1991] uses the conditional variance of the forecast error as a generated regressor to represent the risk premium.

Table 11: Use of proxy variables and generated regressors

Type of Variable	Frequency	
	Proxy Variable	Generated Regressor
(Expected) Risk-free interest rate	4	0
Return on (metals) market portfolio	2	0
Inventory or stocks[1]	2	1
Convenience yield	2	0
Risk premium	0	4
Detrended futures price	0	1
Expected spot price	1	0
Production shock	0	1
Producer transactions price	0	1
Producer price residual	0	1
Metals price trend	1	0
Metals market fundamental characteristics	6	0
Unconditional variance of prices or returns	0	5
Conditional variance of returns	0	2

Note:
1. Includes stock variables in levels, first difference, and the ratio of stocks to consumption trend (fitted value).

4.4 Model Specification

The types of empirical models specified to test the economic hypotheses of interest are listed in Table 12. Each type of model is presented in terms of the number of papers that used the model, and the total number of models specified in those papers. Fifteen classes of specification are considered. The linear regression model is the most frequently used specification, appearing in 24 of the 45 papers. In these 24 papers, a total of 655 linear regression models were specified.

Models of the Autoregressive Conditional Heteroskedasticity (ARCH) and Generalised ARCH (GARCH) family were also popular in terms of the number of papers in which they were used and the total number of models specified. Table 12 deconstructs this class of models into four broad categories, namely symmetric ARCH or GARCH, asymmetric ARCH or GARCH, ARCH in Mean (ARCH-M) and GARCH in Mean (GARCH-M), and Fractionally Integrated GARCH (FIGARCH). Symmetric ARCH or GARCH models appeared in nine papers, and a total of 115 models were specified, while asymmetric models appeared in 11 papers containing 139 models. In general, there appears little support in the literature for asymmetric GARCH modelling of industrial metals returns data, as the returns series are close to symmetric. Six FIGARCH model are specified in one paper for non-ferrous metals spot price data (see Brunetti and Gilbert [1997]).

ARCH-M and GARCH-M models are specified in five papers, for a total of 21 models. GARCH models are frequently applied to modelling the risk premia in financial data. Hall [1991] compares GARCH-M, SGARCH-M, and SGARCH-M with MA(2) errors in modelling the risk premium in metals forward prices on the LME. GARCH-M models provide a framework in which a time-varying risk premium can be estimated and tested. The standard ARCH or GARCH model is assumed to be non-stochastic. Hall [1991] argues that this is an extreme assumption in the context of empirical applications, and suggests the SGARCH-M model may be more appropriate. The stochastic generalisation of the GARCH-M model specifies the GARCH process as stochastic. Hall and Taylor [1989] used a model from the dynamic latent variable class of specifications, called the Dynamic Multiple-Indicator Multiple-Cause (DYMIMIC), for which four empirical specifications were used.

Cointegration models were also reasonably popular, which is to be expected given the presumption that commodity prices are non-stationary. Seven papers specified 57 bivariate cointegration models, five papers specified 10 multivariate cointegration models, and one paper contained five error correction models. These models were estimated by a variety of techniques, to be discussed in the next section.

Seven papers used 23 Autoregressive Moving Average (ARMA) or Autoregressive Integrated Moving Average (ARIMA) models, and two papers specified 14 Vector Autoregression (VAR) models. The nonlinear regression model is only used in one paper, but 12 models are specified in that single paper. Nine linear or nonlinear systems of equations are estimated in four papers. One paper specifies a Tobit model, and one specifies a Probit model. Structural Time Series (STS) models are used in two papers, in which the authors estimate 23 models.

Table 12: Model specification

Model Specification	Number of Papers	Number of Models
Linear Regression	24	655
Nonlinear Regression	1	12
Bivariate Cointegration	7	57
Multivariate Cointegration	5	10
Error Correction	1	5
ARMA or ARIMA	7	23
Vector Autoregression	2	14
Linear or Nonlinear System of Equations	4	9
Symmetric ARCH or GARCH	9	115
Asymmetric ARCH or GARCH	11	139
ARCH in Mean or GARCH in Mean	5	21
Fractionally Integrated GARCH	1	6
DYMIMIC	1	4
Structural Time Series	2	23
Tobit or Probit	2	2
Total[1]	82	

Note:
1. Some papers specified more than one model.

4.5 Methods of Estimation

Table 13 shows the methods of estimation used in the 45 papers. It should be noted that numerous papers used more than one method of estimation. OLS was cited as the most common method of estimation. In eleven cases, authors reported OLS with no adjustment to the covariance matrix. Four times, the method of estimation was not stated, when it was apparent that OLS was used. In ten cases, where the residual was expected to heteroskedastic or serially correlated, OLS was used with a modified covariance matrix. Standard errors were sometimes used, as suggested by White [1982], Newey and West [1987], Hansen and Hoderick [1980] and Hansen [1982]. The Cochrane-Orcutt transformation or iterative method was used in five cases where first-order autoregressive errors were expected.

Table 13: Methods of estimation

Methods of Estimation	Frequency
Ordinary Least Squares (OLS)	11
Presumably OLS	4
OLS with modified covariance matrix	10
Cochrane-Orcutt	5
GLS with modified covariance matrix	2
Feasible Generalised Least Squares	1
2-Stage Least Squares	1
3-Stage Least Squares	1
IV with modified covarince matrix	1
Generalised Instrumental Variable Estimator	3
Heckman 2-Step Estimator	1
Nonlinear Least Squares	1
Presumably Nonlinear Least Squares	1
Johansen ML Method	6
Engle-Granger Method	6
Maximum Likelihood (ML)	7
Presumably ML	5
Full Information ML	1
Phillips-Hansen Fully Modified OLS	1
Kalman Filter	2
Generalised Method of Moments	1
Total[1]	71

Note:
1. Some studies used more than one method of estimation.

The estimation methods applied to the cointegration models discussed in Section 4.4 were the Johansen Maximum Likelihood method in six instances, the Engle-Granger method in 6 instances, and the Phillips-Hansen Fully Modified Ordinary Least Squares method in one case. Moore and Cullen [1995] suggest the Johansen estimation procedure is not appropriate where there is overlapping data. Overlapping spot and futures or forward price data generate moving average errors. When the errors in the cointegrating relationship are characterised by a non-invertible moving average process, the Granger Representation Theorem does not hold. The Phillips-Hansen fully modified OLS estimation procedure can deal with a wider class of serial correlation, which is an advantage for modelling with overlapping data, but it allows only one cointegrating vector. However, as Moore and Cullen [1995] deal only with bivariate models, this does not present a limitation. Estimates of parameters and standard errors are asymptotically equivalent to those produced by maximum likelihood estimation.

Two stage least squares was used on 1 occasion to estimate a model using instruments. The method of estimation was not stated in five instances, for ARMA, linear regression with MA(2) errors, and GARCH models, although the presumption is that maximum likelihood was used. In seven cases, Maximum Likelihood estimation (including Quasi-Maximum Likelihood) is used. Maximum likelihood estimators of the SGARCH-M model are inconsistent, so Hall [1991] uses quasi-maximum likelihood estimation, which provides consistent but not fully efficient estimates.

4.6 Descriptive Statistics

Descriptive statistics, as reported by the authors, are summarised in Table 14. The table indicates the frequency with which various types of descriptive statistics and discrimination criteria are reported. Discrimination criteria refer to goodness of fit measures, such as the coefficient of multiple determination and information criteria. These descriptive statistics assess how well different models fit the data, with some adjustment for parsimony. The coefficient of multiple determination (R^2), including the adjusted R^2 and quasi-R^2 measures, was the most frequently reported descriptive statistic, and was often used by the authors to indicate both the statistical adequacy of a model and to discriminate between models. Numerous authors reported the corrected R^2 when discriminating between models with different numbers of explanatory variables, and the quasi-R^2 was reported for one model only. In some cases, the R^2 was the only statistic of any kind reported for a regression model. The R^2 was most used in the evaluation of competing nested or non-nested models, or predictive ability. Goodness of fit measures, such as the coefficient of multiple determination and information (or discrimination) criteria, assess the goodness of fit of different models, with appropriate adjustments for parsimony. The philosophy in using discrimination criteria to choose between models is that the best predicting model is the closest approximation to the “true” specification. Each model is evaluated only in terms of its own performance, which is the principal disadvantage of discriminating between models on the basis of goodness of fit measures. One model will always be chosen, regardless of whether it can predict the consequences of non-nested alternatives.

Standard errors were reported in 23 papers, often based on a modified covariance matrix, when the residuals were expected to be heteroskedastic or serially correlated. Methods of calculating standard errors and t-statistics include those of White [1982], Newey and West [1987], Hansen and Hodrick [1980], and Hansen [1982]. Bollerslev and Wooldridge [1992] robust t-ratios were sometimes used for GARCH models. A number of measures of the regression error were used to evaluate and compare competing regressor models. In 8 papers, the standard error of the regression was reported as a measure of the statistical adequacy of an estimated regression model. In 11 cases, the in-sample or out-of-sample forecasting ability of (competing) models was compared on the basis of error measures, including root mean squared error, mean absolute error and mean squared error.

Table 14: Reported descriptive statistics

Regression Descriptive Statistics	Reporting Incidence
R^2 (including corrected and quasi-)	31
Standard Error[1]	23
Standard Error of Equation	8
Log-Likelihood	6
Information Criteria	8
Regularity Conditions	1
Correlogram	1
Skewness and Kurtosis of Standardised Residuals	1
Forecast Error Measures	11
Forecast Error Variance	1
No Descriptive Statistics Reported	5
Total[2]	96

Notes:
1. Includes standard errors of the following forms: White, Newey-West, Hansen and Hoderick, Hansen, Bollerslev and Wooldridge, asymptotic and approximate;
2. Some papers reported more than one type of descriptive statistic.

Information criteria are used in 8 cases, including Schwarz’s Bayesian Criterion (SBC) and the Akaike Information Criterion (AIC). SBC is used by Krehbiel and Adkins (1993) to select the optimal lag length for the Johansen test for the number of cointegrating relationships. The value of the log-likelihood function is quoted in six cases. Regularity conditions are important for the interpretation of standard errors and test statistics for GARCH models. The second moment condition is reported by McKenzie et al. [2001] only for the Asymmetric Power GARCH model. However, the regularity condition reported pertains to the GARCH(p,q) model, and not the model actually used! Five papers reported no descriptive statistics whatsoever.

4.7 Diagnostic Testing

Diagnostic testing plays a critical role in assessing the adequacy of empirical economic models. Given the theoretical model, an intermediate specification is needed to obtain an equation suitable for estimation and testing. This involves making numerous auxiliary assumptions about functional forms, relevant variables, and adequate approximations of the "true" variables of interest, and stability of the model. Pesaran and Smith [1985, p.138] note that: "A consequence of this procedure is that one cannot know whether the results of the statistical analyses reflect inferentially on the economic theory or on the auxiliary assumptions." Through diagnostic testing of the auxiliary assumptions, tests of a number of specific null hypotheses, listed below, assess the statistical adequacy of a model. McAleer [1994] considers a linear regression model, and in the context of OLS, lists the following assumptions that require diagnostic testing: i) correct functional form, ii) no heteroskedasticity, iii) no serial correlation, iv) exogeneity of the explanatory variables, v) normality of the errors, vi) parameter consistency, vii) non-nested models (the model is adequate in the presence of non-nested alternative models), and viii) robustness to departures from the auxiliary assumptions.

While having a specific null hypothesis, diagnostic tests generally do not have specific alternatives. For example, tests rejecting the null hypothesis of correct functional form may be due to a number of possibilities, including incorrect functional form, omitted variables, serial correlation, structural change, heteroskedasticity, or sample selection bias. There are important complications for estimation and inference when any of the assumptions is not satisfied [McAleer, 1994].

Table 15 summarises the auxiliary assumptions for which diagnostic tests were reported in the 45 empirical papers. Serial correlation was the most frequently tested auxiliary assumption. The Durbin-Watson (DW) test (or the Cointegrating Regression Durbin-Watson test) was reported in 16 papers. However, the DW test is somewhat limited, typically detecting only first-order serial correlation.

Chang et al. [1990] use OLS to estimate models for copper, platinum and silver, over the full data set and two sub-samples. The problem of first-order serial correlation for 10 silver contracts models is indicated by the Durbin-Watson statistic. While the authors re-estimate these models using the Cochrane-Orcutt iterative process, this results in estimates that are not qualitatively different from the OLS estimates. For this reason, the authors present the original OLS estimates, ignoring the Durbin-Watson statistic. If higher-order serial correlation is present in a model, the Durbin-Watson test will identify the first-order component only. Therefore, the problem with the models in Chang et al. [1990] was possibly higher-order serial correlation. Hence, it might be expected that the Cochrane-Orcutt method will not produce substantially different results.

Four papers report the Box-Pierce Q-statistic. This test is used in Sephton and Cochrane [1990] to test for autocorrelation. Their model of market efficiency precludes third- or higher-order autocorrelation in the forecast error series for non-ferrous metal forward prices on the LME. A potential problem with the Box-Pierce Q-statistic is that it has very poor power in small- to medium-sized samples. The Ljung-Box test also has poor power, but is a superior test in smaller samples, and is used to test for serial correlation in 6 papers. Twelve papers report other tests for serial correlation, such as the Wallis test in Goss [1981, 1983].

The second most frequently tested auxiliary assumption is that of stationarity in the data. Twenty papers report tests for unit roots, most frequently citing the Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) tests, while some also implement the Phillips-Perron test. The Phillips-Peron test for a unit root is able to deal with a wide range of serial correlation and heteroskedasticity in the residuals. However, more than half of the papers analysed do not consider testing for unit roots. In general, spot, forward and futures prices for metals are found to be non-stationary and integrated of order one, particularly for higher frequency data. For some sampling frequencies, and sample periods, unit roots are not detected. Occasionally, tests for non-stationarity are misused, and authors do not properly establish the order of integration in their data. In other cases, non-stationarity is ignored. For example, Sephton and Cochrane [1990] use cointegration, but Sephton and Cochrane [1991] ignore non-stationarity in essentially the same data. Given the evidence for unit roots in metals spot, forward and futures price series, the relationships described in some of the 25 papers that do not test for non-stationarity may be spurious.

Tests for heteroskedasticity and normality in the residuals are each reported in five papers. The tests used for heteroskedasticity are White’s test, a test of unequal variances, and LM and F-tests. Bera-Jarque is the most frequently used test for normality. As can be seen from Table 15, most of the 45 studies models reported do not report diagnostic tests for many of the auxiliary assumptions. In nine papers, there were no diagnostic tests reported at all. Tests for ARCH, linear trend, misspecification, normality, exogeneity, multicollinearity, linear trend, instrument validity, and intercept in a cointegrating vector, are reported in one paper each. One paper also conducts tests for Granger Causality.

Table 15: Reported diagnostic tests

Diagnostic Tests	Reporting Incidence
No Diagnostics Reported	9
Serial Correlation: Durbin-Watson or CRDW	16
Serial Correlation: Box-Pierce Q	4
Serial Correlation: Ljung-Box	6
Serial Correlation: Other Tests	12
Unit Root	20
Structural Change	5
Parameter Stability	4
Linear Trend	1
Mis-specification	1
Normality	5
Heteroscedasticity	5
ARCH	1
Causality	1
Exogeneity	1
Multicollinearity	1
Presumably predictive failure	1
Instrument Validity	1
Intercept in a Cointegrating Vector	1
Total[1]	95

Note:
1. Some studies used more than one type of diagnostic test.

No tests of stochastic seasonality are conducted, and no models include dummy variables for constant seasonal fluctuations. The potential problem of stochastic seasonality is ignored. While metals are not typically subjected to the seasonal patterns expected in agricultural commodities, the cyclical behaviour of metals prices warrants closer inspection of seasonality in the data.

Analysis of expected structural breaks is avoided by most authors. Few papers use methods to test for permanent or transitory structural change in the series, or to model structural change in futures, and particularly in the forward markets of the LME. It can be argued that if the behaviour of a variable during a period of structural change is not understood, it is not possible to understand the variable’s behaviour before and after the structural shift. Thus, it is important to model explicitly structural breaks to determine whether the change is permanent or transitory, whether it is a one-off jump in the series, and whether the order of integration has been changed. Tests for structural change are conducted in five papers. Four papers report parameter stability tests. CUSUM of squares tests presented in Sephton and Cochrane [1991] provide evidence of structural change in the LME nickel, copper, aluminium, lead, and zinc markets, while a structural change could not be detected in the tin market. Heaney [2002a] uses the CUSUM and CUSUMSQ tests.

During 1979-80, the Hunt Brothers' attempted to manipulate prices in the silver market. Krehbiel and Adkins [1993] test whether their results for the silver market are sensitive to suspected structural change caused by this episode. Using the Perron test, the presence of a unit root is rejected for silver futures prices, so that the cointegration model may not be appropriate for analysis of the silver market.

The tin market collapsed in late-1985, and tin contracts were suspended on the LME from October 1985 to June 1989. An inter-governmental agency, the International Tin Council, dominated the tin market and traded in tin to stabilise prices. In trying to support the tin price, the International Tin Council's reserves were exhausted and, on 24 October 1985, the Council ceased operations, with debts of over £900 million. Sephton and Cochrane [1990] delete tin and zinc from their third model since "consistent' series are unavailable for both metals on the LME after 1985. Moore and Cullen [1995] perform unit root tests on a sample from the tin market taken after the resumption of trading, and find that both the forward and spot prices are stationary. They attribute this to the structural change caused by the default of the International Tin Council. The nature of the market had changed, trading was thin, and the market was in contango for most of the period since trading resumed.

Until the end of 1985, zinc contracts on the LME were classified as either zinc standard or zinc high grade. Subsequently, only high grade contracts were traded. Several empirical models, including those of Agbeyegbe [1992] and Hall [1991], use data for the zinc market covering the period of the change without testing for structural change, despite using empirical methods where the results are sensitive to structural change. Sephton and Cochrane [1990] are aware of the change in contracts, and specify their sample up to September 1985, thereby avoiding the period of change. Some researchers appear to be unaware of changes in the contract specification for some data used in their empirical analysis, while others approach this problem by selecting a data set pertaining to one specification of the contract only.

4.8 Nested and Non-Nested Tests

Non-nested (or separate) tests are specification tests with specific alternatives. Hence, the purpose of non-nested tests is to achieve high power against the specified alternative. In the assessment of specific non-nested alternatives, an appropriate philosophy is to test whether the null model can predict the performance of an alternative model "significantly well". The essential difference between discriminating on the basis of descriptive statistics, or information criteria, and on the basis of testing, is that the latter enables classical inferential procedures to be applied.

Table 16 shows the number of papers that reported the use of nested and non-nested tests. No non-nested tests between competing separate alternative models are conducted. Six papers report nested tests. McKenzie et al. [2001] use likelihood ratio (LR) tests among 14 nested GARCH models. McMillan and Speight [2001] also use the LR test between nested GARCH models. Canarella and Pollard [1986], Gilbert [1995], and MacDonald and Taylor [1988a] test between nested models using Wald and LR tests. Canarella and Pollard [1986] also use a nested F-test, while Ben Nowman and Wang test between nested models using the J-statistic. Hypothesis test are also included in Table 16, and indicate that 27 papers reported hypothesis tests on estimates, mostly using t-tests or F-tests.

Table 16: Reported nested and non-nested tests

Nested, Non-Nested and Hypothesis Tests	Reporting Incidence
Nested Tests	6
Non-Nested Tests	0
Hypothesis Tests	27
Total[1]	33

Note:
1. Some studies conducted both nested tests and hypothesis tests.

5 Conclusion

Published empirical research has been evaluated in the light of the type of contract examined, frequency of data used, choice of both dependent and explanatory variables, use of proxy variables, type of model chosen, economic hypotheses tested, methods of estimation and calculation of standard errors for inference, reported descriptive statistics, use of diagnostic tests of auxiliary assumptions, use of nested and non-nested tests, use of information criteria, and empirical implications for non-ferrous metals.

Several conflicting empirical results with regard to futures and forward market models are apparent in the literature. Important empirical issues such as overlapping data, structural change, measurement error, correct use of proxy variables, and non-stationarity of spot futures and forward price series, have frequently been ignored. Diagnostic testing of the auxiliary assumptions is seldom undertaken on a systematic basis, leaving open to question the statistical adequacy of the models presented. Most research does not consider nested or non-nested testing of competing models.

Seen in this light, while a substantial amount of empirical analysis of non-ferrous metals has been conducted over the past two decades, the empirical conclusions of the research that has been reviewed should be interpreted with some caution.

Acknowledgements

The first author would like to acknowledge the financial support of the Department of Economics, University of Western Australia, and the Australian Research Council. The second author is most grateful for the financial support of the Australian Research Council.

References

Agbeyegbe, T., Common stochastic trends: Evidence from the London Metal Exchange, *Bulletin of Economic Research*, 44, 141-151, 1992.

Ben Nowman, K. and H. Wang, Modelling commodity prices using continuous time models, *Applied Economics Letters*, 8, 341-345, 2001.

Bilson, J.F.O., The ‘speculative efficiency’ hypothesis, *Journal of Business*, 54, 435-451, 1981.

Bollerslev, T. and J.M. Wooldridge, Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances, *Econometric Reviews*, 11, 143-172, 1992.

Bracker, K. and K. L. Smith, Detecting and modelling changing volatility in the copper futures market, *Journal of Futures Markets*, 19, 79-100, 1999.

Brenner, R.J. and K.F. Kroner, Arbitrage, cointegration and testing the unbiasedness hypothesis in financial markets, *Journal of Financial and Quantitative Analysis*, 30, 23-42, 1995.

Bresnahan, T.F. and V.Y. Suslow, Inventories as an asset: the volatility of copper prices, *International Economic Review*, 26, 409-424, 1985.

Brunetti, C. and C.L. Gilbert, Metals price volatility, 1972-95, *Resources Policy*, 21, 237-254, 1995.

Canarella, G. and S.K. Pollard, The 'efficiency' of the London Metal Exchange: a test with overlapping and non-overlapping data, *Journal of Banking and Finance*, 10, 575-593, 1986.

Chang, E.C., C. Chen and S.N. Chen, Risk and return in copper, platinum, and silver futures, *Journal of Futures Markets*, 10, 29-39, 1990.

Chowdhury, A.R., Futures market efficiency: evidence from cointegration tests, *Journal of Futures Markets*, 11, 577-589, 1991.

Cox, J., J. Ingersoll, and S. Ross, The relation between forward prices and futures prices, *Journal of Financial Economics*, 9, 321-346, 1981.

Fama, E., *Foundations of Finance*, Basic Books, New York, 1976.

Fama, E., Forward and spot exchange rates, *Journal of Monetary Economics*, 14, 319-338, 1984a.

Fama, E., The information in the term structure, *Journal of Financial Economics*, 13, 509-528, 1984b.

Fama E.F. and K.R. French, Business cycles and the behaviour of metals prices, *Journal of Finance*, 43, 1075-1093, 1988.

Figuerola-Ferretti, I. and C.L. Gilbert, Price variability and marketing method in non-ferrous metals: Slade's analysis revisited, *Resources Policy*, 27, 169-177, 2001.

Franses, P.H. and P. Kofman, An empirical test for parities between metal prices at the LME, *Journal of Futures Markets*, 11, 729-736, 1991.

Fraser, P. and R. MacDonald, Spot and forward metals prices: efficiency and time series behaviour, *Review of Futures Markets*, 11, 24-34, 1992.

French, K.R., A comparison of futures and forward prices, *Journal of Financial Economics*, 12, 311-342, 1983.

Garber, S and S. Kepper, Extending the classical normal errors-in-variable model, *Econometrica*, 48, 1541-1546, 1980.

Gilbert, C.L., Testing the efficient markets hypothesis on averaged data, *Applied Economics*, 18, 1149-1166, 1986.

Gilbert, C.L., Futures trading, storage and price stabilisation, *Review of Futures Markets*, 8, 152-176, 1989.

Gilbert, C.L., Modelling market fundamentals: a model of the aluminium market, *Journal of Applied Econometrics*, 10, 385-410, 1995.

Gilbert, C.L., Manipulation of metals futures: lessons learned from Sumitomo, *Discussion Paper No. 1537*, Centre for Economic Policy Research, 1996.

Goss, B.A., The forward pricing function of the London metal exchange, *Applied Economics*, 13, 133-150, 1981.

Goss, B.A., The semi-strong form efficiency of the London Metal Exchange, *Applied Economics,* 15, 681-698, 1983.

Goss, B. A., Rejection of unbiasedness is not rejection of market efficiency: Reply to Gilbert, *Applied Economics*, 18, 1167-1178, 1986.

Granger, C.W.J., Developments in the study of cointegrated economic variables, *Oxford Bulletin of Economics and Statistics*, 48, 213-228, 1986.

Gross, M., A semi-strong test of the efficiency of the aluminium and copper markets at the LME, *Journal of Futures Markets*, 8, 67-77, 1988.

Hakkio, C.S. and M. Rush, Market efficiency and cointegration: an application to the Sterling and Deutschemark exchange markets, *Journal of International Money and Finance*, 8, 75-88, 1989.

Hall, S. G., An application of the stochastic GARCH-in-mean model to risk premia in the London Metal Exchange, *The Manchester School*, 59, 57-71, 1991.

Hall, S.G. and M.P. Taylor, Modelling risk premia in commodity forward prices: some evidence from the London Metal Exchange, *Review of Futures Markets*, 8, 200-217, 1989.

Hansen, L.P., Large sample properties of generalised method of moments estimators, *Econometrica*, 50, 109-1054, 1982.

Hansen, L.P., and R.J. Hodrick, Forward exchange rates as optimal predictors of future spot rates: an econometric analysis, *Journal of Political Economy*, 88, 829-853, 1980.

Hardouvelis G.A. and D. Kim, Margin requirements, price fluctuations, and market participation in metal futures, *Journal of Money, Credit, and Banking*, 27, 659-671, 1995.

Heaney, R., A test of the cost-of carry relationship using the London Metal Exchange lead contract, *Journal of Futures Markets*, 18, 177-200, 1998.

Heaney, R., Does knowledge of the cost of carry model improve commodity futures price forecasting ability? A case study using the London Metal Exchange lead contract, *International Journal of Forecasting*, 18, 45-65, 2002a.

Heaney, R., Approximation for convenience yield in commodity futures pricing, *Journal of Futures Markets*, 22, 1005-1017, 2002b.

Hill, S. R., N. H. Moore, and S. W. Pruitt, Cold fusion - hot metal: An analysis of the metals futures market reactions to the cold fusion announcement. *Journal of Futures Markets*, 11, 385-397, 1991.

Hsieh, D.A. and N. Kulatilaka, Rational expectations and risk premia in forward markets: primary metals at the London Metal Exchange, *Journal of Finance*, 37, 1199-1207, 1982.

Hussey, R. and J. Quiroz, Copper, futures and Codelco, *Revista de Analisis Economico*, 12, 63-84, 1997.

Kocagil, A. E., Does futures speculation stabilize spot prices? Evidence from metals markets, *Applied Financial Economics*, 7, 115-125, 1997.

Krehbiel, T. and L.C. Adkins, Cointegration tests of the unbiased expectations hypothesis in metals markets, *Journal of Futures Markets*, 13, 753-763, 1993.

Labys, W.C., A. Achouch, and M. Terraza, Metal prices and the business cycle, *Resources Policy*, 25, 229-238, 1999.

Labys, W.C., J.B. Lesourd and D. Badillo, The existence of metal price cycles, *Resources Policy*, 24, 147-155, 1998.

Labys, W.C., H.J.B. Rees and C.M. Elliott, Copper price behaviour and the London Metal Exchange, *Applied Economics*, 3, 99-113, 1971.

MacDonald, R. and M.P. Taylor, Testing rational expectations and efficiency in the London Metal Exchange, *Oxford Bulletin of Economics and Statistics*, 50, 41-52, 1988a.

MacDonald, R. and M.P. Taylor, Metal prices, efficiency and cointegration: Some evidence from the London Metal Exchange, *Bulletin of Economic Research*, 40, 235-239, 1988b.

MacDonald, R. and M.P. Taylor, Rational expectations, risk and efficiency in the London Metal Exchange: an empirical analysis, *Applied Economics*, 21, 143-153, 1989.

MacKinnon, J.G. and N.D. Olewiler, Disequilibrium estimation of the demand for copper, *Bell Journal of Economics*, 11, 197-211, 1980.

McAleer, M., Sherlock Holmes and the search for truth: a diagnostic tale, *Journal of Economic Surveys*, 8, 317-370, 1994. Reprinted in L. Oxley et al. (eds.), *Surveys in Econometrics*, Blackwell, Oxford, 1994, pp 91-138.

McAleer, M. and C.R. McKenzie, Recursive estimation and generated regressors, *Economics Letters*, 39, 1-5, 1992.

McKenzie, C.R. and M. McAleer, On the effects of misspecification errors in models with generated regressors, *Oxford Bulletin of Economics and Statistics*, 56, 441-455, 1994.

McKenzie, C.R. and M. McAleer, On efficient estimation and correct inference in models with generated regressors: a general approach, *Japanese Economic Review*, 48, 368-389, 1997.

McKenzie, M.D., H. Mitchell, R.D. Brooks, and R.W. Faff, Power ARCH modelling of commodity futures data on the London Metal Exchange, *European Journal of Finance*, 7, 22-38, 2001.

McMillan, D.G. and A.E.H. Speight, Non-ferrous metals price volatility: a component analysis, *Resources Policy*, 27, 199-207, 2001.

Moore M. J. and U. Cullen, Speculative efficiency on the London Metal Exchange, *The Manchester School*, 63, 235-256, 1995.

Newey, W.K. and K.D. West, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55 (3), 703-708, 1987.

Ng, V.K. and S.C. Pirrong, Fundamentals and volatility: storage, spreads, and the dynamics of metals prices, *Journal of Business*, 67, 203-230, 1994.

Otani, K., The price determination in the inventory stock market: a disequilibrium analysis, *International Economic Review*, 24, 709-719, 1983.

Oxley, L. and M. McAleer, Econometric issues in macroeconomic models with generated regressors, *Journal of Economic Surveys*, 7(1), 1-40, 1993.

Pagan, A.R., Econometric issues in the analysis of regressions with generated regressors, *International Economic Review*, 25, 221-47, 1984.

Pesaran, M.H. and L.J. Slater, *Dynamic Regression: Theory and Algorithms*, Ellis Horwood, West Sussex, 1980.

Pesaran, M.H. And R.P. Smith, Keynes on Econometrics, in T. Lawson and H. Pesaran (eds), Keynes' Economics: Methodological Issues, Croom Helm, London, 134-150, 1985.

Sephton, P.S. and D.K. Cochrane, A note on the efficiency of the London Metal Exchange, *Economics Letters*, 33, 341-345, 1990a.

Sephton, P.S. and D.K. Cochrane, Metal prices, cointegration, and market efficiency at the LME: another look, *Kentucky Journal of Economics and Business*, 30-37, 1990b.

Sephton, P.S. and D.K. Cochrane, The efficiency of the London Metal Exchange: another look at the evidence, *Applied Economics*, 23, 669-674, 1991.

Shyy, G. and B. Butcher, Price equilibrium and transmission in a controlled economy: a case study of the metal exchange in China, *Journal of Futures Markets*, 14, 877-890, 1994.

Slade, M.E., Market structure, marketing method, and price instability, *Quarterly Journal of Economics*,106, 1309-1339, 1991.

Teyssiere, G.P., C. Brunetti and C.L. Gilbert, Non-ferrous metals price volatility: a fractionally integrated process?, in R. Smith (ed.), *Managing Metals Price Risk*, Risk Publications, London, 141-152, 1997.

Varela, O., Futures and realised cash or settle prices for gold silver and copper, *Review of Financial Economics*, 8, 121-138, 1999.

White, H., Maximum Likelihood Estimation of Misspecified Models, *Econometrica*, 50(1), 1-25, 1982.

Appendix 1

Classification of Empirical Studies

Paper	Author(s)	Year	Journal	Topic	Exchange
1	Agbeyegbe	1992	Bulletin of Economic Research	Investigation of common stochastic trends in the spot prices on the LME	LME
2	Ben Nowman and Wang	2001	Applied Economics Letters	Modelling metals price volatility based on a continuous time theory relating volatility, price level and mean reversion.	LME; KL Tin Exchange; others not stated
3	Bracker and Smith	1999	Journal of Futures Markets	Modelling time-varying volatility, structural breaks and asymmetries in the copper futures market.	Not stated, presumably COMEX
4	Bresnahan and Suslow	1985	International Economic Review	Volatility in copper prices is related to the rate of return to holding copper stocks and implications of low inventory levels.	LME
5	Brunetti and Gilbert	1995	Resources Policy	Metals price volatility is examined for trends over a 24-year period. A model is estimated for price volatility based on the metals balance (stock to consumption ratio)	LME
6	Canarella and Pollard	1986	Journal of Banking and Finance	Three alternative models are used to investigate whether the futures price is an unbiased predictor of the future spot price, and thus provide evidence on the speculative efficiency of the market.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
1	Spot	Copper; lead; zinc	Econometric	Quarterly (seasonally unadjusted) spot prices for copper, lead, zinc	Jan 1972 to Dec 1987
2	Spot	Copper; gold; nickel; silver; tin	Econometric	Monthly spot prices	Dec 1987 to May 1997; Feb 1968 to May 1997; Feb 1976 to Nov 1994; Feb 1970 to May 1997; Mar 1986 to May 1997
3	Futures	Copper	Econometric	Daily open to close, close to open, and close to close futures prices	31 Dec 1974 to 28 Jun 1996
4	Spot	Copper	Econometric	Monthly spot price; British nominal call money interest rate; Producer list price; Return on shares in a copper mining firm	May 1958 to Dec 1979
5	Spot	Aluminium; copper; lead; nickel; tin; zinc	Econometric and non-econometric	Daily spot settlement price (used to create monthly volatility measure); Monthly US Producer Price Index; Monthly stock levels;	Jan 1972 to Dec 1995; Oct 1982 to Dec 195 for aluminium and nickel; Tin is not traded between Nov 1985 and Jul 1989
6	Spot and Futures	Copper; lead; tin; zinc	Econometric	Monthly spot and 3-month futures prices (average of bid and ask). Full data set is overlapping, subsamples are non-overlapping.	Jan 1975 to Dec 1983

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
1	64	None	Common stochastic trends exist in metals prices, indicating that prices in different markets respond to common underlying economic forces.	Multivariate cointegration [1]; Bivariate cointegration [3]
2	115 - 352	None	Metals price volatility depends on the level of prices; the degree of mean reversion is different across markets.	Non-linear system of equations [5]
3	5609	15 consecutive sub-samples	Copper futures returns volatility is time-varying and asymmetric.	GARCH [1]; EGARCH [1]; AGARCH [1]; GJR-GARCH [1]; AR(1) [1]
4	260	None	Do inventory holders earn a competitive rate of return on copper? Low inventory levels introduce asymmetries into metals prices.	Non-linear regression model [12]; Linear regression model [1]
5	288 (monthly)	None	Non-ferrous metals volatility is trend stationary and volatility in metals prices depends on the metals balance (ratio of stocks to consumption).	Linear regression model [6]
6	108	3 sub-samples of 3 monthly non-overlapping data (36 obs)	Speculative efficiency hypothesis: the futures/forward price as an unbiased predictor of the future spot price.	Linear regression model [24]; ARMA model [12]; Bivariate autoregression [8]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
1	Johansen Method	None	Spot price	None
2	GMM	Spot returns	Spot price	None
3	Not stated (presumably ML)	Futures returns	Lagged futures returns	None
4	Not stated (presumably Non-linear Least Squares); OLS	Excess capital gain on holding copper	See proxy variables.	Lagged rate of change between spot price and interest rate adjusted future spot price (for "state of the copper market"); Lagged proportional diference between spot price and producer price (for "state of copper market")
5	Not stated (presumably OLS)	Intra-month standard deviation of daily log returns (measure of volatility)	Lagged intra-month standard deviation of daily log returns; Lagged deflated spot price	Lagged ratio of stock level to consumption trend (consumption trend is a fitted value)
6	OLS; Not stated (presumably ML); FIML	Log spot price; log forecast error; First difference of log spot price; First difference of log futures price	Lagged log futures price; log forecast error; Lagged log forecast error; Lagged first difference of log spot price; Lagged first difference of log futures price	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
1	R2; SEE; Newey-West SE	ADF; Phillips-Perron (with and without trend); CRDW; CRADF; F-test for serial correlation	None	None
2	None reported	None	t-test	J
3	In-sample RMSE	None	t-test	None
4	R2; Asymptotic SE	DW	None	None
5	R2; SEE	DW; ADF(3) without trend;	t-test	None
6	R2; Log-Likelihood; SE; Asymptotic SE	Non-stationarity test (presumably ADF); DW; Ljung-Box; Hosking multivariate portmanteau	See nested and non-nested tests	F-test; Wald; Likelihood ratio

Paper	Acknowledged Modelling Problems	Empirical Implications
1	Choice of quarterly data means there are fewer observations relative to monthly data; argue that a quarterly sampling frequency will better capture the long run relationship between the series.	The presence of unit roots in the variables suggests the unanticipated component of metals prices dominate price movements. Cointegration implies common stochastic trends exist between spot prices. Copper, lead, zinc spot prices have a long run equilibrium relationship.
2	None	Price volatility shows a strong dependence on price level in each market. The degree of mean reversion varies across markets.
3	Future research should consider out-of-sample forecasting ability.	GARCH best describes the data, followed by EGARCH and GJR. The AGARCH and random walk models perform poorly. Evidence supports the proposition that copper futures returns volatility is time varying and symmetric.
4	None	Proxy variables indicating the "state of the copper market" can predict capital losses on copper stock holdings, but not capital gains. Cyclical as well as permanent demand and supply side phenomena are reflected in metals prices. Real side of the economy affects metals prices more directly than financial market shocks.
5	Lack of data on speculative positions means there is no evidence with which to attribute high frequency volatility movements to speculative pressures.	Metals volatility over the sample is generally consistent, volatility has neither increased nor decreased. There are no stochastic trends in monthly volatility measures. However, the presence of deterministic trends in the data was not explicitly considered in the ADF tests. Real prices are I(1), stock to consumption ratios are also I(1), except for aluminium, which may be I(2).
6	None	The speculative efficiency hypothesis is not rejected. This implies that the futures price is an unbiased predictor of future spot prices, and other variables used to forecast future spot prices contain no additional information beyond that contained in the value of the futures price. A strategy designed to enhance long term profitability in trading metals futures is unlikely to succeed.

Paper	Author(s)	Year	Journal	Topic	Exchange
7	Chang, Chen and Chen	1990	Journal of Futures Markets	Keynesian normal backwardation theory and the Sharpe-Lintner CAPM implications for risk are examined for metals futures markets.	COMEX NYMEX
8	Chowdhury	1991	Journal of Futures Markets	Two tests of the efficient markets hypothesis for the LME metals futures markets are conducted using cointegration: (i) between markets, and (ii) between the spot and futures prices in the same market.	LME
9	Fama and French	1988	Journal of Finance	Test of implications of theory of storage as to the behaviour of marginal convenience yield with respect to inventories and the business cycle.	LME; COMEX; NYMEX
10	Franses and Kofman	1991	Journal of Futures Markets	Test for flow parities between forward prices for different metals on the LME using cointegration.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
7	Futures	Copper; platinum; silver	Econometric	Month-end settlement prices used to generate artificial futures contracts with a fixed maturity; Series for 3-, 5-, 7-, 9-, & 12-month maturities are generated for copper and silver; 4-, 7-, 10-month for platinum; CRSP Stock Index; Dow Jones Cash Commodity Index, 1-month US T-bill returns.	Dec 1963 to Dec 1983
8	Spot and Futures	Copper; lead; tin; zinc	Econometric	Monthly average of spot price; Monthly average of 3-month futures price.	Jul 1971 to Jun 1978
9	Forward and Futures	Aluminium, copper; gold; lead; platinum; silver; tin; zinc	Econometric and non-econometric	Daily LME spot and LME 3-month forward prices for aluminium, copper, lead, tin, zinc, and spot, 3-, 6-, 12-month forward prices for silver; COMEX 12-month futures prices for copper and gold; NYMEX 12-month futures prices for platinum.	1972 (Copper, lead, tin, zinc, silver, platinum), 1975 (gold), and 1979 (aluminium), to 1983 (all metals)
10	Forward	Aluminium; copper; lead; nickel; zinc	Econometric	Daily forward prices	5 Jan 1981 to 31 Dec 1981

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
7	240 (Copper, platinum); 180 (silver)	2 sub-samples: pre-1977 and post-1978	Are risk premia in non-ferrous metals futures markets consistent with the Sharpe-Lintner CAPM?	Linear Regression Model [39]
8	64	None	Efficient market hypothesis (using a cointegration test).	Bivariate cointegration model [16]
9	1149 (Al); 2818 (Cu, Pb, Sn, Zn); 611-844 (Au); 226-452 (Pt); 2483-2823 (Ag); 359-919 (Cu Futures)	2 sub-samples based on positive and negative interest-adjusted basis	Theory of Storage: Marginal convenience yield on inventory falls at a decreasing rate as inventory increases. The Theory of Storage explains the cyclical metals price relationship with the business cycle.	Linear regression model [51]
10	251	None	Equilibrium flow parity relationships exist between forward prices of the non-ferrous metals traded on the LME.	Multivariate cointegration model [1]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
7	OLS; Cochrane-Orcutt	Realised futures return	See proxy variables.	Return on market portfolio (combination of two market indices used as proxy); Risk-free interest rate (US Treasury Bill rate used as proxy).
8	Engle-Granger Method	Spot price; Futures price	Spot price; Futures price	None
9	Presumably OLS	Percent change in forward price	Percent change in spot price	None
10	Johansen Method	None	Log of forward price	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
7	R2; SE; SSE	DW	None	None
8	R2	ADF, Phillips-Perron	F-test (presumably) performed for the 4 models of spot and forward price	None
9	R2; SE	None	None	None
10	None	ADF (without trend)	None	None

Paper	Acknowledged Modelling Problems	Empirical Implications
7	Results are sensitive to the proxy selected to represent market returns.	All beta coefficients are significant, but the intercepts are not significant. Metals futures are riskier than common stocks. Large positive and significant systematic risks are identified for each metal within the CAPM framework. Traders of metals futures bear above average risks.
8	Sample size of 64 observations is too small for valid inference based on cointegration	All price series are I(1). All pairs of spot prices are cointegrated, as are all pairs of futures prices except for copper and lead, and copper and zinc. For the single market models, there is cointegration between the spot and futures prices for copper only. Two market models provide evidence of inefficiency. Single market models imply inefficiency for all markets, except copper. The nested test result implies inefficiency in each market. The efficient markets hypothesis is rejected.
9	Tests of rationality of market forecasts failed because variances of unexpected spot price changes are large relative to variances of expected spot price changes, such that expected changes cannot be reliably extracted from observed changes. Tests of rationality lack power as the variances of unexpected spot price changes are large relative to the variances of expected changes.	When inventory is high, spot and forward prices for LME metals have similar variability, but when inventory is low, spot prices are more variable than forward prices. That is, forward prices respond less to demand and supply shocks than spot prices when inventories are low. However, precious metal spot prices are not consistently more variable when inventory is low. Demand shocks arund business cycle peaks reduce inventories and generate positive convenience yields and negative interest adjusted bases.
10	Inconclusiveness in determining optimal sample length to detect stable flow parity relationships, as structural changes that occur in stock parities will effect flow parity relationships	All price series are I(1). One cointegrating relationship exists between the variables, so that one flow parity relationship exists between the five metals on the LME. The existence of cointegration between prices series contradicts the EMH. Efficiency is rejected for the LME (where the definition of the EMH is that a random walk is the best forecasting scheme).

Paper	Author(s)	Year	Journal	Topic	Exchange
11	Fraser and MacDonald	1992	Review of Futures Markets	Cointegration, persistence tests and structural time series modelling are use to examine whether metals spot and futures prices approximate a random walk process. This gives insight into the permanent or transitory nature of shocks, and market efficiency.	LME and others
12	French	1983	Journal of Financial Economics	Comparison of futures and forward prices for the copper and silver markets.	LME COMEX CBOT
13	Gilbert	1986	Applied Economics	Paper examines the empirical work of Goss (1981, 1983), and analyses the same data set using different methods for results that conflict with those of Goss (1981, 1983).	LME
14	Gilbert	1995	Journal of Applied Econometrics	The world aluminium market is modelled using a rational expectations commodity price specification of market fundamentals which reflect the supply-demand balance.	LME and Producer List

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
11	Spot and forward	Aluminium; copper; gold; lead; nickel; platinum; silver; tin; zinc	Econometric	Monthly spot and forward prices for aluminium, copper, lead, nickel, tin, zinc; monthly spot prices for gold, silver, platinum.	15 samples within Jan 1976 to Feb 1989
12	Futures	Copper; silver	Econometric	Presumed monthly silver spot, 3-, 6-, 12-month forward and futures; copper spot, 3-month forward prices; T-bill prices; Federal funds rate; Federal funds returns.	1968 for 1980 3-month contracts; 1974 to 1980 for 6- and 12- month contracts
13	Forward	Copper; lead; tin; zinc	Econometric	Monthly average spot and 3-month forward prices.	Jul 1971 to Jun 1978
14	Spot	Aluminium	Econometric	Annual world refined consumption of primary aluminium; imports minus exports of primary aluminium by western countries; aluminium price; world refined production of primary aluminium; stocks of primary aluminium; OECD construction index; Dollar exchange rate; OECD industrial production index; Dollar real interest rate; US producer prices index	1966 to 1991

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
11	Presumably 115 - 158	Sub-sample used for cointegration (109 obs)	Efficient Market Hypothesis: Do shocks have a permanent or transitory effect, and do time-varying risk premia exist?	Multivariate cointegration model [4]; Structural time series model [15]
12	39 - 112	Subsamples for 3-month contracts: 1968 to 1973; 1974 to 1980.	Interest rates explain the differrence between futures and forward prices.	Linear regression model [28]
13	79	None	Efficient Market Hypothesis: Unbiased expectations hypothesis and both weak-form and semi-strong-form EMH.	Linear regression model [12]
14	26	None	In a rational expectations model jointly estimating production, consumption, stock demand, and price equations, price and stock demand depend on short-run and long-run market imbalance or excess supply variables. Stocks will be held if there is short-term excess supply and long-term excess demand.	Linear regression model [12]; Nonlinear system of equations [1]; AR(2) [1]; VAR [2]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
11	Johansen Method; Kalman Filter (ML using method of scoring)	Spot price; forward price	Spot price; forward price; lagged spot price; lagged forward price	None
12	OLS (with serial correlation consistent covariance matrix)	Variance of T-Bill price; Variance of Federal funds returns; Variance of futures prices; Varinace of converted forward prices; Variance of converted spot prices; Cumulative one-day funds interest rate minus the contract rate	Log of the ratio of futures to forward price	Expected forward price variance (GR); Expected bond return (Actual return from rolling over one-day bonds as proxy)
13	GLS (with analytic error variance matrix); OLS (with analytic error variance matrix, and Hansen and Hodrick (1980) method)	Forecast error	Forecast error; lagged forecast error	None
14	OLS (with White's SE); IV (with heteroscedasticity consistent SE); Iterated Nonlinear 3-stage Least Squares; Presumably ML	Change in production; consumption; stock demand; change in net imports; log spot price adjusted for exchange rates; log change in exchange rate index	Net production and imports; lagged stock demand; log spot price adjusted for exchange rates; industrial production index; trend modified industrial production index; construction index; dummy for post-1986 imports; short-run excess supply; long run excess supply; interest rate; lagged log change in exchange rate index	Production shock (residual from production equation as GR)

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
11	Asymptotic SE; AIC; R2; quasi-R2	DF; variance ratio test for unit root; rescaled range statistic	None	None
12	R2; Asymptotic SE	None	F-test	None
13	SE; Hansen and Hoderick SE	None	Chi-squared test using analytic error variance matrix; Chi-Squared test using Hansen and Hoderick.	None
14	R2; SEE; RMSE	DF; ADF; DW; LM test for serial correlation; LM test for normality; RESET; LM test for hetroskedasticity; LM test for instrument validity	White's t-test	Wald test; LR test; No non-nested tests

Paper	Acknowledged Modelling Problems	Empirical Implications
11	None	All metals are I(1), except for lead and tin. For the I(1) metals, there is no (multivariate) cointegration among the LME metals (spot and forward separately), the precious metals, or the LME and precious metals. Persistence tests and structural time series modelling show that randomness of price movement is a feature of all the prices considered. The hypothesis that metals prices follow a random walk is not rejected.
12	Measurement error; serial correlation; explanatory variable correlated with previous period error term; possibly omitted variables or misspecification of the model.	The CIR type models are not useful in explaining intra-sample variations in the forward-futures price differences. The failure may be due to misspecification of the model, and measurement error. Results are ambiguous. Ratios, and differences of long and short term interest rates, are not useful in explaining the observed differences between futures and forward prices.
13	Averaged data are inappropriate for testing the EMH. Averaged data and overlapping observations produce serial correlation in the errors.	There is evidence of departures from efficiency in the tin and lead markets. The risk premium in the tin market appears to be consistently positive over time, while in the lead market the lagged forward price forecasting error holds information about future errors. There is some indication that this might be the case for the zinc market, but none of the tests provides any evidence of a departure from efficiency for the copper market.
14	In some of the estimated relationships there are indications of misspecification, so a nonlinear specification may improve the fit. Lack of observations prevents recursive modelling.	Restrictions imposed by the model are not rejected for the aluminium market data. A price semi-elasticity of storage is generated by the model. Short term and long term excess supply variables are important in the model. Additions to storage take place when there is substantial excess short term supply in the context of lower or negative long-term excess supply. RE models are only superior to the alternatives considered in forecast performance over short horizons. Implications of the REH may be limited over the long-term because the market does not possess the information relevant to longer-term developments.

Paper	Author(s)	Year	Journal	Topic	Exchange
15	Goss	1981	Applied Economics	Evaluates the performance of current spot and futures prices as predictors for future spot prices.	LME
16	Goss	1983	Applied Economics	Semi-strong-form efficiency: Do LME futures prices fully reflect all relevant publicly available information?	LME
17	Goss	1986	Applied Economics	Reply to Gilbert's [1986] criticism of Goss [1981, 1983]. A larger set of non-overlapping data is re-analysed with respect to unbiased expectations and semi-strong-form efficiency of the LME.	LME
18	Gross	1988	Journal of Futures Markets	Test of the semi-strong-form efficient market hypothesis for the LME aluminium and copper markets. The predictive performance (for the future spot price) of the LME futures markets are tested against 3 forecasting models.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
15	Futures	Copper; lead; tin; zinc	Econometric	Monthly average spot and 3-month forward prices	Jul 1971 to Jun 1978
16	Futures	Copper; lead; tin; zinc	Econometric	Monthly average spot and 3-month futures prices	Jul 1971 to Jun 1979
17	Futures	Copper; lead; tin; zinc	Econometric	Monthly average spot and 3-month forward prices (non-overlapping)	Apr 1966 to Apr 1984
18	Futures	Aluminium; copper	Econometric	Daily 3-month futures prices	First trading day Jan 1983 to last trading day Sep1984.

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
15	83 (81 for GIVE regressions)	None	Speculative Efficiency Hypothesis: Unbiased expectations.	Linear regression model [40]
16	95	None	Efficient Markets Hypothesis (semi-strong-form): Do LME prices reflect as fully as possible, all relevant publicly available information, where the information is defined as the sum of immediately prior forecast errors for all four metals?	Linear regression model [24]
17	55	None	Efficient Market Hypothesis (semi-strong form): Unbiased expectations hypothesis.	Linear regression model [16]
18	439	None	Efficient Market Hypothesis (semi-strong form).	ARIMA model [2]; SURE model [2]; AR(2) model [2]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
15	Cochrane- Orcutt; GIVE	Spot price	Lagged futures price; Lagged spot price	None
16	Cochrane- Orcutt; GIVE	Forecast error	Lagged forecast error	None
17	Cochrane-Orcutt; GIVE	Spot price; forecast error	Futures price; lagged forecast error	None
18	3SLS; GLS; OLS	Futures price	Futures price; lagged futures price	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
15	Adjusted R2; SE	DW; Wallis test (for serial correlation)	t-test; Asymptotic t-test	None
16	Adjusted R2; SE	DW; Wallis test (for serial correlation)	t-test; F-test; Chi-squared test	None
17	R2	F-test for multicolinearity	None	None
18	SE; Correlation between errors; Adjusted R2; MSE; RMSE; MAE	Box-Pierce Q; Chow;	t-test between MSE	None

Paper	Acknowledged Modelling Problems	Empirical Implications
15	Autocorrelated errors in the presence of a lagged endogenous variable mean that OLS will be biased and inconsistent. for this reason Goss also used GIVE. Some hypothesis tests and DW are not strictly valid for GIVE. Where OLS and GIVE conflict, the latter are preferred since GIVE is consistent.	Copper and zinc markets are performing their forward pricing function quite well, and the tin market somewhat better. The lead market is performing less well against the criteria for unbiased prediction. Thus, agents using the copper, zinc and tin markets as a basis for forward pricing have, on average, had a good indication of the spot price in future periods. Those using the lead market would have experienced unexpected profits or losses, so that hedging costs in this market will be higher. Goss observes that the lead market is not necessarily inefficient since the EMH (which presumes unbiasedness) is conditional on information available at the time the futures price is formed.
16	Rejection of the null of no correlation between forecast errors may be due to misspecification of the relevant information set in the model; that is, rejection of the null does not necessarily imply the market is inefficient. OLS with lagged dependent variables and serial correlation results in biased and inconsistent estimates. GIVE is used, but is also biased due to the LDV's. Serial correlation is expected due to the use of monthly averaged data, and misspecification (omitted variables).	Most estimated coefficients are insignificant, providing evidence in favour of the semi-strong EMH. However, coefficients are jointly significant in most cases, which is contrary to the semi-strong EMH. Goss suggests high collinearity between regressors biases the joint tests towards rejection of the null (not jointly significant). Overall, the results support rejection of the semi-strong EMH. A low adjusted R2 would be expected under the semi-strong EMH, but this is not always the case. Rejection of the semi-strong EMH and non-rejection of the forward pricing hypothesis (weak-form EMH) may be explained in terms of the price exclusion of some agents from the more costly units of information, resulting in significant inefficiency of consumption.
17	None	On the basis of non-overlapping data, unbiasedness cannot be rejected for any of the metals (although zinc is marginal). Using non-overlapping data, the semi-strong-form EMH is rejected for copper and zinc, but not for tin and lead.
18	Models formulated are purely forecasting models, and may not have economically meaningful coefficients.	LME futures prices forecast future spot prices more accurately than any of the three models considered (ARIMA, SURE, AR(2)). Thus, EMH cannot be rejected for both the copper and aluminium markets at the LME on the basis of a forecasting model outperforming the futures price. Of the three models, the SURE model performs relatively best for both commodities.

Paper	Author(s)	Year	Journal	Topic	Exchange
19	Hall	1991	The Manchester School	The risk premia in the forward prices of four LME metals are modelled using GARCH-M, SGARCH-M, and SGARCH-M with MA(2) errors.	LME
20	Hall and Taylor	1989	Review of Futures Markets	Modelling of risk premia, conditional on the assumption of rational expectations, (that is the speculative efficiency hypothesis) in forward markets for metals on the LME using ARCH-M, GARCH-M, and DYMIMIC models.	LME
21	Hardouvelis and Kim	1995	Journal of Money, Credit, and Banking	The relationship between futures contract margin requirements, market participation and volatility in futures prices is examined.	COMEX; CBOT; NYMEX.
22	Heaney	1998	Journal of Futures Markets	Cointegration modelling of the LME lead futures price based on the cost-of-carry model	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
19	Forward	Copper; lead; tin; zinc	Econometric	Monthly spot and 3-month forward prices	Apr 1974 to Oct 1985
20	Forward	Copper; lead; tin; zinc	Econometric	Monthly spot and 3-month forward prices	Jan 1976 to Mar 1987 (Oct 1985 for tin)
21	Futures	Aluminium; copper; gold; palladium; platinum; silver	Econometric	Daily trading volume; open interest; margin requirements; futures prices (opening, settlement, highest and lowest)	COMEX: gold: 31 Dec 74 to 31 Oct 90; silver: 29 Jul 71 to 31 Oct 90; copper: 22 Aug 72 to 18 Nov 90; Aluminium 8 Dec 83 to 13 Nov 90; CBOT: gold:12 Apr 84 to 13 Nov 90; silver 7 Sep 74 to 31 Oct 90; NYMEX: platinum: 15 Oct 79 to 30 Jun 89; palladium: 1 Nov 82 to 13 Nov 90.
22	Futures	Lead	Econometric	Quarterly spot price; futures price; LME stock level; UK T-bill 3-month mid-rate; Eurocurrency Sterling 3-month mid-rate	Mar 1976 to Jun 1995

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
19	138	None	Speculative Efficiency Hypothesis: Joint hypothesis of the existence of time-varying risk premia and rational expectations.	GARCH-M (1,1) [4]; Stochastic GARCH-M (1,1) [4]; SGARCH-M (1,1) incorporating MA(2) in mean [4]
20	134 (118 for tin)	None	Speculative Efficiency Hypothesis: Joint hypothesis of the existence of time-varying risk premia and rational expectations.	Linear regression model [4]; ARCH-M(8) with MA(2) in mean [4]; GARCH-M(1,1) with MA(2) in mean [4]; DYMIMIC (with MA(2)) [4];
21	8-620 (several samples)	5 sub-samples considered	Market participation and price volatility are negatively related to futures margin requirements.	Linear regression model [140]
22	77	None	Cost-of-Carry model.	Multivariate cointegration model [1]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
19	ML	Log forecast error	See proxy variable.	Risk premium (conditional variance of the forecast error as GR)
20	OLS (with Hansen (1982) corrected covariance matrix); ML; Kalman filter	Difference in log spot price (3-period); log forecast error	Log forward premium	Risk premium (GR: conditional variance of the forecast error as); Risk premium (GR: from ARMAX)
21	Not Stated (presume OLS).	Log change in volatility; log change in trading volume; log change in open interest; log change in growth rate of open interest	Log change in average margin requirement	Volatility (residual standard deviation from an AR(2) model of daily returns as GR); volatility (Garman-Klass variance estimator)
22	Johansen Method	None	Forward price; spot price	Risk-free interest rate (UK T-bill 3-month mid-rate as proxy (Eurocurrency Sterling 3-month mid-rate also tried)); stock level (LME stocks used as proxy)

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
19	White÷s SE; Log-likelihood	Ljung-Box; Bera-Jarque	White's t-test	None
20	SEE	DF; ADF; CRDW; Ljung-Box for serial correlation; Test for heteroskedasticity based on Ljung-Box statistic; Bera-Jarque	t-test; squared t-statistic	None
21	R2	None	t-test	None
22	AIC; SBC	ADF; Phillips-Perron; Exogeneity; (Chi-squared test of speed of adjustment parameter); Chi-Squared test for serial correlation in residuals of the VAR; LR test for intercept in cointegrating vector	Chi-squared test of restrictions on parameters	None

Paper	Acknowledged Modelling Problems	Empirical Implications
19	The residuals of the stochastic GARCH-M model with MA(2) errors are not normally distributed.	Three models are used to explain time-varying risk premia. The stochastic GARCH-M model with MA(2) error process is statistically adrequate, and preferred to the GARCH-M and SGARCH-M models. The prefered model provides robust evidence of the existence of systematically time-varying risk premia in metals markets.
20	Problem in determining the optimal order for the conditional variance in the ARCH-M model, so it is set arbitrarily. DYMIMIC model is not identified.	Fama-type test for the existence of risk premia based on OLS do not reject the existence of risk premia. Models generating a risk premium variable generally support the existence of time-varying risk premia. The GARCH-M model outperformed the ARCH-M and DYMIMIC models (based on significance of estimated parameters). The DYMIMIC model was not identified, which was interpreted as misspecification. Risk premia exist in forward prices, are time-varying and related to the conditional variance of the forecast errors.
21	One regressor is non-stationary, but it is used in some of the models in levels and in others in first differences. Exchanges raise and lower margin requirements based on their estimates of volatility.	Evidence of a "causal" negative influence from margin requirements to market participation. As margins increase, agents leave the metal market affected and move into similar market unaffected by the increase. Appears to support the competitive hypothesis (that margins restrict rational investors) since margin requirements are positively related to volatility and negatively associated with measures of participation.
22	Unit root tests for spot and forward prices provide inconclusive evidence for non-stationarity in the spot and futures price as the results change with the sampling frequency and statistical test chosen.	Interest rates and stock levels are non-stationary, while the evidence for unit roots in the spot and futures prices is tentative. One cointegrating vector between the variables is found, which is consistent with the cost-of-carry model. The difference between the futures price and the spot price varies with stock levels and the level of interest rates. This result supports the cost-of-carry model for the LME lead market.

Paper	Author(s)	Year	Journal	Topic	Exchange
23	Heaney	2002 (a)	International Journal of Forecasting	Examines the importance of cost-of-carry model variables in forecasting the future spot price.	LME
24	Heaney	2002 (b)	Journal of Futures Markets	A model approximating the convenience yield in commodity futures pricing is developed using a simple trading strategy. The model depends on spot price volatility, futures price volatility and the futures contract time to maturity.	LME
25	Hill, Moore and Pruitt	1991	Journal of Futures Markets	The efficiency of metals futures markets are examined with regard to the informational effects of the cold fusion announcement.	CBOT COMEX NYMEX NCE
26	Hsieh and Kulatilaka	1982	Journal of Finance	A test of whether forward prices equal expected future spot prices at maturity under two models of expectations formation: full information rational expectations and an incomplete information mechanical forecasting rule.	LME
27	Hussey and Quiroz	1997	Revista de Analisis Economico	Examination of the economic dynamics associated with the optimal use of futures markets by firms producing commodities.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
23	Futures	Lead	Econometric	Quarterly spot and futures prices; LME lead inventory; UK Treasury Bill rate	Dec 1964 to Jun1995
24	Futures	Copper; lead; zinc	Econometric	Quarterly spot and 3-month futures prices; Euro-currency (London) Sterling 3-month mid-rate; Euro-currency (London) USD 3-month mid-rate.	Mar 1975 to Sep 2000
25	Futures	Aluminium; copper; gold; palladium; platinum; silver	Econometric	Daily futures prices for the following contracts: gold (Aug 1989), silver (Aug 1989), copper (Sep 1989), aluminium (Sep 1989), platinum (Oct 1989), palladium (Sep 1989).	23 Nov 1988 to 1 Mar 1989
26	Forward	Copper; lead; tin; zinc	Econometric	Monthly spot and 3-month forward prices (settlement prices on first Friday of each month and corresponding spot price on delivery date).	Jan 1971 to Dec 1980
27	Futures	Copper	Econometric and numerical methods	Monthly spot and 3-month futures prices; Monthly Codelco copper production; Monthly LME copper stocks; US producer price index.	Jan 1981 to Nov 1995

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
23	123	1 sub-sample (1975-1995)	Cost-of-Carry model explains the difference between futures prices and subsequent spot prices. Including carrying cost variables improves the forecasting of subsequent spot prices with the curent futures price.	Linear regression model [2]; multivariate cointegration model [3]
24	103	None	Cost-of-carry model: an approximation of the convenience yield explains the interest adjusted basis.	Linear regression model [3]; Recursive linear regression model [3]
25	75	None	Efficient Market Hypothesis.	Presumably linear regression model [2]
26	129	1 sub-sample (72 observations)	Speculative Efficiency Hypothesis: Joint null of rational expectations no risk premium under (i) full information and (ii) incomplete information.	Linear regression model [24]; AR(6) model [4]
27	164	None	What is the effect on a firm's income due to undertaking optimal hedging in futures markets, where spot and futures prices are stochastic?	ARMA(3,2) [1]; ARCH(2) [1]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
23	OLS; Johansen method; Engle-Granger method	Log spot returns; log spot price	Lagged log futures returns; lagged log forward basis (lagged) log spot price; (lagged) log futures price; (lagged) log standard deviation of spot returns; (lagged) first-order auto-correlation coefficient of spot returns	(Lagged) Risk-free interest rate (UK T-Bill 3-month rate as proxy); (lagged) convenience yield (log inventory as proxy)
24	Cochrane-Orcutt; OLS	Interest adjusted basis	See proxy variables	Estimated convenience yield (difference between the value of profitable trading opportunities available under the underlying asset and futures positions as proxy)
25	OLS	Daily futures returns	See proxy variables	Metals market returns (equally weighted index of gold silver, copper and aluminium returns as proxy)
26	OLS (with Hansen and Hodrick [1980] SE); GLS (Hannan efficient); Kalman filter	Realised forward return; log spot price	Lagged realised forward return; lagged forecast error; lagged spot return; lagged log spot price	Forward price (Kalman filter forecast) as proxy for the risk premium.
27	Presumably ML	Log spot price	Lagged log spot price; lagged log futures basis	Log world inventory (LME stocks as proxy)

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
23	R2; ME; MAE; MSE	ADF (with trend); Phillips-Peron; DW; CUSUM; CUSUMSQ; chi square test of serially correlated errors; Poskitt test	t-test; F-test	None
24	R2	Phillips-Perron; DW; Wald test for AR(1); Whites heteroskedasticity test; Chow test for structural change; LM test for serial correlation; LM test for ARCH	t-test; F-test	None
25	None	None	t-test; F-test	None
26	ME; MSE	Chi-square test for serial correlation of forecast errors; Chi-square test for correlation forecast errors with available information	t-test of mean difference between MSE	None
27	SE; Log Likelihood; SBC; AIC	F-test for serial correlation; F-test for heteroscedasticity	t-test	None

Paper	Acknowledged Modelling Problems	Empirical Implications
23	A substantial degree of variation in the spot price is unexplained by the cost-of-carry models, requiring further development of the models to improve their predictive and explanatory power. This may be achieved through more precise modelling of the cost-of-carry model.	Over the sample, lead spot and futures prices are I(1), as are inventories and the UK T-bill rate. The standard deviation of spot returns, and the first-order autocorrelation coefficient of spot returns are stationary. Single equation error correction models support the cost-of-carry model by indicating a significant relationship between carrying cost variables in previous periods and current spot price change. The forecasting ability of models that include carying cost variables is superior to those that do not.
24	Estimates of the convenience yield coefficient may be affected by the fact that the convenience yield variable is an approximation for convenience yield.	Estimated coefficients for the convenience yield variable are negative and significant, as predicted by theory. Recursive estimates approach the value specified by prior theory of -1 as the sampe size is increased. Convenience yield accounts for a substantial portion of the difference between spot and futures prices.
25	None	A market model is estimated as the basis on which to examine returns in metals markets around the period of the cold fusion announcement. Statistics calculated from market data and the market model support efficiency in the metals markets considered. Platinum futures market was unaffected by the cold fusion announcement; gold, silver, copper and aluminium were similarly unaffected; palladium futures reacted temporarily to the announcement in terms of price, volume and volatility, and subsequently returned to previous levels.
26	Forecast errors will be correlated with information on realised spot prices.Overlapping data means error term will be MA(n-1), where n=forward contract maturity period.	Reject joint hypothesis of rational expectations and no risk premium under full information. Forecast errors of forward prices have non-zero means, serial correlation, and correlation with errors from other markets. Kalman filter better forecasts future spot prices than does the forward price. Forward prices contain non-zero risk premia.
27	None	On the basis of diagnostics and information criteria, the ARCH model is preferred to the ARMA model. Optimal hedging based on futures contracts leads to higher average income for the firm, but occasionally generates significant losses over short periods of time. The results indicate that even though there is a possibility of generating profits in futures transactions, the chance of incurring significant economic losses from time to time cannot be eliminated. Such losses seem large enough to limit participation in these markets, on behalf of both state-owned and private enterprises.

Paper	Author(s)	Year	Journal	Topic	Exchange
28	Kocagil	1997	Applied Financial Economics	A rational expectations model is developed to examine the relaionship between the level or intensity of metals futures speculation and the volatility of spot prices for metals.	COMEX
29	Krehbiel and Adkins	1993	Journal of Futures Markets	A cointegration model is used to test the unbiased expectations and no risk premium hypotheses. The cost of hedging is argued to be higher if the futures price is a biased expectation, or contains a systematic risk premium.	COMEX; NYMEX
30	Labys, Lesourd and Badillo	1998	Resources Policy	Cycles in metals prices are investigated using macroeconomic business cycle identification techniques, and their statistical significance is examined.	LME; KL Tin Exchange; and other price quotatons
31	MacDonald and Taylor	1988 (a)	Oxford Bulletin of Economics and Statistics	The "Bivariate Vector Autoregressive approach" used to test the efficient market hypothesis joint propositions of rational expectations and forecast error orthogonalty.	LME
32	MacDonald and Taylor	1988 (b)	Bulletin of Economic Research	The efficient market hypothesis is tested for the spot markets on the LME. Cointegration between prices in different markets implies a rejection of the EMH because Granger Causality must be in at least one direction.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
28	Futures	Aluminium; copper; gold; silver	Econometric	Weekly spot and futures prices (aggregated from daily data)	1980 to 1990 (exact sample period not stated)
29	Futures	Copper; gold; platinum; silver	Econometric	4 monthly (3 monthly) futures price data for copper, gold, silver, (platinum). Spot prices from gold and silver markets, producer price of copper; expiring futures price for platinum.	Jan 1960 to May 1992 (copper); Jun 1975 to Jun 1992 (gold); Jan 1968 to Apr 1992 (platinum); May 1964 to May 1992 (silver)
30	Spot	Aluminium, copper, gold, lead, nickel, silver, tin, tungsten, zinc	Econometric	Monthly average spot prices	Jan 1960 (Jan 1970 for aluminium and gold) to Dec 1995
31	Forward	Copper; lead; tin; zinc	Econometric	Monthly (month-end) spot and 3-month forward prices	Jan 1976 to Mar 1987 (Oct 1985 for tin)
32	Spot	Lead; tin; zinc	Econometric	Monthly (month-end) spot prices	Jan 1976 to Oct 1985

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
28	366 - 568	None	An increased degree of speculation in a futures market leads to lower volatility in the spot price for the underlying asset.	Linear regression model [4]
29	52 - 98	None	Speculative Efficiency Hypothesis: Unbiased expectations hypothesis and the no-risk premium hypothesis.	Bivariate cointegration model [4]
30	Presumably 432 (312 for aluminium and gold)	5 sub-samples	Short-term metals price movements contain substantial and significant cyclical components.	Structural time series model [8]
31	134; 118 (tin)	None	Efficient Markets Hypothesis: assuming risk neutrality, the forward price is an unbased predictor of future spot prices, and forecast errors are orthogonal.	Bivariate vector autoregression [4]
32	118	None	Efficient Markets Hypothesis (using a cointegration test).	Bivariate cointegration model [3]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
28	OLS (with Newey-West SE)	Detrended futures price; first difference of futures price	Lagged first difference of futures price; first difference of forecast error	Lagged detrended futures price (generated by regressing the futures price on a time trend and time trend squared)
29	Johansen Method	None	Log futures price; dummy variable	Log expected spot price (log spot price or expiring futures price as proxy)
30	ML	Spot price	No non-stochastic explanatory varibles	None
31	OLS	Log spot returns; log forward basis	Lagged log spot returns; lagged log forward basis	None
32	Engle-Granger	Spot price	Spot price	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
28	SE; Newey-West SE; R2	DF; DW	Monte Carlo simulation test of significance of the regression coefficients	None
29	SBC	DF; ADF (with and without trend); CRDF; CRDW; Perron test; LR test for linear trend	LR test of parameter restrictions	None
30	R2; Log Likelihood; Forecast error variance	Box-Pierce Q; Unequal variances test for heteroscedasticity; Bowman-Shenton normality test; Perron test.	Weibull distribution test	None
31	AIC, R2	Ljung-Box Q; LR test for serial correlation	See nested and non-nested tests	Wald test; LR test; No non-nested tests.
32	R2	DF, CRDW, CRDF	None	None

Paper	Acknowledged Modelling Problems	Empirical Implications
28	Several assumptions of the underlying theoretical model motivating the empirical analysis are questionable. Noise in daily data means that a weekly frequency must be used.	The hypothesis that an increased intensity of speculation in futures markets leads to decreased volatility in spot markets, is rejected for the aluminium, copper, gold and silver markets.
29	Results are sensitive to structural change in the silver market. Perron test implies silver futures prices are I(0), so that the cointegration model is inappropriate for analysing the silver futures market.	Spot and futures prices are I(1), and in each market there is 1 cointegrating vector. The absence of a risk premium is not rejected for silver, gold, platinum, but is rejected for copper. Unbiased expectations hypothesis is rejected for silver, gold and copper, and is not rejected for platinum. Joint test of unbiased expectations and no risk premium is rejected for all markets except copper. No linear trend is present in any market. Explicitly modelling the silver market structural break affects the test outcomes. Assuming silver is I(1), one cointegrating vector is found. However, the unbiased expectations hypothesis is not rejected.
30	Serial correlation in the errors and heteroscedasticity prevented anlaysis of the full sample.	Results provide evidence for cyclical behaviour in the expansion, contraction, and duration phases for metals prices. The term of cyclical activity is shorter than has been shown in previous studies. Two kinds of cycle are predominant. The first usually has a periodicity of less than 12 months, while the second has a periodicity of greater than one year, which is largely stochastic and time-invariant.
31	None	The EMH joint hypothesis of rational expectations and forecast error orthogonality is not rejected for copper and lead, while the EMH restrictions are rejected for tin and zinc. This result may be rationalised by considering the structure of the respective commodity markets: copper and lead were competitive markets, while there are imperfections in tin and zinc markets (tin was controlled, and the zinc industry was concentrated).
32	Implicitly a system of equations is being considered. However, the analysis is limited to single equation (bivariate) tests for cointegration.	The spot price series are all non-stationary, but in a bivariate setting, there is no cointegration between spot prices. Finding that the spot prices are cointegrated supports the efficient markets hypothesis for spot markets on the LME.

Paper	Author(s)	Year	Journal	Topic	Exchange
33	MacDonald and Taylor	1989	Applied Economics	Investigation for the preence and nature of time-varying risk premia in LME forward prices conditional on the hypothesis of rational expectations. Results are interpreted in the context of the EMH.	LME
34	MacKinnon and Olewiler	1980	Bell Journal of Economics	Estimation of the disequilibrium demand for refined copper in the USA.	LME; Producer List
35	McKenzie, Mitchell, Brooks and Faff	2001	European Journal of Finance	Modelling time-varying volatility of metals futures returns comparing Power ARCH with various symmetric and asymmetric ARCH and GARCH models.	LME
36	McMillan and Speight	2001	Resources Policy	A model of time-varying conditional variance is applied to metals markets that incorporates short-run effects that revert to a long-run process, that is itself mean-reverting.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
33	Forward	Copper; lead; tin; zinc	Econometric	Monthly (month-end) spot and 3-month forward prices	Jan 1976 to Mar 1987 (Oct 1985 for tin)
34	Spot	Copper	Econometric	Quarterly data for US copper consumption; US producer price; spot price (LME); US index of industrial production; Manufacturers inventories of durable goods; Average hourly earnings of metal workers; Mining plant and equipment spending; Zinc price; US refined copper production	Jan 1947 to Dec 1974
35	Futures	Aluminium; aluminium alloy; copper; lead; nickel; tin; zinc	Econometric	Daily 3-, 15- and 27-month futures prices	Five samples within 3 Jan 1989 to 30 Sep 1997
36	Spot	Aluminium; copper; lead; nickel; tin; zinc	Econometric	Daily spot prices	Nov 1971 (Oct 1982 (aluminium and nickel), Jul 1989 (tin)) to Dec 2000

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
33	134; 118 (tin)	2 sub-samples: Jan 1976 to Jun 1981; Jul 1981 to Mar 1987; (Tin: Jan 1976 to Dec 1980; Jan 1981 to Oct 1985)	Efficient Markets Hypothesis: Existence of time-varying risk premia.	Linear regression model [24]
34	90	Disequilibrium (42 observations and equilibrium period (48 observations) subsamples.	Demand for copper is explained by a stochastic demand model in which disequilibrium may occur.	Linear regression model [1]; Modified Tobit model [1]; Simultaneous equation model [1]
35	1200 - 2209	None	Is non-ferrous metals futures returns volatility symmetric or asymmetric, and best modelled using A-PARCH (A-PGARCH) or a nested alternative model?	ARCH [17]; GARCH [17]; Leverage ARCH [17]; Leverage GARCH [17]; GJR-ARCH [17]; GJR-GARCH [17], Taylor ARCH [17]; Taylor GARCH [17]; TARCH [17]; TGARCH [17]; NARCH [17]; Power GARCH [17]; Asymmetric Power ARCH [17]; Asymmetric Power GARCH [17]; All models are of order (1) or (1,1).
36	7361 (4605 (aluminium and nickel), 2881 (tin))	None	Volatility in non-ferrous metals returns is better approximated by a model involving three components: short run financial effects, long run market fundamentals and a common long run trend among markets.	AR(0)-GARCH(1,1) [4]; AR(2)-GARCH(1,1) [1]; AR(3)-GARCH(1,1) [1]; AR(0)-CGARCH(1,1) [4]; AR(2)-CGARCH(1,1) [1}; AR(3)-CGARCH(1,1) [1]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
33	OLS (with Hansen [1982] SE)	Forecast error; change in spot price	Forward basis	None
34	OLS; ML	Consumption; producer price	Lagged producer price; spot price; lagged spot price; index of indusrial production; trend; seasonal dummy variables; deflated average hourly metal worker wage;Deflated mining plant and equipment spending; refined copper production	Change in inventories of copper products (change in manufacturers' inventories of durable goods used as proxy); Metal price trends (Zinc price used as proxy); producer price residual included to test for exogeneity (as GR).
35	ML	Log futures returns	None	None
36	ML with Bollerslev and Wooldridge [1992] SE	Log spot returns	Lagged log spot returns	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
33	SE; R2; SEE	None	None	None
34	SE; SEE; Adjusted R2; Log-likelihood	Chow test; LR test for structural change	None	None
35	SE; Second moment; Corellogram	None	See nested and non-nested tests	LR test (Nested)
36	Bollerslev-Wooldridge SE; SBIC	Ljung-Box; Jarque-Bera	See nested and non-nested tests	LR test (Nested)

Paper	Acknowledged Modelling Problems	Empirical Implications
33	Data are overlapping, and so the error term may be serially correlated.	Empirical tests of the EMH are usually tests of a joint hypothesis of risk neutrality and rational expectations. This paper assumes rational expectations, and tests for the presence of time-varying risk premia. The presence of a risk premium will violate the common joint (EMH) hypothesis of risk neutrality. In general, the evidence for the presence of time-varying risk premia was found to be weak. However, some support was found for the presence of time-varying risk premia in the tin and zinc markets. The time-varying risk premia in these markets behaved in an intuitive manner, particularly a negative covariation with expected price change.
34	None	Results are consistent with institutional evidence on the existence of rationing in the market. They suggest that conventional estimates of the demand for copper, which implicitly assume the market is always in equilibrium, are severely biased. Explicitly recognising the existence of disequilibrium in some periods substantially affects the size of all coefficient estimates, which are biased towards zero when equilibrium is assumed. Taking account of the simultaneity between the US producer price and the demand for copper affects the size of the estimated price elasticities, which are biased towards zero when the US producer price is treated as exogenous.
35	None	Asymmetry does not appear to be present in metals futures data. Models with asymmetric terms are generally outperformed by symmetric models. Leverage effects are not present, or are not strong in metals markets. There is some support for the inclusion of a power term in GARCH models. The standard GARCH model is preferred over other specifications nested in the Asymmetric Power GARCH model, except for Power GARCH. However, the Taylor GARCH model outperforms Power GARCH.
36	None	CGARCH model is superior to GARCH on the basis of diagnostics and the nested LR test. Long-run volatility exhibits long memory, but is stationary and (slowly) mean reverting. The long run volatility trends of some metals markets demonstrate co-movement. Three common elements account for most non-ferrous metals volatility. CGARCH is used to model long-run and short-run volatility.

Paper	Author(s)	Year	Journal	Topic	Exchange
37	Moore and Cullen	1995	The Manchester School	Applies the Phillips-Hansen modified estimation method to estimating and testing a model for the speculative efficiency hypothesis using data for six LME markets. Shows that Johansen procedure does not apply to overlapping spot and futures data.	LME
38	Ng and Pirrong	1994	Journal of Business	Examines the role of supply and demand fundamentals in determining non-ferrous metals spot and forward returns volatility, using implications of the theory of storage.	LME
39	Sephton and Cochrane	1990 (a)	Economics Letters	The efficient market hypothesis is examined for six metals on the LME in terms of a zero mean forecast error that is uncorrelated with past forecast errors. Several aspects of the EMH tests in MacDonald and Taylor (1988) are criticised.	LME
40	Sephton and Cochrane	1990 (b)	Kentucky Journal of Economics and Business	The efficient market hypothesis is examined for six LME metals markets using cointegration methods. Several aspects of the EMH tests in MacDonald and Taylor (1988) are criticised.	LME

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
37	Forward	Aluminium; copper; lead; nickel; tin; zinc	Econometric	Spot and 3-month forward prices for aluminium (weekly), copper (monthly), lead (monthly), nickel (monthly), tin (weekly), zinc (weekly).	13 Oct 1988 to 23 Jan 1992 (aluminium); Feb 1979 to Jan 1992 (copper); Feb 1979 to Jan 1992 (lead); Aug 1979 to Jan 1992 (nickel); 1 Jun 1989 to 23 Jan 1992 (tin); 1 Dec 1988 to 30 Jan 1992 (zinc)
38	Spot and forward	Aluminium; copper; lead; zinc; silver	Econometric	Daily spot and 3-month forward prices; Eurosterling rate; Eurodollar rate; LME warehousing cost; Silver storage cost	1 Sep 1986 to 15 Sep 1992; 27 Aug 1987 to 15 Sep 1992 (aluminium)
39	Forward	Aluminium; copper; lead; nickel; tin; zinc	Econometric	Monthly spot and 3-month forward prices	Jan 1976 to Feb 1989 (copper and lead); Jan1976 to Sep 1985 (zinc); Jan1976 to Oct 1985 (tin); Jan 1979 to Dec 1988 (aluminium); Sept 1979 to Feb1989 (nickel)
40	Spot and forward	Aluminium; copper; lead; nickel; tin; zinc	Econometric	Monthly spot and 3-month forward prices	Jan 1976 to Feb 1989 (copper and lead); Jan1976 to Sep 1985 (zinc); Jan1976 to Oct 1985 (tin); Jan 1979 to Dec 1988 (aluminium); Sept 1979 to Feb1989 (nickel)

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
37	139 - 172	None	Speculative Efficiency Hypothesis.	Bivariate Cointegration Model [5]
38	1517; 1267 (aluminium)	None	Theory of storage implications for the variances and correlations of commodity spot and forward prices, and the spread between spot and forward prices.	Error-Correction model [5] (consisting of linear regression model and augmented bivariate GARCH model estimated as a system)
39	114 - 158	2 sub-samples	Efficient Market Hypothesis.	Linear regression model [16]
40	114 - 158	3 sub-samples	Efficient Market Hypothesis.	Bivariate cointegration models [24]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
37	Phillips-Hansen Fully Modified OLS	Log spot price	Log forward price	None
38	Engle-Granger	Log spot returns; log futures returns; spot return conditional variance; forward return conditional variance; spot and forward return covariance	Lagged spot returns; lagged forward returns; lagged interest and storage adjusted spread	Lagged spot return conditional variance; lagged forward return conditional variance; lagged spot return unconditional variance; lagged forward return unconditional variance (unconditional variances are generated from OLS, conditional variances generated from a bivariate GARCH model)
39	OLS (with Hansen [1982] standard errors)	Forecast error	Lagged forecast error	None
40	Engle-Granger	Spot price	Spot price; presumably forward price	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
37	None	Phillips-Perron test (with and without trend); Hansen test for parameter instability; Chow test for structural change	t-test of linear restrictions on coefficients	None
38	Adjusted R2; Log likelihood;	Ljung-Box; ADF	t-test; F-test	None
39	None	Box-Pierce Q	t-test	None
40	R2	DF; ADF; CRDW; CRDF	None	None

Paper	Acknowledged Modelling Problems	Empirical Implications
37	Sample for tin is too small for meaningful analysis, given proximity to the collapse of the market.	All variables are I(1), except for tin spot and forward prices. Spot prices are cointegrated with forward prices in the aluminium, copper, lead, nickel, and zinc markets. Tests of significance of parameter estimates support the speculative efficiency model for aluminium, copper, lead, zinc, and results for the nickel market reject the model. Stability tests support the speculative efficiency model for aluminium, copper, lead, zinc, and reject the model for nickel. Long run speculative efficiency cannot be rejected for aluminium, copper, lead and zinc. Speculative efficiency is rejected for nickel.
38	None	Results for industrial metals provide clear support for the theory of storage, and are consistent with the hypothesis that spot-and-forward-return dynamics are strongly related to variations in fundamental supply and demand conditions. In particular: spot and futures returns volatility varies directly with the squared spread, forward returns are less volatile than spot returns, volatility of forward returns declines relative to spot returns as the squared adjusted spread increases, correlations between spot and forward returns vary inversely with the spread; volatility of the spread returns varies directly with the spread; forward price elasticities increase as the adjusted spread narrows; hedge ratios vary directly with the squared spread. In contrast, the adjusted spread for silver does not explain the dynamics of silver prices. These results arise because marginal storers of silver hold stocks as a store of value, rather than to smooth consumption and production of the commodity over time, so that the adjusted spread is small and invariant.
39	More lagged values could have been included the model.	The EMH is not rejected for single metal market models. Multiple market models including reject efficiency for zinc, tin, lead, and aluminium. Efficiency in the copper and nickel markets is not rejected. Multiple market models (excluding tin and zinc from the analysis) reject efficiency for copper and nickel, while lead and aluminium appear efficient.
40	Overlapping data imply MA(2) errors	Spot price series are I(1). The cointegration tests for sub-samples 1 and 2 find that no pairs of metals spot prices are cointegrated, thereby providing support for the EMH. For sub-sample 3, nickel and copper, copper and aluminium, lead and nickel, are (bivariate) cointegrated. This evidence rejects the EMH. Paper presumes cointegration between spot prices, and spot and forward prices to imply inefficiency.

Paper	Author(s)	Year	Journal	Topic	Exchange
41	Sephton and Cochrane	1991	Applied Economics	The efficient market hypothesis is examined for six metals on the LME. Several aspects of the EMH tests in MacDonald and Taylor (1988) are criticised.	LME
42	Shyy and Butcher	1994	Journal of Futures Markets	Price behaviour of copper forward contracts on the SHME and the lead-lag relationship with prices on the LME.	LME; SHME
43	Slade	1991	Quarterly Journal of Economics	The relationship between the organisation of markets and the behaviour of prices is examined with respect to the determinants of price instability (volatility in returns). The effects of concentration on the production side, and hedging and speculation on the consumption side, are evaluated.	LME and Producer List

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
41	Forward	Aluminium; copper; lead; nickel; tin; zinc	Econometric	Monthly spot and 3-month forward prices	Jan 1976 to Feb 1989 (copper and lead); Jan1976 to Jun 1985 (zinc); Jan1976 to Jul 1985 (tin); Jan 1979 to Sept 1988 (aluminium); Sept 1979 to Feb1989 (nickel)
42	Forward	Copper	Econometric	Daily spot and 3-month forward prices; swap Reminbi/USD price; GBP/USD exchange rate	1 Jun 1992 to 14 Oct 1993
43	Spot	Aluminium; copper; lead; nickel; silver; zinc	Econometric and non-econometric	Monthly average LME spot prices and producer prices (no producer price for silver, no transactions producer price for aluminium); Alcan annual aluminium sales volume and revenues; Hirschman-Herfindahl US concentration index (HHI); HCW qualitative measure of world industry concentration (HCW); secondary metal recovery from new and old scrap; qualitative measures for substitutability in downstream production	1970 (1979, aluminium and nickel) to 1986

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
41	114 - 158	None	Efficient Market Hypothesis: Joint hypothesis of risk neutrality and rational expectations.	Linear regression model [17]
42	Not stated (presumably 213)	None	Equilibrium price parity and lead-lag relationships between LME and SHME (spot and futures) prices.	Bivariate cointegration model [2]
43	Not stated (presumably 204)	2 sub-samples (1970s and 1980s)	Volatility in returns on non-ferrous metals is explained by industry organisation and marketing method variables: prices are more stable in concentrated industries; prices are more stable in markets where buyers are consumers relative to those with consumers, hedgers and speculators.	Linear regression model [18]; probit model [1]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
41	Not stated (presumbly OLS)	Forecast error; change in spot price	Forward basis	None
42	Engle-Granger	Log spot price; log forward price	Log spot price; log forward price	None
43	OLS (with White's SE); Heckman's 2-Step Estimator; Two Stage Least Squares	Variance in log returns; producer price dummy;	Producer price dummy; 1980s dummy; silver market bubble dummy; simultaneous exchange and producer pricing dummy	Aluminium producer transactions price (GR); horizontal industry concentration (HHI and HCW as proxies); recycling activity (secondary recovery from new and old scrap as proxy); substitutability in downstream production (qualitative proxy); importance of by-product production (qualitative proxy); degree of vertical industry integration (proxy)

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
41	SE; R2; SEE	LR test for parameter stability; CUSUMSQ	None	None
42	R2	ADF; Granger Causality	t-test	None
43	R2; Adjusted R2	Presumably LR test for predictive failure	t-test; White's t-test; F-test	None

Paper	Acknowledged Modelling Problems	Empirical Implications
41	Market efficiency test methodologies are inadequate, and tests using cointegration may provide a superior analysis of the efficient markets hypothesis. Overlapping data imply MA(2) errors.	Previous empirical tests of market efficiency applied to the LME have serious deficiencies in terms of data and methodology. Using a similar methodology, results contradicting previous studies are produced. Risk premia exist in the copper and tin markets, so that the authors conclude these markets are inefficient. Stability tests show that previous results by MacDonald and Taylor [1988] on the efficiency of metals marets on the LME (except for the tin market) are questionable (or invalid).
42	Short sample period. As more data become available, investigation of the relationship between Chinese prices futures prices and world futures prices will be more feasible. Other Chinese markets can be included in the analysis. Chinese metals market is relatively young at the time of analysis.	Copper spot and forward prices on LME and SHME are I(1). LME spot (futures) prices and SHME spot (futures) prices are cointegrated. The spot market cointegrating relationship is more significant than the forward market relationship. For the spot market, LME Granger-causes SHME, and the SHME does not Granger-cause LME. In the forward market there is also unidirectional Granger causality from LME to SHME.
43	No data on metals production cost variability available. Simultaneity leads to bias in OLS.	Horizonal market structure and contractual arrangement are significant determinants of volatility in returns. However, only contractual arrangement (producer or exchange pricing) is important (in terms of having a substantial effect on voatility). The increase in metals price instability over the 1980s is claimed to be entirely explained by changes in underlying market structure and organisation variables. Prices in markets with hedging and speculation are more volatile than prices in markets, where all buyers are consumers of metal. It is proposed that this may be explained by the low margin requirements of the LME.

Paper	Author(s)	Year	Journal	Topic	Exchange
44	Teyssiere, Gilbert and Brunetti	1997	Managing Metals Price Risk	A FIGARCH process is used to model non-ferrous metals return volatility.	LME
45	Varela	1999	Review of Financial Economics	Relationship between futures prices and realised cash prices is modelled using data corresponding with the first, middle and last day of the delivery month, over various contract maturities.	NYMEX

Paper	Futures or Spot Focus	Metal(s)	Econometric or Non-Econometric Analysis	Data	Sample
44	Spot	Aluminium; copper; lead; nickel; tin; zinc	Econometric and non-econometric	Daily spot (settlement) prices	Jan 1972 (Oct 1982 (aluminium and nickel) to Dec 1995 (no tin data for Nov 1985 to Jul 1989)
45	Futures	Copper; gold; silver	Econometric	Daily 15-, 30-, 45-, 60-day futures and delivery (realised) prices	Jul 1971 (silver), Dec 1974 (gold), Aug 1988 (copper) to Sep 1995

Paper	Number of Observations	Sub-samples	Economic Hypotheses Tested	Model Specification [Number Estimated]
44	3343 - 6056	None	Long memory and long run dependencies due to market fundamentals in non-ferrous metals returns are approximated by FIGARCH.	AR(1)-FIGARCH(1,d,0) [6]
45	41 - 151	None	Speculative Efficiency Hypothesis: Futures prices are unbiased predictors of future realised delivery prices.	Linear regression model [144]

Paper	Method of Estimation	Dependent Variable(s)	Explanatory Variables	Proxy Variables and Generated Regressors
44	ML	Log spot returns	Lagged log spot returns	None
45	OLS	Spot price; Log spot price; Futures price; Log futures price	Futures price; Log futures price	None

Paper	Descriptive Statistics	Diagnostic Tests	Hypothesis Testing	Nested and Non-Nested Tests
44	SE; Skewness and excess kurtosis of standardised residuals	Box-Pierce Q	None	None
45	R2	DW; ADF; Phillips-Perron	t-test	None

Paper	Acknowledged Modelling Problems	Empirical Implications
44	Non-normality in financial returns data means that standard errors based on the assumption of normality should be interpreted with caution. Estimates are quasi-ML.	The estimated FIGARCH models adequately account for the long-range dependence and persistence in non-ferrous metals returns volatility. Metals returns volatility is fractionally integrated, and the volatility processes of the six LME metals are similar. In the long run, returns volaitlity is stationary and mean reverting.
45	None	Unbiased expectations hypothesis is not rejected for: 15- and 30-day gold for first, middle and last delivery day prices; 15-day silver; and 15-, 30- 45- and 60-day copper for first and middle delivery prices. The hypothesis was rejected for: last delivery day prices in silver and copper; for 45- and 60-day gold; and 30-, 45- and 60-day silver. Cash prices respond less than futures prices in these instances.