Competing Models of Firm Profitability

by

Margaret E. Slade[1]

Department of Economics
University of Warwick
Coventry CV4 7AL
UK

December 2003

Abstract:

In this paper, I look at four models of firm profitability: two taken from Industrial Organization, one from Finance, and one from the Economics of Exhaustible Resources. Only one predicts that there will be a positive relationship between firm profitability and the structure of the market in which the firm operates, and only that one views high profits as an indication of monopoly power. Nevertheless, most antitrust authorities base their policies on a belief in those relationships. Using panel data from 14 nonferrous–metal mining and refining markets, I find strong empirical support only for the market–structure model.

Journal of Economic Literature classification numbers: D43, G13, L13, L71

Keywords: Firm profits, market structure, market share, risk premia, resource exhaustion, principal components.

[1] This paper was prepared for my Presidential Address to the members of the European Association for Research in Industrial Economics (EARIE) at their annual meeting in Helsinki, Finland in August of 2003. The research was supported by a grant from the Hampton Foundation at the University of British Columbia. I would like to thank Simon Anderson, Stephen Martin, and Henry Thille for helpful comments and Kazuko Kano for valuable research assistance.

1 Introduction

Virtually all branches of economics embrace the notion that firms attempt to maximize profits. Nevertheless, although some firms are substantially more profitable than others, most earn only a competitive rate of return. Given those facts, it is not surprising that economists from various subdisciplines have developed models that predict which firms will earn high rates of return and how those rates can be sustained in a world in which profits attract entry.

In this paper, I summarize four such models: two that originate in the field of industrial organization (IO), one that comes from financial economics, and one that has its origins in the economics of exhaustible resources. Each singles out a different factor as the principal determinant of profitability. In particular, the first emphasizes the structure of the market in which the firm operates, the second points to the firm's share of that market, the third focuses on the firm's risk class, and the last emphasizes resource scarcity and intertemporal arbitrage.

The four models of firm profitability originate in traditions that view the economic world in very different lights. Nevertheless, although some are better characterized as searches for empirical regularities, all can be rationalized by economic theory. Only the first, however, predicts a causal relationship between industry concentration and firm profitability. Furthermore, only the first interprets the existence of high profits as evidence of monopoly power.

Nevertheless, antitrust agencies in most Western countries implicitly assume that a profits/concentration connection exists. In particular, most agencies require that product and geographic markets be defined and indices of concentration be calculated before any monopolization or merger case can be brought forward.[2] This practice is adhered to in spite of the fact that it is not only impossible to prove unambiguously that a causal relationship links profits to concentration,[3] but also many theories attribute high profits to superior efficiency, not market power.

Given an abundance of theories and predictions, it is natural to turn to data in an attempt to disentangle and assess predicted effects. This exercise is not new. Rather, it is a very old tradition in IO that has fallen out of fashion.[4] One reason for the fall from fashion is that it is difficult to move beyond mere correlation to a determination of causality. This difficulty arises because most of the variables that have been used in empirical tests are potentially endogenous.

[2] Many other sorts of antitrust cases also rely heavily on concentration indices.

[3] In particular, under some assumptions, one can demonstrate that no such connection exists.

[4] The empirical IO literature has moved from an examination of many markets using the same model to an emphasis on case studies that are fine tuned to fit particular markets and are much more data intensive.

Rather than attempting to determine causality, I take a descriptive approach.[5] In particular, I look at equilibrium rather than causal relationships. Those relationships are evaluated by calculating the principal components of a matrix of endogenous variables, interpreting the components that have the highest explanatory power, and assessing how the original variables are related to those components. In other words, I seek to uncover the important relationships — those that explain a large fraction of the cross–sectional and time–series variation in the data.

This technique is applied to data from nonferrous–metal mining and refining industries. Those industries produce homogeneous commodities that have well defined product and geographic markets. In particular, fourteen commodity markets are examined in the application. The data consist of an unbalanced panel of the principal mining and refining firms and include the production of each of those firms in each market.[6]

The industries were chosen because they satisfy the assumptions that underlie the theoretical models. In other words, the theories are examined in a context in which the probability that they are valid is high. In particular, accurate market definition is extremely important for assessing models of firm profitability. Indeed, when markets are delineated too broadly or too narrowly, statistical tests are biased against finding relationships between profits, market structure, market share, and market risk. Fortunately, the measures of market structure and market share that are used here are unusually accurate.

The industries were also chosen because a number of high–profile mergers between mining firms have been proposed or consummated in recent years (e.g., Alcan/Pechiney). The arguments that favor those mergers rely on alleged efficiencies, moderate horizontal concentration, and the inability of achieving market power when products are homogeneous and are traded in highly liquid commodity–futures markets.[7] It is therefore important to assess whether market structure matters in that context.

The organization of the paper is as follows. The next section describes the competing models. That description is followed by a discussion of the industries, the data, the empirical model, and the results of the principal–component analysis. Finally, I suggest how one might interpret the findings.

[5] The exercise that is performed here is most closely related to the work of Schmalensee (1985).

[6] The panel is unbalanced due to mergers and acquisitions.

[7] For example, Allaz and Villa (1993) show that when homogenous commodities are traded continuously, marginal–cost pricing results regardless of market structure.

2 Models of Firm Profitability

I begin by discussing two models that have been developed by researchers in the field of industrial organization. I then develop the principal contenders from the fields of finance and natural–resource economics. All four models are very simple. However, my object is to uncover robust and stable equilibrium relationships.

2.1 Models from Industrial Organization

Product–Market Structure

The structure–conduct–performance (SCP) paradigm that dominated IO until the early 1980s held that market structure (the number and size distribution of firms in an industry) determines market conduct (the way in which the firms in that industry interact), which in turn determines firm performance (profitability). Academics from that tradition claimed that market structure was principally influenced by technological factors such as economies of scale and scope, and that the existence of high profit levels in an industry was evidence that the firms in that industry possessed monopoly power.

Researchers in the SCP tradition, which was principally an attempt to assess empirical regularities, often based their assessments on cross-sectional data for "markets" (usually Standard Industrial Classifications or SICs).[8] Typically, they regressed average profit rates on a number of market–wide variables such as indices of horizontal concentration, measures of economies of scale and other barriers to entry, and R&D and advertising intensities. I make no attempt to survey the results of the early studies but instead simply note that the relationship between market structure and firm profitability was generally found to be positive but not necessarily strong.[9]

That literature, which is vast, came under attack in the early 1980s on both theoretical and empirical fronts. Among other things, empiricists pointed out that all of the variables were potentially endogenous and that the models therefore produced correlations that could not be given a structural or causal interpretation. In addition, broad SICs were not really markets, either because they were defined too broadly and contained firms that operated in industries with very different structures, or because the firms that were assigned to each SIC had substantial operations in other SICs. Finally, the accounting data that were typically used to measure profitability were thought to be poor proxies for economic profits.

Theorists on the other hand criticized the SCP paradigm on the grounds that it was not derived from models of optimal decision making of economic agents. More-

[8] See, e.g., Bain (1951, 1956).

[9] For a summary, see Weiss (1974).

over, the game–theoretic revolution that swept the field questioned the notion that market structure could be considered exogenous. Instead, theorists often made assumptions about exogenous conduct (e.g., Cournot or Bertrand behavior) in an attempt to endogenize entry and firm performance. Unfortunately, from an empirical point of view, the predictions obtained from game–theoretic models were very sensitive to the assumptions made concerning conduct, and there were few robust predictions from those models that could be taken to cross–sectional data.

In spite of the fact that the SCP tradition did not originate in a rigorous theory, it is possible to build a theoretical model that yields the SCP predictions. To illustrate, Stigler's (1964) theory of oligopoly is based on the idea that, when firms collude either tacitly or overtly, uncertainty inhibits detection of secret price cuts and facilitates cartel instability. Furthermore, he showed that uncertainty in the percentage gain in sales from undetected price cutting increases with the number of firms in the industry and falls as the inequality of firm market shares rises.[10] Finally, he demonstrated that, if one aggregates the variances of firms' shares of sales, that variable is inversely proportional to the Hirschman Herfidahl index (HHI) of industry concentration.

There were many later studies that attempted to remedy the flaws that characterized the early cross–sectional models.[11] Typically, those studies were based on regression equations of the form

$$\pi_j = \alpha HHI_i + \beta^T x_j + u_j, \tag{1}$$

where π_j is the profit rate of firm j, HHI_i is an index of concentration for the relevant market (i.e., firm j is in market i), and x_j is a vector of other explanatory variables such as measures of minimum efficient scale and advertising and R&D intensities.[12]

Nevertheless, partially as a result of the attacks, the structure–conduct–performance literature has virtually disappeared from mainstream IO research.[13] Indeed, when I searched EconLit for SCP, although I uncovered more than 100 entries dating from the last 15 years, most were published in books or relatively obscure journals. Furthermore, unlike the earlier studies, most applications were not based on a wide range of SICs. Instead, they were heavily concentrated in a few sectors such as agriculture and banking. Finally, I found that, as they became less fashionable in IO, SCP models moved to other disciplines. For example, they have been applied to the public

[10] The latter effect is due to information pooling. In other words, a large firm is similar to several small firms that share information.

[11] For summaries, see Schmalensee (1989) and Scherer and Ross (1990).

[12] Equation (1) is a cross-sectional equation. However, it extends easily to a panel.

[13] In a history of thought piece, Blaug (2001, p.45) states that *The structure–conduct–performance paradigm of yesteryear ... has since been superseded by game theory and transaction cost on the one hand and the Chicago School ... on the other.*

sector (e.g., to local governments) as well as to economies in transition or in the early stages of development.

Many alternative explanations for the positive correlation between market structure and firm profitability have been proposed in the IO literature, and I do not survey them all. However, one explanation – firm market share or size – has received more attention than any of the others. I therefore elaborate on that criticism.

Firm Market Share

In the 1970s, a number of "Chicago–school" economists criticized the SCP paradigm on the basis that its proponents had the causality backwards.[14] Indeed, SCP critics claimed that profits determine concentration rather than the other way around. Their story runs as follows. Markets are workably competitive, but firms are not equally productive. Efficient firms grow and capture large shares of their markets, whereas inefficient firms shrink and eventually exit theirs. As a consequence, the industries in which efficiency differences are greatest have the most asymmetric market structures and the highest horizontal concentration. Moreover, since the large firms in those industries both dominate the market and are more profitable, there is a positive correlation between concentration and profitability. That correlation, however, does not result from the exercise of market power. Quite the contrary, it is a sign that markets are evolving efficiently.

Although the original Chicago criticism was not based on a rigorous theory, later models embody the basic idea. For example, Jovanovic (1982) builds a theoretical model of a competitive market in which firms, who differ in their productivities, enter the market without knowing how productive they are. As they gain experience, they learn about their costs. When expectations are revised upwards, firms grow, whereas when the news is unfavorable, they shrink and eventually exit. The distribution of efficiencies and the maturity of the market therefore determine its structure.

The predictions of market–share and market–power models are thus very different from one another. At one end of the spectrum, conditional on market structure, market share shouldn't matter and horizontal concentration alone should determine profitability. At the other end of the spectrum, conditional on a firm's share of the market, the extent of horizontal concentration in that market should have no effect on its profitability.[15] Nevertheless, one might expect to find empirical support for both theories. Indeed, the simple Cournot model of homogenous products predicts that a firm's price/cost margin will be directly proportional to its market share, and

[14] See, e.g., Demsetz (1973) and Peltzman (1977).

[15] There is an alternative model that predicts a positive relationship between market share and profitability. Bain (1956) argued that small firms might be less profitable because they are unable to take full advantage of economies of scale.

that an index of industry price/cost margins will be directly proportional to the HHI (see, Cowling and Waterson 1976).

In the late 1970s and early 80s, empiricists were able to obtain more disaggregate data that enabled them to assess the relative importance of market structure and market share.[16] A typical study was based on a regression equation of the form

$$\pi_{ji} = \alpha_1 HHI_i + \alpha_2 SHARE_{ji} + \beta^T x_{ji} + u_{ji}, \tag{2}$$

where π_{ji} is the profit rate of firm j in market i, and $SHARE_{ji}$ is the firm's share of that market.

The results of those tests varied; some researchers found that market–structure effects were dominant, whereas others found market–share effects to be more important.[17] Furthermore, although the issue of which effect dominated was hotly debated, it was never satisfactorily resolved.[18]

The differences in opinion can be illustrated by the following quotes:

Profitability is positively associated with a seller's own market share, but there is little evidence, at least in recent richly disaggregate data, of a positive association between profitability and indices of seller concentration independent of the profit — market–share correlation. (Scherer and Ross 1990, p. 446)

Estimates supporting (the correlation between market share and profitability of US firms) may be dominated by a small number of industries with unusually strong positive relations between share and profitability. (Schmalensee 1989, pp. 984-985, text in parentheses added.)

2.2 A Financial Model

Financial economists have also attempted to explain and assess firm profitability. Like members of the Chicago school, they tend to construct models in which markets are workably competitive. Nevertheless, in their models, returns to investing in assets such as firms vary considerably depending on the firms' characteristics. One characteristic that has been emphasized is systematic risk, where systematic means

[16] Many tests were undertaken using the US Federal Trade Commission's line–of–business data, which contains statistics on firms' operations in each of their lines of business.

[17] See Schmalensee (1989) and Scherer and Ross (1990, chapter 11) for summaries of the empirical evidence.

[18] Some researchers replaced firm and industry variables, such as firm market shares and industry concentration ratios, with firm and industry fixed effects (see, e.g., Scott and Pascoe 1986). However, although one normally finds large differences across firms and markets, the use of fixed effects does not allow one to determine the sources of the differences.

risk that is not diversifiable. Specifically, an asset with higher systematic risk should command a higher return.

The simplest model that embodies that notion is the capital asset pricing model or CAPM.[19] According to the CAPM, investors will be willing to hold the jth asset only if the expected return to holding that asset, r_j, equals the risk–free rate of return, r_f, plus an asset–specific risk premium. When the asset is a firm, the return to ownership is the firm's profit, π_j. Furthermore, the risk premium is a linear function of the difference between the rate of return on the market portfolio, r_m, and the risk–free rate, where the market portfolio is a portfolio of all assets in the economy. Formally,

$$\pi_j = r_f + \beta_j(r_m - r_f) + u_j, \tag{3}$$

where u_j is an i.i.d. random variable that captures diversifiable or unsystematic risk.

Equation (3) is clearly a regression equation. The regression coefficient can therefore be interpreted using the regression formula, $\beta_j = \mathrm{COV}(r_m, \pi_j)/\mathrm{VAR}(r_m)$.[20] This formula shows that the risk premium does not depend on the riskiness of the asset *per se*, which is measured by $\mathrm{VAR}(\pi_j)$. Instead it depends on the covariation of the return on asset j with the market portfolio. In other words, only systematic or undiversifiable risk matters. Idiosyncratic risk, u_j, is irrelevant because investors can insure against such risk by diversifying their portfolios.

The CAPM explains why highly risky assets such as gold need not command high rates of return. Gold is a real asset whose return is not highly correlated with, for example, the return to holding a portfolio of stocks. Indeed, when the stock market is expected to plummet, it is not uncommon for investors to switch from financial into real assets. Such behavior leads to low risk premia and can even cause the return to holding real assets to be negatively correlated with r_m. When this is the case, risk premia are negative.

The CAPM predicts that a firm's risk class, not the structure of the market within which it operates, determines profit rates. If there is any possibility that market structure or market share is correlated with systematic risk, it is important to use measures of risk as conditioning variables in tests of IO models.[21] In addition, it is interesting to assess the relationship between risk and return in its own right.

Many researchers from finance have attempted to assess the testable predictions of the CAPM. They tend to reject the model in its simplest form but find support for

[19] See Sharpe (1964) and Lintner (1965). Unlike the early IO models, the CAPM is derived from an equilibrium model of optimal decisions taken by economic agents.

[20] It is assumed that r_f is not a random variable.

[21] Bothwell and Keeler (1976) is perhaps the first study to include systematic risk in an SCP model.

modified versions that include additional explanatory factors.[22] Most of the factors that have been considered, however, are economy wide rather than market or firm specific.

2.3 An Exhaustible–Resource Model

The firms that are used in my assessment of models of profitability are engaged in mining and refining nonferrous metals. In other words, they operate in exhaustible–resource markets. Natural–resource economists have also developed theories of firm profitability.[23] Under the assumption that resource deposits are homogeneous, those theories yield a number of sharp and sometimes surprising predictions. For example, they predict that the profit on the marginal unit should increase exponentially over time, and that there should be *no* systematic relationship between market structure and firm profitability.

It is possible to illustrate the two predictions using a simple stylized model. Suppose that a firm owns a property that contains 100 gold nuggets that can be extracted at zero marginal cost. In a world of certainty, market equilibrium via intertemporal arbitrage requires that the owner be indifferent between extracting a nugget today and putting the proceeds in the bank where it will earn r_f, and waiting to extract the nugget in a future period. This can only be true if the profit on that nugget (the nugget's price in the example) also increases at the rate of interest, r_f. Specifically,

$$\pi_{jt} = (1 + r_f)\pi_{jt-1} = (1 + r_f)^t \pi_{j0}, \tag{4}$$

where π_{jt} is the profit on the marginal unit in mine j in period t.

To understand the second prediction, consider two possibilities: With the first the owner of the property is a monopolist, whereas with the second, the industry is competitive (the 100 nuggets are owned by a large number of small firms). Let monopoly and competitive extraction be Q_m and Q_c, which can be plotted as functions of time. In any equilibrium, the area under each extraction path must equal 100 (since all nuggets must eventually be sold). This implies that the two paths must cross at least once. Moreover, since price will adjust so that consumers are willing to purchase the quantity that is put on the market in every period, monopoly and competitive price paths must also cross. In particular, if the monopolist initially sets a higher price and sells fewer units than the competitors, he will eventually set a lower price and sell more units. Furthermore, it is possible for monopoly and competitive price

[22] For a summary of more general asset–pricing models, see Brennan (1987), and for a survey of empirical tests of more general models, see Huberman (1987).

[23] In particular, see the seminal article by Hotelling (1931).

paths to coincide. For example, Stiglitz (1976) shows that, in the context of the above simple model, when the industry price elasticity of demand is constant, the two paths coincide. When the elasticity of demand increases (falls) over time, however, the monopolist initially charges higher (lower) prices than the competitors.

When deposits are heterogeneous, the situation becomes more complex. In particular, it is not always the case that all ore will eventually be extracted. Nevertheless, in periods in which there is neither rapid technical change nor large discoveries and when exhaustion is not imminent, the profit on the marginal unit should increase at a rate that is positive but less than the rate of interest. The lower rate of increase depends on the extent to which current extraction affects future extraction costs (i.e., on a depletion effect).[24] Furthermore, as with the homogeneous-deposit case, there is no simple relationship between firm profitability and the structure of the market in which the firm operates. In particular, with heterogeneous deposits, if exhaustion is complete, monopoly and competitive price and extraction paths must cross.[25]

There have been many attempts to test the implications of resource scarcity for profitability, and support for the model is limited at best. Those tests tend to be either time–series examinations of prices and profits (e.g., Heal and Barrow 1980, Smith 1981, and Slade 1982) or more structural dynamic models (e.g., Halvorsen and Smith 1984, Farrow 1985, Young 1992, Slade and Thille 1997, and Ellis and Halvorsen 2002).[26] The structural studies, however, require much more detailed data.

There is clearly an abundance of theories of firm profitability. The predicted equilibrium relationships are summarized in table 1. In that table, rows are theories, columns are variables that can be constructed from data, and entries indicate signs of predicted correlations with firm profits. It is useful to adopt a unified empirical framework in which to assess the summarized predictions. I now turn to such an assessment.

3 The Nonferrous–Metal Industries

The production of nonferrous metals consists of several phases. In simple terms, the exploration process produces reserves of metal ore, mining extracts that ore from the ground, smelting and refining produce metal from the extracted ore, and downstream firms (e.g., mills) produce metal products (e.g., sheet and tube) from the output of the refineries and smelters. With the exception of the last phase, most of the large min-

[24] See Levhari and Levitan (1977).

[25] Exhaustion will be complete if, for example, the demand choke price is infinite and if extraction costs are multiplicatively separable in current extraction and remaining reserves.

[26] The Slade and Thille paper combines a test of the theory of exhaustible resources with the CAPM, whereas the Ellis and Halvorsen paper combines that theory with a test of market power.

ing companies are vertically integrated and are involved in all phases of production. However, the exploration phase is much more fragmented than the others. Indeed, many small firms engage only in exploration and sell the deposits that they discover to mining companies. For this reason, in my empirical analysis I limit attention to phases two and three — mining and refining.[27]

The principal nonferrous metals are aluminum, copper, gold, lead, nickel, silver, tin, and zinc. With the exception of aluminum companies, most large mining firms operate in several commodity markets. Multimarket production is due to two factors. First, many ores contain more than one metal (e.g., lead and zinc), and production must be joint. Second, the technology of mining is similar across commodities, which means that a firm with experience with one commodity can easily transfer that experience to another.

There are substantial economies of scale in mining and even more in refining. As a consequence, mining firms are included in lists of the world's largest. However, the markets in which they operate are also large. Indeed, nonferrous–metal markets are worldwide. Due to their geographic size, markets for individual commodities are not highly concentrated. Instead, they range from moderately concentrated to workably competitive.

In addition to economies of scale, entry barriers include the need to possess a scarce resource. In particular, one cannot mine unless one owns an ore deposit. Furthermore, it is uneconomical to enter the market if the deposit that one owns is very small.

Technological breakthroughs revolutionized the mining industry in the late nineteenth and early twentieth centuries. To illustrate, the availability of cheap sources of electricity made aluminum smelting possible, the discovery of froth flotation made recovery of metals from sulfide ores economical, and the advent of large earth–moving equipment made strip mining of low–grade surface deposits inexpensive. In recent years, however, there have been few breakthroughs. Nevertheless, there are constant small, gradual improvements in the technology of mining.

On the production side, mineral commodities are thus similar to one another in many respects. To reiterate, i) they are homogenous commodities that are sold in world markets, ii) their technologies exhibit economies of scale, but minimum efficient scales are not usually large relative to market sizes, iii) they are industrial goods that are not subject to rapid technological change, and iv) their technologies are capital intensive with similar rates of depreciation across commodities.

On the consumption side, in contrast, the principal uses for a given commodity are

[27] Henceforth, I use the word refining to denote smelting and refining and sometimes use the word mining (e.g., mining firm) to denote mining, smelting, and refining.

quite different from those for the others. For example, the largest use of aluminum is in the container and food–packaging sector, whereas the largest use of lead is in storage batteries. This means that each commodity is sold in a different product market. The combination of these factors implies that both geographic and product markets are well defined. In particular, there are 16 world markets corresponding to the eight commodities at two stages of production.

These industries are thus ideal for testing market–structure and market–share models. Indeed, when products are homogeneous and markets are accurately defined, those models are most apt to be capable of capturing cross–sectional and time–series variation in firm profitability. Unfortunately, the industries are less well suited to testing the relationship between risk and return. As we shall see, the problem arises because, although the markets are highly risky, that risk is principally unsystematic.

4 The Data

Prices of each commodity come from the London Metal Exchange (aluminum, copper, lead, nickel, tin, and zinc) and from COMEX (gold and silver).[28] Both yearly and monthly averages of the daily data were used in constructing commodity-market and firm variables.

The data on firms — output of each commodity, company profits, and assets at each stage of production — were obtained from the Raw Materials Group (RMG), a consulting firm in Sweden.[29] These data are somewhat unusual. Indeed, most data–collection agencies publish commodity statistics by geographic region, and those data contain no information on market structure. The Raw Materials Group, in contrast, keeps track of the activities of mining companies. In particular, it tracks mergers and other changes in the complex linkages among mining and refining firms and is consequently a unique source of information on who owns and controls whom.

RMG does not keep track of refining of silver and gold. This omission is perhaps due to the fact that flows of newly refined silver and gold are small relative to the stocks of those commodities that are in circulation. In other words, silver and gold are not discarded but remain in use or are recycled. This lack of data reduces the total number of markets to 14, eight mining and six refining markets.

Most of the data are annual for the 1994-1998 period. 1994 is the first year for which profit data are available, and 1998 was the last year available at the time of purchase of the data. Although 5 years is not long enough to estimate a dynamic model, it is the approximate length of a business cycle.

[28] COMEX is now part of NYMEX.

[29] For more information on the Raw Materials Group, see their web page at www.rmg.se.

An observation is a firm, $i, i = 1, \ldots, I_j$, at a stage of production $j, j = 1, 2$, in a year $t, t = t_i, \ldots, T_i$. Although there is much overlap, the number and identities of the mining and the refining firms differ (thus the j subscript on I). Furthermore, due to mergers and acquisitions, the panel is unbalanced (thus the subscript i on t).

82 mining and 54 refining companies were selected for the analysis. The selection procedure was as follows. For each commodity, year, and phase of production, the 20 largest companies were chosen.[30] All companies that made the cut for at least one year and commodity were initially considered. However, companies were dropped if no profit data were available, which reduced the number of companies in each set. In addition, observations were dropped from the sample if either profit or output data were incomplete, which is another reason why the panel is unbalanced. The final sample contains 320 observations on mining and 210 observations on refining firms.

A number of variables were constructed from the raw data. First consider the *profit rates*. Nominal profits net of taxes are recorded in $US for each firm at each stage of production. To make nominal profits comparable across companies, they were normalized in two ways: they were divided by the firm's total revenue and by its assets (also in $US) and multiplied by 100. Real profit rates were then created by subtracting the annual rate of inflation in OECD countries from the nominal profit rates. The real profit rate in percentage is denoted $RPROFIT_{it}$ when divided by revenue and $APROFIT_{it}$ when divided by assets. For ease of notation, the phase of production has been suppressed here and in what follows.

Four *commodity–market variables* were constructed. The first is the Hirschman Herfindahl index of concentration, $CHHI_{kt}$, which is the sum of the squared shares of output of the firms that operate in commodity market k in year t, multiplied by 10,000. The second is the four-firm concentration ratio, $CCR4_{kt}$, which is the percentage of industry output that is attributable to the four largest firms in that industry and year. The third is the commodity beta, $CBETA_k$, which is calculated as $\mathrm{COV}(r_k, r_m)/\mathrm{VAR}(r_m)$. For this calculation, r_k is the rate of price appreciation for commodity k,[31] and r_m is an average of the total returns (capital gains plus dividends) to holding a number of prominent stock–market indices.[32] Monthly data for the 1990-1998 period were used in estimating the betas. The final commodity–market variable, $CCUMEX_{kt}$, measures cumulative extraction of each commodity.

[30] I limit attention to large companies because, with those companies, economies of scale are apt to be exhausted. In other words, I want to eliminate the possibility that finding a positive relationship between size and profitability is due to economies of scale.

[31] r_k is the capital gain that is associated with holding the processed commodity. A better measure of risk would use shadow rather than market prices (see, e.g., Slade and Thille 1997).

[32] Specifically, the indices are Standard and Poors 500 (US), the Dow Jones Industrial Average (US), FTSE 100 (UK), CAC 40 (France), Hang Seng (Hong Kong), TSE 300 (Canada), Nikkei 500 (Japan), and DAX 30 (Germany).

Firm variables other than profit rates are weighted averages, where the weights are revenue or output shares. In other words, w_{ikt} (W_{ikt}) is the fraction of firm i's revenue (output) that comes from its operations in commodity market k in year t. The first two firm variable are its HHI and its CR4, which are weighted averages of the commodity–market HHIs and CR4s,

$$FHHI_{it} = \sum_k w_{ikt} CHHI_{kt} \tag{5}$$

$$FCR4_{it} = \sum_k w_{ikt} CCR4_{kt}. \tag{6}$$

The third firm variable is i's market share, $FMKTSH_{it}$, which is a weighted average of its shares of the commodity markets in which it operates,

$$FMKTSH_{it} = \sum_k w_{ikt} SHARE_{ikt}, \tag{7}$$

where $SHARE_{ikt}$ is firm i's share of the output of commodity market k in year t. This variable is in percentage. In addition, an alternative measure of market share, $FMAXSH_{it}$, was constructed. This measure is the largest share of any of the markets in which a firm operates. In other words, if a firm were to produce three metals, and if its shares of its three commodity markets were 0.5%, 10%, and 5%, the new measure of market share would be 10.

The fifth firm variable, which is a weighted average of the commodity betas, is a proxy for firm i's risk premium,[33]

$$FBETA_{it} = \sum_k w_{ikt} CBETA_k. \tag{8}$$

Finally, an exponential trend equal to $exp(0.05t)$ was created. A 5% annual real rate of interest was assumed in creating that variable, which is denoted $TREND_t$. This variable should be correlated with profits when deposits are homogeneous. In addition, the variable $CCUMEX_{kt}$ that measures cumulative extraction for each commodity was weighted by the output weights, W_{ikt}, to obtain cumulative extraction for each firm, $FCUMEX_{it}$. This variable is a better measure of depletion when deposits are heterogeneous.

A number of comments on the data are in order. First, much has been written about the relative merits of various measures of profitability. My preferred measure is $RPROFIT$, which is often equated with the price/cost margin in the SCP literature. The principal drawback to using that measure is related to the fact that

[33] An alternative measure of the risk premium would use company profits as r_i. However, there are only five observations from which to construct that variable.

rates of depreciation and competitive rates of return differ across industries (see, e.g., Schmalensee 1989). Within the mining sector, however, this problem is not severe. An advantage of normalizing by sales is that data on sales are much more reliable than data on capital stocks. More generally, however, the use of any accounting data has its limitations, which are no more or less troubling here than in other applications.

Profits are recorded by firm rather than by each market in which the firm operates. This lack of detail is less troublesome than it might at first appear. Indeed, it is always difficult to allocate joint costs. However, the problem is particularly acute in the mining industry. To illustrate, when an ore is extracted, it can be difficult to know *a priori* which metals the ore will contain. Moreover, even when the metals are known, the grades of each can usually only be determined *ex post*. Furthermore, when mining firms allocate joint costs across commodities, they often assign all joint costs to the primary or high–revenue commodity. This procedure makes economic sense since, at the margin, capacity expansions are usually driven by a desire to produce more of the more profitable commodity. However, for the purpose of this study, the practice has the drawback of implying that, as the price of a metal increases, its costs can also appear to increase, and low–value metals (byproducts) appear to be produced at very low cost. For this reason, even if available, profits by line of business would be unreliable. Revenues, in contrast, can be accurately constructed for each market, and those revenues are used to form the weights that enter into the construction of firm variables.

In contrast to the measure of profitability, the measures of market structure that are used here are unusually accurate. Indeed, both product and geographic markets are well defined, and the raw data on firm output is recorded at the appropriate market level. The data were collected at two phases of production. Although refining markets are clearly important, due to the prevalence of vertical integration, one might question the relevance of mining markets. For this reason, the empirical model is estimated for each phase of production separately.

SCP studies typically employ many other control variables. For example, it is common to include measures of R&D and advertising intensities and dummy variables that indicate whether the product is durable and whether it is a consumer or a producer good. With my data, due to the similarity of the markets, there is no need to control for those factors.

Finally, one problem with the use of a single cross section is that the relationships of interest are apt to vary over the business cycle. Using a panel that is at least as long as a typical business cycle, as is done here, eliminates that problem.

Table 2 gives summary statistics by phase of production for each of the four endogenous variables that are used in the analysis. The variables have been scaled so that they have similar means. In particular, the HHI has been divided by 100 so that it ranges between 0 for perfect competition and 100 for monopoly, and the betas have been multiplied by 100.

The table shows that, at least in mining, the variation in profit rates is substantially greater than the variation in the other variables. Moreover, in spite of the fact that mining and refining are highly risky activities, the betas are small, which is not unexpected. The table also shows that the largest firms in the sample (as measured by FMKTSH) control just over 20% of the markets in which they operate. Finally, markets vary between fairly competitive (the usual HHI = 180) to moderately concentrated (HHI = 1880).

5 The Empirical Model

Much of the variation in the data is cross sectional, and differences across firms are relatively stable. This means that one can interpret cross sectional variation in the firm variables, RPROFIT, FHHI, FMKTSH, and FBETA, as long–run stable differences. Unfortunately, this also means that all of those variables are potentially endogenous. Moreover, without firm–level exogenous variables, such as firm–specific factor prices, it is difficult to find instruments that could be used in a regression analysis. For this reason, a descriptive approach is adopted. In particular, I use principal components to analyze variations in the data.

Principal components transforms a set of variables, X, into a new set of variables, Z, that are pairwise uncorrelated. Furthermore, the first of the Z variables (or components) has the maximum possible variance, the second has the maximum among those that are uncorrelated with the first, and so forth. If there are M variables in the original data, $x_\ell, \ell = 1, \ldots, M$, and if those variables are not perfectly collinear, it is always possible to find M linearly–independent components, $z_\ell, \ell = 1, \ldots, M$, that explain all of the variation in X.[34] However, it is often the case that $m < M$ components explain a very large fraction of the variation.

The components are linear combinations of the original variables, and it is not always possible to give each an economic interpretation. However, sometimes one can. Whether or not one can interpret the components, it is possible to calculate correlation coefficients between each component and each of the original variables. One can also calculate the proportion of the variation in each x_ℓ that is associated with each component.

The procedure adopted here is to consider a matrix X that consists of the four

[34] For the expression "the variation in X" to make sense, the constituent vectors should be measured in the same units. The variables in this study can be loosely interpreted as percentages.

endogenous firm variables that are shown in table 2 plus the exponential trend ($M = 5$). Five principal components are computed. The number m that is retained is chosen as the smallest number of components that explains at least 95% of the variation in X. An attempt is then made to interpret the retained components, and the correlations between the original variables and those components are calculated.

6 Results

6.1 The Principal–Component Analysis

Table 3 shows the contribution of the first 3 principal components to the total variation in X, both individually and jointly, for both phases of production. One can see that the first component accounts for more than 60% of the variation, the second for about 10–30%, and the third for 4–7%. There is therefore very little variation left for the last two components to explain. Moreover, given the rule for retention of components, the first two are retained for mining and the first three are retained for refining.

Table 4 contains correlation coefficients between each variable in X and the retained components for both phases of production. The first thing to notice is that the first component is essentially profit, $RPROFIT$, the second is essentially the risk premium, $FBETA$, and, for refining, the third is essentially market share, $FMKTSH$. Indeed, the correlation between those variables and the respective components is at least 0.95 in all five cases. In particular, this means that $RPROFIT$ and $FMKTSH$ are orthogonal to one another in refining. In other words, in those industries there is no systematic relationship between a firm's market share and its profitability, and, within a market, smaller firms are just as profitable as larger ones.

This finding is contrary to what the proponents of "market–share" models predict. In particular, even though the firms in the sample are all large, they control between 0.1 and 21% of the markets in which they operate. Variation in $FMKTSH$ is therefore substantial. That variation, however, is uncorrelated with profitability. Furthermore, this finding implies that, if there is a correlation between profitability and market structure in those markets, it cannot be due to a failure to control for the common causal factor, market share.

Since one can interpret the first three components as profit, systematic risk, and market share respectively, it is possible to investigate how those variables are related to the remaining variables. In particular, the correlations in table 4 reemphasize the fact that a firm's market share is not correlated with its profitability as embodied in the first component.

Table 4 shows that the market–structure variable, $FHHI$, is positively correlated with all of the retained components. Moreover, those correlations are significant at 1%. The implications are that firms that operate in concentrated markets are more profitable and command higher risk premia, and concentrated refining markets have larger firms (in relative terms). The first of these conclusions gives support to the traditional structure–conduct–performance paradigm. In other words, market structure matters. The third, which is more of an accounting identity, simply implies that if some firms control large shares of a market, that market is apt to be concentrated.

The measure of systematic risk, $FBETA$, which is proportional to the firm's risk premium, is positively correlated with profits in mining but not in refining. The results are therefore mixed, and the CAPM receives only partial support. This finding, however, might simply be due to the fact that all of the estimated betas are small, which means that none of the firms in the sample commands a high rate of return as compensation for bearing risk.

Finally, profits do not exhibit an upward trend. In fact, there appears to be a downward trend in mining profitability.

6.2 Sensitivity Analysis

The robustness of the results was assessed in a number of ways. In particular, the calculations were redone using profit divided by assets, $APROFIT$, as an alternative measure of profitability, using the four–firm concentration ratio, $FCR4$, as an alternative measure of market structure, the variable, $FMAXSH$, as an alternative measure of market share, and with $FCUMEX$ as an alternative measure of scarcity. In addition, I created a firm–specific trend that depends on the firm's risk premium, $FBETA$. Finally, the commodity risk premium, $CBETA$, was allowed to vary by year. None of the alternative specifications, however, changed the qualitative nature of the conclusions.

For the final specification test, I performed the calculations for each year in the data. When I did this, I found that the relationship between profits and market structure is stronger in economic downturns. This finding might indicate that firms in all industries do well in upturns, and in fact I find that profits are positively related to industrial production. Firms in concentrated industries, however, might have better methods of disciplining each other and might therefore be less likely to suffer profit losses in periods when demand is falling.

7 Conclusions

Four models of firm profitability were surveyed, each originating in a different tradition and each focusing on a different determinant of profitability. The first, the familiar structure–conduct–performance (SCP) model, predicts that the structure of the market in which the firm operates will be the most important determinant of its profits. The second predicts that firms with large market shares will be profitable and that failure to condition on market share will bias the results of tests of SCP models. The third singles out a firm's risk class and predicts that firms that must bear more systematic risk will earn higher rates of return. Finally, the fourth predicts a temporal increase in profits as the firm's reserves are depleted and scarcity rents are earned. Although at times mutually contradictory, the predictions from all four models can be rationalized by rigorous theories. The question of what actually determines firm profitability is therefore an empirical issue.

Using panel data from nonferrous mining and refining markets, I am able to assess each prediction. I find strong support for the first (SCP) model. Indeed, firms' profits are positively and significantly related to the structures of their markets, and this relationship holds in all specifications that were estimated. A firm's market share, in contrast, is found to be uncorrelated with its profitability. This means that, not only is market share not an important determinant of profitability in nonferrous–metal markets, but also that the correlation between market structure and profits is not spurious. A firm's risk premium, which is a measure of the systematic risk that it must bear, is found to be positively correlated with profitability in mining but not in refining. The latter finding, however, is perhaps due to the fact that in the industries studied, there is not much variation in risk premia across firms and time periods. Finally, there is no evidence of a temporal increase in profits. These empirical regularities are summarized in the last row of table 1 under the heading of 'In the Data.'

It is perhaps worth elaborating on the second finding. It is somewhat surprising that market share does not affect profitability in these industries. In particular, since low-cost firms should expand more in an upturn and contract less in a downturn, most economic models predict that size, even when not the principal determinant of profitability, will be positively correlated with it. My counterintuitive finding can be explained, however, if some firms are capacity constrained while others are not. Under that hypothesis, output expansions will be undertaken by the unconstrained firms, which need not be the low-marginal-cost firms. Furthermore, this is not just a short–run phenomenon since, in exhaustible–resource industries, capacities are more apt to be determined by reserves than by marginal costs.

I began by stating that competition authorities rely heavily on concentration indices as tools for determining potential competitive harm. The issue of whether this practice is justified is clearly not resolved. Most economists would agree, however, that it does not work well when products are differentiated. For this reason, authorities are starting to supplement the conventional analysis with merger simulations and other more sophisticated evaluation methods. This practice, however, is also problematic. For example, the results of merger simulations are sensitive to the choice of functional form for demand equations. Furthermore, the predictions that are obtained depend heavily on the cross–price elasticities that are used. Unfortunately, the magnitudes of estimated elasticities are sensitive to the choice of the outside good as well as to the number of alternatives considered.[35] When products are homogeneous, shares may thus be as good indicators as simulations. It is therefore unlikely that the traditional approach will be totally abandoned in the near future.

[35] A general discussion of the alternatives is beyond the scope of this paper. Interested readers can look at, e.g., Slade (2003).

References Cited

Allaz, B. and Villa, J-L. (1993) "Cournot Competition, Forward Markets, and Efficiency," *Journal of Economic Theory*, 59: 1-16.

Bain, J.S. (1951) "Relation of Profit Rate to Industry Concentration: American Manufacturing," *Quarterly Journal of Economics*, 65: 293–324.

Bain, J.S. (1956) *Barriers to New Competition*, Harvard University Press.

Blaug, M. (2001) "Is Competition Such a Good Thing? Static Efficiency versus Dynamic Efficiency," *Review of Industrial Organization*, 19: 37–48.

Bothwell, J.L. and Keeler, T.E. (1976) "Profits, Market Structure, and Portfolio Risk," in *Essays in Industrial Organization in Honor of Joe S. Bain*, R.T. Masson and P.D. Qualls (eds.) Ballinger.

Brennan, M.J. (1987) "Capital Asset Pricing Model," in *The New Palgrave, Finance*, Eatwell, J., Milgate. M., and Newman, P. (eds) New York: W.W. Norton.

Cowling, K. and Waterson, M. (1976) "Price–Cost Margins and Market Structure," *Economica*, 43: 267–274.

Demsetz, H. (1973) "Industry Structure, Market Rivalry, and Public Policy," *Journal of Law and Economics*, 16: 1–10.

Ellis, G.M. and Halvorsen, R. (2002) "Estimation of Market Power in a Nonrenewable Resource Industry," *Journal of Political Economy*, 110: 883–899.

Farrow, S. (1985) "Testing the Efficiency of Extraction from a Stock Resource," *Journal of Political Economy*, 93: 452-487.

Halvorsen, R. and Smith, (1984) "On Measuring Natural-Resource Scarcity," *Journal of Political Economy*, 92: 954-964.

Heal, G. and Barrow, M. (1980) "The Relationship Between Interest Rates and Metal-Price Movements," *Review of Economic Studies*, 47: 161-182.

Hotelling, H. (1931) "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39: 137–175.

Huberman, G, (1987) "Arbitrage Pricing Theory," in *The New Palgrave, Finance*, Eatwell, J., Milgate. M., and Newman, P. (eds) New York: W.W. Norton.

Jovanovic, B. (1982) "Selection and the Evolution of Industry," *Econometrica*, 50: 649–670.

Levhari, and Levitan, (1977) "Notes on Hotelling's Economics of Exhaustible Resources," *Canadian Journal of Economics*, 10: 177–192.

Lintner, J. (1965) "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 47: 13–47.

Peltzman, S. (1977) "The Gains and Losses from Industrial Concentration," *Journal of Law and Economics*, 20: 229–263.

Schmalensee, R. (1985) "Do Markets Matter Much?" *American Economic Review*, 75: 341–351.

Schmalensee, R. (1989) "Inter–Industry Studies of Structure and Performance," in *Handbook of Industrial Organization*, R. Schmalensee and R.D. Willig (eds.) Elsevier, 951–1009.

Scherer, F.M. and Ross, D. (1990) *Industrial Market Structure and Economic Performance*, Third Edition, Houghton Mifflin.

Scott, J.T. and Pascoe, G. (1986) "Beyond Firm and Industry Effects on Profitability in Imperfect Markets," *Review of Economics and Statistics*, 68: 284–292.

Sharpe, W.F. (1964) "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance*, 19: 425–442.

Slade, M.E. (1982) "Trends in Natural-Resource Commodity Prices: An Analysis of the Time Domain," *Journal of Environmental Economics and Management*, 9: 122-137.

Slade, M.E. and Thille, H. (1997) "Hotelling Confronts CAPM: A Test of the Theory of Exhaustible Resources", *Canadian Journal of Economics*, 30: 685–708.

Slade, M.E. (2003) "Market Power and Joint Dominance in UK Brewing," *Journal of Industrial Economics*, forthcoming.

Smith, V.K. (1981) "The Empirical Relevance of Hotelling's Model for Natural Resources," *Resources and Energy*, 3: 105-117.

Stigler, G.J. (1964) "A Theory of Oligopoly," *Journal of Political Economy*, 72:

Stiglitz, J. (1976) "Monopoly and the Rate of Extraction of Exhaustible Resources," *American Economic Review*, 66: 655–661.

Weiss, L.W. (1974) "The Concentration Profits Relationship and Antitrust,"in *Industrial Concentration: The New Learning*, H.J. Goldschmidt and J.F. Weston (eds.) Little Brown.

Young, D. (1992) "Cost Specification and Firm Behavior in a Hotelling Model of Resource Extraction," *Canadian Journal of Economics*," 25: 41-59.

Table 1: **Equilibrium Predictions**
Relationship with Firm Profit

Model Variable	HHI	Market Share	Risk	Trend	Cumulative Production
SCP	+	0	0	0	0
Bain (1951, 56)					
Firm Efficiency	+	+	0	0	0
Demsetz (73), Peltzman (77)					
CAPM	0	0	+	0	0
Sharpe (64), Lintner (65)					
Exhaustible Resource					
Hotelling (29)					
Homogeneous Deposits	0	0	0	+	0
Stiglitz (76)					
Heterogeneous Deposits	0	0	0	+	+
Levhari & Levitan (77)					
In the Data	++	0	+?	0	0

Table 2: **Summary Statistics for Firm Variables**

Mining

Variable	Mean	S.D.	Min	Max	Coef. of Variation
Real Profit Rate (RPROFIT)	2.5	18.5	-91.3	71.8	7.40
Firm HHI (FHHI)	5.1	2.8	1.8	18.8	0.55
Firm Market Share (FMKTSH)	3.1	3.9	0.1	21.0	1.26
Firm Beta (FBETA)	2.3	6.7	-7.4	17.7	2.91

Refining

Variable	Mean	S.D.	Min	Max	Coef. of Variation
Real Profit Rate (RPROFIT)	3.8	11.0	-46.2	62.3	2.89
Firm HHI (FHHI)	5.3	2.5	2.8	15.2	0.47
Firm Market Share (FMKTSH)	4.4	3.7	0.2	21.2	0.84
Firm Beta (FBETA)	3.4	7.5	-7.5	17.7	2.21

Profit rate in percentage.
HHI is divided by 100.
Size in percentage of market.
Beta is multiplied by 100.
Coefficient of variation = standard deviation/mean.

Table 3: **Contribution of Principal Components to Variation**

Mining			
Component	1	2	3
Contribution	0.84	0.12	0.04
Cumulative Contribution	0.84	0.96	0.99
Refining			
Component	1	2	3
Contribution	0.62	0.29	0.07
Cumulative Contribution	0.62	0.91	0.98

Table 4: **Correlation Coefficients**

Mining

	Component	
Variable	1	2
$RPROFIT$	0.996**	-.028
$FHHI$	0.217**	0.674**
$FMKTSH$	0.101	0.457**
$FBETA$	0.177**	0.967**
$TREND$	-.143*	0.042

Refining

	Component		
Variable	1	2	3
$RPROFIT$	0.995**	0.070	0.004
$FHHI$	0.271**	0.461**	0.252**
$FMKTSH$	-.049	0.075	0.991**
$FBETA$	0.141	0.989**	-.031
$TREND$	0.014	-.018	0.079

* denotes significance at 5%.
** denotes significance at 1%.