Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data

by

Tor Jakob Klette [z)]
(Department of Economics, University of Oslo and Statistics Norway)

May 25, 1998

Abstract: This paper presents a somewhat new econometric framework that permits simultaneous estimation of price-cost margins, scale economies and productivity from a panel of establishment data. The econometric model contains only a few, economically interesting parameters to be estimated, but it is nevertheless consistent with a °exible (translog) underlying technology, quasi-¯xed capital and the presence of persistent di®erences in productivity between establishments. The econometric framework is applied to study market power, scale economies and productivity di®erences in a number of manufacturing industries in Norway. The results reveal statistically signi¯cant, but quite small, margins between price and marginal costs in most manufacturing industries. No industry exhibits increasing returns to scale; the average plant in most industries seems to face constant or moderately decreasing returns to scale. There is more variation in market power within the fairly narrow industry groups investigated compared to the variation between the industry groups. The results show that ¯rms with higher market power tend to be less productive.

JEL classi¯cation: D40, C23.
Keywords: Market power, scale economies, productivity, plant level data, GMM, panel data.

---

z Department of Economics, University of Oslo, P.O. Box 1095 Blindern, N { 0317 Oslo, Norway. E-mail: t.j.klette@econ.uio.no. Tel: + 47 22 85 51 34. Fax: + 47 22 85 50 35.
This paper has bene¯tted from comments from Bruno Crepon, Erik Bi¿rn, Rolf Golombek, Erling Holm¿y, Torbj¿rn H½geland, Jacques Mairesse, Arne Melchior, Michael Waterson, and in particular Jarle M¿en. Special thanks to Zvi Griliches for advice and encouragement. The research was ¯nanced by the Research Council of Norway.

1 Introduction

Theoretical studies of the nature and consequences of imperfect competition and scale economies are central throughout the economic discipline. Still, the appropriate methodology to study the empirical signi¯cance of scale economies and price-cost margins remains an unsettled issue in econometrics despite its long history; see e.g. Griliches and Mairesse (1998) and Hyde and Perlo® (1995). This paper presents an econometric framework - drawing on Hall (1988, 1990) - to simultaneously estimate price-cost margins and scale economies using a panel of ¯rm or plant level data.

The empirical part of this paper examines the importance of market power and scale economies in Norwegian manufacturing. The manufacturing sector in Norway is highly exposed to competition in export markets and from imports in domestic markets. On the other hand, it has been noticed in several case studies that regulations and anti-competitive behavior have seriously restricted competition for a number of important products such as fertilizers, cement, ships, oil rigs and other manufacturing products, at least in the domestic markets and several case studies have identi¯ed markets in Norwegian manufacturing with signi¯cant market power[1]. To the extent that trade is restricted, one might expect potential monopoly rents will induce excessive entry and therefore unexploited scale economies.

The case studies have focused on narrowly de¯ned markets chosen because they are expected to be most seriously a®ected by imperfect competition, and these market segments do not seem to be representative for the degree of market power in the Norwegian manufacturing sector, according to the empirical results presented below. The main empirical ¯nding in this study is that problems with market power and unexploited scale economies seem small on average in most Norwegian manufacturing industries. The econometric model is estimated on a comprehensive panel of establishment data covering most of manufacturing over the period 1980-90. The data set permits extensive testing of the validity of the econometric model. The preferred estimates reveal small, but statistically signi¯cant market power in a majority of the 14 industries considered. Most of the margins belong to the interval 5-10 percent. The results show little evidence of scale economies. None of the industries reveal estimates of the scale elasticity above one, while the plants in several of the industries appear to face moderate decreasing returns to scale.

However, the empirical results just discussed refer to the average price-cost margins and scale

[1]See e.g. Gabrielsen's (1989) study of the domestic market for fertilizers and S¿rgard's (1997, ch. 5.5) study of the cement industry. Holm¿y et al. (1993) and S¿rgard's book (1997) give references also to some other case studies. See section 5.3 for further comments on the case studies.

elasticities within each industry. It is likely that the price-cost margins and scale coefficients vary, perhaps substantially, within each of the industries analyzed. Consider a well known example from the U.S.; concerns about market power in the software industry is focused on Microsoft rather than on the average software producer. To examine within-industry variations in market power, I use a random coefficient framework that allows for differences in market power and scale economies across firms within each industry. The estimates reveal more variation in market power and scale economies within an average industry, as compared to variations between industries. Interestingly, I find that firms with higher market power also tend to be less productive, as was found by Nickell (1996) in a recent study. This suggests that lack of competition does not only create inefficient price setting, but also productive inefficiencies and slack.

As mentioned, the framework presented in this paper has been inspired by Hall (1988, 1990). Hyde and Perloff (1995) have argued that "the key weakness of Hall's approach ... is that one must maintain the assumption of constant returns to scale". The present paper shows how Hall's approach to estimation of market power can easily be extended to account for scale economies, and also the quasi-fixity of capital. When estimating price cost margins, it is essential to adjust for scale economies, as the estimate of scale economies will tend to be tightly linked to the estimate of the ratio of price and marginal costs. For instance, with price and average costs as the observable point of departure, overestimating the scale economies will imply underestimated marginal costs, providing an overestimated price-marginal cost ratio. Considering the large order of magnitude of Hall's estimate of scale economies (Hall, 1990), keeping constant returns as a maintained hypothesis in his study of price cost margins (Hall, 1988, 1990) questions the consistency of the estimates.

Most studies following Hall (1988) are based on industry level data[2]. However, micro level data are essential for a simultaneous study of price-cost margins and scale economies, since scale economies at the industry level are affected by externalities[3], entry and exit. These are phenomena that have little to do with the scale economies relevant for the firms' price setting decisions. The use of plant or firm level panel data also has an additional benefit compared to studies based on industry level data as the model is implemented at the level for which it is constructed. This eliminates the well-known and important problem of aggregation and allows one to control for permanent productivity differences between plants (by "fixed effects"). Permanent productivity differences between plants are known to be present in most data sets

[2] Levinsohn (1993) and Harrison (1994) are two notable exceptions.

[3] Bartelsman, Caballero and Lyon (1991) interpret Hall's (1990) scale estimates in terms of external economies.

on establishments and ¯rms[4], and their presence seriously questions the interpretation of results from aggregate data that are based on the notion of a representative ¯rm.

More generally, the framework and analysis presented here goes beyond Hall's studies and the related studies on four accounts: (i) The framework is extended to allow for scale economies in the estimation of the margin between price and marginal costs, as discussed above. (ii) The model draws on the index literature for productivity measurement to allow for a °exible (translog) technology. (iii) The estimates are obtained from microdata accounting for persistent productivity di®erences between plants, as discussed above. (iv) The instrumental variables used in the GMM-estimation are extensively tested and di®er entirely from the instruments used by Hall which have been seriously questioned[5].

Section 2 spells out the theoretical framework. Section 3 presents the construction of the data set and variables. Stochastic assumptions, speci¯cation testing and other econometric issues are considered in section 4. The empirical results on market power and scale economies are presented and discussed in section 5. Section 6 provides the analysis of cross sectional heterogeneity in market power and productivity within industries. Section 7 gives some concluding remarks.

2 The theoretical model

2.1 Price-cost margins, scale economies and quasi-¯xed capital

The ¯rms[6] within an industry are assumed to be constrained by a production function $Q_{it} = A_{it}F_t(X_{it})$, where Q_{it} and X_{it} represent output and a vector of inputs for ¯rm i in year t. A_{it} is a ¯rm-speci¯c productivity factor, while $F_t(\cdot)$ is a part of the production function common to all ¯rms. The time subscript on the F- function indicates that the function can change freely between years. That is, the model does not impose constraints on the form of technical progress that is common across the ¯rms within the industry, and the model is consequently consistent e.g. with factor augmenting technical progress. In section 4.2 I will introduce constraints on the idiosyncratic changes in technology, i.e. the ¯rm-speci¯c changes that deviate from the industry wide changes in technology.

Using a version of the multivariate, generalized mean value theorem[7], the production function

[4] See Baily, Hulten and Campbell (1992) for a study of the di®erences and dynamics of plant level productivity in U.S. manufacturing and M¿en (1998) for a similar study for Norwegian manufacturing. See also Griliches and Mairesse (1998).

[5] See Abbott, Griliches and Hausmann (1988) for a critiscm of Hall's instrument set.

[6] I will in the theoretical section use the term ¯rm rather than plant, even though the plant is the unit of observation in the empirical analysis. As can be seen from Table 1, a large majority of the plants belong to single-plant ¯rms.

[7] Cf. Berck and Syds½ter (1991, p. 11) for a statement of the generalized mean value theorem. The extension to the multivariate case is straight forward, as suggested in e.g. Thomas (1968, p.545).

relationship can be expressed in terms of logarithmic deviations from a point of reference. This point of reference can be thought of as the level of output and inputs for the representative firm. Rewriting the production function relationship in terms of logarithmic deviations from the representative firm, we have

$$\hat{q}_{it} = \hat{a}_{it} + \sum_{j \in M} \bar{\alpha}^j_{it} \hat{x}^j_{it}; \tag{1}$$

where a lower case letter with a hat is the logarithmic deviation from the point of reference of the corresponding upper case letter. E.g., $\hat{q}_{it} \equiv \ln(Q_{it}) - \ln(Q_t)$, where Q_t is the level of output for the representative firm, i.e. at the reference point. In the empirical application, this reference point has been chosen as the year specific average value of output within the industry. A similar (industry-year) average value is used as a reference point for each of the inputs. I will denote this reference vector for the inputs by $X_t = \{X^1_t, X^2_t, \cdots, X^m_t\}$. M denotes the set of (the m) inputs. $\bar{\alpha}^j_{it}$ is the output elasticity for factor j [8] evaluated at an internal point ($\bar{X}_{it}$) between X_{it} and the reference point X_t [9]. I use the notation that a bar over a variable such as $\bar{\alpha}^j_{it}$ indicates that it is evaluated at the internal point.

Let me briefly explain the motivation behind the use of a mean value theorem rather than a first or second order Taylor approximation in the derivation above. Equation (1) is a relationship in terms of cross sectional differences in outputs and inputs between firms, and such cross sectional differences in outputs and inputs can be of the magnitude of several hundred percent in many industries. Truncating a Taylor approximation after the first or second order term might be problematic with such large differences in inputs[10]. Equation (1), which is derived by using the mean value theorem, is a priori suitable for samples with any size of the cross sectional differences in output, productivity and inputs ($\hat{q}_{it}, \hat{a}_{it}$ and $\hat{x}^j_{it}$). I will return to this issue in section 4.

According to basic producer theory, profit maximizing behavior requires that marginal costs should be equal to the marginal revenue product. I assume that the firm has some market power

---

[8] That is

$$\alpha^j_{it} \equiv \left(\frac{X^j_{it}}{F_t(X_{it})} \frac{\partial F_t(X_{it})}{\partial X^j_{it}} \right)_{X_{it} = \bar{X}_{it}} :$$

where the point $\bar{X}_{it}$ will be defined below.

[9] That is, the point ($\bar{X}_{it}$) belongs to the convex hull spanned by the coordinates $\{X_{it}, (X^1_t, X^2_{it}, X^3_{it} \cdots, X^m_{it}), (X^1_{it}, X^2_t, X^3_{it} \cdots, X^m_{it}), \cdots (X^1_{it}, X^2_{it}, X^3_{it} \cdots, X^m_t), X_t\}$. Cf. e.g. Thomas (1968, p.545).

[10] Consider the case with only one input X, i.e. $Y = F(X)$, which can be rewritten in terms of log output and input as $y = f(x)$: Take a Taylor expansion from a reference firm with (y_0, x_0) as output and input, we get $y = y_0 + f'(x_0)(x - x_0) + 1/2 f''(x_0)(x - x_0)^2 + \cdots$. Now, $(x - x_0) = \ln(X/X_0)$ is a number which can exceed one in the cross-sectional dimension, and it is clear that strong restrictions on the derivatives of the f-function is needed for a first or second order Taylor expansion to be an adequate approximation.

in the output markets, while the firm act as a price taker in the input markets when determining its factor inputs. Notice that this assumption is perfectly consistent with a bargaining situation where the firm and the union bargain over the wage rate, while the firm unilaterally determines the number of hours employed. Such a bargaining model has been widely considered as the appropriate model for studies of wage formation in Norwegian manufacturing[11]. First order conditions with these behavioral assumptions imply that

$$A_{it}\frac{\partial F_t(X_{it})}{\partial X_{it}^j} = \frac{W_{it}^j}{(1 - 1/\epsilon_{it})P_{it}} \tag{2}$$

where W_{it}^j is the factor price for input j, while the denominator on the right hand side is marginal revenue. That is, P_{it} is the price of output, while ϵ_{it} is the (conjectured) price elasticity of demand[12]. According to the theory of imperfect competition, the factor $(1 - 1/\epsilon_{it})^{-1}$ represents the ratio of price and marginal costs. Denoting this ratio between price and marginal costs by μ_{it}, and using the set of first order conditions in equation (2), we have that

$$\begin{aligned}\tilde{\alpha}_{it}^j &= \mu_{it}\frac{\tilde{W}_{it}^j\tilde{X}_{it}^j}{\tilde{P}_{it}\tilde{Q}_{it}} \\ &= \mu_{it}\,\tilde{s}_{it}^j,\end{aligned} \tag{3}$$

where $\tilde{s}_{it}^j$ is the cost share of input j relative to total revenue.

Various rigidities make it dubious to assume that (3) holds for capital, i.e. to impute the marginal product of capital from observed prices on new equipment, tax rules, interest and depreciation rates[13]. This problem can be handled as follows: The elasticity of scale in production is defined by

$$\tilde{\mu}_{it} = \sum_{j \in M}\tilde{\alpha}_{it}^j. \tag{4}$$

Using (3), it follows that

$$\tilde{\alpha}_{it}^K = \tilde{\mu}_{it} - \sum_{j \neq K}\mu_{it}\tilde{s}_{it}^j \tag{5}$$

[11] This bargaining framework was first introduced as a model for wage formation in Norwegian manufacturing by Hoel and Nymoen (1988).

[12] This price elasticity should be interpreted in a broad sense, incorporating the "conjectured price and quantity responses" of the competitors. Bresnahan (1989) has emphasized the generality of this formulation in empirical work.

[13] Similarly, one could clearly argue that adjustment costs for labor should also be accounted for. To my knowledge, attempts to estimate the adjustment costs for labor in a production function relationship have had little empirical success, and have not been explored within this study.

Notice that the output elasticity of capital as constructed in (5) will vary across firms and over time. If we for the moment neglect the randomness in μ_{it} and μ_{it}, equation (5) has the implication that the capital elasticity will *cet. par.* be high when e.g. the labor elasticity is low, and vice versa. This is quite sensible as a low labor elasticity tend to reflect shortage of capital, i.e. a situation with a high capital elasticity.

Applying (3) for the non-capital inputs and (5) for capital, it follows that (1) can be rewritten

$$\hat{q}_{it} = \hat{a}_{it} + \sum_{j \neq K} \mu_{it} s^j_{it} (\hat{x}^j_{it} - \hat{x}^K_{it}) + \eta_{it} \hat{x}^K_{it} \tag{6}$$

Using this relationship and adding the stochastic assumptions to be presented below give the econometric model to be estimated.

To summarize; only mild regularity conditions are imposed on the production technology in order to derive (6). The model is consistent with non-constant returns to scale and the presence of market power as price can exceed marginal costs. The model allows for the possibility that capital is not fully adjusted to its equilibrium value, but is considered (quasi-) fixed while the firm solves its short run profit maximizing problem. η_{it} and μ_{it} have the interpretation of the scale elasticity and the ratio of price to marginal costs.

2.2 A few remarks on related studies

In a recent study, Roeger (1995) has provided "an alternative method for estimating a markup of prices over marginal costs that avoids certain difficulties inherent in [Hall's] method of estimation". Roeger's estimating procedure can be derived as follows: Consider the markup (μ) of price (P) over marginal cost (C_Q): $P = \mu C_Q$. Assuming constant returns to scale, we have that marginal cost is equal to average cost, i.e. $C_Q = C/Q$. Combining these two expressions, it follows that the markup can be written

$$\mu = \frac{PQ}{C}. \tag{7}$$

Roeger considers a cost function with wages (W) and capital rental costs (R) as its arguments, i.e. $C = C(W; R)$. Instead of using (7) directly, he rewrites (7) as

$$\Delta y = B \Delta x; \tag{8}$$

where $\Delta y = \Delta q + \Delta p - s^L(\Delta l + \Delta w) - (1 - s^L)(\Delta k + \Delta r)$, $\Delta x = \Delta q + \Delta p - (\Delta k + \Delta r)$, and $B = 1 - 1/\mu$; where a Δ in front of a variable corresponds to its logarithmic difference, e.g. $\Delta q = dQ/Q$, and s^L is labor's cost share[14]. Roeger argues that estimating (8) is advantageous,

[14] To derive (8), one must use the relationship $dC(W; R)/C = 1/C(C_W\, dW + C_R\, dR)$. With Sheppards Lemma (i.e. $C_W = L$ and $C_R = K$), it follows that $dC(W; R)/C = s^L\, dW/W + (1 - s^L)\, dR/R$.

as (8) "does not require the strong identifying assumptions found in Hall's analysis", in particular the exogeneity assumptions for the instrumental variables. However, Roeger does not point out that he could have disposed of estimation altogether, by focusing directly on (7). From (7) we can directly calculate the markup, given the assumptions maintained by Roeger that (i) constant returns to scale prevail, (ii) we can impute the rental costs for capital, and (iii) capital is fully adjusted to the rental costs. These three assumptions are all relaxed in the present framework.

The framework put forward in this paper can be used to study interfirm differences in productivity and technical change, as illustrated in Klette (1996). Indeed, the productivity measure a_{it} in (6) is an extension of 'the multilateral total factor productivity index' proposed by Caves, Christensen and Diewert (1982) for multilateral comparisons of productivity. The productivity index of Caves et al. is also based on the concept of the representative firm as a benchmark for comparing productivity differences across a number of firms. The multilateral total factor productivity index is, however, based on the restrictive assumptions of constant returns to scale and competitive output markets, while these assumptions are not needed to analyze interfirm differences in productivity on the basis of the framework presented in this paper[15].

3 The data

The sample covers almost all manufacturing industries for the period 1980-90[16]. The sample is based on the annual census carried out by Statistics Norway[17]. Separate estimates are presented for 14 different industry groups corresponding to 2/3-digit ISIC classes. As mentioned above, the unit of observation is an establishment.

In the current study, only operating establishments with at least five employees have been included. All observations that did not report the variables required have been eliminated. I also removed observations with an extreme value added per unit of labor input or extreme value added per unit of capital[18]. Establishments that existed for less than three consecutive years were eliminated. These trimming procedures together reduced the sample sizes by 5-10 percent.

[15] Baltagi, Griffin and Rich (1995) have emphasized the importance of accounting for scale economies in the measurement of firm-specific indexes of technical change.

[16] I have left out the sector "Manufacture of food, beverages and tobacco" (ISIC 31), partly since it is very large, with almost 50 000 observations for the period considered, and partly because it is heavily regulated, questioning the validity of the behavioral model applied above. The industry "Other manufacturing" (ISIC 39) has also been eliminated as it is a rather small and heterogeneous collection of plants.

[17] See Halvorsen, Jensen and Foyn (1991) for documentation and Manufacturing Statistics from Statistics Norway (several years) which reports a variety of summary statistics from the manufacturing census.

[18] Extreme values were defined as logarithmic deviations from the median exceeding 3 in absolute value for each year and each 5-digit industry.

Output and inputs are measured relative to the median values for the industry (at the 5-digit ISIC-code level) to which the ¯rm belongs[19]. The industry median values are estimated separately for each year, which is required as we want to allow the technology (cf. $F_t(¢)$) to change freely over time. Shifting the normalization (i.e. the reference point) each year has the additional bene¯t that it eliminates the need for de°ating the nominal variables. De°ators for inputs and outputs are in many, if not most, manufacturing industries heavily contaminated by noise, not least due to the problems of dealing with goods undergoing important quality changes over time.

All costs and revenues are adjusted for taxes and subsidies, re°ecting the prices facing the ¯rm[20]. The output variable is nominal output adjusted for changes in inventories[21] and measured net of sales taxes and subsidies. Four inputs are treated separately in this study: Capital, energy, labor and materials. Details on the construction of the labor and capital variables are presented in appendix A. The wage payments incorporate salaries and wages in cash and kind, social security and other costs incurred by the employer. The capital variable is constructed on the basis of ¯re insurance values for buildings and machinery[22]. Table 1 reports summary statistics for each industry in the sample for 1985.

The present study incorporates the cost contribution of material and energy inputs, in contrast to Hall's analysis (1988, 1990). Hyde and Perlo® (1995) found that \the markup estimate is sensitive to the choice of input factors included ... [Higher and incorrect markups appear] if we ... use only labor and capital (ignore materials and energy)". Norrbin (1993) has made the same observation.

4 The econometric issues

4.1 Constructing the shares

The theoretical model presented in section 2 includes the factor costs' share in the value of total output, evaluated at some internal point in the domain between the reference point, i.e.

[19] Using the median rather than the mean as the reference point was based on the observation that the median is less in°uenced by extreme observations. However, in most industries and for most variables the di®erences between the two statistics are small and therefore unlikely to be important for the results.

[20] See Manufacturing Statistics from Statistics Norway (several years), and Halvorsen et al. (1991) for details about these adjustments.

[21] At least in principle, the employed output measure also accounts for repair works for customers, investment activities done by the plants' workers and a number of other (minor) outputs; see Halvorsen et al. (1991, ch. VI.5) for details.

[22] This help us to overcome the criticism to scale estimates based on accounting measures of capital, raised by Friedman (1955). Friedman argued that accounting measures of capital would imply constant return by de¯nition. See Griliches and Ringstad (1971, ch. 3.3 and p.59) for further remarks on the pros and cons of the use of ¯re insurance values to construct the capital variable.

the industry-year median values, and the observed level of operation for the establishment in question. Since the location of this particular point and the corresponding shares are unknown, I have approximated the shares by taking the average value of the share for the observed establishment and the time-industry median share. The \Quadratic approximation lemma" in Diewert (1976) shows that using this average cost share of factor j to replace α^{j}_{it} in (1) will introduce no approximation error if the underlying technology is of the translog type. Hence, the empirical model is exact for a translog technology which may vary from industry to industry and year to year.

The framework is consistent with the widely recognized pattern that di®erent ¯rms within an industry face di®erent wages. In particular, it has been documented in a number of studies that larger ¯rms tend to pay higher wages and hire more high-skilled workers[23]. The model presented here captures these phenomena in two ways: (i) The ¯xed e®ects will capture di®erences in productivity levels between ¯rms due to di®erences in labor quality. (ii) Using the factor shares of individual establishments in the way described above, the model is also consistent with variations in the output elasticity of labor and the other factors of production across observations[24].

In constructing the labor share using (2), it is appropriate to use the marginal wage rate, which might di®er from the average wage rate when overtime work is the marginal labor input. However, our data set does not contain information about overtime work or overtime pay, so I have used the average wage rate as is done in most econometric ¯rm level studies. Since the average wage rate is lower than the wage rate at the margin for ¯rms using overtime labor, there might be a downward bias in the shares, which will bias the estimated markups upwards.

4.2 Fixed e®ects

As mentioned above (cf. footnote 4), productivity di®erences between ¯rms tend to be highly persistent over time. These productivity di®erences are important determinants of growth and exit[25]. The term $\hat{a}_{it}$ will be represented by an error component structure;

$$\hat{a}_{it} = a_i + u_{it}; \tag{9}$$

where a_i is treated as a ¯xed (correlated) e®ect, while u_{it} is a random error term. Treating a_i as a ¯xed e®ect means that we allow the cross sectional di®erences in productivity between establishments to be freely correlated with all the variables in the estimating equation, i.e.

[23] See e.g. Brown and Medo® (1989) for an empirical analysis of the employer size-wage relationship.

[24] The most common panel data model of production seems to be the Cobb-Douglas speci¯cation with ¯xed e®ects. A Cobb-Douglas model with ¯xed e®ect is consistent with (i), but not (ii).

[25] Klette and Mathiassen (1996) show that measured productivity is an important determinant of plant survival in Norwegian manufacturing. See also Olley and Pakes (1996).

output and all factor inputs. Initial tests for the presence of ¯xed versus random (uncorrelated) e®ects strongly rejected the hypothesis of random e®ects, as is widely experienced with these kinds of data (Griliches and Mairesse, 1998). Notice that technical change common across plants within an industry is captured by measuring all variables as deviations from time-industry averages.

There can be several explanations for the presence of ¯xed e®ects as captured by a_i. Establishments might di®er in the e®ectiveness of the management, labor quality, the vintage of the capital and so fourth. Such di®erences will emerge as variations in productivity. More to the point, these productivity di®erences will tend to be positively correlated with size, in the sense that more productive establishments will gain larger market shares. Another possible explanation for ¯xed e®ects is that some establishments do not have their own headquarter activities, while others do. This will show up in measured productivity. Furthermore, if there is a correlation between establishment size and the frequency of establishments incorporating their own headquarter services, the estimates will be inconsistent unless ¯xed (correlated) e®ects are incorporated into the estimated model. Whatever the reason, the model and the data require a ¯xed e®ect formulation. To eliminate the ¯xed e®ect, the model is estimated in terms of ¯rst di®erences (see below).

The scale coe±cients presented in this paper are long-run scale elasticities, as they incorporate changes in both variable factors (materials, energy and working hours) and capital. But the approach focuses on changes in the level of operation in the longitudinal dimension, and disregards the cross-sectional information about e±ciency di®erences in small versus large plants. Some people have argued that cross sectional comparisons of establishments is more relevant to understand long-run scale economies. However, the comparison in e±ciency between small and large plants raises the question of causality: Are large plants more e±cient because they are large (which would support claims about scale economies), or have they grown larger than other plants because they are more e±cient (due to e.g. better technology or better management)[26]? This question raises doubt about whether cross-sectional di®erences in e±ciency can be interpreted as evidence on scale economies[27].

[26] This issue was raised in the empirical production function literature by Marschak and Andrews (1944), and has also been empasized in the controversy on whether concentration is desirable or not, raised by Demsetz (1974) and others in the theoretical I.O.-literature.

[27] The di®erences between cross sectional and panel data studies of production functions is an old and extensively discussed issue; see Griliches and Mairesse (1998) and references cited there.

4.3 The orthogonality conditions and GMM-estimation

Inserting (9) into (6), and taking first differences to eliminate the fixed effect (a_i), we obtain the estimating equation

$$\Delta \hat{q}_{it} = \mu \Delta \hat{x}^V_{it} + \gamma \Delta \hat{x}^K_{it} + \Delta v_{it}; \tag{10}$$

where I have defined the variable $\hat{x}^V_{it} = \sum_{j \neq K} \hat{s}^j_{it}(\hat{x}^j_{it} - \hat{x}^K_{it})$ and $\Delta \hat{q}_{it} = \hat{q}_{it} - \hat{q}_{i,t-1}$ and so fourth. Δv_{it} is given as $v_{it} - v_{it-1}$, where

$$v_{it} = u_{it} + (\mu_{it} - \mu)\hat{x}^V_{it} + (\gamma_{it} - \gamma)\hat{x}^K_{it} \tag{11}$$

Equation (10) can not be consistently estimated by OLS for two reasons. First, allowing for fixed effects by estimating the model in growth rates might not solve the whole problem of correlation between the productivity differences, u_{it}, and the differences in firms' choices of factor inputs. To the extent that a firm experiences changes in productivity over time relative to the average firm, a productivity shock might be correlated with changes in factor inputs to the extent that the shock is anticipated before the factor demands are determined[28]. This will create a correlation between the right hand side variables and the error term in (10). Second, errors-in-variables due to reporting errors will create an endogeneity problem. The errors-in-variables problem is well known to be augmented when estimating the model in first differences (cf. the discussion in Griliches and Mairesse, 1998).

The model has been estimated using orthogonality assumptions between Δv_{it} and alternative sets of instruments:

$$E(\Delta v_{it} Z_{is}) = 0; \tag{12}$$

where Z_{is} is a vector of instruments dated s.

Two steps have been taken to ensure that the instrument set is chosen so that the condition (12) is fulfilled. First, the instrument set has been restricted to two variables: the capital variable and the number of employees. These variables are less responsive than the inputs materials, energy and man-hours, to temporary changes in productivity[29]; see Biørn and Klette (1996) for some econometric support to this claim. Second, within this set of instruments,

[28] Olley and Pakes (1996) have also addressed this problem, but in a different way than I do. See also Griliches and Mairesse (1998) for a discussion of the problem.

[29] Lagged values of $\hat{x}^V_{it}$ could also been considered as instruments, but these additional instruments did not significantly improve the precision of the estimates.

alternative orthogonality assumptions have been tested. I have tested the assumptions whether the instruments are strictly exogenous, predetermined or (only) contemporaneously correlated with the errors. That is to say, I have tested whether condition (12) holds for:

(I) All values of t and s.

(II) Only for non-contemporaneous instruments, i.e. $|t - s| \geq 1$. Such an instruments set is interesting when the error term in (10) is not autocorrelated beyond an MA(1) structure[30].

(III)-(V) Only predetermined instruments, i.e. $t - s \geq l$, for three different values of l; $l = 0; 1$ or 2:

There are five alternative specifications of the orthogonality conditions and the procedure used to discriminate between these specifications will be discussed below.

GMM provides the optimal way to combine the set of orthogonality conditions (12). The GMM estimator $(\hat{\alpha}; \hat{\beta})$ minimizes

$$J = N \left(\frac{(\Delta v)' Z}{N} \hat{V}^{-1} \frac{Z' \Delta v}{N} \right); \tag{13}$$

where I have stacked all the Δv_{it}'s in (10) into a single vector Δv, and Z is a matrix with all the instruments[31]. N is the total number of observations. $\hat{V}$ is a consistent estimator of the covariance matrix of $(Z' \Delta v)$. The main results presented below is based on the one step GMM -estimator, where $\hat{V}$ is replaced with $N^{-1} \sum_i Z' H Z$ and H is a square matrix with twos in the main diagonal, minus ones in the first subdiagonals and zeros otherwise[32].

The various specifications have been tested by means of the overidentification test based on the minimized value of J in (13), which asymptotically has a Chi-square distribution with degrees of freedom given as the number of orthogonality conditions minus the number of parameters (see e.g. Newey (1985) and Arellano and Bond (1991)). A more powerful test of nested orthogonality assumptions can be based on differences in the J-values for the competing specifications, as explained by Arellano and Bond. I will refer to such a test as a J-difference test. Details of the test procedure is given in appendix B.

[30] Griliches and Hausman (1986) and Biørn and Klette (1996) have discussed and applied similar instrument sets.

[31] The care needed in stacking the instrument vector in the presence of an unbalanced set of panel data has been discussed by Arellano and Bond (1988, 1991). See also Biørn and Klette (1996). The GMM-estimates presented in the current study have been obtained using the GAUSS-program "DPD" documented in Arrelano and Bond (1988).

[32] I have reported one step rather than two step GMM estimates since the standard errors associated with the two step estimates tend to be seriously downward biased, as noticed by Arellano and Bond (1991) and others. The two step estimates are reported in a previous version of this paper, cf. Klette (1994).

4.4 Some additional remarks on the choice of IVs

Both Hall and the present study apply an instrumental variable approach to the estimation of the markups and other parameters of interest. Hall pointed out the need for instruments due to the correlation between productivity shocks buried in the residual and factor demands, as discussed in the previous section. This correlation has motivated the choice of the number of employees and capital as instruments in the present study. Let me emphasize that the instrument set used here is entirely di®erent from Hall's instrument set which consisted of the oil price, military spending and a dummy for the party of the president.

Abbott, Griliches and Hausman (1988) have argued forcefully that the oil price is not a valid instrument. They emphasized the omission of adjustment for capacity utilization in the models estimated by Hall. Their point is that this omitted variable problem creates biases since Hall's instruments, in particular the oil price, are correlated with a left out variable; the degree of capacity utilization. Hall adjusts for changes in capacity utilization of capital by using a residual share to impute the output elasticity of capital, but his procedure is only correct to the extent that constant returns to scale is a valid maintained hypothesis. In the present study, the constant returns to scale hypothesis is rejected in several industries and relaxing this hypothesis signi¯cantly reduces the markup estimates. Both Hall and the present study use man-hours as the measure of labor inputs. This should reduce the need to adjust for changes in utilization of the work force[33]. Finally, notice that to the extent my instruments are invalid, they are likely to bias the parameter-estimates upward since e.g. a positive productivity shock will typically stimulate investment and hiring.

5 Estimates of markups and scale elasticities

5.1 Speci¯cation testing

The results from the speci¯cation tests are presented in Table 2. The ¯rst 5 rows present the overidenti¯cation tests (the J-values) for each instrument set separately, while rows 6-10 present the outcomes of the J-di®erence tests. None of the instrument sets are rejected for the industries with ISIC-codes 321, 332, 341, 351-2, 36, 37, 383 and 384. I have consequently chosen the most extensive instrument set, I, as the preferred speci¯cation for these industries. For the industries 322-4, 331 and 381, the instrument set II has been chosen. For the industries 331 and 381 both instrument sets III and IV are rejected on the basis of the overidenti¯cation tests. Instrument set III is rejected on the basis of the J-di®erence test for the industry 322-4, and I will present

[33] See Hall (1990) for a detailed discussion of di®erent kinds of misspeci¯cation related to this point.

the results based on instrument set II for this industry, but comment also on the results based on instrument set IV[34]. For the industries 355-6 and 382 the instrument sets III and IV, respectively, are preferred. None of the instrument sets have been accepted for industry 342. As we shall see below, this industry stands out in several respects[35].

5.2 Average markups and scale elasticities

The ¯rst row in Table 3 shows the markups estimates and the second row shows the estimated scale elasticities. Perhaps the most striking aspect of the results reported in Table 3 is that the markups and the scale elasticities are close to one, and that there are few statistically signi¯cant di®erences in market power between the industries considered. Still, a majority of the industries considered in this study reveal small, but statistically signi¯cant market power. Ordered according to market power, the seven industries with (statistically) signi¯cant market power are: Metals (37), Paper products (341), Wood products (331), Metal products (381)[36], Clothing (322-4), Furniture (332), and Textiles (321). The margins between price and marginal costs for these industries range from 1.09 for Metals to 1.05 for Textiles. Six industries exhibit no signi¯cant market power: Chemicals (351-2), Plastics (355-6), Mineral products (36)[37], Machinery (382), Electrical equipment (383) and Transport equipment (384).

Printing (342) obtained a very low markup estimate; the reported estimate is implausible and the estimates based on the alternative instrument sets are very similar. But this result should be neglected as the overidenti¯cation test presented above reveals that the model is misspeci¯ed. It is somewhat comforting that the overidenti¯cation test is able to identify this industry/sample as suspect, given the implausible parameter estimates.

Even if the di®erences in market power across industries are small and to a large extent not statistically signi¯cant, it is natural to examine whether the estimated di®erences across industries in market power can be related to some external measures of market power. I have therefore considered the correlation between the markups in Table 3 and di®erences in a He¯ndahl index of industry concentration, an index of import penetration and di®erences in export shares[38],

[34] Cf. the discussion on the choice of non-nested speci¯cations in appendix B. For the three industries - 322-4, 331 and 381 - the instrument set I has not been rejected in a direct test against II. But the instrument set III, which is strictly larger than I, has been rejected. Hence, the instrument set I should also be rejected.

[35] A number of experiments, such as splitting the industry up into ¯ner industry categories and restricting the instrument set further, have been run on this industry without success. The results were poorest for the subindustry group \Printing and bookbinding" (ISIC 3421). A more detailed investigation of the industry 342 is left for future research.

[36] Choosing instruments based on observations dated $t ¡ 1$ and earlier, provided an estimated price-cost margin at 1.005 (std.err: 0.012) and a scale elasticity at 0.930 (std.err: 0.010).

[37] The alternative choice of instruments, based on predetermined instruments dated $t ¡ 2$ and earlier, provided an estimated price-cost margin at 1.137 (std.err: 0.032) and a scale elasticity at 0.963 (std.err: 0.026).

[38] I have considered both correlations in levels and in rankings.

but these analyses revealed no signi¯cant relationship. This is perhaps not too surprising for at least two reasons. First, such an analysis should control for product di®erentiation and market segmentation as all the industries considered produce a number of di®erent products and varieties[39]. However, empirically useful measures of product di®erentiation and market segmentation are di±cult to derive even in principle, not to mention the practical problems with data availability. Second, the Her¯ndahl indices referred to above is based on detailed output observations at the ¯rm level[40]. But ¯rm-level data is not adequate for this purpose, as a number of the manufacturing ¯rms in Norway are integrated into interlocking groups of ¯rms. Unfortunately, the available data sets do not contain information on these ownership structures. Consider as an example of the di±culties involved, Paper products (341) which is the industry with the second highest markup. This industry is not particularly concentrated when we consider the Her¯ndahl index based on ¯rm level data, but most of the largest ¯rms in this industry are organized into an interlocking group of ¯rms (Norske Skog). Furthermore, domestic concentration is not an obvious proxy for market power in this case as the industry is highly export oriented with an export share exceeding 50 percent in 1985. The metal industry (37) which has the highest markup in Table 3, has a similar industry structure with a high export share and high concentration, but this is also true for the industry with the lowest markup; Chemicals (351-352) producing a large range of industrial chemicals.

Turning to the scale elasticities, none of the industries reveal signi¯cant scale economies and most industries do not reject constant returns to scale. The industries not rejecting constant returns are Textiles (321), Clothing (322-4), Wood products (331), Furniture (332), Paper products (341), Metals (37), Metal products (381), Electrical equipment (383) and Transport equipment (384). Four industries reveal moderate decreasing returns to scale with scale elasticities in the range 0.89-0.96. These are Chemicals (351-2), Plastics (355-6), Mineral products (36) and Machinery (382). The results for Printing should be ignored, as discussed above.

[39] Even in industries producing basic metals, product di®erentiation can be an important source of market power. In comenting on the market position for the Norwegian ¯rm Falconbridge, a leading world producer of cobolt, its director states: \Not only are we a large producer of cobolt, we are also making the best cobolt in the world which is in high demand for jet enginees and super turbines." (Dagens N½ringsliv, 26.3.98. My translation.).

[40] More precisely, the Her¯ndahl index of concentration was constructed on the basis of observations for each business unit, using output reported at the line of business level for each individual ¯rm. Separate indexes were constructed for more than one hundred 5-digit industries, and then aggregated to the 3 digit level with industry output as weights.

5.3 A comparison to related results

Martins, Scarpetta and Pilat (1996) have used both Roeger's (1995) and Hall's (1988) procedures to estimate the market power in 36 manufacturing industries for 14 OECD countries, including Norway, over the 1970-92 period. Their analysis is based on industry level data. Considering their estimates for Norwegian manufacturing for the period 1980-92, they tend to be somewhat higher than the estimates presented in Table 3, varying from 1.08 (in Chemicals) to 1.45 (in O±ce and computing machinery), with an average value slightly less than 1.20 . Unfortunately, Martins et al. do not present standard errors of their parameter estimates, so the precision of their estimated markups is unclear. Their estimates are based on constant returns to scale as a maintained hypothesis, and this assumption is likely to explain their higher estimates for the price cost margins to a large extent. Comparing the ranking of market power across industries, the correlation between the present study and Martins et al. is positive but not particularly strong[41].

Few other microeconometric studies have recently examined market power and (or) scale economies in Norwegian manufacturing. Griliches and Ringstad (1971) used a cross section of establishments from 1963 to estimate scale economies in Norwegian manufacturing. They found scale elasticities around 1.05-1.06 for total manufacturing and mining. The results in Griliches and Ringstad (1971) di®er substantially from the ¯ndings presented in this paper, as I do not ¯nd any presence of increasing returns. However, since the study of Griliches and Ringstad, it has become a widely held view that scale estimates from cross sectional studies are upward biased, as they do not account for persistent di®erences in e±ciency between plants; see Griliches and Mairesse (1998).

As discussed in the introduction, it is di±cult to relate my results on market power to case studies such as Gabrielsen (1989) and the studies surveyed in S¿rgard (1997), as these case studies typically study markets covering very small and not representative parts of the markets faced by the manufacturing ¯rms in the industries I consider. For instance, in his study of the fertilizer market in Norway, Gabrielsen found signi¯cant market power for the leading producer (Norsk Hydro). Production of fertilizers is a part of the Chemical industry (351-2), where I ¯nd little evidence of market power. The results are not necessarily contradictory, however, since about 80 percent of the Norwegian production of fertilizers is exported, and production of fertilizers is only about a ¯fth of total production in the Chemicals industry. Intraindustry variations in market power will be examined in section 6.

[41]Comparing the estimates is not entirely straightforward as the levels of industry aggregation do not perfeclty match between the two studies.

5.4 Additional remarks on the econometric speci¯cation

Let me add a few comments on the test-statistics for ¯rst and second order autocorrelation in the residuals in Table 3. The presence of signi¯cant ¯rst order autocorrelation is expected given that the model is estimated in ¯rst di®erences, but the presence of signi¯cant second order autocorrelation questions the use of lagged values of the regressors as instruments[42]. Cyclical errors could for instance be due to more rapid response of output than labor input to cyclical shocks, even with working hours as the employed measure of labor input. Notice, however, that my instruments are not exactly lagged regressors but lagged (and in some cases leaded) values of capital inputs and the number of employees as discussed in section 4.3 above. To examine the importance of potential endogeneity problems, I have in the speci¯cation testing considered estimates based on restricted instrument sets, and Table 3 reports these estimates when the speci¯cation tests or the parameter estimates suggest that estimates based on restricted instrument sets are preferable.

A direct comparison of the parameter estimates based on di®erent instrument sets con¯rm that the estimates in Table 3 are not strongly a®ected by restricting the instrument set further. The estimates based on the most restricted instrument set are not signi¯cantly di®erent from the preferred estimates reported in Table 3[43]. A possible exception is the point estimate for the metal industry (ISIC 37) which is somewhat lower with the most restricted instrument set, but the standard error of the markup estimate is high with this instrument set.

The markup and the scale coe±cient estimates presented in Table 3 share a problem with most microeconometric studies of market power and scale elasticities, according to an argument put forward by Klette and Griliches (1996). Klette and Griliches identi¯ed a downward bias in the estimation of scale elasticities caused by replacing real output by de°ated sales, where de°ation is based on an industry-wide de°ator. Such a de°ating procedure is essentially equivalent to the normalization approach used in the present study. The point is that if idiosyncratic productivity shocks are important determinants of ¯rm growth, growth in de°ated sales will be a systematically biased indicator for growth in real output. This bias in the growth rate in the output measure tends to create a downward bias in estimated scale coe±cients, and such a downward bias might be present in the estimates of the scale coe±cients (and consequently in

[42] Most of the reported test statistics are signi¯cant at the 5 percent level, while four of the test statistics are signi¯cant at the 1 percent level. Given the large sample sizes in most industries, a conservative 1 percent signi¯cant level might be considered appropriate.

[43] Consider in particular the four industries with the highest test statistics for second order autocorrelation as reported in Table 3. The markup estimates for these industries based on the most restricted instrument set are: 1.10 (.030) for ISIC 331; 1.13 (.040) for ISIC 341; 0.986 (.070) for ISIC 37; 1.020 (.042) for ISIC 382.

the markups) presented above.

6 Heterogeneity in market power, scale economies and productivity

6.1 Measuring heterogeneity in market power, scale economies and productivity

The markups presented above are not very large, and suggests little reason for worries about large welfare losses. But concerns about market power is often focused on one or a few leading firms in an industry while the average competitor has little market power, as argued in the introduction. The price-cost margins and scale elasticities presented above represent averages for different industries, and it is interesting to examine whether there are large variations in market power and scale economies within each industry. If we are willing to impose a few additional assumptions, these variations can be estimated by a method suggested by Hildreth and Houck (1968)[44]. Consider the residuals

$$\begin{aligned} \hat{\varepsilon}_{it} &\equiv \hat{q}_{it} - \hat{\mu}\hat{x}^V_{it} - \hat{\eta}\hat{x}^K_{it} \\ &= \hat{a}_{it} + \hat{v}_{it} \\ &= a_i + (\mu_i - \mu)\hat{x}^V_{it} + (\eta_i - \eta)\hat{x}^K_{it} + \omega_{it}. \end{aligned} \tag{14}$$

ω_{it} captures sampling error and the annual fluctuations in the markup and scale coefficients. Let us now assume that ω_{it} is uncorrelated with $\hat{x}_{is}$, when $|t - s| > l$, for some value of l (=1,2,3..). Then, one can show that

$$\begin{aligned} E(\hat{\varepsilon}_{it}\hat{\varepsilon}_{is}|\hat{x}^V_{it};\hat{x}^K_{it};\hat{x}^V_{is};\hat{x}^K_{is}) = \sigma^2_a + \sigma^2_\mu\left(\hat{x}^V_{it}\hat{x}^V_{is}\right) + \sigma^2_\eta\left(\hat{x}^K_{it}\hat{x}^K_{is}\right) + \sigma^2_{a\mu}(\hat{x}^V_{it} + \hat{x}^V_{is}) \\ + \sigma^2_{a\eta}(\hat{x}^K_{it} + \hat{x}^K_{is}) + \sigma^2_{\mu\eta}(\hat{x}^V_{it}\hat{x}^K_{is} + \hat{x}^K_{it}\hat{x}^V_{is}) \qquad |t - s| > l; \end{aligned} \tag{15}$$

where σ^2_a is the variance of the a_i's, i.e. the variance of the permanent productivity differences between plants. σ^2_μ and σ^2_η are the variances of the price-cost margins and the scale elasticities; $\sigma^2_{a\mu}$; $\sigma^2_{a\eta}$ and $\sigma^2_{\mu\eta}$ represent the covariances between the differences in productivity, the price-cost margins and the scale elasticities. From (15) it follows that

$$\begin{aligned} (\hat{\varepsilon}_{it}\hat{\varepsilon}_{is}) = \; & \sigma^2_a + \sigma^2_\mu\left(\hat{x}^V_{it}\hat{x}^V_{is}\right) + \sigma^2_\eta\left(\hat{x}^K_{it}\hat{x}^K_{is}\right) + \sigma^2_{a\mu}(\hat{x}^V_{it} + \hat{x}^V_{is}) + \sigma^2_{a\eta}(\hat{x}^K_{it} + \hat{x}^K_{is}) \\ & + \sigma^2_{\mu\eta}(\hat{x}^V_{it}\hat{x}^K_{is} + \hat{x}^K_{it}\hat{x}^V_{is}) + e_{its} \end{aligned} \tag{16}$$

[44] See Mairesse and Griliches (1990) for further references to the subsequent literature on the random coefficient model.

where $E(e_{its} | \hat{\chi}^V_{it}; \hat{\chi}^K_{it}; \hat{\chi}^V_{is}; \hat{\chi}^K_{is}) = 0$ when $|t - s| > l$. Equation (15) has the form of a familiar linear regression model and the variances and covariances σ^2_a, σ^2_μ, σ^2_η, $\sigma^2_{a\mu}$, $\sigma^2_{a\eta}$ and $\sigma^2_{\mu\eta}$ can consequently be estimated by regressing cross-products of the residual term on the $\hat{\chi}_{it}$'s and their squares and cross-products. There are many possible choices of t and s; the most efficient estimates of the σ^2s are obtained by pooling the estimates by combining all permissible combinations and Table 4 reports the outcome of such an estimation procedure[45]. It turned out that the estimated covariances were insensitive to the choice of l in (15), and all the estimates in Table 4 are based on l=1.

6.2 Empirical findings on heterogeneity

Table 4 shows that the variation in price-cost margins is largest in Plastics (355-6) and smallest in Textiles (321) and Paper products (341). The average value of these variances is 0.004[46], which is 4 times the variance of the markups across industries presented in Table 3. This finding confirms the argument that market power tend to vary more across firms within the same industry, than across the average firm in different industries.

Firms with higher productivity tend to set lower markups, as can be seen from the negative values of $\sigma_{a\mu}$ in Table 4. In other words, plants in more competitive niches of an industry tend to be forced to be more productive and to charge a lower markup. This result is in accordance with the finding recently reported by Nickell (1996) that market power not only create inefficient pricing, but also reduces incentives for efficient organization of production. Nickell concludes on the basis of a study of 670 U.K. companies, "that competition ... is associated with higher rates of total factor productivity growth". Harrison (1994) has also presented related evidence on the links between trade protection, productivity and market power.

Turning to heterogeneity in the scale elasticities, the largest variance is in Plastics (355-6) and the smallest is in Metal products (381). The weighted average variance is 0.003, which is 3 times as large as the variance of the scale elasticities across industries presented in Table 3. In this sense, there are larger differences also in the scale coefficient within each industry as compared to between industries.

The results in Table 4 reveal significant differences in permanent productivity levels (cf. the fixed effects) across plants within all industries except one. The largest variance appears in Paper products (341), with $\sigma^2_a = 0{:}043$, and the smallest in Plastics, with $\sigma^2_a = 0{:}001$ (not significantly different from zero). The weighted average of these variances across industries is

[45] The heteroskedasticity-robust standard errors reported in Table 4 allow for correlation in the residuals as I am combining various ts and ss for each plant.

[46] This average is calculated using the inverse of the standard errors as weights.

0.013, using the inverse of the standard errors as weights.

Finally, notice that although this section has emphasized the presence of large differences in market power across firms in the same industry, it has not identified what firm or product characteristics (e.g. R&D intensity, product quality[47] or firm size) explain these differences. At an early stage of this research project, I examined the relationship between the markup, scale economies and firm size, but I found little systematic relationship between size and the other measures. I believe this finding reflects a problem related to the difficulty of defining the size of the relevant market. In Norway as in most other countries, the larger manufacturing firms are often more export oriented than smaller firms in the same industry and the large firms may consequently face stronger market competition. This is true even when the industries are narrowly defined. In future work it would be interesting to search for industries where independent measures of market power can be obtained, and incorporate these into the specification of the markup in the econometric framework presented above[48].

6.3 Remarks on parameter stability

Panel data studies of productivity and scale economies using plant level data are often based on a Cobb-Douglas specification of the technology. Mairesse and Griliches (1990) found substantial heterogeneity and instability in the coefficients of their estimated Cobb-Douglas production functions for US, French and Japanese firms. The findings by Griliches and Mairesse suggest that a more flexible specification of the technology is desirable, and the framework presented in this paper provides such a flexible framework (cf. section 4.1). The parameter estimates in the present study are substantially more stable than the parameters estimated by Mairesse and Griliches (1990): The variances of the productivity differences (cf. σ_a^2) and the scale elasticities (or the markups) are between one and two order of magnitudes smaller than the coefficient variances found by Mairesse and Griliches (1990, see in particular Table 6).

Perhaps a natural response to the findings by Griliches and Mairesse would be to focus on flexible assumptions about functional forms, but estimation of flexible functional forms often create problems with the concavity conditions when fixed effects are allowed for in the econometric analysis of panel data. The simple parametric structure, yet flexible specification of the econometric model presented above is a significant advantage in panel data studies compared to more heavily parametrized econometric models.

[47] Notice that higher prices on products with higher quality produced by using more inputs (cf. leather seats and more horse power in cars) will not show up as market power within the present model.

[48] That is, one could model the markup in terms of observable firm, product and market characteristics, M_{it}, such that $\mu_{it} = \mu(M_{it})$. Such characteristics could include indicators for market concentration, firm size, product characteristics and business cycle conditions.

7 Concluding remarks

This paper has presented a framework for estimating price-cost margins and scale economies from a panel of plant level data, allowing for considerable °exibility and heterogeneity in the technological constraints facing di®erent ¯rms. Applying this econometric framework to 14 industries in Norwegian manufacturing, I found that:

- Estimated (average) margins between price and marginal costs are statistically signi¯cant, but small in economic terms. Price exceeds marginal costs by between 5 and 10 percent in most of the industries considered.
- Increasing returns to scale is not a widespread phenomenon in Norwegian manufacturing. Rather, the average ¯rm in most industries seems to face constant or moderately decreasing return to scale.
- There is more within-industry variation in market power compared to the variation in market power between the industries considered.
- Plants (¯rms) with higher market power tend to be less productive.

Given the large interest in non-competitive models and theories about increasing returns in the economic discipline, the results presented in this paper are perhaps somewhat surprising. However, the model with persistent productivity di®erences and parameter estimates suggesting moderate decreasing returns is consistent with Lucas' famous `span-of-control' model (Lucas, 1978) which has been the basis for much theoretical and empirical work on ¯rm growth, ¯rm heterogeneity and industry evolution[49].

[49] See Jovanovic (1982) and the large subsequent empirical and theoretical literature.

References:

Abbott, T.A., Z. Griliches and J.A. Hausman (1988): Short Run Movements in Productivity: Market Power versus Capacity Utilization. Paper presented at the NBER Summer Institute 1988.

Arellano, M. and S.R. Bond (1988): Dynamic Panel Data Estimation Using DPD - A Guide for Users. Working Paper 88/15, Institute for Fiscal Studies.

Arellano, M. and S.R. Bond (1991): Some Tests of Speci¯cation for Panel data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies 58, 277-97.

Baily, M., C. Hulten, D. Campbell (1992): Productivity Dynamics in Manufacturing Plants, Brookings Papers on Economic Activity, Microeconomics, 187-249.

Baltagi, H.B., J.M. Gri±n and D.P. Rich (1995): The Measurement of Firm-Speci¯c Indexes of Technical Change, Review of Economics and Statistics 77, 654-63.

Bartelsman, E.J, R.J. Caballero and R.K. Lyons (1991): Short and Long Run Externalities. NBER Working Paper no. 3810.

Berck, P. and K. Syds½ter (1991): Economists' Mathematical Manual. Springer-Verlag (Heidelberg).

Bi¿rn, E. and T.J. Klette (1996): The Labour Input Response to Permanent Changes in Output: Errors in Variables Econometrics Based on Panel Data. Memorandum no. 14, Department of Economics, University of Oslo.

Bresnahan, T.F. (1989): Empirical Studies of Industries with Market Power, in R. Schmalensee and R. Willig (eds.): Handbook of Industrial Organization, Vol. 2, North Holland Publ. Co. (Amsterdam)

Brown, C. and J. Medo® (1989): The Employer Size-Wage E®ect. Journal of Political Economy 97, 1027-59.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982): Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers, Economic Journal 92, 73-86.

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

Diewert, E. (1976): Exact and Superlative Index Numbers. Journal of Econometrics 4, 115-45.

Friedman, M. (1955): Comment, in C.A. Smith (ed.): Business Concentration and Price Policy. A report of the National Bureau of Economic Research. Princeton University Press (Princeton).

Gabrielsen, T.S. (1989): Âkonomisk tilpasning i det norske markedet for kunstgj¿dsel (NPK). SAF-rapport nr. 24/89 (Bergen).

Griliches, Z. and J.A. Hausman (1986): Errors in Variables in Panel Data, Journal of Econometrics 18, 93-118.

Griliches, Z. and J. Mairesse (1998): Production functions: The Search for Identi¯cation.in Steinar Str¿m (ed.): The Ragnar Frisch Centennial Symposium. Cambridge University Press (Cambridge, UK). Forthcoming.

Griliches, Z. and V. Ringstad (1971): Economies of Scale and the Form of the Production Function. North-Holland Publ. Co. (Amsterdam).

Hall, R.E. (1988): The Relationship between Price and Marginal Cost in U.S. Industry. Journal of Political Economy 96, 921-47.

Hall, R.E. (1990): Invariance Properties of Solow's Productivity Residual. in P. Diamond (ed.): Growth, Productivity, Unemployment. MIT Press (Cambridge, US.).

Halvorsen,R., R. Jensen and F. Foyn (1991): Dokumentasjon av industristatistikkens tidsseriebase ("Documentation of the Panel Data Base for Manufacturing"). Mimeo, Statistics Norway. (In Norwegian.)

Harrison, A.E. (1994): Productivity, Imperfect Competition and Trade Reform: Theory and Evidence. Journal of International Economics 36, 53-74.

Hildreth, C. and J.P. Houck (1968): Some Estimators for a Linear Model with Random Coef-¯cients, Journal of the American Statistical Association 63, 584-95.

Hoel, M. and R. Nymoen (1988): Wage formation in Norwegian manufacturing - An empirical application of a theoretical bargaining model. European Economic Review 32, 977-997.

Holm¿y, E., T. H½geland, Â. Olsen and B. Str¿m (1993): E®ektive satser for n½ringst¿tte. (\E®ective Rates of Assistance"). Reports 93/31. Statistics Norway (Oslo).

Hyde, C.E. and J.M. Perlo® (1995): Can Market Power Be Estimated? Review of Industrial Organization 10, 465-85.

Jovanovic, B. (1982): Selection and the Evolution of Industries, Econometrica 50, 649-70.

Judge, G.G. et al. (1985): The Theory and Practice of Econometrics, Second Edition. John Wiley and Sons (New York).

Klette, T.J. (1994) Estimating Price-cost Margins and Scale Economies from A Panel of Microdata. Discussion Papers 130, Statistics Norway.

Klette, T.J. (1996): R&D, Scope Economies, and Plant Performance, RAND Journal of Economics 27, 502-22.

Klette, T.J. and Z. Griliches (1996): The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous. Journal of Applied Econometrics 11, 343-61.

Klette, T.J. and A. Mathiassen (1996): Vekst og fall blant norske industribedrifter. (\Growth and Turnover among Norwegian Manufacturing Firms"). Social and Economic Studies no. 95. Statistics Norway (Oslo).

Levinsohn, J. (1993): Testing the Import-as-Market-Discipline Hypothesis. Journal of International Economics 35, 1-22.

Lucas, R.E. (1978): On the Size Distribution of Business Firms, Bell Journal of Economics 9, 508-23.

Mairesse, J. and Z. Griliches (1990): Heterogeneity in Panel Data: Are There Stable Production Functions?. in P. Champsaur et al. (eds.): Essays in Honour of Edmond Malinvaud, Vol. 3: Empirical Economics. MIT Press. (Cambridge, U.S.).

Marschak, J. and W. Andrews (1944): Random Simultaneous Equations and the Theory of Production. Econometrica 12, 143-205,.

Martins, J.O., S. Scarpetta and D. Pilat (1996): Mark-up Ratios in Manufacturing Industries. Estimates for 14 OECD Countries. Working Papers no. 162, Economics Department, OECD.

M¿en, J. (1998), Produktivitetsutviklingen i Norsk Industri 1980-1990 - En analyse av dynamikken basert på mikrodata (Productivity Dynamics in Norwegian Manufacturing 1980-1990 - A Study Based on Plant Level Data), Reports xx/98, Statistics Norway (Oslo), forthcoming.

Newey, W.K. (1985): Generalized Methods of Moments Speci¯cation Testing. Journal of Econometrics 29, 229-56.

Nickell, S. (1996): Competition and Corporate Performance. Journal of Political Economy 104, 724-46.

Norrbin, S.C. (1993): The Relation between Price and Marginal Cost in U.S. Industry: A Contradiction. Journal of Political Economy 101, 1149-64.

Olley, S. and A. Pakes (1996): The Dynamics of Productivity in the Telecommunications Equipment Industry, Econometrica 64, 1263-97.

Roeger, W. (1995): Can Imperfect Competition Explain the Di®erence between Primal and Dual Productivity Measures? Journal of Political Economy 103, 316-30.

S¿rgard, L. (1997): Konkurransestrategier (Competitive strategies), Fagbokforlaget.

Thomas, G.B. (1968): Calculus and Analytic Geometry. Addison Wesley Publ. Co. (Reading, Mass.) .

8 Appendix A: Details on the construction of the labor and capital variables

This appendix presents details about the construction of the labor and capital input variables used in the present study.

Before 1982, man-hours referred to blue collar workers only. Following Griliches and Ringstad (1971, p.24), total labor input (X_{it}^L) was estimated according to the formula

$$X_{it}^L = H_{it}\left(1 + \frac{C_{it}^{wc}}{C_{it}^{bc}}\right); \tag{17}$$

where H_{it} is man-hours for blue collar workers. C_{it}^{wc} and C_{it}^{bc} refer to total wage costs for white collar and blue collar workers. After 1982, the total number of man-hours was reported (while man-hours for blue collar workers alone were not), and used as the labor input variable.

As in most studies, capital inputs are perhaps the most problematic of the variables used in my analysis. My sample has an advantage to most other production data sets, in that the establishments report total fire insurance values for machinery and buildings (separately). Rental costs for rented capital are also reported. One of the problems with the fire insurance values is that there are a lot of missing values. Also, these variables have not been used by Statistics Norway, and little effort has been put into identifying and correcting erroneous reports. Once more, I have followed Griliches and Ringstad (1971, p.27) and estimated the capital services as

$$X_{it}^K = R_{it} + (\rho + \delta^M)V_{it}^M + (\rho + \delta^B)V_{it}^B \tag{18}$$

where R_{it} is rental costs, ρ is a real rate of return, and δ^M and δ^B are depreciation rates for machinery and buildings. ρ is chosen as the average real rate of return to physical capital in manufacturing (0.07), and the depreciation rates are taken from the Norwegian National Accounts (0.06 and 0.02 for machinery and buildings, respectively). V_{it}^M and V_{it}^B are the fire insurance values for machinery and buildings at the beginning of the year. $\rho; \delta^M$ and δ^B are to be considered as rough weights, and the validity of these weights varies substantially across plants and years. An interesting topic for future work would be to estimate the weights as an integrated part of the econometric modeling.

To avoid losing too many observations due to missing fire insurance values, and to eliminate some noise, three different estimates of the fire insurance value were calculated for each observation (plant-year). In addition to the reported fire insurance values for year t, the fire insurance values were also estimated by a perpetual inventory method on the basis of investment figures and fire insurance values for the years $t+1$, t and $t-1$ (if available). The mean value of the three different estimates was used as the final estimate.

9 Appendix B: The specification testing

The minimized value of J in (13) has asymptotically a Chi-square distribution with degrees of freedom given as the number of orthogonality conditions minus the number of parameters (see e.g. Newey, 1985, and Arellano and Bond, 1991). I have used this J-statistic to test the validity of the various specifications discussed above. Only models with sufficiently low J-values have been considered as acceptable. Newey (1985) has pointed out that even though the J-statistic comes closest to be an omnibus test for misspecification for models estimated by GMM, it has some limitations and he shows how the J-statistic may fail to detect misspecified. models.

A more focused specification test that considers a specific subset of a priori suspect instruments is desirable. Arellano and Bond (1991) have shown that this can easily be done for nested hypotheses on the basis of the J-statistics: Denote the J-statistic for the extended instrument set by J_E, and consider a subset of instruments, with J-statistic J_M, that is considered valid under the maintained hypothesis. In that case, $J_E - J_M$ has a Chi-square distribution with degrees of freedom given by the difference in the number of orthogonality conditions between the two sets of instruments (see Arellano and Bond (1991) for a formal derivation of this result). I refer to such a test as a J-difference test.

The test scheme between the alternative sets of orthogonality conditions are presented in figure 1. The procedure has been as follows: I have started at the top, with model V. If that specification is accepted on the basis of its J-statistic, the next model (with additional orthogonality conditions) has been considered; see model IV (neglect model II for the moment). That next model has been preferred if it does not fail on the basis of its J-statistic, or on the basis of the J-difference test for the two models.

As shown in figure 1, model I corresponds to using all leads and lags (of capital and the number of employees) as instruments. Model II uses only non-contemporaneous instruments, i.e. when instrumenting for growth rates from $t-1$ to t, only variables dated $t-2$ and earlier and $t+1$ and later are used as instruments. Model III restricts the instrument set only to predetermined variables, dated t and earlier, while model IV and V restrict the instrument sets to respectively $t-1$ and $t-2$ and earlier.

There is one problem with this procedure. An instrument set based on predetermined variables dated, say, $t - 1$ and earlier (see model IV in fig. 1) does not nest the instrument set based on non-contemporaneous variables (model II in fig. 1). There is no clear cut answer to which model to prefer when two such models are competing. In the present case, the estimates based on the non-contemporaneous instruments have been reported in the tables, but I have also discussed the estimates of the alternative specification in the text.