

Erling Holmøy and Torbjørn Hægeland

Aggregate Productivity Effects of Technology Shocks in a Model of Heterogeneous Firms: The Importance of Equilibrium Adjustments

Abstract:

This paper studies how productivity shifts at the level of the firm are transmitted to aggregate industry productivity in a model of heterogeneous firms. We analyse both uniform productivity shifts, and catching up by reducing the productivity differentials between firms. The two kinds of shifts affect aggregate productivity in different ways and through different mechanisms. Endogenous equilibrium adjustments play a crucial role for the influence on aggregate productivity. Moreover, when firms sell their output to several markets, and their market power differs between markets, aggregate productivity may be inversely related to productivity at the firm level. A by-product of the analysis is to demonstrate that productivity heterogeneity can be incorporated in the standard model of monopolistic competition at a low cost in terms of analytical tractability.

Keywords: Productivity, Heterogeneity, Aggregation, Monopolistic competition

JEL classification: D24, L11

Acknowledgement: The authors thank Tor Jakob Klette and seminar participants at the CPB Netherlands Bureau for Economic Policy Analysis, the Hague, for valuable comments and discussions.

Address: Erling Holmøy, Statistics Norway, Research Department. E-mail: erl@ssb.no
Torbjørn Hægeland, Statistics Norway, Research Department. E-mail: thd@ssb.no

* Corrected version, August 1997.

Discussion Papers

comprises research papers intended for international journals or books. As a preprint a Discussion Paper can be longer and more elaborated than a usual article by including intermediate calculation and background material etc.

Abstracts with downloadable postscript files of
Discussion Papers are available on the Internet: <http://www.ssb.no>

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
P.O. Box 8131 Dep
N-0033 Oslo

Telephone: +47 22 00 44 80
Telefax: +47 22 86 49 76
E-mail: Salg-abonnement@ssb.no

1. Introduction

The distinction and relationship between production functions for individual firms and the aggregate production function for an industry have long been recognised. Johansen (1972) still stands out as one of the most elaborate and rigorous studies of how these concepts should be interpreted. This study built on the insight provided by Houthakker (1955-56), Johansen (1959) and Salter (1960)¹, which recognised and focused on two fundamental points. First, firms even within narrowly defined industries, differ with respect to size, performance and productivity. This observation would neither be sensational nor challenging if such inter-firm differences were modest and temporary. However, there has long been strong empirical evidence that such differences are both substantial and persistent. Recent research on micro data has strengthened this conclusion, see for example Sutton (1997), Baily, Hulten and Campbell (1992), Klette (1994), Klette and Mathiassen (1995, 1996). Second, relationships between aggregate output and inputs reflect economic equilibrium adjustments as well as properties of the micro production functions. Aggregate industry variables will be weighted sums of the corresponding micro variables. More concretely, aggregate factor productivities are by definition output weighted averages of the productivity in the individual firms. Obviously, the autonomy, and thereby the value of such average concepts as «deep» structural parameters in aggregate analyses, deteriorates when the weights in the average are likely to be sensitive to shifts in different variables. The more inter-firm variation in productivity and endogenous variation of the relative firm size, the more serious are the violations of the conditions for exact aggregation, and the less is the adequacy and relevance of the notion of the «representative firm».

Despite massive empirical evidence of technology heterogeneity and the theoretical contributions pointing out the need for replacing the representative firm by a more sophisticated model of aggregate producer behaviour, cost heterogeneity has typically been ruled out by assumption in popular models of aggregate industry behaviour. On the other hand, the role of product differentiation has been given considerable attention, especially in the new literature of international trade and economic growth. One of the most popular analytical tools in these fields of research is the model of the Large Group case of Chamberlinian Monopolistic Competition (LGMC), see e.g. Helpman and Krugman (1985) and Helpman and Grossman (1992). However, this literature generally assumes that all firms have the same production function, even if the seminal work of Chamberlin (1933) and Stigler (1949) argued

¹ Also in Marshall's theory of value, as laid out in Frisch (1950), the argument for assuming decreasing returns to scale at the aggregate industry level relies on the existence of productivity differentials between firms combined with a selection mechanism, which ensures that only profitable firms are active.

that product differentiation is unlikely to exist without non-uniform costs. One exception from this tradition is Montagna (1995).

The present paper contributes to clarify how aggregate industry productivity depends both on productivity of individual firms, and on the market structure determining equilibrium output of firms. The very presence of productivity differentials between firms in the model, allows us to study the aggregate effects of technical change through catching up, i.e. reduction of the productivity gaps between firms. We compare these effects with those resulting from uniform productivity shifts in all firms. The market structure influences the aggregate productivity effect of the shifts by determining the changes in equilibrium firm outputs, and thereby the weights in the relevant weights in the average productivity calculation. We consider a model where firms have access to two segmented markets differing in competitiveness. This enables us to identify the aggregate productivity effect of endogenous adjustments of the allocation of output between different markets. In this respect we distinguish what we refer to as a perfectly competitive export market from a monopolistically competitive domestic market. In the export market all firms face a perfectly elastic demand function represented by an exogenous world price. Domestic consumers are, on the other hand, assumed to recognise products from different firms as imperfect substitutes, and the equilibrium in the domestic product markets is described according to the LGMC model. However, our version of the LGMC model becomes asymmetric due to productivity differentials among firms. A by-product of our study of the transmission mechanism of firm-specific productivity shifts into equilibrium changes in aggregate productivity is that we demonstrate a formal specification of productivity heterogeneity in the LGMC framework, that does not make the model intractable with respect to closed form analytical results.

By considering markets where both demand and supply simultaneously determine equilibrium prices and quantities, our approach differs drastically from the one taken in Houthakker (1955-56) and Johansen (1972). The latter studies presume that firms have different Leontief micro production functions up to an exogenous capacity level. Thus, profitable firms will produce at their exogenous capacity level and non-profitable firms will not produce at all. In this model, the endogeneity of aggregate productivity is due to the endogeneity of which firms that will be profitable. Prices of both factors and products are exogenous, giving no role for the demand side to play in the determination of profits. Our model framework has much more in common with Montagna (1995), but differs by imposing more structure on the inter-firm productivity differentials, yielding a return in terms of

analytical tractability. However, our focus on aggregate productivity differs from the issues analysed by Montagna.

The main results of the paper can be summarised as follows: Uniform productivity shifts are found to affect aggregate productivity in a way fundamentally different from productivity shifts through catching up. The change in aggregate productivity due to technology shocks at the level of the firm is substantially affected by endogenous equilibrium adjustments. The deviation of aggregate productivity from the productivity of the most efficient firm is confined to an interval whose width is determined by the substitutability of the products of the firms in the industry. The equilibrium adjustments may also cause aggregate productivity to be inversely related to productivity at the firm level.

Especially transparent results are obtained by analysing an approximate solution, where we neglect that the entry-exit condition causes the set active firms to be finite. A closer examination of the accuracy of this approximation shows that it will be very high if one accepts the standard assumptions underlying the LGMC case.

The paper is organised as follows. Section 2 outlines the partial equilibrium model of the industry. Section 3 derives comparative statics results for both a uniform shift in firms' productivities and catching up through a reduction of the productivity gaps between firms. Section 4 concludes.

2. The model of a heterogeneous industry

2.1. The general model

The model is designed to capture the fact that most firms sell to several markets in which they have different degrees of market power. In applied models, the most important distinction is typically made between the export market and the domestic market. Empirical studies of firm data suggest that the export share of output is positively correlated with productivity, cf. the survey of Bartelsman and Doms (1997). In order to capture these patterns, we assume the export and the domestic markets to be segmented from each other. Moreover, we assume that variable costs of export deliveries are additively separable from variable costs of domestic deliveries. In each of the n active firms in the industry, the total cost function takes the form

$$(1) \quad C_i = c_i \left[(X_i^W)^{1/s} + (X_i^H)^{1/s} \right] + F$$

where $i \in [0, n]$, C_i is total costs in firm no. i . X_i^H and X_i^W are deliveries to the domestic and the export market respectively, $0 < s \leq 1$ is the scale elasticity, c_i is the inverse of a productivity parameter specific to firm i , and F is fixed costs, associated with e.g. product development. The variable cost function is assumed to take the same convex form for both kinds of deliveries. The gain from this restrictive assumption is a simple representation of two phenomena that we want to capture: (i) The possibility of decreasing returns to scale at the level of the firm, and (ii) that specialising in deliveries to a single market may be costly. Appendix A shows how (1) can be derived from explicit assumptions about the underlying production function.

Whereas the fixed cost element is assumed to be identical for all firms, productivity differentials cause variable costs to differ between firms. We rank firms according to productivity, so that firm 0 is the most efficient firm². Productivity heterogeneity is formalised in a simple way by assuming the relative productivity differentials between any two adjacent firms to be constant. For analytical convenience, the set of firms and varieties is treated as a continuum. We also ignore the problem that i should be an integer:

$$(2) \quad \frac{dc_i}{di} = tc_i \Leftrightarrow c_i = ce^{ti}, \quad t > 0,$$

where $(-t)$ is proportional to the relative productivity differential between firms³. $c_0 = c$ is exogenous. The reward for restricting productivity heterogeneity to be represented by an exponential structure is a tractable model of aggregate behaviour. As will be shown in the following, equilibrium solutions for all variables associated with the heterogeneous firms turn out to be related by exponential functions of i . This makes it straightforward to carry out the relevant integrals defining the corresponding aggregate variables⁴.

² As will be evident from the subsequent analysis, the relevant variable for ranking firms is profits. However, in this model profits are an increasing function of productivity, so this is a point of no practical significance.

³ More precisely, the variable cost function is derived from a production function of the form $X = aV^s$, where a is the productivity parameter and V is a variable input. When the productivity differentials take the exponential form $\dot{a}_i = \gamma a_i$, it follows that $t = -\gamma/s$.

⁴ In this respect, our approach has much in common to the derivation of aggregates in the perpetual youth overlapping generation model used by Blanchard (1985) and Yaari (1965). In the perpetual youth model, the assumption that the probability of dying is constant, independent of age, implies an exponential structure between the variables associated with each cohort. This is the key to the tractability of the aggregate model.

The model makes it possible to study two kinds of productivity shifts: Catching up can be represented by a decrease in the productivity differentials, t , whereas a uniform improvement of the productivity of all firms may be represented by a decline in c .

As noted above, the rationale for specifying two markets is that the market structure is important for how productivity shifts at the level of the firm are transmitted to aggregate productivity. In the foreign market, we assume that the price is exogenous and identical for all firms. On the other hand, producer behaviour in the domestic market is characterised by monopolistic price setting. The difference in market power between the two markets is consistent with several empirical studies, see e.g. Aukrust (1970) and Bowitz and Cappelen (1994) for studies on Norwegian data. In the domestic market, consumers consider the products of the n firms in the industry to be imperfect substitutes. Still, products supplied by different firms within the industry have so much in common that they constitute a separable differentiated good in the demand structure. Consumer preferences over the products belonging to the differentiated good are assumed to take the symmetric CES-form, usually referred to as Spence-Dixit-Stiglitz preferences. Equilibrium in the domestic market for each variety then implies:

$$(3) \quad X_i^H = \left(\frac{P_i^H}{P} \right)^{-\sigma} A P^{-\varepsilon},$$

where P_i^H is the price of domestic deliveries of variety i , P is the consumer price index in the home country for this composite industry good, $\sigma > 1$ is the elasticity of substitution between the varieties constituting the differentiated product, $-\varepsilon$ is the own price elasticity of the composite good, and A is a constant. We assume that $\sigma > \varepsilon$. The price index, P , consistent with the CES subutility function, becomes:

$$(4) \quad P = \left(\int_0^n (P_i^H)^{1-\sigma} di \right)^{1/(1-\sigma)}$$

Firms maximise profits, defined by

$$(5) \quad \pi_i = P_i^H X_i^H - c_i (X_i^H)^{1/s} + P^W X_i^W - c_i (X_i^W)^{1/s} - F$$

with respect to P_i^H and X_i^W , subject to the common exogenous world price level and the perceived demand schedule in the home market. The own price elasticity of perceived demand equals $-\sigma$ for all firms, since each firm neglects the influence of its own price on the price index. Because of the separability of variable costs, the profit maximisation problem is separable. Operating profits from domestic sales, $\pi_i^H = P_i^H X_i^H - c_i (X_i^H)^{1/s}$, is maximised with respect to P_i^H subject to the perceived demand elasticity in the domestic market. Operating profits from exports, $\pi_i^W = P^W X_i^W - c_i (X_i^W)^{1/s}$ is maximised with respect to X_i^W .

The optimal price setting in the domestic market follows the familiar mark-up rule

$$(6) \quad P_i^H = m \frac{c_i}{s} (X_i^H)^\lambda,$$

where $m = \sigma/(\sigma - 1)$ is the mark-up factor and $\lambda = 1/s - 1$ is the elasticity of marginal costs with respect to output. Note that in the present asymmetric model, the positive bias of the perceived price elasticity compared to the true elasticity differs between firms. The combination of cost differentials and mark-up pricing will be shown to cause the budget share of variety i to be a decreasing function of i since $\sigma > 1$. Thus, for a fixed number of active firms, the most efficient firm underestimates its own market power to a greater extent in this asymmetric model compared to the corresponding misperception made by the representative firm in the symmetric model. On the other hand, for the least productive active firm, $-\sigma$ will be closer to the true demand elasticity than in the symmetric model.

Optimal export deliveries implies equality between the world price and marginal cost of export deliveries

$$(7) \quad P^W = \frac{c_i}{s} (X_i^W)^\lambda.$$

An industry equilibrium where incentives to enter or exit have been eliminated, requires that operating profits equals fixed cost for the marginal firm:

$$(8) \quad \pi_n = F.$$

In order to derive industry aggregates, we utilise that the solutions for the individual variables of the same category are related to each other by exponential functions. To see this, substitute (6) into the domestic product market equilibrium condition (3). By exploiting the exponential structure of productivity differentials assumed in (2), it can be verified that the domestic market deliveries will differ between adjacent firms at a constant rate:

$$(9) \quad X_i^H = X_0^H e^{\left(\frac{-\sigma}{1+\sigma\lambda}\right)i},$$

where it can be shown that

$$(10) \quad X_0^H = \left[A \left(\frac{mc}{s} \right)^{-\sigma} P^{\sigma-\varepsilon} \right]^{-\frac{1}{1+\sigma\lambda}}.$$

The structure of domestic prices becomes

$$(11) \quad P_i^H = P_0^H e^{\frac{t}{1+\sigma\lambda}i},$$

where $0 < \frac{t}{1+\sigma\lambda} < t$ and $P_0^H = \frac{mc}{s} (X_0^H)^\lambda$.

The exponential structure of variety prices makes it possible to find a closed form expression for P by inserting (11) into (4). After integration, one can solve the resulting equation with respect to the

domestic price index P , utilising that $-\frac{\sigma t}{1+\sigma\lambda} = -\frac{mt}{m/s-1} < 0$:

$$(12) \quad P = \left\{ A^\lambda \left(\frac{mc}{s} \right) \left[\left(\frac{m/s-1}{t} \right) (1-e^N) \right]^{\frac{1+\sigma\lambda}{1-\sigma}} \right\}^{\frac{1}{1+\varepsilon\lambda}}$$

where $N = -\frac{nt}{m/s-1}$.

$0 < e^{mN} < 1$ equals domestic deliveries from the marginal firm relative to those from the most efficient firm. By combining (2) and (7) deliveries to the export market can also be shown to vary exponentially between firms with different productivity.

$$(13) \quad X_i^W = X_0^W e^{-(t/\lambda)i},$$

$$\text{where } X_0^W = \left(\frac{sP^W}{c} \right)^{1/\lambda}.$$

Thus, export is a decreasing function of i , that is, the lower the productivity, the lower is output. It is easy to show that maximised operating profits obtained by firm i in the domestic market and the export market, respectively, becomes:

$$(14) \quad \pi_i^H = \left(\frac{m}{s} - 1 \right) c_i (X_i^H)^{1/s}, \quad \pi_i^W = (1-s)P^W X_i^W.$$

Despite the separability of variable costs for export and domestic deliveries for an individual firm, the two markets are connected at the aggregate industry level through the determination of the number of active firms. Using (14) the entry-exit condition (8) can be written:

$$(8') \quad \left(\frac{m}{s} - 1 \right) c_n (X_n^H)^{1/s} + (1-s)P^W X_n^W = F.$$

Using the expressions (9) - (13), (8') implicitly determines a unique solution for n in terms of exogenous variables, since total operating profits can be shown to be a monotonically decreasing function of n .

The primary purpose of this paper is to explore the nature of the relationship between aggregate industry productivity and the set of firm specific productivities. We choose the productivity of variable inputs as our productivity concept. Since factor prices are fixed, this concept can equivalently be measured by the aggregate variable unit costs defined as

$$(15) \quad \bar{C} = \frac{C}{X}$$

where aggregate industry costs, C , and output, X , is defined as

$$(16) \quad X \equiv X^H + X^W = \int_{i=0}^n X_i^H di + \int_{i=0}^n X_i^W di$$

$$(17) \quad C \equiv C^H + C^W = \int_{i=0}^n c_i (X_i^H)^{1/s} di + \int_{i=0}^n c_i (X_i^W)^{1/s} di$$

Alternatively, we could have included the aggregate fixed costs, nF , in our productivity measure. nF will adjust endogenously due to entry or exit at the aggregate level. As a matter of fact, the subsequent analysis will provide all information necessary to assess the changes in nF/X . We have two reasons for disregarding fixed costs. First, the relative importance of these costs is of course strongly dependent on the size of the exogenous F . The LGMC model presumes all equilibria to have many firms, which is inconsistent with a large value of F , see (8'). Consequently, the changes in these two productivity concepts will be approximately the same. Second, we have found the forces determining the changes in variable productivity to be most interesting from an analytical point of view. Our choice of productivity concept sharpens the focus on these forces.

3. The links between productivity parameters at the firm level and aggregate productivity

3.1. Decomposing aggregate productivity

Aggregate industry productivity is a weighted average in two dimensions in our model. First, it includes an average of productivities related to export and domestic deliveries. This is clearly seen from the following decomposition of aggregate total average costs

$$(18) \quad \bar{C} = \left(\frac{X^H}{X} \right) \bar{C}^H + \left(\frac{X^W}{X} \right) \bar{C}^W,$$

where $\bar{C}^j = C^j/X^j$, $j = H, W$ is the aggregate variable average costs associated with deliveries to market j . Second, aggregate industry productivity includes a weighted average of firm specific costs. We have

$$(19) \quad \bar{C}^j = \int_0^n \left(\frac{X_i^j}{X^j} \right) \left(\frac{C_i^j}{X_i^j} \right) di.$$

We utilise the exponential structure of the equilibrium solutions for prices and deliveries, to express aggregate deliveries and costs as

$$(20) \quad X^H = X_0^H \left(\frac{m/s-1}{mt} \right) (1 - e^{mN}).$$

$$(21) \quad X^W = X_0^W \frac{\lambda}{t} (1 - e^{-(t/\lambda)n}).$$

$$(22) \quad C^H = \left(\frac{m/s-1}{t} \right) c (X_0^H)^{1/s} (1 - e^N)$$

$$(23) \quad C^W = \frac{\lambda}{t} c (X_0^W)^{1/s} (1 - e^{-(\lambda/t)n})$$

The rest of this section is concerned with a close examination of the components of these averages.

3.2. The scope for heterogeneity to make aggregate productivity differ from productivity in the most efficient firm

It is straightforward to verify that the equilibrium solution for aggregate export productivity takes the simple form

$$(24) \quad \bar{C}^W = sP^W.$$

Thus, \bar{C}^W is independent of c and t . The intuition is simple: Each firm equates marginal cost of export deliveries to the fixed world price P^W . For each firm, average variable costs of exports constitute a fraction s of the corresponding marginal cost, so variable average cost defined over all firms must equal sP^W . Thus, aggregate productivity for deliveries to a market where all firms face the same exogenous price is unaffected by the degree of heterogeneity in the productivity parameter c when the scale elasticity is constant and common for all firms. It follows that a pure exporting

industry will never experience productivity growth at the aggregate level as long as sP^W remains constant.

Aggregate productivity of domestic deliveries is determined in a more complex way. In terms of unit costs, it can be expressed as

$$(25) \quad \bar{C}^H = \frac{c(X_0^H)^{1/s} \int_0^n e^{-\left[1 - \frac{\sigma}{s(1+\sigma\lambda)}\right]ti} di}{X_0^H \int_0^n e^{-\frac{\sigma}{s(1+\sigma\lambda)}i} di} = m\bar{C}_0^H \left(\frac{e^N - 1}{e^{mN} - 1} \right) = mc(X_0^H)^\lambda \left(\frac{e^N - 1}{e^{mN} - 1} \right).$$

The scope for heterogeneity to cause \bar{C}^H to deviate from \bar{C}_0^H is seen from (25) to depend on the term $mg(N)$, where $g(N) = (1 - e^N)/(1 - e^{mN})$.

The reason why m turns up in (25) is that $m\bar{C}_0^H$ can be seen to represent the ratio between the integrals defining C^H and X^H when the integrals are calculated to infinity instead of n . When n grows beyond all limits, \bar{C}^H / \bar{C}_0^H is dominated by the ratio between the growth rate of X_i^H and the growth rate of C_i^H wrt. i . The growth rates of X_i^H and C_i^H wrt. i , equal $-\sigma t/(1 + \sigma\lambda)$ and $t - \sigma t/[s(1 + \sigma\lambda)]$, respectively. Using $m = \sigma/(\sigma - 1)$ and $\lambda = 1/s - 1$, the fraction between these negative growth rates equals $[-\sigma t/(1 + \sigma\lambda)] / \{t - \sigma t/[s(1 + \sigma\lambda)]\} = m$. This transformation underlies the second equality in (25). Note that the mark-up ratio itself is no primary determinant to \bar{C}^H , but under the price setting assumptions in the LGMC model, it compactly summarises the effects on \bar{C}^H / \bar{C}_0^H of the equilibrium adjustments of demand to the price differentials between the differentiated products.

Turning to why $g(N)$ turns up in (25), this factor captures the correction that must be made when the number of firms is finite compared to the hypothetical limiting case. It is easily seen that $g(N)$ converges to unity when N increases beyond limits, either because of entry or increased heterogeneity. Moreover, this function $g(N)$ is strictly increasing and concave with and $\lim_{N \rightarrow \infty} g'(N) = 0$, so the convergence is from below. Using L'Hopital's rule it can be shown that $\lim_{N \rightarrow 0} g(N) = 1/m$, which confirms the obvious intuitive result that \bar{C}^H and \bar{C}_0^H coincides if firm no. 0 is the only active firm. We also have $\lim_{N \rightarrow 0} g'(N) = \infty$ indicating that $g(N)$ converges relatively fast towards the asymptotic

value 1 when n increases from 0 into the range where the LGMC model is an appropriate model of market behaviour.

We are now able to draw a surprisingly sharp and general conclusion about the scope for heterogeneity to cause \bar{C}^H to deviate from \bar{C}_0^H , i.e. to create aggregate scale diseconomies through entry and exit. It follows from (25) and the properties of the $g(N)$ function that \bar{C}^H is bound to lie within the interval

$$(26) \quad \bar{C}_0^H < \bar{C}^H < m\bar{C}_0^H .$$

The relative width of this interval is given by $m = \sigma/(\sigma - 1) > 0$, which is smaller the higher is the substitutability among the differentiated products. This result was also shown in Holmøy (1997) in a model with more restrictive assumptions. Note that if the LGMC is an appropriate model, the range of variation will be very narrow for two reasons. First, the model supposes that the products are imperfect but close substitutes, which implies that m is relatively close to unity. Second, sticking to the price setting behaviour as in the LGMC model for all firms, we have implicitly assumed that even the largest firm, i.e. no. 0, has an insignificant market share. Then N is large, and the extremely concave shape of $g(N)$ implies that variations in N takes place within a range in the interval where $g'(N)$ is close to 0. Consequently, the LGMC model of the domestic market contains endogenous mechanisms which causes the productivity in the most efficient firm to be a very good predictor of the average productivity of output delivered to this market. Neither the degree of heterogeneity between adjacent firms, measured by t , nor the absolute gap between the most efficient and the infinitely inefficient firm will affect the relative width of the interval in this model. It is also independent of the scale elasticity of the individual production functions (provided of course non-increasing returns to scale in variable inputs).

The provocative part of (26) is of course the existence of the upper asymptote for \bar{C}^H despite the fact that the entering firms becomes successively less efficient when n increases beyond limits. The principal reason for this result is that the equilibrium level C_i^H is not increasing, but decreasing in i . This is indeed a condition for the integral defining C^H to converge when n approaches infinity. Formally, this is easily seen from inspection of the growth rate for C_i^H wrt. i , which equals

$t - \sigma t/[s(1 + \sigma\lambda)] = mt/(1 - m/s) < 0$. The intuitive explanation is that the negative demand response to an upward shift in the marginal cost curve is strong enough to reduce total variable costs of domestic deliveries.

When firm technologies exhibit constant returns to scale, i.e. $\lambda = 0$ and $\bar{C}_0^H = c$, then not only the range of potential variation in \bar{C}^H/\bar{C}_0^H is limited. We can also conclude that the range of potential variation in the absolute level of \bar{C}^H is limited.

It follows that the existence of and variation in heterogeneity, will have a very little influence on the determination of the aggregate productivity level, unless \bar{C}_0^H is affected. Such an impact is proportional to λ , for a given impact on X_0^H . From (9)-(11) we find that X_0^H is related to exogenous variables and n by

$$(27) \quad X_0^H = \left\{ A \left(\frac{mc}{s} \right)^{-\varepsilon} \left[\left(\frac{m/s-1}{t} \right) (1 - e^N) \right]^{-(\sigma-\varepsilon)/(\sigma-1)} \right\}^{1/(1+\varepsilon\lambda)}.$$

The degree of heterogeneity, measured by t , influences $X_0^H = (P_0^H/P)^{-\sigma} AP^{-\varepsilon}$ through the price index P . P is proportional to the term in the square brackets in (27) raised by $-1/(\sigma-1)$. For a fixed n , a rise in t will have a positive effect on each price in terms of P_0^H . This will reduce P_0^H/P . A closer study of how \bar{C}_0^H will change due to equilibrium adjustments of X_0^H is one of the tasks carried out in the next two sections.

3.3. Approximate results in the case where the marginal firm is negligible

In order to clarify the mechanisms that turn out to be the strongest determinants of \bar{C} , we will in this section first consider approximate solutions where the effects working through N are neglected. In section 3.4, we return to an interpretation and assessment of the deviations between the approximate solutions and the true ones.

From the formal expressions for the terms in (18), it is clear that the degree of inaccuracy caused by this simplifying approximation is small when N is large. Since $N = nt/(1 - m/s) < 0$, setting $e^N = 0$ is a

tolerable approximation when we restrict the analysis to the case where the marginal firm is of negligible size relative to the industry as a whole in terms of output and costs. We will therefore refer to this case as the «Negligible Marginal Firm» (NMF) case. This case is relevant when there is a large number of firms, which has already been assumed. In section 3.4 we formally show the intuitive result that n is a decreasing function of F . Thus, the smaller the entry barrier, the more relevant is the NMF case and the LGMC model. The accuracy of the NMF case is also larger the more asymmetric is the equilibrium. More asymmetry is created by increasing t , which makes a firm $i > 0$ smaller relative to the most efficient one.

The inaccuracies introduced when moving from the true solution to the NMF approximation, are of two distinct kinds. First, since the marginal firm is insignificant wrt. output and costs relative to the industry as a whole, the deliveries and costs in the unprofitable inactive firms will be even less significant. Consequently, we overestimate, *cet. par.*, aggregate deliveries and costs, but to a negligible extent, when contributions from these firms are included when adding over firms. Equivalently, but more formally, there will be a negligible difference between the finite and the infinite integrals required to find aggregate productivity.

The second kind of inaccuracy turns up in the price index P . This price index also involves a finite integral of product prices, which is replaced by an infinite integral in the NMF approximation. The effect is that the negative love-of variety effect on P is maximised for a given t . Thus, when we use the NMF approximation to study shifts in c and t , we disregard changes in the strength of this effect. We will explain in more detail in section 3.4 that the underestimation of P in the NMF case contributes to a negative bias in the determination of the relative size of the most efficient firm.

It is obvious from (24) that the aggregate productivity of exports is unaffected by the NMF approximation. Aggregate export deliveries will be biased upwards in the NMF solution because exports from non-active firms are included whereas P is irrelevant. For \bar{C}^H and X^H the two kinds of biases have opposite sign, so neglecting both improves the accuracy of the NMF approximation compared to neglecting only one of them.

Note that when $\lambda = 0$, the approximate NMF solution for \bar{C}^H , denoted by $\bar{C}^H|_{NMF}$, will unambiguously overestimate the true solution. In this case the approximate solution will be equal to

the upper limit of the interval restricting the true solution, i.e. $\bar{C}^H|_{NMF} = mc$. There is no modifying cost effect from the negative bias in X_0^H .

Accepting the NMF approximation, it is clear that aggregate productivity of domestic deliveries can be studied by analysing the productivity of the most efficient firm⁵. In the NMF case the following closed form solution for this equilibrium productivity can be derived from (25) and (27):

$$(28) \quad \bar{C}_0^H|_{NMF} = c \left\{ A \left(\frac{mc}{s} \right)^{-\varepsilon} \left(\frac{m/s-1}{t} \right)^{-(\sigma-\varepsilon)/(\sigma-1)} \right\}^{\lambda/(1+\varepsilon\lambda)}.$$

From (28) we get the elasticities in the NMF case:

$$(29) \quad 0 < \text{El}_c \left(\bar{C}_0^H|_{NMF} \right) = \frac{1}{1+\varepsilon\lambda} \leq 1,$$

$$(30) \quad \text{El}_t \left(\bar{C}_0^H|_{NMF} \right) = \left(\frac{\sigma-\varepsilon}{\sigma-1} \right) \left(\frac{\lambda}{1+\varepsilon\lambda} \right) = \lambda \left(\frac{\sigma-\varepsilon}{\sigma-1} \right) \text{El}_c \left(\bar{C}_0^H|_{NMF} \right) > 0.$$

The intuition behind (29) can be acquired through the following argument. If X_0^H were constant, raising c by 1 percent contributes directly to a 1 percent increase in \bar{C}_0^H . However, if there are strictly decreasing returns to scale ($\lambda > 0$), the equilibrium adjustment of X_0^H implies a negative modification of \bar{C}_0^H . The reason is that the marginal cost functions of all firms will shift upwards by 1 percent. As long as domestic deliveries from the infinite mass of firms are fixed, P will increase by 1 percent, causing a drop in demand equal to ε percent. As long as t is constant, the relative distribution of the domestic deliveries will remain constant, so that the decline in aggregate demand is met by the same relative reduction in the delivery supplied by each firm. When $\lambda > 0$, this delivery adjustment reduces

⁵ As a matter of fact, the NMF assumption of an infinite number of firms makes the model work in exactly the same way with respect to shifts in c as a partial equilibrium model where the supply side consists of one single firm facing a demand function of the form $X = AP^{-\varepsilon}$. In such a model optimal producer behaviour would imply price setting according to $P = (mc/s)X^\lambda$. It is straightforward to verify that average cost in equilibrium becomes $\bar{C} = bc^{1/[1+\varepsilon\lambda]}$, where b is a positive constant.

marginal costs in all firms. This will be carried forward to domestic prices and P . In equilibrium the contribution from the adjustment of X_0^H on \bar{C}_0^H equals $-\varepsilon\lambda/(1 + \varepsilon\lambda)$ percent. Adding the 1 percent direct shift in the cost function yields (29).

The increase in \bar{C}_0^H caused by increasing t , is generated by a different mechanism. In this case the most efficient firm is the only firm that does not experience a shift in its cost function. A positive impact on \bar{C}_0^H therefore requires strictly decreasing returns to scale combined with an increase in X_0^H . This is caused by the dominance of what we will refer to as an *internal* substitution effect. More precisely, the shift in marginal domestic costs implied by increasing t makes all but product 0 more expensive for domestic consumers. They will redirect their demand towards product 0. This internal substitution effect is proportional to the rise in P/P_0^H , which in the NMF case equals $[(m/s-1)/t]^{1/(\sigma-1)}$, so the term $1/(\sigma-1)$ in (30) is the elasticity of P/P_0^H wrt. t . If total domestic demand towards the industry were constant, a one percent increase in P/P_0^H would cause a relative expansion of X_0^H equal to $\sigma/(\sigma-1)$. However, this internal substitution effect is modified by the reduction in aggregate domestic demand through what we may refer to as an *external* substitution effect. The net effect equals $(\sigma - \varepsilon)/(\sigma - 1)$, which is positive since the varieties within the industry are closer substitutes to each other than to products classified in another industry. The modification introduced through the term $\lambda/(1 + \varepsilon\lambda)$ is due to the marginal cost effect of adjusting X_0^H .

The last equality in (30) establishes the relationship between aggregate domestic productivity effects from c and t . Productivity improvements induced by catching up instead of uniform shifts in c , depend critically on the dominance of the internal over the external substitution effect, and on diseconomies of scale within each firm. In the special case of NMF combined with constant returns to scale at the firm level, it is impossible to improve aggregate domestic productivity through catching up; reducing c is the only potent source. Again, this result is generated by the endogenous equilibrium adjustment of the domestic deliveries at the micro level, which constitute the relevant weights when calculating average domestic productivity.

We now turn to a closer examination of the changes in the output shares in (18). It follows from (9), (10) and (13) that, in the NMF case, aggregate deliveries can be written

$$(31) \quad X^H|_{NMF} = \left(\frac{m/s-1}{mt} \right) \left\{ A \left(\frac{mc}{s} \right)^{-\varepsilon} \left(\frac{m/s-1}{t} \right)^{-(\sigma-\varepsilon)/(\sigma-1)} \right\}^{1/(1+\varepsilon\lambda)},$$

$$(32) \quad X^W|_{NMF} = \frac{\lambda}{t} \left(\frac{sP^W}{c} \right)^{1/\lambda}.$$

From (31) and (32) we find the NMF-case elasticities of aggregate deliveries:

$$(33) \quad \text{El}_c \left(X^W|_{NMF} \right) = -\frac{1}{\lambda}$$

$$(34) \quad \text{El}_t \left(X^W|_{NMF} \right) = -1$$

$$(35) \quad 0 > \text{El}_c \left(X^H|_{NMF} \right) = -\frac{\varepsilon}{1+\varepsilon\lambda} > -\frac{1}{\lambda}$$

$$(36) \quad \text{El}_t \left(X^H|_{NMF} \right) = -1 + \left(\frac{\sigma-\varepsilon}{\sigma-1} \right) \left(\frac{1}{1+\varepsilon\lambda} \right) > -1$$

A comparison of (33) and (35) shows that reductions in c will increase aggregate export deliveries relatively more than aggregate domestic deliveries. A similar result holds when t is changed. This implies that aggregate productivity may be inversely related to shifts in the productivity parameters. This possibility can also be seen directly from (18). For example, in the case of export productivity being lower than productivity of domestic deliveries, aggregate productivity will decline if shifts in c and t cause a sufficiently strong increase in the export share of total output. The possibility and strength of such an effect have nothing to do with productivity heterogeneity among firms. Such a “weighing effect” will exist even if the industry consists of firms with identical technology, or a representative firm. However, the magnitude of the adjustments of productivity components and weights entering (18) are, in different degrees, affected by the existence of productivity differentials.

We have calibrated a numerical version of the model to real data for manufacturing of chemical and mineral products in Norway in 1992. The computations displayed in Figure 1 illustrate the possibility

of an inverse relationship between the productivity parameters and aggregate productivity performance, and show how large the shifts in c must be to bring about such an inverse relationship when parameters and variables have reasonable initial values. The figure shows the effects on aggregate average variable costs, of 50 reductions, each by 1 percent, in c . For smaller reductions in c , average costs for the industry as a whole decreases (industry productivity increases). But as c is reduced beyond a critical level, industry productivity growth is diminished, and eventually even reversed, due to the weighing effect. Average variable costs of exports are independent of c and t , whereas average variable costs of domestic deliveries decline steadily. Export deliveries increase faster than domestic deliveries, see Figure 2, causing the weights in the average industry costs to shift in favour of exports for which productivity remains constant. Eventually, this weighing effect outweighs the increased productivity of domestic deliveries. Simulations of changes in t reveal a similar pattern.

Figure 1. Aggregate average costs

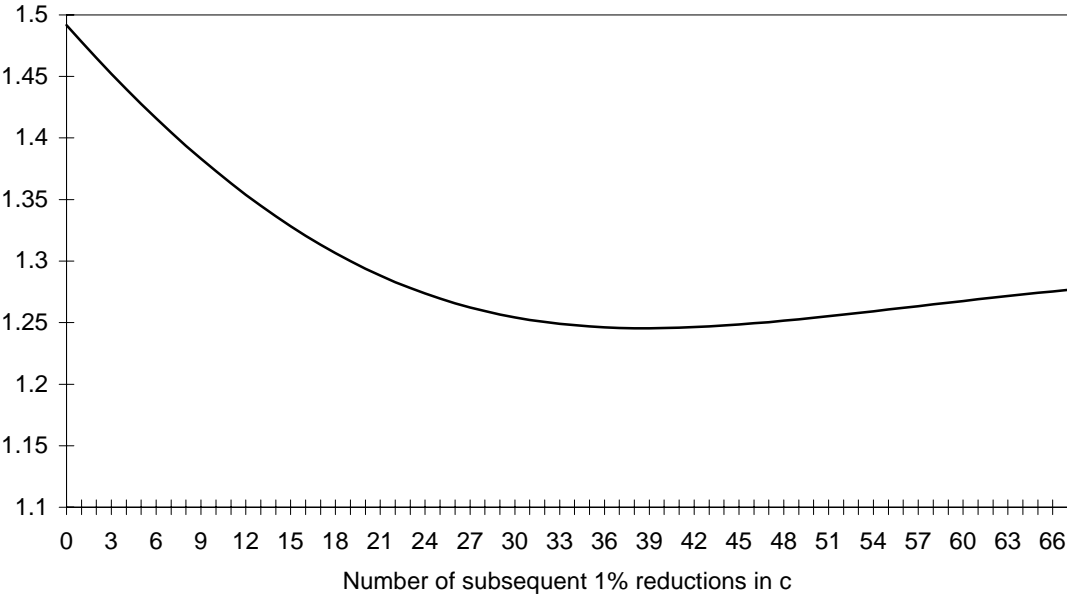
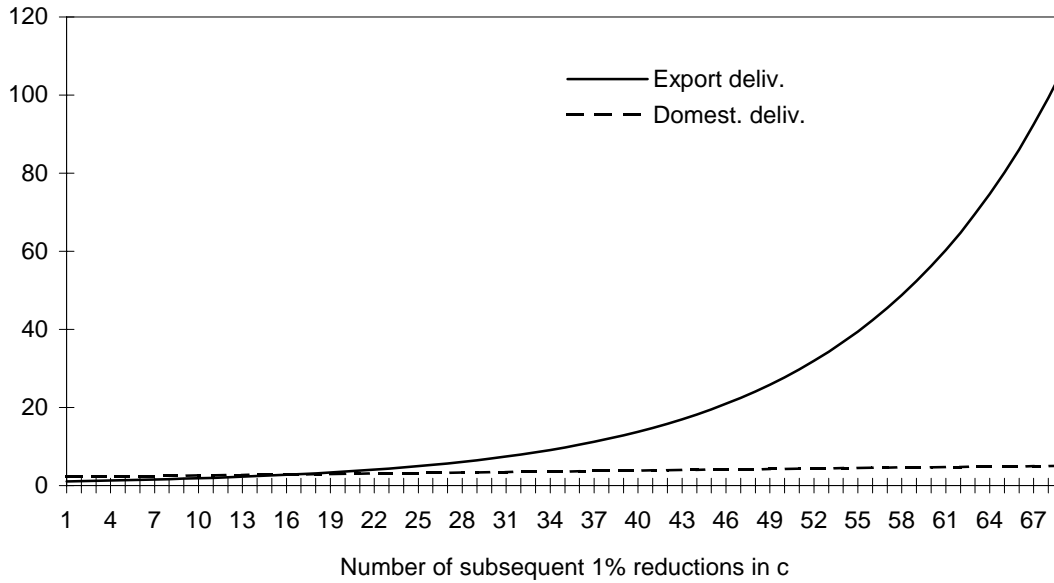


Figure 2. Aggregate output (billions NOK)



To summarise the analysis of the NMF case, we have found that in both markets there is a fixed relationship between aggregate productivity and the productivity of the most efficient firm. This relationship is independent of both the degree of productivity heterogeneity and the scale elasticity. We have also seen that shifts in c and t have fundamentally different effects on aggregate productivity in the domestic market. Shifts in t have no effects on aggregate productivity if there are constant returns to scale. Compared to the effects of shifts in c , the magnitude of the effects of t depends critically on the strength of the internal substitution effect. Aggregate export deliveries are more sensitive to shifts in c and t than domestic deliveries. Combined with the effects of market structure on productivity in the different markets, this implies that aggregate productivity may be inversely related to the productivity parameters at the level of the firm, if ε is sufficiently low. Productivity heterogeneity is not necessary for such a possibility to exist, but it generally affects the strength of the effect.

3.4. A closer examination of the bias of the approximate results

In this section we examine in more detail how the results of the NMF case will be modified when the marginal firm contributes significantly to the determination of the industry aggregates. For \bar{C}^H , it is clear from (26) that the relevance of such an examination is increasing in m (and thus decreasing in σ).

Determination of the number of firms

We start by analysing how changes in c and t will affect the equilibrium number of active firms. Since the fixed cost requirement is the same for all firms, variable profits of the new marginal firm will equal variable profits of the old one. We use the labels «old» and «new» marginal firm to distinguish the marginal firm ex ante changes in n (old) from the marginal firm ex post changes in n (new). These labels should of course not be treated too literally since n is treated as a continuous variable. Since the single entry decision implies deliveries to both the export and the domestic market simultaneously, the entry incentive is a weighted average of the market specific entry incentives, being i) profits earned on exports and ii) profits from domestic sales. This was formalised in (8') where n is a function of c and t .

Logarithmic differentiation of (8') wrt. c yields the total equilibrium elasticity of n wrt. c

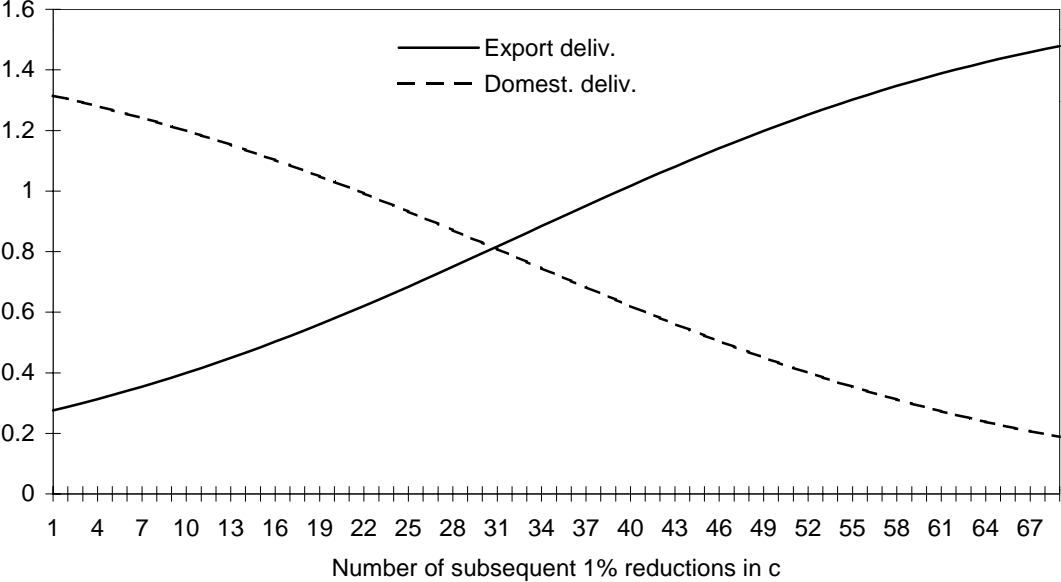
$$(37) \quad \text{El}_c n = - \frac{\pi_n^H \left[1 - \frac{\varepsilon}{s(1+\varepsilon\lambda)} \right] + \pi_n^W \left(-\frac{1}{\lambda} \right)}{Z},$$

$$\text{where } Z = \pi_n^H N \left[1 - \left(\frac{(\sigma - \varepsilon)}{(\sigma - 1)s(1 + \varepsilon\lambda)} \right) \left(\frac{1}{1 - e^{-N}} \right) \right] + \pi_n^W (-nt/\lambda).$$

The denominator Z measures the relative impact on operating profits earned by the marginal firm of a partial increase in n . For stability in the Marshallian sense, this impact should be negative. $-nt/\lambda$ captures the relative difference in operating profits earned from exports between the old and the new marginal firm. It is negative because the new firm exports at a lower scale to compensate for its lower productivity. Since $N < 0$, and $-nt/\lambda < 0$, a sufficient, but not necessary, condition for Marshallian stability is that the term in square brackets is positive. This condition is met because $1/(1 - e^{-N}) < 0$ and $\sigma > \varepsilon$. The bracketed term captures the relative difference in variable costs related to domestic deliveries between the old and the new marginal firm. Recall that there is a fixed proportion between domestic profits and variable costs of domestic deliveries. N is the relative impact on C_n^H of increasing n when X_0^H is fixed. It is negative because the marginal firm experiences decreasing domestic output, which dominates the positive profit effect from higher marginal costs. The other expression in the square brackets captures the cost effect due to equilibrium adjustments in X_0^H .

The partial elasticity wrt. c of profits earned in the domestic market shows up in (37) as the term $1 - \varepsilon/[s(1 + \varepsilon\lambda)]$. Since we always have $1 \geq 1 - \varepsilon/[s(1 + \varepsilon\lambda)] = (1 - \varepsilon)/(\lambda\varepsilon + 1) > -1/\lambda$, the relative impact of c will always be less negative on profits earned abroad than on profits earned in the domestic market. In fact, the sign of the partial elasticity of domestic profits wrt. c is ambiguous. It is decreasing from 1 as ε increases from 0. It passes 0 when $\varepsilon=1$, and approaches $1/(1-m/s) > -1/\lambda$ as ε approaches σ . The intuitive explanation is straightforward: A one percent partial increase in c will be carried forward by the price setting producers to a 1 percent rise in the prices of all pre-existing products and thereby the price index P . This will reduce domestic demand by more than one percent when $\varepsilon > 1$. Thus, sales income and profits for the old marginal firm go down by the same proportion. When $\varepsilon < 1$, the old marginal firm will experience higher income and profits from sales in the domestic market. Thus, due to the presence of profits from exports, there exists a critical value $\varepsilon^* < 1$, depending negatively on the export share for the marginal firm, for which the number of firms will be insensitive to a change in c . If ε less (greater) than ε^* , reductions in c will reduce (increase) total operating profits in the old marginal firm and firms will exit (enter) the industry.

Figure 3. Profits from exports and domestic deliveries in the marginal firm



The numerator in (37) is a weighted sum of the relative impact of c on profits in the old marginal firm earned on the domestic and the export market respectively. For the old marginal firm the partial elasticity of profits from exports wrt. c will be equal to the elasticity of exports, which equals

$-1/\lambda < 0$. Figure 3 shows how profits from exports and domestic deliveries in the marginal firm respond to changes in c in the numerical simulations.

(37) includes of course only first order effects on n . Since the determinants of $El_c n$ in (37) include endogenous variables, the relationship between c and n will in general not be log-linear. We just showed that the relative impact of c will always be less negative on export profits than on domestic profits. Moreover, the relationship between c and n may be non-monotonic when one takes into account that ε may be endogenous. This is the case in the numerical version of our model, where the demand for the differentiated product is modelled as a two step separable budgeting process. Formally, we have $P^{-\varepsilon} = (P/Q)^{-\delta}$, where Q is a linearly homogenous price function of P and the price indexes for composites supplied by other industries and imports. δ is the elasticity of substitution between the differentiated industry product and other composite goods. The own price elasticity of the composite differentiated product then becomes $\varepsilon = (1 - \theta)\delta$, where the budget share of the composite industry good in total domestic expenditure, θ , will vary when $\delta \neq 1$. This effect causes ε to decline with productivity growth when $\delta > 1$, which is the case in our numerical simulations, where we also have $\varepsilon > 1$ at the initial values for the productivity parameters. Thus, it is possible to shift c in such a way that they induce entry within a limited range and exit when c generates a model solution where $\varepsilon < \varepsilon^*$. In the numerical simulations there is a positive monotonic relationship between c and n in the productivity spectre that has explored. The basic determinant of this result is the endogeneity of the profit shares. The share of total operating profits earned from exports will increase as productivity improves. Therefore, although the entry incentive from the domestic market diminishes and eventually becomes negative, this effect is more than outweighed by the increasingly strong positive entry incentive from the export market.

The total equilibrium elasticity of n wrt. t , $El_t n$, is found from (8²):

$$(38) \quad El_t n = -\frac{\pi_n^H N \left[1 - \left(\frac{m(1-\varepsilon) + \varepsilon}{s(1-\varepsilon) + \varepsilon} \right) \left(\frac{1}{1-e^{-N}} - \frac{1}{N} \right) \right] + \pi_n^W (-nt/\lambda)}{Z} = -1 + \frac{\pi_n^H \left(\frac{m(1-\varepsilon) + \varepsilon}{s(1-\varepsilon) + \varepsilon} \right)}{(-Z)} > -1$$

The reason why $El_t n > -1$ can be explained by showing why $El_t n = -1$ is inconsistent with equilibrium.

If $El_t n = -1$, $N = nt/(e^{mN} - 1)$ and $\pi_n^W = (1-s)P^W (sP^W/c)^{1/\lambda} e^{-nt/\lambda}$ would have been unchanged.

However $\pi_n^H = (m/s - 1)c(X_0^H)^{1/s} e^N$ will increase because X_0^H is increasing in t , see (31) and (36). In other words, a 1 percent reduction of n as a response to a 1 percent increase in t would imply that the new marginal firm earned more profits in the domestic market and the same profits in the export market, compared to the old marginal firm. Consequently, profits in the new marginal firm would exceed F . Therefore the equilibrium multiplier of n wrt. t will be smaller than unity, and this modification is proportional to initial profits from the domestic market.

How changes in n modifies the solutions of the NMF case

Having examined how and why n will adjust to reductions in c and t , we proceed with a closer study of how changes in n influence the average aggregate productivity. As pointed out in the previous section, we should keep in mind that this influence should represent a relatively minor modification compared to the solution of the NMF case. Recall that the modifications will not apply to \bar{C}^W , but to \bar{C}^H , X^H and X^W in (18). Moreover, it should be noted that \bar{C}^H and X^H are modified by n through N , and that N depends symmetrically on n and t . The subsequent discussion of the impact of changes in n , will consequently also apply to the impact on N of changing t , which was also neglected in the NMF case. Above, we found that the equilibrium number of active firms could adjust in both directions as response to changes in c and t . However, we will confine the discussion to the case where n increases when c or t is reduced.

According to (25) and (27) the modifications due to changes in n can be decomposed into changes in the multiplier $0 < g(N) = (1 - e^N)/(1 - e^{mN}) < 1$, and changes in \bar{C}_0^H due to modifications in X_0^H through the term $(1 - e^N)$. Recall that $N = nt/(1 - m/s) < 0$, implying a unique correspondence between relative changes in n (and t) and N . Since we have found that $g(N)$ is monotonically increasing in n between the limit values $1/m$ and 1 , an increase in n implies that the true solution of \bar{C}^H will move upwards within the interval $[\bar{C}_0^H, m\bar{C}_0^H]$, i.e. in the direction of the approximate solution of the NMF case.

The range of the interval containing the true solution of \bar{C}^H is fixed by m , but the position of the whole interval will move because of changes in \bar{C}_0^H . An increase in n or t will increase $(1 - e^N)$, which reduces X_0^H and \bar{C}_0^H when $\lambda > 0$. The intuition behind this effect is that $(1 - e^N)$ is the correction of P relative to the NMF solution. When n increases P is reduced because of the «love of

variety» property of the utility function. This induces both the internal and external substitution away from product 0 as discussed in the previous section. The «love of variety» effect is well known in the symmetric LGMC model. It is somewhat modified in our asymmetric model because prices charged by entrants will be successively higher, due to their successively lower productivity. But although the price index includes higher prices as n grows, the valuation of increased variety still causes P to decline in n .

Consequently, the net effect on \bar{C}^H of taking into account that the marginal firm may be significant, is ambiguous. On the one hand, the true change in \bar{C}^H becomes more accurately approximated by the elasticity calculated for the NMF case. Since the solution of the NMF case for \bar{C}^H/\bar{C}_0^H includes a positive bias, this modification implies that \bar{C}^H/\bar{C}_0^H is not reduced as much as found in the NMF case when c and t declines. On the other hand, the NMF solution underestimates the reduction in \bar{C}_0^H , which implies a modification in the opposite direction on \bar{C}^H . The fact that the two modifications of the NMF solution counteract each other, strengthens the argument for the NMF case being a good approximation of the true solution.

The ambiguity of the elasticity of \bar{C}^H wrt. n can be seen explicitly by the formal elasticity expression. From (25) one obtains

$$(39) \quad \text{El}_n \bar{C}^H = \left[\left(\frac{me^{mN}}{1-e^{mN}} \right) - d \left(\frac{e^N}{1-e^N} \right) \right] N,$$

where $d \equiv [1 + m\lambda(\varepsilon - 1)] / (1 + \varepsilon\lambda) < 1$ since $\varepsilon < \sigma$ and $\lambda > 0$. The properties of the function $g(N)$ implies that the term in the square bracket is negative if d were equal to unity. However, since $d < 1$, the sign of the bracketed term in (39) is ambiguous. d captures the negative scale effect on variable unit costs within the pre-existing firms due to the redistribution of demand from pre-existing to less productive entrants⁶.

⁶ The empirical magnitude of the difference between d and unity is modest in our numerical simulations. Here ε is well above unity while $s = .8$. Since $d(1) = s$, and $d'(\varepsilon) > 0$, it follows that d is close to unity in these simulations. Accordingly, the sign of the term inside the brackets in (37) will then still be negative. Since $N < 0$, it follows that the partial elasticity of aggregate variable cost per unit of domestic deliveries with respect to n is positive

It follows that if the technology of each firm exhibits constant returns to scale, expanding domestic deliveries from the industry through entry of firms implies decreasing returns to scale at the aggregate industry level. The main reason for this is the ranking of firms according to diminishing productivity. However, this conclusion may be reversed if there is a strong degree of decreasing returns to scale within each firm. The positive contribution to aggregate productivity from redistributing the most costly units of supplies from the pre-existing firms to the entrants, may be large enough to dominate the negative contribution implied by the lower efficiency of these entrants.

From (21) it follows that the relative change in aggregate exports is modified compared to the NMF case by the term $\left[(nt/\lambda) / (e^{nt/\lambda} - 1) \right] (\hat{t} + \hat{n})$, where \hat{t} and \hat{n} denote the relative changes in t and n . Still considering the case where n increases when c declines, the true equilibrium elasticity of aggregate exports wrt. c is then seen to be greater than in the NMF case. This is consistent with the claim that the NMF case represents an overestimation of the true equilibrium level for aggregate exports which is decreasing in n and t . Turning to shifts in t , \hat{t} and El_n have opposite sign, but from (38) it follows that $\hat{t} + El_n < 0$. Therefore, the true elasticity of aggregate exports wrt. t is somewhat overestimated in the NMF case.

Whether the true solution for X^H is greater or smaller than in the NMF case is ambiguous for the same reasons as those causing the ambiguity of modification of \bar{C}^H . Taking into account that $e^N > 0$, rather than 0, (20) implies a smaller value for X^H , given X_0^H , since the NMF solution includes deliveries from the non-active firms. However, inspection of (20) reveals that the true solution for X_0^H is larger than the corresponding solution in the NMF case. As mentioned, the reason is that the price index P is underestimated in the NMF case where the presence of infinitely many products maximises the love-of-variety effect.

Formally, the net ambiguous impact of N on X^H is captured by the term $(1 - e^{mN}) / (1 - e^N)^{(\sigma - \varepsilon) / [(\sigma - 1)(1 + \varepsilon\lambda)]}$. We then have

$$(40) \quad El_n X^H = El_N X^H = \frac{mNe^{mN}}{e^{mN} - 1} - \left(\frac{\sigma - \varepsilon}{\sigma - 1} \right) \left(\frac{1}{1 + \varepsilon\lambda} \right) \left(\frac{Ne^N}{e^N - 1} \right)$$

which is clearly analogous to (39). Since $mNe^{mN} / (e^{mN} - 1) < Ne^N / (e^N - 1)$, this elasticity is surely negative if $(\sigma - \varepsilon) / [(\sigma - 1)(1 + \varepsilon\lambda)] > 1$, or equivalently if $\varepsilon < 1 / [1 + \lambda(\sigma - 1)]$, that is, if the external

substitution effect is sufficiently weak. In the special case where $\varepsilon = 1$ and $\lambda = 0$, the influence on X^H of N degenerates to $(1 - e^{mN})/(1 - e^N) > 1$, which implies that the true solution of X^H is larger than the solution obtained in the NMF case. However, when $\varepsilon = 1$ and $\lambda = 0$, the ratio of the NMF solution to the true solution of X^H will exceed $1/m$, which is the minimum limit of the ratio when N approaches 0 from above. For a fixed ε , $(1 - e^{mN})/(1 - e^N)^{(\sigma - \varepsilon)/[(\sigma - 1)(1 + \varepsilon\lambda)]}$ is increasing in $\lambda > 0$, which implies that the degree of underestimation is reduced, and for a sufficiently high λ the NMF solution will overestimate the true value of X^H . The same change in impact will take place when ε increases from 0. When $\varepsilon = 0$, $(1 - e^{mN})/(1 - e^N)^{(\sigma - \varepsilon)/[(\sigma - 1)(1 + \varepsilon\lambda)]} = (1 - e^{mN})/(1 - e^N)^m > 1$ representing the minimum values of this function wrt. to ε . (This minimum value is in turn increasing in both N and m .)

Finally, we consider the modifications through changes in N in the case where t is reduced. From (38) we know that $\hat{t} + \text{El}_t n < 0$, making N less negative. Consequently, accounting for changes in N , will modify the true equilibrium solution of X^H in opposite directions when we compare the equilibrium adjustments to shifts in c and t as long as we assume $\text{El}_t n > 0$. A less negative value of N implies that reductions in t inevitably makes the NMF case a less accurate approximation. Still, it can not be decided theoretically whether the impact of t on the true solution of X^H will be smaller or greater than the adjustments calculated for the NMF case, where X^H could change in both directions. However, if the external substitution effect is sufficiently weak to make the NMF elasticity of X^H wrt. t positive, cf. (36), we have shown that then $\text{El}_t X^H = \text{El}_N X^H < 0$. Since the equilibrium value of $|N|$ is increasing in t , the *relative change* in N will be positive when t is increased. Therefore, the modifications due to changes in N will reduce the true elasticity of X^H wrt. t compared to the corresponding NMF elasticity.

To summarise, the inaccuracy of the NMF case compared to the true solution is due to the fact that the NMF case implicitly includes deliveries and prices of infinitely many non-active firms. The inclusion of deliveries of non-active firms contributes to bias the NMF solution of \bar{C}^H , X^H and X^W upwards. The inclusion of prices of non-active firms biases P downwards to the level where the «love-of-variety» effect is maximised in all equilibria. This in turn contributes to negative biases on \bar{C}^H and X^H , through underestimation of \bar{C}_0^H and X_0^H . The fact that the NMF modifications case work in opposite directions, strengthens the argument that the NMF solution is a reasonable approximation to the true solution.

4. Concluding remarks

The analysis presented in this paper has shown that endogenous equilibrium adjustments are of crucial importance to the aggregate productivity effects of firm specific productivity shifts. These adjustments are particularly important for the quantity weights used in the determination of aggregate productivity. The analysis has confirmed the view that it may be problematic to interpret changes in aggregate productivity as an indicator of technological change, since the link between technological shocks and aggregate productivity is heavily influenced by market structure and equilibrium adjustments.

By incorporating equilibrium adjustments through simultaneous interactions between supply and demand incentives, our analysis represents an extension compared to the works of Johansen(1959,1972) and Houthakker (1955-56). When determinants of aggregate productivity depend on the outcome of market equilibria, it will in general depend on the organisation and competitive structure in the relevant markets. Our analysis has demonstrated this dependence by considering two kinds of market structure, one perfectly competitive «export» market where the product price is fixed exogenously, and one monopolistically competitive «domestic» market.

The most interesting and counterintuitive results were obtained in the case of monopolistic competition. Here, the equilibrating forces determining the firm specific contributions to aggregate productivity, cause the relative weights of successively less productive firms to decline faster than their productivity level. Our analysis shows that aggregate productivity of deliveries to this market will be less than, but very close to the product of the mark-up factor and the productivity of the most efficient firm. Thus productivity heterogeneity will not contribute significantly to decreasing returns to scale at the industry level. When products are close substitutes, as assumed in the LGMC model, this implies that the equilibrium gap between aggregate productivity and the productivity of the most efficient firm will be modest.

The equilibrium outcome for the aggregate productivity related to domestic deliveries will have qualitative similarities to the outcome of the much more transparent process determining the equilibrium level of productivity for export deliveries. In the export market, all firms will equal their marginal costs to the fixed world price. This generates a common equilibrium productivity level in each firm, which is independent of firm specific productivity parameters. However, we have shown that the underlying mechanisms generating these results are very different.

Our framework makes it possible to study firm specific as well as uniform technology shifts. We have demonstrated that the aggregate productivity effect of uniform productivity shifts is fundamentally different from the one caused by reduction of the productivity differentials through catching up. An important by-product of our study is that we demonstrate how it is possible to incorporate heterogeneity into the standard LGMC model at a low cost in terms of analytical tractability.

Some qualifications should be made explicit. First, although the equilibrium quantities used as weights in aggregate productivity measures are determined in a widely accepted modelling framework, one may question if the results rely on too price sensitive demand responses. Numerical experiments suggest that it is difficult to reproduce consistency between observed size distributions of firms within industries, productivity differentials between them and an elasticity of substitution calibrated to estimated mark-up factors. This suggests that it may be fruitful to study the effects of heterogeneity and productivity shifts under other assumptions of market structure.

Second, our model relies on some restrictive assumptions with respect to the structure of productivity differentials. Especially, the exponential structure made it possible to summarise the aggregate industry structure by a small number of parameters. This would not be possible if the relative productivity differentials varied significantly when moving from the most efficient to successively less efficient firms. However, most of our local comparative statics results would probably hold.

References

- Aukrust (1970): PRIM 1. A Model of the Price and Income Distribution Mechanisms of an Open Economy, *Review of Income and Wealth* **16**, 51-78.
- Baily, N.M., C. Hulten and D. Campbell (1992): Productivity Dynamics in Manufacturing Plants, *Brookings Papers on Economic Activity: Microeconomics*, 187-268.
- Bartelsman and Doms (1997): Understanding Productivity: Lessons from Longitudinal Micro Datasets, mimeo.
- Blanchard, O. J. (1985): Debt, Deficits and Finite Horizons, *Journal of Political Economy* **93**, 1045-1076.
- Bowitz, E. and Å. Cappelen (1994): *Price Formation and Factor Demand in Norwegian Industries*. Social and Economic Studies 85, Statistics Norway. (In Norwegian).

- Chamberlin, E. H. (1933): *The Theory of Monopolistic Competition*, Cambridge: Harvard University Press.
- Devarajan S., J. D. Lewis and S. Robinson (1990): Policy Lessons from Trade-Focused, Two-Sector Models, *Journal of Policy Modeling* **12**, 625-657.
- Frisch, R. (1950): Alfred Marshall's Theory of Value, *Quarterly Journal of Economics* **64**, 495-524.
- Helpman, E. and G. M. Grossman (1992): *Innovation and Growth in the Global Economy*, Cambridge: MIT Press.
- Helpman, E. and P. Krugman (1985): *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy*, Cambridge: MIT Press.
- Holmøy, E. (1997): Aggregate Industry Behaviour in a Model of Asymmetric Monopolistic Competition, forthcoming in the series Discussion Papers, Statistics Norway.
- Houthakker, H. (1955/56): The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis, *Review of Economic Studies* **23**, 27-31.
- Johansen, L. (1959): Substitution versus fixed coefficients in the theory of economic growth, *Econometrica* **27**, 157-176.
- Johansen, L. (1972): *Production Functions*, Amsterdam: North-Holland.
- Klette, T.J. (1994): Estimating Price-Cost Margins and Scale Economies from a Panel of Microdata, Discussion Papers 130, Statistics Norway.
- Klette, T.J. and A. Mathiassen (1995): Job Creation, Job Destruction and Plant Turn-over in Norwegian Manufacturing, *Annales d'économie et de Statistique* 41/41, 97-125.
- Klette, T.J. and A. Mathiassen (1996): *Growth and Turnover among Norwegian Manufacturing Plants*. Social and Economic Studies 95, Statistics Norway. (In Norwegian).
- Montagna, C. (1995): Monopolistic competition with Firm-Specific Costs, *Oxford Economic Papers* **47**, 318-328.
- Salter, W. (1960): *Productivity and Technical Change*, London: Cambridge University Press.
- Stigler, G. J. (1949): "Monopolistic Competition in Retrospect" in *Five Lectures on Economic Problems*, London: London School of Economics.
- Sutton, J. (1996): Gibrat's Legacy, *Journal of Economic Literature* **XXXV**, 40-59.
- Yaari, M. E. (1965): Uncertain Lifetime, Life Insurance, and the Theory of the Consumer, *Review of Economic Studies* **32**, 137-150.

Deriving the additively separable cost function from a separable production function

The firm technology is a special case of the general production function

$$(A1) \quad F(X^W, X^H, V) = 0,$$

where V is a variable input. In this appendix we specify the restrictions, which imply that the variable cost function can be separated additively as in (1). The first restriction is that input is additively separable from an aggregate of outputs:

$$(A2) \quad G(X^W, X^H) = V.$$

The second restriction is that the $G(\cdot)$ function takes the symmetric form

$$(A3) \quad G(X^W, X^H) = \left[(X^W)^\rho + (X^H)^\rho \right]^\mu,$$

which is a homothetic transformation function. The elasticity of transformation, σ_W is given by $\sigma_W = 1/(\rho - 1) > 0$. The third and final restriction is that $\mu = 1$. The technology can then be written

$$(A4) \quad (X^W)^\rho + (X^H)^\rho = V.$$

In order to establish the connection to the cost function in (1), define a variable s instead of μ so that $s\rho\mu = s\rho = 1$. (A4) can then be written

$$(A5) \quad \left[(X^W)^\rho + (X^H)^\rho \right]^{1/\rho} = V^s,$$

where the left hand side is a linearly homogenous transformation function with constant elasticity of transformation, σ_w . This is the CET function motivated and discussed in Devarajan, Lewis and Robinson (1990). It is natural to interpret this CET function of export and domestic deliveries as an aggregation function for output, X . (A5) then becomes

$$(A6a) \quad X = \left[(X^W)^\rho + (X^H)^\rho \right]^{1/\rho}$$

$$(A6b) \quad X = V^s,$$

where s now has the interpretation as the elasticity of scale in the production function transforming (aggregate) input into aggregate output. The cost function dual to this technology is of course

$$(A7) \quad C^* = qV = qX^{1/s},$$

where q is the factor price of V .

$$(A8) \quad C^* = q \left[(X^W)^\rho + (X^H)^\rho \right]^{1/s\rho}.$$

Since $s\rho = 1$, the exponent in (A8) vanishes. Redefining units in which V is measured enables us to write (A8) as

$$(A9) \quad C^* = c \left[(X^W)^{1/s} + (X^H)^{1/s} \right],$$

which is the variable part of the cost function (1).

Recent publications in the series Discussion Papers

- 107 S. Kverndokk (1994): Depletion of Fossil Fuels and the Impacts of Global Warming
- 108 K.A. Magnussen (1994): Precautionary Saving and Old-Age Pensions
- 109 F. Johansen (1994): Investment and Financial Constraints: An Empirical Analysis of Norwegian Firms
- 110 K.A. Brekke and P. Børing (1994): The Volatility of Oil Wealth under Uncertainty about Parameter Values
- 111 M.J. Simpson (1994): Foreign Control and Norwegian Manufacturing Performance
- 112 Y. Willassen and T.J. Klette (1994): Correlated Measurement Errors, Bound on Parameters, and a Model of Producer Behavior
- 113 D. Wetterwald (1994): Car ownership and private car use. A microeconomic analysis based on Norwegian data
- 114 K.E. Rosendahl (1994): Does Improved Environmental Policy Enhance Economic Growth? Endogenous Growth Theory Applied to Developing Countries
- 115 L. Andreassen, D. Fredriksen and O. Ljones (1994): The Future Burden of Public Pension Benefits. A Microsimulation Study
- 116 A. Brendemoen (1994): Car Ownership Decisions in Norwegian Households.
- 117 A. Langørgen (1994): A Macromodel of Local Government Spending Behaviour in Norway
- 118 K.A. Brekke (1994): Utilitarianism, Equivalence Scales and Logarithmic Utility
- 119 K.A. Brekke, H. Lurås and K. Nyborg (1994): Sufficient Welfare Indicators: Allowing Disagreement in Evaluations of Social Welfare
- 120 T.J. Klette (1994): R&D, Scope Economies and Company Structure: A "Not-so-Fixed Effect" Model of Plant Performance
- 121 Y. Willassen (1994): A Generalization of Hall's Specification of the Consumption function
- 122 E. Holmøy, T. Hægeland and Ø. Olsen (1994): Effective Rates of Assistance for Norwegian Industries
- 123 K. Mohn (1994): On Equity and Public Pricing in Developing Countries
- 124 J. Aasness, E. Eide and T. Skjerpen (1994): Criminometrics, Latent Variables, Panel Data, and Different Types of Crime
- 125 E. Biørn and T.J. Klette (1994): Errors in Variables and Panel Data: The Labour Demand Response to Permanent Changes in Output
- 126 I. Svendsen (1994): Do Norwegian Firms Form Extrapolative Expectations?
- 127 T.J. Klette and Z. Griliches (1994): The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous
- 128 K.E. Rosendahl (1994): Carbon Taxes and the Petroleum Wealth
- 129 S. Johansen and A. Rygh Swensen (1994): Testing Rational Expectations in Vector Autoregressive Models
- 130 T.J. Klette (1994): Estimating Price-Cost Margins and Scale Economies from a Panel of Microdata
- 131 L. A. Grünfeld (1994): Monetary Aspects of Business Cycles in Norway: An Exploratory Study Based on Historical Data
- 132 K.-G. Lindquist (1994): Testing for Market Power in the Norwegian Primary Aluminium Industry
- 133 T. J. Klette (1994): R&D, Spillovers and Performance among Heterogenous Firms. An Empirical Study Using Microdata
- 134 K.A. Brekke and H.A. Gravningsmyhr (1994): Adjusting NNP for instrumental or defensive expenditures. An analytical approach
- 135 T.O. Thoresen (1995): Distributional and Behavioural Effects of Child Care Subsidies
- 136 T. J. Klette and A. Mathiassen (1995): Job Creation, Job Destruction and Plant Turnover in Norwegian Manufacturing
- 137 K. Nyborg (1995): Project Evaluations and Decision Processes
- 138 L. Andreassen (1995): A Framework for Estimating Disequilibrium Models with Many Markets
- 139 L. Andreassen (1995): Aggregation when Markets do not Clear
- 140 T. Skjerpen (1995): Is there a Business Cycle Component in Norwegian Macroeconomic Quarterly Time Series?
- 141 J.K. Dagsvik (1995): Probabilistic Choice Models for Uncertain Outcomes
- 142 M. Rønsen (1995): Maternal employment in Norway, A Parity-Specific Analysis of the Return to Full-Time and Part-Time Work after Birth
- 143 A. Bruvoll, S. Glomsrød and H. Vennemo (1995): The Environmental Drag on Long- Term Economic Performance: Evidence from Norway
- 144 T. Bye and T. A. Johnsen (1995): Prospects for a Common, Deregulated Nordic Electricity Market
- 145 B. Bye (1995): A Dynamic Equilibrium Analysis of a Carbon Tax
- 146 T. O. Thoresen (1995): The Distributional Impact of the Norwegian Tax Reform Measured by Disproportionality
- 147 E. Holmøy and T. Hægeland (1995): Effective Rates of Assistance for Norwegian Industries
- 148 J. Aasness, T. Bye and H.T. Mysen (1995): Welfare Effects of Emission Taxes in Norway
- 149 J. Aasness, E. Biørn and Terje Skjerpen (1995): Distribution of Preferences and Measurement Errors in a Disaggregated Expenditure System
- 150 E. Bowitz, T. Fæhn, L. A. Grünfeld and K. Moum (1995): Transitory Adjustment Costs and Long Term Welfare Effects of an EU-membership – The Norwegian Case
- 151 I. Svendsen (1995): Dynamic Modelling of Domestic Prices with Time-varying Elasticities and Rational Expectations
- 152 I. Svendsen (1995): Forward- and Backward Looking Models for Norwegian Export Prices
- 153 A. Langørgen (1995): On the Simultaneous Determination of Current Expenditure, Real Capital, Fee Income, and Public Debt in Norwegian Local Government

- 154 A. Katz and T. Bye(1995): Returns to Publicly Owned Transport Infrastructure Investment. A Cost Function/ Cost Share Approach for Norway, 1971-1991
- 155 K. O. Aarbu (1995): Some Issues About the Norwegian Capital Income Imputation Model
- 156 P. Boug, K. A. Mork and T. Tjemsland (1995): Financial Deregulation and Consumer Behavior: the Norwegian Experience
- 157 B.E. Naug and R. Nymoene (1995): Import Price Formation and Pricing to Market: A Test on Norwegian Data
- 158 R. Aaberge (1995): Choosing Measures of Inequality for Empirical Applications
- 159 T.J. Klette and S.E. Førre (1995): Innovation and Job Creation in a Small Open Economy: Evidence from Norwegian Manufacturing Plants 1982-92
- 160 S. Holden, D. Kolsrud and B. Vikøren (1995): Noisy Signals in Target Zone Regimes: Theory and Monte Carlo Experiments
- 161 T. Hægeland (1996): Monopolistic Competition, Resource Allocation and the Effects of Industrial Policy
- 162 S. Grepperud (1996): Poverty, Land Degradation and Climatic Uncertainty
- 163 S. Grepperud (1996): Soil Conservation as an Investment in Land
- 164 K.A. Brekke, V. Iversen and J. Aune (1996): Soil Wealth in Tanzania
- 165 J.K. Dagsvik, D.G. Wetterwald and R. Aaberge (1996): Potential Demand for Alternative Fuel Vehicles
- 166 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part I: Theory
- 167 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part II: Inference
- 168 R. Aaberge, A. Björklund, M. Jäntti, M. Palme, P. J. Pedersen, N. Smith and T. Wennemo (1996): Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States
- 169 K. Nyborg (1996): Some Norwegian Politicians' Use of Cost-Benefit Analysis
- 170 E. Berg, S. Kverndokk and K. E. Rosendahl (1996): Market Power, International CO₂ Taxation and Petroleum Wealth
- 171 R. Aaberge, U. Colombino and S. Strøm (1996): Welfare Effects of Proportional Taxation: Empirical Evidence from Italy, Norway and Sweden
- 172 J.K. Dagsvik (1996): Dynamic Choice, Multistate Duration Models and Stochastic Structure
- 173 J.K. Dagsvik (1996): Aggregation in Matching Markets
- 174 H.C. Bjørnland (1996): The Dynamic Effects of Aggregate Demand, Supply and Oil Price Shocks
- 175 A. Bruvoll and K. Ibenholt (1996): Future Waste Generation. Forecasts Based on a Macroeconomic Model
- 176 T. Fæhn and L. A. Grünfeld (1996) Recent Leaps Towards Free Trade. The Impact on Norwegian Industry and Trade Patterns
- 177 R. Barrell and K. A. Magnussen (1996): Counterfactual Analyses of Oil price Shocks using a World Model
- 178 E. Bowitz and S. I. Hove (1996): Business cycles and fiscal policy: Norway 1973-93
- 179 H.C. Bjørnland (1996): Sources of Business Cycles in Energy Producing Economies: The case of Norway and United Kingdom
- 180 K. Nyborg (1996): The Political Man and Contingent Valuation: Motives Do Count
- 181 E. Berg, S. Kverndokk and K.E. Rosendahl (1996): Gains from Cartelisation in the Oil Market
- 182 R. Aaberge and I. Aslaksen (1996): Decomposition of the Gini Coefficient by Income Components: Various Types of Applications and Interpretations
- 183 B. Bye (1996): Taxation, Unemployment and Growth: Dynamic Welfare Effects of "Green" Policies
- 184 T.J. Klette and F. Johansen (1996): Accumulation of R&D Capital and Dynamic Firm Performance: A Not-so-fixed Effect Model
- 185 B. Bye (1996): Environmental Tax Reform and Producer Foresight: An Intertemporal Computable General Equilibrium Analysis
- 186 S. Grepperud (1997): Soil Depletion Choices under Production and Price Uncertainty
- 187 N.M. Stølen and T. Åvitsland (1997): Has Growth in Supply of Educated Persons Been Important for the Composition of Employment?
- 188 T.J. Klette and Z. Griliches (1997): Empirical Patterns of Firm Growth and R&D Investment: a Quality Ladder Model Interpretation
- 189 J. Aune, S. Glomsrød, V. Iversen and H. Wiig (1997): Structural Adjustment and Soil Degradation in Tanzania. A CGE-model Approach with Endogenous Soil Productivity
- 190 E. Biørn and T.J. Klette (1997): Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation
- 191 L. Belsby and B.K. Wold (1997): Primary Schooling in Zambia Squeezed at Community and Household Level
- 192 E. Bowitz and Å. Cappelen (1997): Incomes Policies and the Norwegian Economy 1973-93
- 193 S. Glomsrød, M.D Monge A. and H. Vennemo (1997): Structural Adjustment and Deforestation in Nicaragua
- 194 F. Johansen and T.J. Klette (1997): Wage and Employment Effects of Payroll Taxes and Investment Subsidies
- 195 T. Fæhn (1997): Non-Tariff Barriers - the Achilles' Heel of Trade Policy Analysis
- 196 R. Aaberge and A. Langørgen (1997): Fiscal and Spending Behavior of Local Governments: An Empirical Analysis Based on Norwegian Data
- 197 A.C. Hansen and H.K. Selte (1997): Air Pollution and Sick-leaves - is there a Connection? A Case Study using Air Pollution Data from Oslo
- 198 E. Holmøy and T. Hægeland (1997): Aggregate Productivity Effects of Technology Shocks in a Model of Heterogeneous Firms: The Importance of Equilibrium Adjustments