Mahlberg, Bernhard; Url, Thomas

Working Paper

Effects of the Single Market on the Austrian Insurance Industry

WIFO Working Papers, No. 97

Provided in Cooperation with:
Austrian Institute of Economic Research (WIFO), Vienna

Effects of the Single Market on the Austrian Insurance Industry

February, 1998
revised October, 1998

Bernhard Mahlberg
Vienna University of Economics and Business Administration
Research Institute for European Affairs
e-mail: Bernhard.Mahlberg@wu-wien.ac.at

Thomas Url
Austrian Institute of Economic Research
e-mail: url@wsr.ac.at

Abstract

The integration of European services markets and the associated deregulation measures open up formerly segmented markets to competition from the Single Market. The purpose of this study is to propose a method of measuring the effects of liberalization on technical efficiency, the existence of economies of scale and diversification, and the dynamic development of productivity for the insurance industry. As an example we construct efficiency frontiers for the years 1992 through 1996 for the Austrian insurance industry using a Data Envelopment Analysis. The results provide yardsticks for efficiency comparisons from a consumer perspective and can be used to construct a Malmquist productivity index for the transition period.

Keywords: European Integration, Insurance Industry, Data Envelopment Analysis
JEL-classification: L50, G22, C14

The authors gratefully acknowledge valuable comments from David Cummins, Shawna Grosskopf, Michael Theil and seminar participants at the University of Vienna, the University of Business Administration Vienna, the Austrian Institute of Economic Research, the Austrian Insurance Association, and the SOR-Conference Jena, the EOR Conference Brussels, and the 25th Geneva Association Meeting Vienna.

Introduction

The integration of Europe into a Single Market is largely realized within goods markets but still lagging within the service markets. A low degree of foreign penetration prevailed particularly within the European Economic Area's (EEA) service sector. As in other member countries, deregulation measures were applied to the Austrian insurance industry in preparation for joining the Single Market. Although the deregulation phase started as early as 1987 with the establishment of the European Union (EU) Solvency Rules for Austrian firms, it took until the beginning of 1994 for all of the three EU-directives on insurance to be fully implemented and thus establish freedom to open up branches and provide services under a single license and prudential control. The most prominent effect expected from deepening integration in the services markets was a higher level of competition from outsiders and more contestable markets. The expected response of domestic firms ranged from increasing their productivity level by creating new and more diversified products (Brandner, 1994), through adapting distribution channels, to merging into larger units in order to enjoy scale economies.

Several characteristics of insurance markets have been mentioned which will dampen the effects of foreign competition on domestic firms. Most important is the lack of cross national distribution channels (Szopo, 1990). Although the big European firms operate across national borders they prefer to set up subsidiaries, rather than branches or independent insurance agents. In the Austrian insurance market, foreign companies own about half of the firms. Given that the industry is characterized by asymmetric information between an insurer and the insured with respect to the individual risk exposure or country specific risk distributions, it probably pays off to have a local representative, who closely monitors clients. Austrian insurance companies solve this problem accordingly by hiring direct employees and exclusive agents, rather than using independent agents. Moreover, European insurance markets are still segregated by widely differing national laws and languages, thus a trustworthy local intermediate is, in addition to the benefit from services in claims settlement, also in the interest of the customer. Whereas these barriers will weigh heavily for households and small firms they may turn out to be less important for firms of medium and large size.

Another important reason for the lack of large scale cross national activities is the structure of European insurance markets. Since national markets are usually dominated by a relatively small number of firms entry deterrence behavior or even predatory pricing is likely to prevent foreign outsiders from entering the local market. Especially because the run-up costs in the insurance business are considerable. For example, in Austria a fierce competition with rebates has emerged within motor third party insurance since 1994. This reduced expected incomes for new entrants and created a multiplicity of product characteristics and additional costs from information collection for consumers. On the other hand, profit margins are falling due to this strategy and thus firms are forced to improve their cost structure.

A wave of mergers and co-operations established new distribution channels with the biggest effects expected to occur in the life insurance business. New channels of distribution, especially bankassurance and direct marketing, transformed the industry and created financial conglomerates that integrate insurance services into traditional banking products. The cross selling of insurance and banking products is supposed to increase the efficiency of banking outlets and insurance companies.

This environment puts pressure on less efficient firms to improve productivity and assess their position relative to the market. We will propose a new method to measure the efficiency of insurance companies from a consumer's perspective, based on a comparison of individual firms to the best practice technology. This measure should be distinguished from a pure shareholder value perspective, which would focus on measures like the return on investment, the dividend yield or similar indices. Our measure of efficiency also allows us to track the development over time and thus to get an impression of productivity change in the industry and hence of the consequences of the Single Market. As the formation of rating agencies in Europe demonstrates, detailed information on efficiency is not only of interest for insurance companies but also for independent insurance agents and consumer protection agencies.

In this paper we will use a panel of Austrian insurance companies over the period 1992 through 1996 published by the Austrian Insurance Supervision Agency to analyze the Austrian insurance industry by means of a Data Envelopment Analysis (DEA). In contrast to conventional econometric techniques, DEA approximates production frontiers within a linear programming problem and thus represents a non-parametric approach. The results of the DEA will allow us to assess the technical efficiency of individual firms with regard to a set of best practice or benchmark firms. Furthermore, we will be able to decompose technical inefficiency into pure technical inefficiency and scale inefficiency. This should give us an idea of the benefits of future merger activities in the Austrian insurance industry. The efficiency scores from DEA can be used in subsequent regressions to reveal whether differences in technical efficiency are more likely due to scale or due to diversification effects. Finally, the analysis of a panel before and after the start of the EEA allows us to assess dynamic aspects of the integration into the Single Market by means of a Malmquist productivity index.

Our paper is organized as follows. We discuss the theory of efficiency measurement and describe the linear programming method (DEA) for measuring efficiency, economies of scale and scope, and technical progress in Section 2. In the following section we discuss our method of modeling insurance production and present the data. The empirical results are presented in Section 4 and, finally, we summarize our findings and conclude.

Theoretical Background

Efficiency

In economic theory several concepts of efficiency are known. The most common measure is cost efficiency, which suggests that profit maximizing behavior drives firms to choose a combination of inputs such that the costs to produce a given level of output are minimized. In Figure 1 technical and economic constraints are combined. The curve QQ' represents all technically feasible combinations of inputs x_1 and x_2 which allow a firm to produce a given level of output y. This curve is called an isoquant. The line WW', on the other hand, represents the budget constraint faced by the firm. The constant slope of the cost curve is determined by the relative price between both inputs. Given the shape of the cost curve and the isoquant the only cost efficient input combination is located at E, where the firm produces an output level y at minimal cost. For production plans off the cost frontier we can distinguish two components: technical inefficiency and allocative inefficiency.

Technical inefficiency is defined as a combination of inputs to the right of or above QQ'. This case is illustrated at point A, where the firm could either increase production using the same quantity of inputs or decrease inputs holding output constant. A firm operating at B uses the available technology optimally, i.e. it is technically efficient. The ratio OB/OA serves as our measure of technical efficiency and its value varies over the range [0,1]. Obviously a ratio of 1 indicates technical efficiency, because A and B coincide in this case. For technically inefficient firms this ratio will be smaller than one. The fraction (1-OB/OA), on the other hand, shows the potential cost savings that a shift to technically efficient production would bring about.

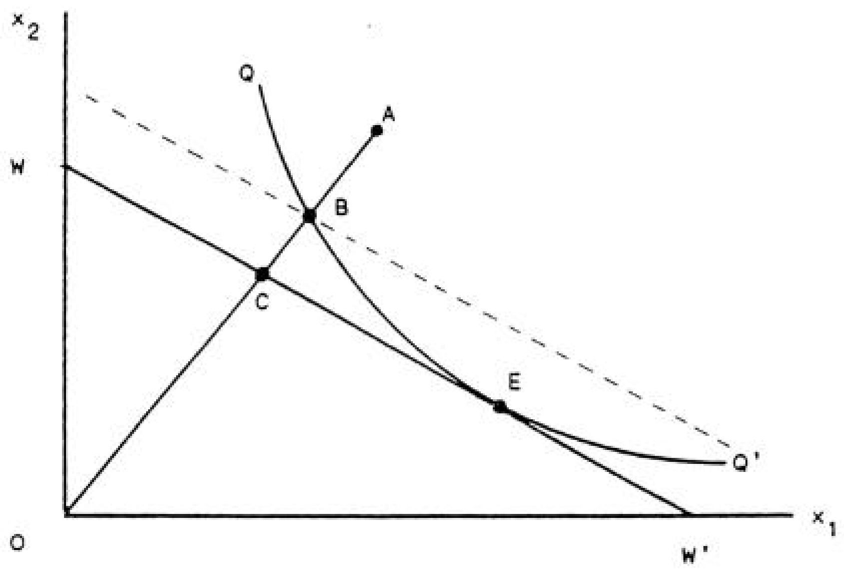

Figure 1: Different measures of efficiency

At point B an insurance company is said to be technically efficient yet not cost efficient. By substituting input x_1 for x_2 up to point E the firm changes the allocation of inputs according to the relative price and further reduces operating costs. This can be seen by the position of the dashed cost curve associated with production at point B. The solid cost curve lies to the right of WW', which represents the expenditures the firm would incur if it produced fully cost efficiently at E. Hence, the ratio 0C/0A corresponds to the degree of cost efficiency of the firm. From Figure 1 it should be clear that: 0C/0A = (0B/0A) (0C/0B). Accordingly, (1-0C/0B) provides the relative cost saving potential. The ratio 0C/0B can serve as a measure of allocative inefficiency. Further efficiency concepts include revenue and profit efficiency (Lovell, 1993; Berger, Hancock, and Humphrey, 1993), but because reliable data on input and output prices are not available we will confine ourselves to the measurement of technical efficiency.

Economies of Scale and Economies of Scope

A formal definition of economies of scale is based on the concept of a production function $y=f(K,L)$, which relates the inputs capital, K, and labor, L, to output y. If all inputs are multiplied by the same positive constant $m>1$ and the production level increases by more than m we have increasing returns to scale, i.e. $f(mK, mL) > mf(K, L)$. Constant returns to scale prevail if $f(mK, mL) = mf(K, L)$ and decreasing returns to scale occur when $f(mK, mL) < mf(K, L)$. The economic intuition of scale economies is related to the behavior of average costs: if scale economies exist, average costs decrease with an increasing level of production either up to an optimal firm size or over the whole range of production.

Economies of scope arise when a firm has a potential to save costs by producing more than one good. Whereas scale economies are economies of specialization and result from increasing the level of production of a single good, economies of scope result from diversification into several product lines. A measure for economies of scope is given by Baumol et al. (1988):

$$SCOPE(y) = \frac{C(y_1,0;w) + C(0,y_2;w) - C(y_1,y_2;w)}{C(y_1,y_2;w)},$$

where a firm can produce two outputs y_1 and y_2 with given inputs w. The measure compares the cost of two specialized firms, each producing either y_1 or y_2 with the costs of production of a single firm producing both outputs simultaneously. A value of SCOPE>0 indicates economies of scope, whereas in the case of diseconomies of scope the measure is smaller than zero.

Methodology

The Data Envelopment Analysis

The basic problem with measuring the productive efficiency of micro units such as insurance companies is to establish a benchmark that can be used to compare the performance of companies. Farrell (1957) introduced for this purpose the concept of the efficiency frontier, which delineates the technological limits of what a firm can achieve with a given level of inputs. This frontier corresponds to the isoquant QQ' in Figure 1 and allows us to classify firms into technically efficient units if they produce at the isoquant and into inefficient units if they produce to the right or above the isoquant. An inefficient firm could either produce more with the available inputs or decrease the levels of inputs while keeping production unchanged. The indicator of inefficiency is then given by the relative distance between the actual observed production and the nearest benchmark production.

To estimate this indicator of efficiency for individual companies we apply a Data Envelopment Analysis (DEA) developed by Charnes, Cooper and Rhodes (1978). This is a non-parametric approach that uses a linear programming technique to construct an envelope for the observed input-output combinations of all market participants under the constraint that all best practice firms support the envelope, whereas - in terms of Figure 1 - all inefficient firms are kept off the frontier. The envelope can be interpreted as an efficiency frontier and does not rely on any restriction on the functional form of the production function. The use of the DEA appears to be especially appropriate in service industries such as the insurance business, since the production function is unknown and we are typically confronted with multiproduct firms.

Measuring technical efficiency

We chose two different procedures to measure efficiency. The first one allows for variable returns to scale (VRS) and thus permits the coexistence of economies and diseconomies of scale. An envelope which fulfills this condition for the one input/one output case is shown as the dashed curve in Figure 2. The second procedure relaxes the convexity constraint and efficiency is measured under the assumption of constant returns to scale (CRS). The solid line in Figure 2 shows an envelope which fulfills this condition. The computation of the envelope can be reduced to a linear program for each individual firm in which the following optimization problem is solved:

$$\min \theta_0 - \varepsilon \left(\sum_r s_r + \sum_i e_i \right)$$

s.t. $\sum_j y_{rj} \lambda_j - s_r = y_{r0}$ *(for each of the r outputs)*

$\theta_0 x_{i0} - \sum_j x_{ij} \lambda_j - e_i = 0$ *(for each of the i inputs)*

θ_o *free,* $e_i \geq 0$, $s_r = \geq 0, \lambda_j \geq 0$,

$$\sum_j \lambda_j = 1 \qquad \textit{(in case of VRS)}$$

$$\sum_j \lambda_j = free \qquad \textit{(in case of CRS)}$$

$$\sum_j \lambda_j \leq 1 \qquad \textit{(in case of NIRS)}$$

where

θ_o ... *efficiency score,*
ε ... *non-Archimedean variable* ($\varepsilon=10^8$)

s_r ... *slack variable of the r-th output,*	$r = 1,...,s,$	s ... *number of outputs,*	$s = 9$
e_i ... *slack variable of the i-th input,*	$i = 1,...,m,$	m ... *number of inputs,*	$m = 2$
y_{rj} ... *r-th output of the j-th firm,*	$j = 1,...,n,$	n ... *number of firms,*	$n = 60$

x_{ij} ... *i-th input of the j-th firm,*
λ_j ... *weight of the j-th firm,*

This procedure minimizes the efficiency score θ_o of a single firm and must be repeated for every firm in the sample.

Determining Economies of Scale

The convexity constraint $\Sigma_j \lambda_j = 1$ provides the basis for measuring economies of scale within the DEA concept (cf. Grosskopf 1986). It determines how closely the production frontier envelops the observed input-output combinations. If the sum of weights is restricted to one, the closest fit will be achieved and a variable returns to scale (VRS) technology is assumed. Non increasing returns to scale (NIRS) result, if the sum of weights is restricted to be equal or less than one. Relaxing the convexity restriction corresponds to tightening the scale restriction, thereby forcing the production frontier to run through the origin of the data space and thus loosening its fit around the observations. A free sum of weights implies a CRS technology. Under CRS the distance of some firms to the efficiency frontier will increase and in the case of an input based measurement the average efficiency score drops. Thus by comparing VRS and CRS frontiers one can identify scale inefficiencies of firms. The scale efficiency θ_o^S measures the distance between the frontiers under VRS and CRS. It can be interpreted as the potential cost saving of a firm adjusting to the optimal size. The scale efficiency is defined by the ratio $\theta_o^{CRS}/\theta_o^{VRS}$, where θ_o^{CRS} is the efficiency score under CRS and θ_o^{VRS} is the efficiency score under the assumption of VRS. Because the linear program under VRS is more constrained than under CRS $\theta_o^S = \theta_o^{CRS}/\theta_o^{VRS} \leq 1$ must hold by definition.

Scale inefficiencies can occur in the range of small and large firms and will reduce the efficiency score compared to the VRS case. Following Färe, Grosskopf and Lovell (1985) allows also to distinguish between firms operating under increasing or decreasing returns to scale. If the efficiency scores under CRS and NIRS are identical a firm can be classified as operating under increasing returns to scale. On the other hand firms having bigger efficiency scores under NIRS compared to the CRS case are operating within the decreasing returns to scale region. The consequence of increasing and decreasing returns to scale can be seen in Figure 2. All efficient firms at low output levels like A and B are able to lower average costs by increasing their size towards an output level like C. At this higher output level the firm would operate at minimal average costs but increasing output further would bring the firm into the range of decreasing returns to scale or increasing average costs, respectively. Thus firms like D and H would be able to improve their cost position by reducing output. This procedure has to be repeated for every firm in the sample in order to characterize the scale economies of each firm.

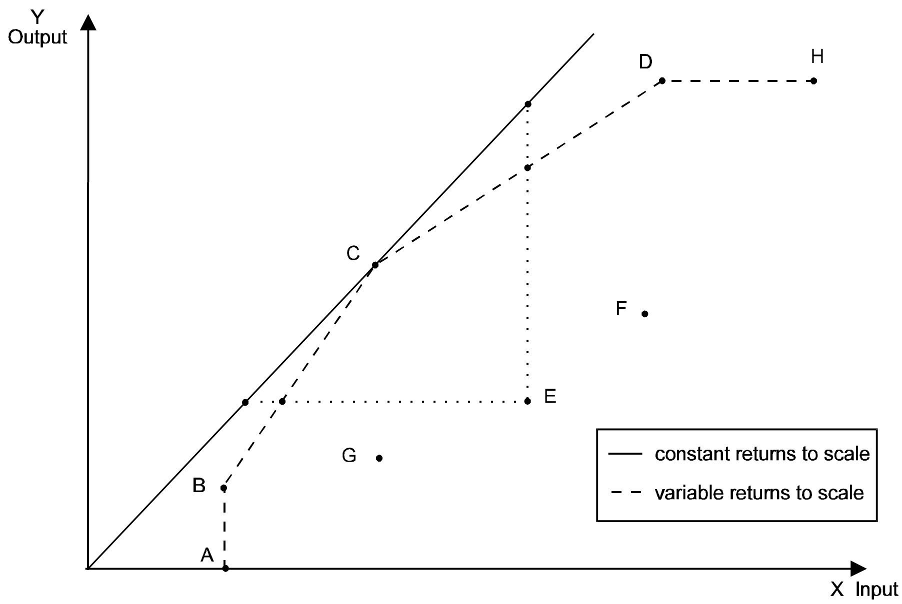

Figure 2: Determining economies of scale

Measuring Economies of Scope

The measurement of economies of scope within the Data Envelopment Analysis is rather difficult. Useful approaches to this aspect are very rare in the literature. Ferrier et al. (1993) developed a DEA approach based on the concept of economies of diversification. Their concept is based on a measure of cost efficiency which we do not apply here. Färe, Grosskopf, and Lovell (1994), discuss the mathematical fundamentals of economies of scope and develop a measure based on cost functions. Since we do not have data on input prices we cannot compute measures of cost efficiency and thus we will rely in this issue again on a second step regression technique and relate the efficiency scores to an indicator of diversification.

Technical progress

Technical progress in the financial services sector is notoriously badly measured since changes in quality are often hard to identify and inputs are mostly used in a joint production process for several products. Thus, real changes in outputs are not easily obtained and individual inputs cannot be attributed to a specific output. Under these circumstances the concept of efficiency scores can be used as a remedy, since it combines several inputs and outputs and always uses the best practice technology as the reference value. Several authors suggest the Malmquist productivity index, which represents a combination of efficiency scores with respect to different production frontiers as a valid method for the measurement of productivity increases. This approach is particularly popular for the banking industry (cf. Berger, 1997). For the European insurance industry, however, we know only of applications to Italian (Cummins et al., 1996) and Spanish insurance companies (Cummins and Rubic-Misas, 1998). In this section we will briefly discuss the methodology for computing the Malmquist index.

The Malmquist productivity index (TFP) is based on the outcome of several DEA computations, where the input bundles of a single firm are compared in pairs to production frontiers from different periods of time. In order to understand the concept of the Malmquist index it is useful to decompose it multiplicatively into two components:

$$\text{TFP} = \text{TP} * \text{EP}.$$

The first represents the relative shift of the production frontier itself between the base period and the second period (TP). This change indicates a general improvement of productivity in the industry. On the other hand, the second component measures the change in the relative position of a firm with respect to the production frontiers (EP). This can be interpreted as a catching up effect towards the best practice firms.

The amount of technical change (TP) can be easily derived from the ratio of efficiency scores TE_{tk} with respect to different production frontiers Q_k:

$$TP = \frac{TE_{11}}{TE_{12}} = \frac{x_{11}/x_1}{x_{12}/x_1} = \frac{x_{11}}{x_{12}},$$

where TE_{tk} represents the technical efficiency of a firm in period t, relative to the production frontier Q_k and k indicates the reference technology's measurement period. For example in Figure 3 the TE_{12} is the technical efficiency of firm A in period 1 relative to the efficiency frontier of the second period Q_2 or, equivalently, the ratio of the distances $0x_{12}$ to the distance $0x_1$. By changing the base period for the reference technology k and relating those ratios, we can measure the general increase in productivity change (TP). If the value for TP>1 the production frontier of the second period Q_2 lies to the left of Q_1 and we confirm technical progress between the first and second periods. If TP<1 the production frontier Q_2 lies to the right of Q_1 and we find a market wide technical regress.

The efficiency progress (EP) of an individual firm can be similarly defined as the ratio between efficiency scores:

$$EP = \frac{TE_{22}}{TE_{11}} = \frac{x_{22}/x_2}{x_{11}/x_1},$$

which are now evaluated at different input bundles but still refer to the productivity frontier of the first and second periods. If EP > 1 holds, we can conclude that the firm improved its relative position with respect to the productivity frontier over time. In Figure

3 this would correspond to a catching up process where A_2 is closer to Q_2 than A_1 is to Q_1. In the case EP < 1 the firm lost touch with competitors and its relative position deteriorated. Translated into Figure 3 this would imply that A_2 lies further away from Q_2 than A_1 does from Q_1.

The Malmquist productivity index is then easily computed as a multiplicative combination of general technical change and individual catching up:

$$\begin{aligned} TFP &= TP * EP \\ &= \frac{TP_{22}}{TP_{11}} \frac{TP_{11}}{TP_{12}} \\ &= \frac{x_{11}/x_1}{x_{12}/x_1} \frac{x_{22}/x_2}{x_{11}/x_1} \\ &= \frac{x_{22}}{x_2} \frac{x_{12}}{x_1} \end{aligned}$$

and corresponds to the ratio of the relative distances of A_1 and A_2 with respect to the production frontier Q_2. If TFP > 1 holds, the total factor productivity of the considered firm has been increased between periods 1 and 2, whereas a TFP < 1 indicates a decline in total factor productivity over time.

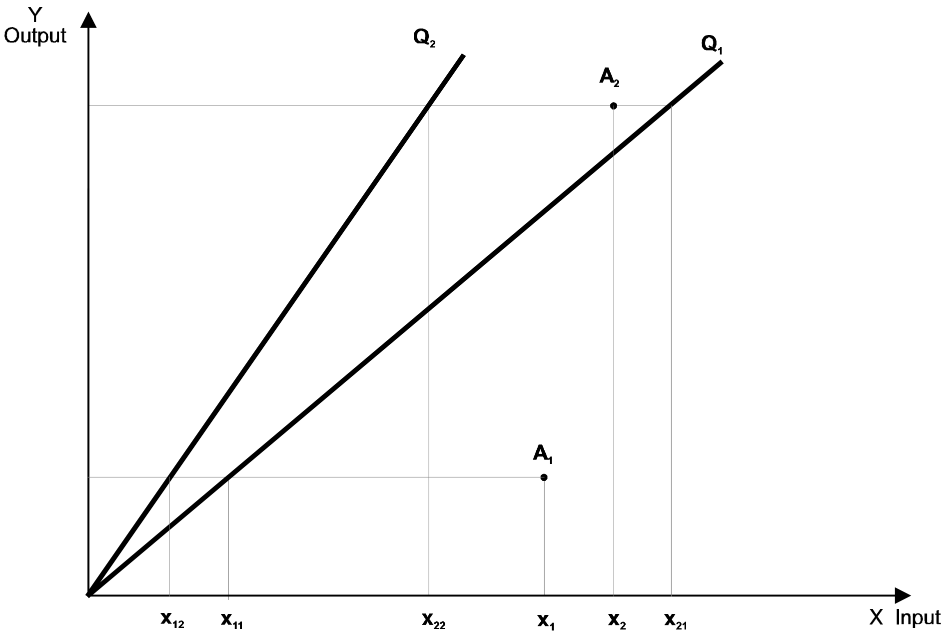

Figure 3: Efficiency frontier in periods 1 (Q_1) and 2 (Q_2) (constant returns to scale).

Measures of Insurance Production and Data

A Model of Insurance

There are numerous measures suggested for inputs as well as outputs of insurance companies. Actual measurement suffers from the fact that insurance companies provide financial services that mainly consist of financial transactions and outputs are not easy to pin down. We intend to base our measurement of the inputs and outputs of insurance companies on the theory of contingent markets. This allows us to embed the analysis into the various possible strategies of risk sharing. Besides the social security system, the stock market, futures, and other derivatives, private insurance companies offer a way for a society to handle individual risks (Laffont, 1989). The concept of contingent markets offers a framework in which risks can be analyzed that harm individuals specifically but do not affect the society as a whole. The basic idea is that through the law of large numbers individuals can form risk pools, such that individual losses can be averaged out within each pool without the necessity of creating contingent markets.

To be specific, consider an exchange economy with N identical consumers and a single good. The endowment of that good available to each consumer, *w*, is stochastically determined. There are two states of nature: with probability (1-π) the consumer will have the full amount *w* at his disposal and with probability π an accident will reduce his endowment by the loss *L*. So we can distinguish between two states of nature for the consumer: the "no-accident" state and the "accident" state. The probability of having an accident is not consumer dependent and if the number of consumers is large enough, i.e. tends toward infinity, we can apply the law of large numbers to get the following per capita endowment for the *i*th consumer:

$$x_s^i = w(1-\pi) + (w-L)\pi = w - \pi L, \forall i = 1,\dots,N, \forall s = 1,\dots,S.$$

where *S* represents the number of possible states of nature ($S=2^N$) and x_s^i is the consumption level in state *s*. By using a von Neuman-Morgenstern utility function, *u(.)*, we can analyze the welfare consequences of stochastic accidents. If we let the usual conditions of positive first and negative second derivatives apply to the utility function, then it can be shown that a competitive and ex ante Pareto optimal Arrow-Debreu equilibrium exists (Laffont, 1989). Within the Arrow and Debreu world the institutional structure would imply that *S* contingent markets must be cleared with *S* equilibrium prices. Transaction costs under such a large number of markets are likely to be enormous.

Insurance companies provide a society with a much simpler institutional structure to achieve a Pareto optimal allocation of goods. This can be achieved by collecting a premium payment, α, and reimbursing all agents that suffer an accident by the amount of their loss *L*. If there are no transaction costs and the zero profit condition holds for insurance companies the actuarially fair premium requires that the individual premium, α, is equal to the expected value of the loss ($\alpha=\pi L$). The endowment of an insured consumer is equal to the expected average endowment and for concave utility functions every consumer has an incentive to join the insurance pool and protect himself fully against losses.

This is illustrated in Figure 4, where the optimal allocation A is characterized by the tangency between the utility function and the budget line provided by the insurance company. Without insurance the consumer would receive *w* in the good state and *w-L* in the bad state, but by paying the premium a an income of *w*-a can be fixed in advance. For risk averse consumers, i.e. for those having concave utility functions, the insured solution A is clearly preferable to the non-insured allocation 0, because the endowments stream is smoothed over time.

By introducing transaction costs, *T*, on the part of the insurance company one can achieve a more realistic picture, where costs arise from organizing the insurance pool. The maximization problem of the household is slightly changed in this case, because we will depart from the actuarially fair premium. This can be shown in a simple example with two states of nature, good and bad. For a given price of the insurance policy, *q*, a representative household chooses its level of consumption in both states and the amount of coverage, $z=(1\text{-q})L$, where q is the share of uncovered losses, by solving the problem:

$$\underset{x_1 x_2 z}{\text{Max}}[(1-\text{p})u(x_1)+\text{p}u(x_2)]$$

s.t.

$$x_1 = w - qz$$
$$x_2 = w - qz - L + z$$

which can be reduced to

$$\underset{z}{\text{Max}}[(1-\text{p})u(w-qz)+\text{p}u(w-qz-L+z)]$$

The first-order condition for utility maximization is then

$$(1-\text{p})u'(w-qz)q = \text{p}u'(w-qz+z-L)(1-q)$$

where *u'(.)* indicates the first derivative of the utility function with respect to the degree of coverage. It can be shown that for zero transaction costs *q*=p and that the slope of the indifference curve at the diagonal is equal to the slope of the budget line running from B to A. In Figure 4 this corresponds to a slope of -(1-p)/p. At this actuarially fair price we will find full coverage (q=*0*). If, due to positive transaction costs, the price for insurance exceeds the fair price, *q*>p, the slope of the budget line gets flatter and the optimal solution for the consumer will be somewhere around C, where another indifference curve reflecting a lower level of utility is tangent to the new budget line. The slope of the budget line can be derived by substituting *q*=p*(1+T)* into the first order condition and it corresponds to the right hand side of the following equation:

$$\frac{(1-\text{p})u'(w-\text{p}(1+T)z)}{\text{p}u'(w-\text{p}(1+T)z-L+z)} = \frac{1-\text{p}(1+T)}{\text{p}(1+T)}$$

The wedge between the fair and the gross premium reduces the chosen coverage and therefore lowers the utility which the household may achieve. Introducing insurance taxes or incomplete competition on the side of the insurance industry makes it more difficult for the households to achieve complete insurance.

The activity of insurance companies can therefore be summarized as providing an alternative institutional setup to the Arrow-Debreu competitive equilibrium with its large number of contingent markets. Within this much simpler institutional framework we can introduce transaction costs for creating and administrating the risk pool, which will result in incomplete coverage of private households. From this perspective we can identify transaction costs as the monetary equivalent of the inputs of an insurance company. The output on the other hand is the provision of smoothed income streams for households, conditional on the occurrence of a loss. The monetary correspondence to the smoothed income stream is therefore not only the actual payment of claims but also the build up of provisions for future losses, which will enable the insurance company to fulfill its future obligations.

A measure of efficiency or productivity based on this definition of inputs and outputs will naturally encompass a consumer's view of insurance activity rather than a shareholder's view. The lower the transaction costs are and the less money is transferred into dividend pay-outs, the higher the relation between outputs and inputs and thus measured efficiency will be. On the other hand, being a low cost insurance company does not rule out high dividend payments, whereas high cost firms are unlikely to show up high profits. Because the premium charges will be large relative to their claims payments, consumers are likely

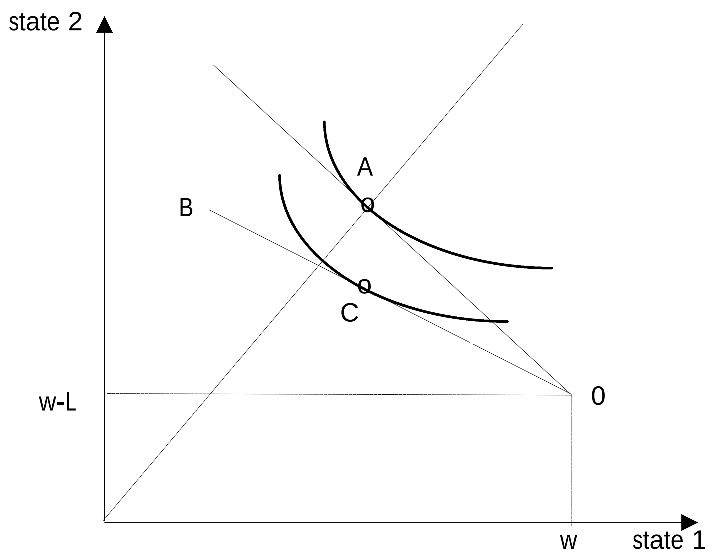

Figure 4: The choice of insurance under transaction costs (Laffont, 1989)

to choose a competitor's offer and our measure should be positively related to standard profit oriented indices. A major advantage of our efficiency measurement is its independence from tax oriented manipulations of the balance sheet.

Data

For the empirical application we will use the definition of inputs and outputs developed in the previous section, i. e. we will use only flow variables to model insurance activities. Since insurance companies are financial intermediaries most of their financial flows are just transfers of money from the insurance pool to agents facing a loss and only a small part of the premium payment is used to cover the costs of organizing the insurance pool. These costs can be interpreted as transaction costs, since they put a wedge between the fair insurance premium and actual premiums. We will regard administration costs, distribution costs net of reinsurance together with costs of capital management as transaction costs or, equivalently, as monetary equivalent to inputs for each insurance company. Because book accounting strategies of individual companies blur distribution and administration costs those two positions will be combined into one input. Administration costs encompass items like expenditures on labor, material, energy, and depreciation, while distribution costs consist mainly of commissions paid to intermediaries.

On the output side we consider three outputs: expenditures on claims, the dotation of insurance reserves, and the amount of returned premiums. This definition of outputs corresponds closely to the concept of smoothed income streams. Claims payments, for example, are direct compensations of income losses. But claims payments alone do not suffice to measure smoothed income completely because for many product lines there is long period between premium and claims payment. Life insurance is a classic example, where contracts expire after periods as long as 30 or 35 years. For this reason the inclusion of dotations of insurance reserves is necessary, since this is the book accounting equivalent of future claims payments. Thus firms with low claims payments will nevertheless show a high level of efficiency as long as they build up equivalent dotations. The third output is returned premiums which on the one hand directly reduce the premium for insured agents and on the other hand represent a participation in profits from good financial management[1].

All inputs and outputs are available for the three main classes of risks in the Austrian insurance business: health insurance, life insurance, and property-liability insurance. Since most of the Austrian insurance companies offer at least two product lines there is no clear separation between expenditures for inputs across classes of risks possible. Consequently, we aggregate the cost components over all business lines into two inputs. We consider nine outputs and two inputs for each firm. Only seven companies offer all three lines of insurance business. This implies that most insurance companies in our sample have zero entries in at least one output line.

[1] One may argue that returned premia occur only if a damage is not claimed by the insured and should therefore be excluded from the output space. But the major part of returned premia consists of the participation in financial profits for life insurance contracts which are transfered into book reserves. Additionally, our definition of outputs relies on the concept of smoothed income streams. The insurance company bears the risk of an indemnity even if there was no actual damage claimed. If claims payments do not occur, less risk is realized in the book accounts of an insurer. Returning premia to the insured reveals those hidden risks and thus must be regarded as an output (c.f. Figure 4).

To achieve full comparability of input and output variables we subtract the reinsurance part from all variables. The dotation of insurance reserves contains the dotation to level premium reserves, other reserves and is an amount less liquidations. All inputs and outputs are measured in Austrian Schillings (ATS). The data comes from the Austrian insurance statistics for the years 1992 through 1996 published by the Austrian Insurance Supervisory Agency (Ministry of Finance). This source contains individual firms' data. Summary statistics for each input and output are given in Table 1. For a detailed study of market characteristics, product structure, and profit related variables see Brandner (1994).

Table 1 gives the minimum, maximum, mean values, and the standard deviation of the variables for the year 1996. A familiar problem for panel data is the exit and entry of firms over time. In order to achieve a complete panel we have eliminated those insurance companies with missing values for one of the years in the observation period. This reduces the sample by 13 observations to 61 insurance companies. Since all removed companies are either foreign insurers that work under single licence since 1994 or domestic firms entering during the sample period we avoid a survivor bias in our analysis. Five more firms are eliminated due their outlier properties.

Table 1: Summary statistics of inputs and outputs for Austrian insurance companies (1996) in ATS 1000

Variable	Description	Mean Value	Standard-Deviation	Minimum	Maximum
Input 1	administration costs (all lines of products)	578178	841883	1014	3847308
Input 2	costs of capital investments (all lines of products)	52826	99776	273	670559
Output 1	claims (health insurance)	213048	713094	0	3736991
Output 2	dotation of insurance reserves (health insurance)	34048	115030	0	626455
Output 3	returned premiums (health insurance)	13291	46584	0	302026
Output 4	claims (life insurance)	496310	911534	0	4185493
Output 5	dotation of insurance reserves (life insurance)	473837	805477	0	4331151
Output 6	returned premiums (life insurance)	232956	399281	0	1683000
Output 7	claims (property-liability insurance)	717275	1269799	0	4708185
Output 8	dotation of insurance reserves (property-liability insurance)	5170	19717	0	138170
Output 9	returned premiums (property-liability insurance)	10430	22725	0	109251

Note: sample size is 56.

Results

The core results of the DEA are the efficiency scores. Those scores are the result of solving the linear program for each insurance company and provide information on the relative position of this firm to all other market participants. Efficiency scores can be calculated under two assumptions on the applied technology: constant or variable returns to scale. In the following we will apply a two step procedure: computing the efficiency scores from the DEA for each year forms the first step, which is followed by a regression analysis as the second step. The regression analysis will provide explanatory variables for the level of firm specific efficiency scores. Another interesting result of DEA can be achieved by comparing constant with variable returns. Conditional on the size of individual firms, we can assess whether a firm lies in the range of increasing, constant, or decreasing returns to scale.

A sensitivity analysis of the optimization outcome is suggested by the relatively high dimension of our output space. With nine outputs it is very likely that we will find many firms at the efficiency frontier because an individual firm is more likely to have one output at the boundary. To check for such effects we compare the efficiency scores with balance sheet information from the sample. By and large the efficiency scores are positively correlated with output-input ratios and the ratio of claims to premium payments. Another measure is the cost ratio which relates administration and distribution costs to premium payments. In this case the correlation is negative, as lower average costs indicate higher efficiency. There are three firms with high cost ratios and efficiency scores of one. Those firms are small companies and the result is probably an outcome of small boundary problems of the DEA.

Further checks of the results can be done by reducing the output space, i. e. by aggregating different outputs into one single variable. For this purpose we chose to aggregate over different risks. Combining health and life insurance, two similar businesses, reduces the output space to six and brings about only negligible changes in the results. If we further aggregate over all business lines, such that the output space is reduced to three dimensions, we can find stronger reactions. In this case the efficiency scores of about 15 firms drop. For this reason, we conclude that our results suffer only to a small extent from a bias towards high efficiency scores.

Table 2 presents efficiency scores under both assumptions on the production technology for insurance companies. The statistics for the variable returns to scale (VRS) case indicate higher efficiency scores compared to the results under the constant returns to scale (CRS) assumption. This is to be expected from the construction of the optimization problem under VRS, since in this case the efficiency boundary is curved and more points of support are required to fix the tighter envelope. Under VRS the average efficiency score is about 77%, indicating that the average firm has a potential for cost cutting of around 23 percentage points. Under CRS the potential for efficiency gains is even larger and fluctuates between 37 and 51 percentage points. As can be seen in Table 2, the difference is mainly due to the relatively large number of fully efficient firms under VRS.

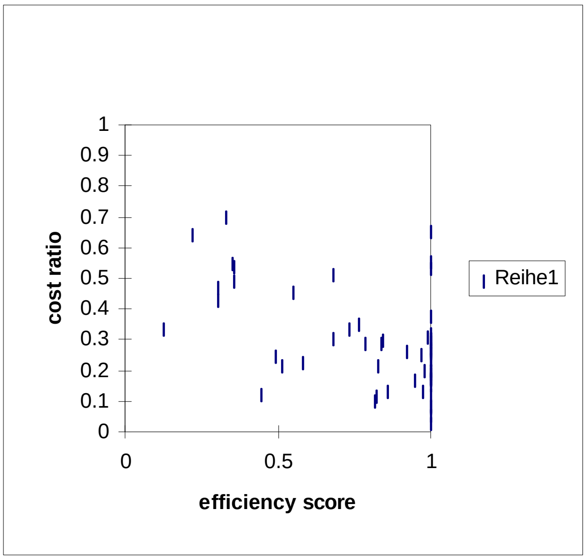

Figure 5: A comparison between the efficiency score and the cost ratio for 1996

The distribution of efficiency scores is given in Figure 6 where 10% classes are distinguished. Actually, very few firms show extremely low efficiency scores under VRS. There are only 4 firms with scores below 50% of the best practice firm. Out of the remaining 52 companies, 33 are concentrated in the range between 91 and 100%. Under CRS technology the number of highly inefficient firms rises somehow towards 11% (or: to 26). Out of the five biggest insurance companies four show over the whole sample period a value for the efficiency score of one. Complete efficiency under VRS holds also if all members of a group are consolidated into a single firm. Under CRS, however, only one of the large companies remains fully efficient, indicating diseconomies of scale.

Table 2: Technical efficiency under Variable Returns to Scale (VRS) and Constant Returns to Scale (CRS)

	1992	1993	1994	1995	1996	average
VRS						
geom. mean	0.747	0.758	0.754	0.811	0.792	0.772
standard deviation	0.250	0.247	0.247	0.213	0.228	0.237
maximum	1.000	1.000	1.000	1.000	1.000	1.000
minimum	0.198	0.206	0.136	0.281	0.305	0.225
number of efficient companies	25	25	24	29	29	26
efficient companies (in percent)	42	42	41	49	49	45
CRS						
geom. mean	0.524	0.493	0.604	0.585	0.634	0.565
standard deviation	0.312	0.300	0.279	0.261	0.264	0.283
maximum	1.000	1.000	1.000	1.000	1.000	1.000
minimum	0.122	0.121	0.126	0.217	0.215	0.160
number of efficient companies	15	11	14	13	15	14
efficient companies (in percent)	25	19	24	22	25	23

Scale efficiency						
geom. mean	0.701	0.650	0.800	0.722	0.800	0.748
standard deviation	0.260	0.265	0.195	0.214	0.194	0.226
maximum	1.000	1.000	1.000	1.000	1.000	1.000
minimum	0.198	0.163	0.382	0.360	0.379	0.296
number of efficient companies	15	11	14	13	16	14
number of efficient companies in %	25	19	24	22	27	23
number of firms in the range of						
decreasing returns to scale	26	28	25	30	30	
constant returns to scale	15	11	14	13	16	
increasing returns to scale	15	17	17	13	10	

Note: sample size is 56.

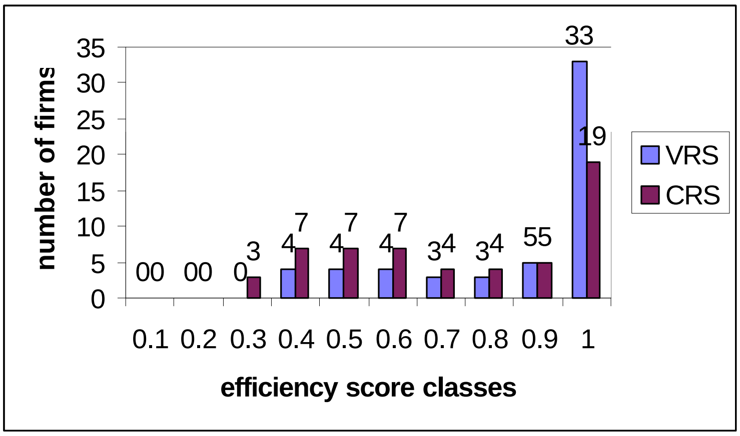

Figure 6: Frequency distribution of technical efficiency scores under VRS 1996

The divergence between the VRS and CRS results reflects scale effects, because the difference between the CRS and the VRS envelope can be interpreted as an inefficiency due to the wrong scale of a firm. Firms may be too large, so that they fall into the region of diseconomies of scale (cf. Figure 2). In this case firms would improve their efficiency by reducing size. On the other hand, a company may be too small so that it falls into the region of increasing returns to scale. For those firms an increase of size would improve efficiency.

Given the industry wide phenomenon of mergers and acquisitions it is interesting to prove whether firms actually can improve their efficiency by increasing their size. Table 2 presents average scale efficiencies for all years, i. e. the relative distance between the CRS and the VRS envelope. The average value of 75% indicates that insurance companies would be able to improve their efficiency on average by 25% by adjusting to the right size. The minimum value of 16 percent indicates that there are firms in the Austrian insurance market which heavily deviate from their optimal size. These are the very small firms with several zero outputs. Counting the number of firms within each range of scale reveals comparatively more firms in the decreasing returns area, for example 30 in 1996 (c.f. Table 2). But a detailed inspection of scale efficiencies according to the size of the firm leads to the conclusion that small units deviate more heavily from the optimum value as compared to larger units. This implies a steep VRS envelope in the region of small inputs and outputs which, for larger firms, takes almost the same slope as the CRS envelope.

Transferred into average cost curves we find an almost L-shaped average cost curve, which is steeply falling for small units but takes only a small positive slope after reaching the lower turning point.

In a second step we try to explain the efficiency scores of individual companies by exogenous variables. Since we already analyze a measure based on a combination of variables under the control of the decision maker it is important to be very careful about the choice of further explanatory variables. In order to minimize the loss in estimation efficiency one must search for variables that are not associated with the decision making process already modeled in the first step (cf. Lovell, 1993). Whether the regression is successful in explaining the variation in efficiency scores or not depends largely on the impact of managerial competence on the efficiency of individual firms. Since managerial skills are an important asset for each company the regressions are likely to show a loose fit.

The distribution of efficiency scores in Figure 6 already indicates a bunching of observations around the full efficiency value of 1. Since this value corresponds to the benchmark firm it is the highest observable value for an efficiency score. The underlying productivity of a firm can be interpreted as a latent variable with unlimited range. The efficiency score may then be interpreted as a truncated observation of the underlying productivity. Under this assumption a Tobit model is appropriate to estimate the impact of exogenous variables on the efficiency scores.

Our regression analysis is oriented towards a general to specific way of model selection. That means we include all possible explanatory variables to identify general factors for high or low efficiency values. Subsequently we eliminate insignificant variables by Likelihood Ratio tests. To get hold of the correlation among exogenous variables we introduce interaction terms and reverse our strategy towards specific to general by adding variables once more. The choice of variables is motivated by our discussion of the development of the European insurance industry over the last couple of years. It is also motivated by our interest in the effects of several indicators related to the size of insurance companies, the degree of diversification, profitability, the ownership structure, and the distribution channel. Since we are analyzing individual years it seems appropriate to add a measure of risk clustering, which indicates a year with large losses and extraordinary business activities. For this purpose we use the claims ratio (CLAIMS), i. e. the ratio between claims and premium payments.

Another argument in favor of a regression analysis is that it will allow us to assess the impact of the firm size on the efficiency score without any assumption about the underlying technology of the firm. Economies of scale are approximated by two variables. First of all, the SIZE of a company indicates scale effects and is measured as the log of the sum over all outputs. Second, we introduce a dummy variable (GROUP), which indicates whether an insurance company belongs to a group (1) or not (0). The variable GROUP is also an indicator of scope economies, since members of a group can be supposed to engage in cross selling of products from befriended firms. A variable for measuring economies of diversification is the concentration of premiums over the three lines of business. We measure diversification by a Herfindahl Index (HERF), which increases with the level of premium concentrated within a single product line (Tirol, 1989).

The impact of new distribution channels is captured by a dummy variable indicating companies which use the banking distribution channel (BANK). Most of the other firms rely mainly on fixed employees or exclusive agents. Another interesting explanatory variable is the profitability of individual companies measured as the ratio of profits from regular business activity to the total premiums (PROFIT). One would expect profitable companies to have a high productivity level, which is reflected in large efficiency scores. As an alternative there can be substantial redistribution between the risk pool and the capital owners. Another variable that has an impact on the profitability of insurance companies is the age of the firm (AGE). Since there are high entry costs during the build up phase of the risk pool, older companies face lower average costs. Finally, the ownership structure may be important in the Austrian market since about half of the insurance companies belong to foreign groups. To capture this effect we introduce a dummy variable which is zero for foreign and one for domestic firms (FOREIGN).

There are two ways to estimate a Tobit model: first we can estimate a censored Tobit model by maximum likelihood and second, we can apply a two step procedure where in the first step we estimate a binomial Probit model to distinguish efficient from inefficient firms and in the second step we estimate a truncated Tobit model to determine the factors affecting inefficient firms. The second set up allows the explanatory variables to have different parameter values or even different signs in the first and the second step. Both models can be tested against each other by a Likelihood Ratio test (Cragg, 1971).

After the elimination of insignificant variables we are left with a model for the efficiency scores in 1996 including four significant variables: the size of a firm, the diversification of its products, its claims ratio, and the ownership structure. As the final model we can identify the censored Tobit model, because the value of the Likelihood Ratio test statistic is at 11.98, while the 5% critical value of the Chi-2 distribution with five degrees of freedom is 11.1. The results from maximum likelihood estimation for efficiency scores under VRS for the year 1996 are as follows:

$$EFFVRS = \underset{(-1.89)}{-1.49} + \underset{(2.33)}{0.77}\,HERF + \underset{(3.04)}{0.12}\,SIZE - \underset{(-2.09)}{0.21}\,FOREIGN + \underset{(1.86)}{0.47}\,CLAIMS + u_{VRSPRO}$$

The claims ratio in these equations serves as a conditioning variable, which indicates companies heavily affected by claims in 1996. As to be expected from the construction of the efficiency scores the parameter is positive. The ownership variable is also relevant for explaining productivity levels. Insurance companies belonging to a foreign group are significantly less efficient compared to domestic firms. This result confirms results for the Austrian insurance industry in Szopo (1987), who found significantly higher average costs for foreign dominated firms.

Both, SIZE and HERF, are significant at the 5% level. Contrary to other studies for the Italian life and non-life insurance (no scope economies, Cummins et al., 1996), the Canadian insurance industry (positive scope effects, Kelly and Siklos, 1996) and the US Life insurance industry (Meador et al., 1997), we find a negative effect from product differentiation. The Herfindahl-Index rises with the concentration in one product line, thus a positive parameter indicates that highly specialized firms are more likely to be fully efficient. This result is more in line with Ferrier et al. (1993) who draw similar conclusions with respect to the US banking industry. The coefficient of the size variable suggests that large firms do indeed have an advantage in achieving high efficiency levels.

The corresponding equation for the efficiency scores under CRS for the year 1996 brings about roughly equivalent results. The claims ratio drops out of the equation due to insignificance and we prefer a censored Tobit model over the two step approach, because the Chi-2 distributed test statistic is at 16.5, which is way above the critical value for 4 degrees of freedom (9.49). The results from a maximum likelihood estimate of the censored Tobit model are:

$$EFFCRS = \underset{(-2.59)}{-1.30} + \underset{(4.69)}{1.04}\,HERF + \underset{(3.47)}{0.09}\,SIZE - \underset{(-3.41)}{0.25}\,FOREIGN + u_{CRS}$$

A comparison over the sample period from 1992 to 1996 and thus an assessment of efficiency improvements during the first years of the Single Market can be made by computing a Malmquist productivity index. This measure compares efficiency scores from different years and allows us to disentangle the overall increase in productivity for each firm into a part resulting from a firm specific catching up in efficiency (efficiency progress) and another part emerging from industry wide improvements of productivity (technical progress). For a firm which lies in each period on the efficiency frontier the Malmquist index follows directly from comparing the relative position of each frontier. For inefficient firms we also have to consider catching up and falling behind movements, respectively. For this reason the base year of the comparison is important. There are basically two approaches to comparing efficiency scores between different years (Berg et al., 1992): on the one hand we can perform the calculation for each successive pair of years, but in this case the reference technology changes every year and, as can be seen from the construction of the Malmquist index, we cannot chain the index over a longer period of time. This means that an index for periods 1 to 3 is not the product of two indices ranging from 1 to 2 and from 2 to 3. As an alternative one can fix the reference technology and relate each successive period to the reference technology of the base year. In this case the circular relation is fulfilled and we can assess the productivity change over the whole observation period.

To give an impression of the different outcomes of both approaches we present both results in Table 3. For the computation we transfer all series into real 1992-levels by applying the subindex for health insurance products in the CPI to deflate outputs in the health insurance branch. We use total CPI to achieve real life insurance outputs. In the property-liability branch we use a weighted average of CPI subindices related to premia in motor third party and household insurance. Costs of capital investments are deflated by the GDP-deflator and for administartion costs we use the wage index for indoor employees. By applying different price indices we take care of shifts in relative prices among inputs and outputs. Our deflators, however, are only rough approximations of the price development over time and using a uniform index such as the GDP-deflator for all inputs and outputs changes the results only slightly.

In our case we have access to the efficiency scores for the period 1992 through 1996. This period comprises the years in which the second and third generations of insurance directives were incorporated into national law and is of particular interest with respect to the impacts on productivity changes. In the period after 1994 the Single Market was completed and thus full market access was possible by using the right to establish branches or by using the freedom to provide services on the basis of a single license and prudential control. The Austrian market, however, is still dominated by firms under domestic regulation and direct sales activities of Austrian insurance companies are still in their infancy (Url, 1997).

The geometric mean of firm specific Malmquist indices of the period 1992 through 1993 indicates on average a drop in the individual catching up component and an increase in general technical progress. Total factor productivity went up by 10 percentage points. As can be seen by the development during the following period 1993 through 1994, the average firm was able to improve its relative position to the benchmark by 22 percentage points, while the general technical progress slumped. The period between 1994 and 1995 can be characterized as one of stagnation and finally a comparison of 1996 with 1995 shows a further catch up of individual efficiency positions accompanied by an impressive improvement in general productivity. Over the whole sample period we can clearly observe a catching up of the average firm towards the best practice, which is strengthened by productivity growth (TP) of 10 percentage points. Removing the two worst and the two best performing firms leaves the results almost unchanged but decreases the standard deviation strongly.

Table 3: The Malmquist Index decomposed into efficiency progress and technical progress

	Efficiency Progress (EP)	Technical Progress	Malmquist Index (TFP)
1992 to 1993			
geom. mean	0.942	1.172	1.104
standard deviation	0.355	0.388	0.543
maximum	2.117	2.793	3.460
minimum	0.163	0.664	0.212
1993 to 1994			
geom. mean	1.224	0.510	0.624
standard deviation	0.915	0.225	0.488
maximum	6.135	1.037	2.653
minimum	0.354	0.107	0.103
1994 to 1995			
geom. mean	0.969	1.002	0.971
standard deviation	0.351	0.349	0.491
maximum	2.376	1.957	2.806
minimum	0.475	0.230	0.141
1995 to 1996			
geom. mean	1.083	1.209	1.310
standard deviation	0.368	0.596	1.119
maximum	2.494	3.436	6.369
minimum	0.294	0.308	0.380
1992 to 1996			
geom. mean	1.210	1.099	1.330

standard deviation	1.108	0.602	1.100
maximum	8.197	4.491	7.246
minimum	0.221	0.410	0.150
1992-1996: outlier corrected			
geom. mean	1.207	1.123	1.355
standard deviation	0.501	0.613	0.695
maximum	2.564	4.491	4.118
minimum	0.501	0.410	0.410

Note: sample size is 56, for outlier corrected results 52.

Again we apply a second step regression analysis in which we search for explanatory variables of changes in productivity. Given the special situation during the years towards the Single Market we expect especially those firms that are exposed to foreign competition to improve their efficiency level. Therefore, variables indicating higher competitive pressures should provide significant explanatory variables. If there is really entry deterrence or predatory pricing behavior in the industry, we would expect premiums to react immediately at the appearance of powerful foreign competitors; maybe even be lowered in advance of market entries. This puts insurance companies under a cost pressure, since lower intakes of premiums have to cover a relatively constant amount of claims. Moreover, the cost position of individual companies cannot be changed very quickly.

Not all branches of the insurance business will be exposed to the same level of foreign penetration. As Szopo (1990) already proposed, we should expect foreign companies to be active in insuring enterprises rather than private households. Within the insurance of private households we expect the mandatory car liability insurance to be contested as it is the main market. For this reason we construct a risk structure index (RISKSTRU), which reflects the exposure of individual insurance companies with respect to those risks as a ratio of premiums in exposed product lines to total premiums[2]. Again we look for the effects of scale and diversification economies by testing for the significance of the log of aggregated premiums (SIZE) and the concentration of premiums in one product line (HERF). Besides the size we also consider membership to a group (GROUP) as a potential way to reduce costs of overhead expenses. Also the age (AGE) of a company might be of relevance for its productivity development, because set up costs arise in the beginning and hidden reserves allow a smoothing of revenue figures.

Since the setup of a banking distribution channel can be regarded as a response to the new competitive climate and because cross selling is widely considered to improve the efficiency level in the financial services industry as a whole, we also include a dummy variable in the regression indicating bankassurance companies (BANK). Additionally, we expect those insurance companies that were lagging in 1992 to make the greatest efforts to improve their efficiency and thus include the efficiency scores from 1992 (THETA92) in the regression. During the sample period the life insurance business experienced a boom in premiums intakes as well as claims payments due to changes in the tax code and to a wave of expiring contracts. Both developments will affect the productivity development of companies with an emphasis on the life insurance business. Since the sales force in life

[2] These product lines are motor third party, airplane, industrial fire, production interuptions due to fire, machinery and production interuptions liability, and other transport casualty insurance.

insurance business is rewarded in the year in which the contract is signed, but no additional sales costs emerge during the following years, an upswing in premiums of 30% as in 1996, increases costs disproportionally. On the other hand, a boom in claims increases productivity. Since both developments are a response to changes in the tax code, we regard the increase in output not as an efficiency increase, rather it may be viewed as a policy induced exogenous shock. For this reason we include a dummy for companies engaged in the life insurance business (LIFE). Foreign or domestic ownership (FOREIGN) may be another relevant variable, because foreign companies may be more flexible but may at the same time face hindrances in raising capital for productivity enhancing investments.

We follow once more a general to specific modeling strategy but we distinguish between sources affecting the total factor productivity (TFP), i.e. the Malmquist index as a whole, and factors determining the change in the individual position with respect to the benchmark firm (EP). Because neither index has an upper boundary where a large portion of the observations are concentrated, we conduct OLS-regressions with 52 observations for the change between 1992 and 1996 (4 outliers have been removed from the sample). The explanatory variables are measured as averages over this sample period. The following models emerge after elimination of insignificant variables:

$$EP = \underset{(0.78)}{0.21} + \underset{(2.52)}{0.89}\, HERF + \underset{(2.10)}{0.09}\, SIZE - \underset{(-7.34)}{1.39}\, THETA92 + u_{EP}$$

$$TFP = \underset{(12.63)}{1.99} + \underset{(4.28)}{0.88}\, BANK - \underset{(-4.89)}{1.18}\, THETA92 + u_{TFP}$$

The remaining significant variables explain about 50% of the variance of the efficiency progress (EP) and 36% of the variance in the Malmquist index (TFP) and indicate that the general productivity development can be surprisingly well explained. Only one variable appears in both models. It is the starting position in 1992. The negative sign for THETA92 indicates that a productivity lag in 1992 motivated insurance companies to increase their relative position towards the benchmark, which in turn allowed firms to improve their total factor productivity.

Bigger and more specialized firms were able to improve their relative position to the bench mark during the period between 1992 and 1996. Dropping the AGE variable in the search for a parsimonious model shows that neither the change in the efficiency position of individual firms nor the one in total factor productivity is related to a longer record of business activity. Also the dummy variable for life insurance is insignificant.

One of the most interesting results is the significant positive impact of the bankassurance variable (BANK) on total factor productivity. Pursuing a policy of bankassurance significantly allowed insurance companies to expand their business and utilize their inputs in a better way over the period 1992 through 1996. This result may be due to the fact that a close cooperation with banks opens cross selling opportunities in the insurance business and the associated fees for intermediary services charged by banks do not outweight the increase in claims and provisions of reserves. This is the more remarkable since most of the firms applying bankassurance are young and concentrated in the life insurance business. Whereas the industry wide average share of life insurance in total premium intakes is at 37.5% (1996), bankassurance companies get 97.7% (1996) of their revenues from this line of business. The fee structure in the life insurance business is of an up-front character which implies that a significant part of the early premium payments goes to the intermediary and affects the provision of reserves.

The risk structure of individual companies does not have a significant effect on total factor productivity. Thus the increase in competition within certain business lines enforced neither efficiency nor technical progress. Only by excluding the years 1995 and 1996, which are heavily affected by the surge in the life insurance industry[3], from the computation of the Malmquist Index, can we find a positive impact of the risk structure. Firms more exposed to competitive insurance markets tended to have a better productivity record between 1992 and 1994. Neither the size nor the diversification of activities of a company plays a significant role in the explanation of the change in total factor productivity, although large and spezialized insurers are more likely to catch up to the best practice. This surprising result may be due to higher adjustment costs faced by bigger companies due to their active response to the opportunities of the Single Market. Widespread merging activites and the expansion into neighboring markets in Germany, Italy and especially Eastern Europe may provide a source for unfavorable cost developments.

[3] The growth rates of premium intakes were at 15% (1995) and 29.6% (1996).

Summary and Conclusions

In this paper we use a panel of Austrian insurance companies over the period 1992 through 1996 to analyze the adjustment to the challenges of the Single Market for the insurance industry by means of a Data Envelopment Analysis (DEA). This allows us to assess the technical efficiency of individual firms in relation to a set of best practice or benchmark firms. We measure technical efficiency from the consumer's perspective, i. e. we are looking for firms that minimize their cost, while maximizing their pay-outs to the insured. Although this measure deviates from conventional profit oriented indices, low cost companies should show up high efficiency measures under both concepts. Moreover, our efficiency measure avoids the problems of tax oriented manipulations of reported profits. A comprehensive assessment of insurance companies must add a measurement of solvency as a complement.

Our efficiency measure is based on the average consumer who possesses average risk characteristics. If it is possible to differentiate between risk classes insurance companies are likely to pursue a strategy of cream skimming by charging lower premiums to identifiably low risk consumers. If a company specializes in this area it will face higher input costs due to intensive screening activities and at the same time have low claims pay-outs because of the favorable risk structure. Our efficiency measure is likely to identify such insurance companies as inefficient relative to other companies, although they may offer favorable terms to consumers. This outcome, however, is due to the restriction of favorable terms to specific risk classes, which are not offered to the average consumer.

We compare expenditures on inputs with payments for claims, dotations of insurance reserves, and the amount of returned premiums, and compute efficiency scores under the assumption of constant and variable returns to scale. From this we decompose technical inefficiency into pure technical inefficiency and scale inefficiency. The extent of scale inefficiencies gives a hint on the possible benefits from future merger activities in the Austrian insurance industry. In a second step we use the efficiency scores from the DEA in subsequent regressions to reveal whether differences in technical efficiency are more likely due to scale or scope effects. Finally, by analyzing our panel before and after the start of the Single Market we assess the impact of integration by means of a Malmquist productivity index.

The results indicate higher efficiency scores for the variable returns to scale (VRS) case compared to the assumption of constant returns to scale (CRS). This outcome is mainly due to the large number of fully efficient firms under VRS. Under VRS the average efficiency score is about 75 percentage points, indicating that the average firm has a potential for cost cutting of around 25 percentage points. Under CRS the potential for efficiency gains is even larger and fluctuates between 44 and 60 percentage points. This is to be expected from the construction of the optimization problem and indicates diseconomies of scale of considerable size.

According to the difference between the CRS and VRS envelopes, the average insurance company would be able to save 30 percentage points of its costs by adjusting to the right size. We do find asymmetries between firms in the increasing and decreasing returns to scale area. First of all, large firms tend to be closer to or at the efficiency frontier regardless of reference technology, whereas small firms deviate on average strongly from the best practice. Moreover, the increasing returns to scale branch of the VRS-frontier is farther apart from the CRS-frontier than its corresponding decreasing returns branch. This result indicates asymmetries in the average cost curve of Austrian insurance companies. The average cost curve is steeply falling for small units and gets just a slightly positive slope for large firms. This almost resembles the common L-shaped average cost curve for the insurance industry and provides a rationale for the industry wide phenomenon of mergers and acquisitions.

To allow for different explanatory factors of the scores of individual companies we run regressions on exogenous variables which indicate size, diversification, and other firm specific characteristics. The second step regressions reveal that the size of a company and a stronger specialization in one of the three distinguished branches significantly improves technical efficiency. These results hold regardless of whether we analyze efficiency scores under VRS or CRS assumptions. We conclude therefore that there are significant economies of scale in the Austrian insurance business that may serve as a justification for recent mergers, but that mergers should keep the range of products small.

The growth of productivity of an insurance company can be analyzed by computing the Malmquist productivity index. This measure distinguishes between a firm specific catching up in efficiency (efficiency change) and the industry-wide improvements of productivity (technical change). Malmquist indices over the period 1992 through 1996 indicate on average an increase in general productivity by 13 percentage points. Insurance companies were able to close the average efficiency gap towards the benchmark firms by 16 percentage points. The combined effect of individual catch up to the best practice and overall productivity growth resulted in an increase of total factor productivity by 31 percentage points.

In second step regressions we relate the Malmquist indices to exogenous factors. We distinguish between the firm specific catching up progress and the general technical change. Firms with a poor efficiency record in 1992 put more emphasis on improving their productivity performance during the period 1992 through 1996. This response is to be expected for any firm regardless from considerations of the Single Market. Again, specialized large firms had a better record in correcting their relative position to the best practice. An important contribution to the growth in total factor productivity does not result from competitive pressures in specific product lines but from the use of new distribution channels and the expansion of the life insurance business due to tax reforms and the debate on the social security system which mask the effects of the Single Market on total factor productivity. Only by excluding the boom years in life insurance from the sample (1995 and 1996) we are able to detect a relation between competitive pressure and improvements in productivity. Insurance companies exposed to business lines with more intense foreign penetration or a high cross selling capacity showed a better performance during the period 1992 through 1994.

References

Baumol, W. J., Panzar, J. C., Willig, R. D. (1988): "Contestable Markets and the Theory of Industry Structure ", 2. ed., Harcourt Brace Jovanovich Publishers, San Diego, New York.

Berg S. A., Försund F. R., Janson E. S. (1992): "Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980-89", Scandinavian Journal of Economics, pp. 211-28.

Berger, A. N., Hancock, D., Humphrey, D. B. (1995): "Bank Efficiency Derived From the Profit Function", Journal of Banking and Finance, 17, pp. 317-47.

Berger, A. N., Humphrey, D. B. (1997): "Efficiency of Financial Institutions: International Survey and Direction for Future Research", European Journal of Operational Research, 98, pp. 175-91.

Brandner, P. (1994): "Versicherungen", in: Breuss, F.: "Die Auswirkungen des Binnenmarktes auf den Dienstleistungssektor in Österreich", Study by the Austrian Institute of Economic Research (WIFO), Vienna, pp. 67-80.

Charnes A., Cooper W. W., Rhodes E. (1978): "Measuring the Efficiency of Decision Making Units", European Journal of Operational Research 2, pp. 429-44.

Cragg, J. G. (1971): "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", Econometrica, Vol. 39, No. 5, pp. 829-44.

Cummins, J. D., Turchetti, G., Weiss, M. A. (1996): "Productivity and Technical Efficiency in the Italian Insurance Industry", Wharton School Financial Institutions Center Working Paper No. 96-10.

Cummins, J., Rubic-Misas, (1998): "Productivity and Technical Efficiency in the Spanish Insurance Industry", mimeo, Wharton School Financial Institutions Center.

Färe, R., Grosskopf S., Lovell C. A. K. (1994): "Production frontiers", Cambridge (Cambridge University Press).

Farrell, M. J. (1957): "The Measurement of Productive Efficiency", Journal of the Royal Statistical Society, Series A, Vol. 120, pp. 253-81.

Ferrier, G. D., Grosskopf, S., Hayes, K. J. and Yaisawarng, S. (1993): "Economies of Diversification in the Banking Industry: A Frontier Approach", Journal of Monetary Economics 31, pp. 229-49.

Grosskopf, S. (1986): "The Rule of the Reference Technology in Measuring Productive Efficiency", Economic Journal, Vol. 96, pp. 499-513.

Kelly, G., Siklos, P. L. (1996): "Evaluating Cost Economies in Canada's Property and Casualty Insurance Industry", mimeo Wilfrid Laurier University, Waterloo, Ontario.

Laffont J.-J. (1989): "The Economics of Uncertainty and Information", MIT-Press, Cambridge MA.

Lovell, C. A. K. (1993): "Production Frontiers and Productive Efficiency ", in H. O. Fried, C. A. K. Lovell, S. S. Schmidt „The Measurement of Productive Efficiency", Oxford University Press, Oxford.

Meador J. W., Ryan, H. E., Schellhorn, C.D. (1997): "Product Focus versus Diversification: Estimates of X-Efficiency for the US Life Insurance Industry ", Wharton School Financial Institutions Center Working Paper pp. 97-116, University of Pennsylvania.

Szopo, P. (1987): "Skalenerträge in der österreichischen Lebensversicherungswirtschaft. Ein Beitrag zur Interpretation des Versicherungsaufsichtsgesetzes", in: Empirica 14 (1), pp. 77-97.

Szopo, P. (1990): "Die Österreichische Versicherung im EG-Binnenmarkt", Study of the Austrian Institute of Economic Research (WIFO), Vienna.

Tirol, J. (1989): "The Theory of Industrial Organization", MIT Press, Cambridge MA.

Url, T. (1997): "Strukturanpassungsgesetz bewirkt 1996 Wachstumsschub in der Versicherungswirtschaft", Austrian Institute of Economic Research (WIFO), Monthly Reports , 70, Vienna, pp. 753-60.