Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 12-19

Performance of the life insurance industry under pressure: efficiency, competition and consolidation

Jacob A. Bikker

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Kriekenpitplein 21-22
3584 EC Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the author

Jacob A. Bikker
De Nederlandsche Bank (DNB)
Supervisory Policy Division
Strategy Department
P.O. Box 98
NL-1000 AB Amsterdam
The Netherlands
E-mail: j.a.bikker@dnb.nl
Utrecht University
Utrecht School of Economics
Kriekenpitplein 21-22
3584 TC Utrecht
The Netherlands.

This paper can be downloaded at: http:// www.uu.nl/rebo/economie/discussionpapers

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 12-19

Performance of the life insurance industry under pressure: efficiency, competition and consolidation

Jacob A. Bikker

De Nederlandsche Bank (DNB)

Utrecht School of Economics
Utrecht University

November 2012

Abstract

A well-performing life insurance industry benefits consumers, producers and insurance firm stockholders alike. Unfavourable market conditions stress the need for life insurers to perform well in order to remain solvent. Using a unique supervisory data set, this paper investigates competition and efficiency in the Dutch life insurance market by estimating unused scale economies and measuring efficiency-market share dynamics during 1995-2010. Large unused scale economies exist for small and medium-sized life insurers, indicating that further consolidation would reduce costs. Over time average scale economies decrease but substantial differences between small and large insurers remain. A direct measure of competition confirms that competitive pressures are at a lower level than in other markets. We do not observe any impact of increased competition from banks, the so-called investment policy crisis or the credit crisis, apart from lower returns in 2008. Investigation of product submarkets reveals that competition is higher on the collective policy market, while the opposite is true for the unit-linked market, where the role of intermediary agents is largest.

Keywords: Life insurance, competition, efficiency, Performance-Conduct-Structure model, Boone indicator, concentration, economies of scale.

JEL classification: G22; L1.

Acknowledgements

The author is grateful to Paul Cavelaars and Janko Gorter for useful comment and Jack Bekooij for excellent research assistance.

1. Introduction

This paper investigates efficiency and competitive behaviour on the Dutch life insurance market.[2] In the Netherlands, the life insurance sector is important with a business volume of € 22 billion in terms of annual premiums paid, invested assets of € 337 billion and insured capital of € 990 billion at end-2010. This market provides important financial products such as endowment insurance, annuities, term insurance and burial funds, of frequently sizeable value to consumers. Financial planning of many households depends on the proper functioning of this market. The complexity of the products and dependency on future investment returns make many life insurance products rather opaque. Therefore, competition and efficiency in this sector are important for consumers (Bikker and Spierdijk, 2010). Most life insurance policies have long life spans, which makes consumers sensitive to the reliability of the respective firms. Life insurance firms need to remain in a financially sound condition over decades in order to be able to pay out the promised benefits.

In recent years, the life insurance sector has been confronted with several major challenges. First, the ongoing long-lasting decline in interest rates, world-wide but particularly in the euro area, has reduced insurers' income which is – among other things – needed to cover future benefits to policyholders. Second, the credit and government debt crises have lowered the value of stocks and PIIGS countries bonds, which have impaired insurance firms' buffers. In the Netherlands, two additional problems for life insurers have emerged. In order to increase competition between banks and insurers, tax privileges for insurance products, such as old-age savings and redemption plans for mortgage loans, have also been made available for comparable banking products. The impact of this tax reform has been huge: more than half of the new production on the respective insurance markets has been gained by the banks. Finally, the insurers face the so-called investment policy crisis. Around 2006, public awareness increased that various types of unit-linked saving policies (such as annuities and mortgage redemption saving plans), which were based on capital market investment at the risk of policyholders, carried high operational costs and relatively high premiums on included life risk policies, eating 50-60% of the invested premiums. Under public pressure, insurers agreed to pay compensation to policy holders for incurred and future costs on the respective policies, estimated at € 2.5 to 4.5 billion,[3] while potential claims may come to a multiple of that amount. One of the consequences of this crisis is that consumer trust in insurance firms and the volume of new production have decreased

Competition and efficiency in financial markets is difficult to measure, particularly due to the unavailability of data with respect to costs and prices of individual financial products (Bikker, 2010).

---

[2] Different from most other countries, in the Netherlands, health, disability and accident insurance is not included in the life sector, but in 'non-life'.

[3] http://www.verzekeraars.nl/sitewide/general/nieuws.aspx?action=view&nieuwsid=880.

The solution in the literature has been to assume a single insurance (or banking) product and to use balance sheet and profit and loss data of entire financial institutions. As a first measure, this paper estimates unused scale economies, which is a proxy of inefficiency, but at the same time an indirect measure of competition: where competition is high, insurers are forced to reduce their cost level wherever possible. Further, we apply a measure of competition which has to date been rarely used in the literature, namely a Performance-Conduct-Structure (PCS) model, also known as the Boone indicator, developed by Hay and Liu (1997) and Boone (2000, 2008). Where the well-known Structure-Conduct-Performance (SCP) paradigm of Mason (1939) and Bain (1951) explains performance via conduct from market structure, this alternative model explains market structure via conduct or competition from performance, as in the so-called efficiency hypotheses (Smirlock, 1985; Goldberg and Rai, 1996). This approach is based on the notion that competition rewards efficiency and punishes inefficiency. In competitive markets, efficient firms perform better – in terms of market share and hence profit – than inefficient firms. The PCS indicator measures the extent to which efficiency differences between firms are translated into performance differences. The more competitive the market is, the stronger is the relationship between efficiency and performance. The PCS indicator is usually measured over time, giving a picture of the development of competition. Other measures of competition, such as the Lerner index and the Panzar and Rosse model, are less suitable because the required data (output prices, cost price, profit margin) are lacking, while the Concentration index and the SCP model have serious shortcomings, see Bikker and Bos (2008). Another advantage of the PCS indicator is that it requires only a small number of data series. We combine the two measures of efficiency and competition, scale economy and PCS indicator, to find out whether they match or differ.

Earlier research on the life insurance market has revealed that the efficiency of life insurers tends to to be poor (Cummins and Weiss, 2012) and that competition between insurers is not strong (Bikker and Van Leuvensteijn, 2008). Unused scale economies point to weak competition: stronger pressure on cost efficiency would lead to further consolidation. The four abovementioned problems facing the Dutch life insurance sector all have in common that cost efficiency would help (i) to maintain market shares in the competition struggle against banks, and (ii) either to restore profitability and impaired buffers, or to reduce the hidden costs in unit-linked products (or both).

Life insurance firms sell several different products through various distribution channels, thereby creating several submarkets. The degree of efficiency and competition varies across these submarkets. For instance, the submarket where parties negotiate collective contracts (mainly employer-provided pension schemes) is expected to be more competitive than the submarkets for individual policy holders. Our data sets allow the subdivision of insurance policies into collective and individual contracts and, for each submarket, a split into unit-linked policies (where investment results are for the risk of policyholders) and policies guaranteeing benefit payouts in euro. Collective unit-linked policies consist mainly of annuities where individual policies in euro (or fixed benefit policies) usually take the form of endowment policies. Therefore, the two approaches measuring scale economies and competition will be estimates for the four submarkets too. Data on submarkets enables us also to further investigate the structure of the life market: do insurers, over time, go for specialization or do they, on the other hand, tend to sell all types of life insurance products in order to take full advantage of scope economies? Furthermore, we pay attention to developments in efficiency and competition over time, in order to assess how insurers have responded to the challenges mentioned earlier. We relate this to the structure of the market and to entry and exit of insurers.

This paper is structured as follows. Section two presents a short review of the production of life insurances, followed by a survey on the literature on efficiency and competition in the insurance industry. Section four provides a theoretical background to the measures of efficiency and competition used, while the next section highlights stylized facts of the Dutch insurance markets and its developments over time. Sections six and seven present the empirical findings for, respectively, scale economies and the PCS indicator of competition, and compare them with the results of the other studies in the field. The last section summarizes and concludes.

2. The production of life insurances

Life insurance covers deviations in the timing and size of predetermined cash flows due to (non-) accidental death or disability. While some life insurance products pay out only in the incident of death (level term insurance and burial funds), others do so either in the incident of death or at the end of a specified term or number of terms in the absence of death (endowment insurance). A typical annuity policy pays an annual amount starting from a given date (if a specific person is still alive) and continues to do so until that person passes away. The nominal amount of the benefits can be guaranteed upfront so that the insurance firm bears the risk that invested premiums may not cover the promised payments. Such guaranteed benefits may be accompanied by some kind of profit sharing, *e.g.* depending on indices of bonds or shares. Alternatively, the insurance benefits may be linked to capital market investments, *e.g.* a basket of shares, so that the insurance firm bears no investment risk at all. Such policies are usually referred to as unit-linked funds. We also observe mixed products, *e.g.* unit linked funds with guaranteed minimum investment returns.

A major feature of life insurance is its long-term character, often continuing for decades. In the Netherlands, the average modified duration of insurers in 2010 was 12 years, which, in a going concern insurance firm, points to long-term policies.[4] This makes policy holders vulnerable to the viability, reliability and efficiency of the insurer, and insurers sensitive to their reputation. Life insurers need large provisions to cover their calculated insurance liabilities. These provisions are financed by – annual or single – insurance premiums and invested mainly on the capital market. The major risk of life insurers concerns mismatches between liabilities and assets. Idiosyncratic life risk is negligible as it can be adequately diversified. Systematic life risk, however, such as increasing life expectancy, can also pose a threat to life insurers, depending on their policy portfolio. Note that the risk of annuities increases with longevity, whereas the risk of term insurance and endowment policies decreases with longevity. Hence, the dominant risk will generally be investment risk. The main services which life insurance firms provide to their customers are life (and disability) risk pooling and financial intermediation. Significant expenditures include sales expenses, whether in the form of direct sales costs or agency, administrative costs, asset management and product development

3. Literature on performance in the life insurance industry

In the literature, direct measurements of competition on the life insurance market are virtually absent. Bikker and Van Leuvensteijn (2008) use the PCS indicator (earlier referred to as the Boone index) to measure competition in the Dutch life insurance sector. They conclude that life insurance competition is weaker than competition in the industrial markets. It has long been common practice to measure efficiency in the life insurance industry on the basis of economies of scale. Using a Flexible Fourier (FF) function, Fenn *et al*. (2008) examine 14 major European countries over 1995-2001 and find large unused economies of scale ranging from 40% for the smallest life insurers to 10% for larger firms. Cummins and Rubio-Misas (2006) focus on Spain and find scale inefficiencies ranging from 20% for small firms via 5% for larger firms to 12% for the largest firms. The latter increase is due to diseconomies of scale for the largest 60% of firms. Strong consolidation, on average, brings decreased scale inefficiencies over time. Using a Translog Cost Function (TCF), Kasman and Turgutlu (2009) observe scale economies in Turkey over 2000-2004 of, on average, 30%. All three studies present higher scale inefficiency for small insurers, lower inefficiency for medium-sized and larger firms, while the largest companies show again more scale inefficiency, pointing to a certain optimal scale. This optimal scale has not been found by Bikker and Van Leuvensteijn (2008) for the Netherlands, but their range of economies of scale from small insurers (30%) to large firms (10%) is similar. Grace and Timme (1992) and Yuengert (1993) have found also increasing returns to scale that most US life insurers, but the latter observes constant returns to scale for the largest ones. Finally, Fecher *et al.* (1991) report that French insurers have unused scale economies of 15%.

[4] Modified duration is the procentual portfolio value change due to a parallel shift in the interest rate term structure and is an estimate of the weighted average term of future cash flows.

The literature on cost efficiency of life insurers is large, as becomes clear from *e.g.* the thorough overview of Cummins and Weiss (2012, Table 6). Remarkable is the huge spread in cost (and profit) inefficiencies, due to the variation across countries and sample periods, but also to the different parametric and non-parametric measurement approaches of inefficiency and varying definitions (*e.g.* allocative versus technical inefficiency). A general problem is that inefficiency cannot always be precisely distinguished from model and measurement errors, particularly where production functions are approximative only, due to data limitations. Typically half of the studies into cost efficiency in the life sector find that up to 50% of costs could be avoided by applying best practices,[5] while the other half state even higher percentages.[6]

Other studies focus on the impact of organizational form on cost performance. Since utility-maximizing managers have a preference to spend more on salaries, staff, office furniture and other perquisites, mechanisms are needed to control managerial opportunism. Agency theory hypothesizes that stock ownership can prove more effective to control owner-manager conflicts than mutuals, the so-called expense preference hypothesis (EPH), see Mester (1991). Erhemjamts and Leverty (2010) find evidence for the EPH in the US life insurance sector. Efficiency, taxes and access to capital explain the increasing share of stock insurers. Cummins, Rubio-Misas and Zi (2004) do not find support for the EPH. Stock insurers and mutuals have different production functions, so that comparative cost advantages determine the dominance of each type on the various submarkets, in line with the efficient structure hypothesis. Many studies do not find evidence for the EHP, *e.g.* Bikker and Gorter (2011) for the Netherlands, Gardner and Grace (1993) and Cummins and Zi (1998) for the US, and Fukuyama (1997) for Japan, while for France, Fecher *et al.* (1993) find higher efficiency for stock life insurers than for mutuals.

4. The measurement of inefficiency and competition

We will estimate scale inefficiency to obtain a measure of inefficiency. At the same time we interpret the existence of unused scale economies as lack of competitive pressure to push down costs. Furthermore, we will measure competition directly using the PCS indicator.

[5] See Pottier (2011); Jeng *et al.* (2007), both applying Data Envelop Analysis (DEA) for the US, Kasman and Turgutlu (2009; Turkey); Fenn *et al.* (2008, EU); Bikker and Van Leuvensteijn (2008, the Netherlands); Klumpes (2004, UK), all applying Stochastic Cost Frontier Analysis (SCFA).

[6] See Mahlberg and Url (2010, Germany), Cummins *et al.* (2010, US); Cummins and Rubio-Misas (2006, Spain), Cummins *et al.* (2004, Spain); Erhemjamst and Leverty (2010, US); all applying DEA, Elling and Luhnen (2010), applying SCFA.

4.1. Economies of scale

The intuition behind using economies of scale to measure inefficiency is that a highly competitive market is expected to force life insurance companies to improve their efficiency in order to be able to survive and gain sustainable profit. As a competitive market forces firms to be efficient, scale economies may be used as an indicator of competition intensity. The existence of unused economies of scale across size classes also indicates what level of consolidation would be optimal from an efficiency point of view (see Bikker and Van Leuvensteijn, 2008). The translog cost function (TCF), introduced by Christensen, Jorgenson and Lau (1973), has long been used extensively to measure economies of scale and is regarded as one of the most effective models. The TCF is based on a U-shaped unit cost function that reflects all the underlying assumptions of the general cost minimization model.[7] It places no prior restrictions on the substitution possibilities among production factors, and thus permits economies of scale to vary with output and input proportions. The TCF is flexible and less complex than alternatives such as the Fourier function. In light of this, the TCF will be used in order to develop an accurate estimate of economies of scale in the Dutch insurance industry. The same TCF will also be used to estimate the marginal costs that will be incorporated as a proxy of efficiency in the PCS model of competition.

A classical expression of a cost function is $C = f(Y, P, t)$, where *Y* is the output vector, *C* total cost, *P* a vector of input prices, and *t* time. Such a cost function is a dual of the related production function. In our empirical application we do not have insurer-specific input prices available,[8] so that the TCF simplifies into:

$$\ln OC_{st} = \beta_0 + \sum_{i=1}^{N} \beta_{Yi} (\ln Y_{ist} - \ln Y_{i\bullet\bullet}) + \frac{1}{2} \sum_{i=1}^{N} \sum_{k=1}^{N} \beta_{Yik} (\ln Y_{ist} - \ln Y_{i\bullet\bullet})(\ln Y_{kst} - \ln Y_{k\bullet\bullet}) + \sum_{j=1}^{L} \gamma_s X_{jst} + \varepsilon_{st} \quad (1)$$

OC_{st} is operational cost of firm *s* in year *t*,[9] Y_{ist} is output of type *i* (i=1,…,N), and X_{jst} represents either insurance firm-specific control variables, or market-wide variables which only change over time (j=1,…,L). Operational costs and outputs are expressed in logarithms. ε_{st} is the random error term. The model contains squares and cross-terms of output components. All output terms (in logarithms) are expressed as deviations from their averages (in logarithms),[10] calculated over all insurer-year

[7] The assumptions are: input demand is downward sloping, cross price effects are symmetric, the shift in marginal cost with respect to an input price is equal to the shift in the input's demand with respect to output, the sum of own and cross price elasticities is equal to zero, and a proportional increase in all input prices must shift cost by the same amount, holding output constant.
[8] But we do have a wage index of the financial industry to explain developments over time.
[9] Operating costs are the sum of management costs and acquisition costs. We do not include investment costs, which may in our dataset be unreliable, see Section 5.
[10] Note that $\ln Y_{k\bullet\bullet}$ is the logarithm of the geometric, not the arithmetic, average of the insurers output measure *k*.

combinations, *cf.* the Taylor series expansion.[11] The average for output type *i* is denoted as $\ln Y_{i\bullet\bullet}$, which dots for the sub-indices over time and across insurance firms. The variables expressed as deviations from their averages help to split linear and quadratic effects of output on costs and simplify the interpretation of the coefficients, as explained below.

Usually, scale economies are defined as the relative increase in output level resulting from a proportional increase in factor inputs. However, Hanoch (1975) has argued that looking at economies of scale as an association between output level and total costs – where costs are minimized for every output level and input prices are constant – is the most appropriate method. Hence, scale economies can be derived from a proportional increase in total costs resulting from a proportional increase in the output level, that is, the elasticity of total costs with respect to output level. In light of this, the ray economies of scale (*EoS*) for firm *s* in year *t* can be defined as a unit minus this cost elasticity:

$$EoS_{st} = 1 - \sum_{i=1}^{N} \frac{\partial \ln OC_{st}}{\partial \ln Y_{ist}} = 1 - \sum_{i=1}^{N} (\beta_{Yi} + \frac{1}{2} \sum_{k=1}^{N} \beta_{Yik} (\ln Y_{kst} - \ln Y_{k\bullet\bullet})) \qquad (2)$$

A positive *EoS* value refers to economies of scale whereas a negative *EoS* indicates diseconomies of scale. *EoS* equal to 1 reflects Constant Returns to Scale (CRS). As *EoS* depends on variables that are expressed in natural logarithm form, its interpretation will be in percentage terms. The *EoS* for the (geometric) average of all life insurers is $1 - \sum_i \beta_{Yi}$, hence based on the linear output elasticities only. This is an example of easier interpretation where output is expressed in deviation from its (geometric) sample average. Similarly, the *EoS* can be calculated for, for instance, the (geometric) mean of the 10 percent smallest insurance firms, or the 10 percent largest ones.

The PCS indicator as measure of competition will be based on the relationship between profits or market shares, on the one hand, and marginal costs on the other. Since marginal costs are unobserved, they need to be estimated. We base them on the translog cost function of Equation (1). The marginal costs of producing output *i* for firm *s* in year *t* are defined as:

$$MC_{sit} = (\frac{\partial \ln OC_{st}}{\partial \ln Y_{ist}})(OC_{\mathrm{st}} / Y_{ist}) = (\beta_{Yi} + \sum_{k=1}^{N} \beta_{Yik} (\ln Y_{kst} - \ln Y_{k\bullet\bullet}))(OC_{\mathrm{st}} / Y_{ist}) \qquad (3)$$

The average MC over the entire sample or average MC numbers of size classes follow from Eq. (3) and are calculated as the sum over all output types (that is, over *k*) of the averages over *s* or subsets of *s*, respectively, of the product of (i) the cost elasticity, (ii) the output in deviation of its overall average, and (iii) the respective cost-output share.

[11] White (1980) and Shaffer (1998, p. 95) explain that this specification also helps to avoid multicollinearity.

4.1.1. Definition of output

In the life insurance sector, output is intangible. Many efficiency studies choose premiums as output proxy (Cummins and Weiss, 2012, Table 4). Yuengert (1993) has criticized this as premiums represent price times output quantity, not output itself. Systematic price differences across large and small firms may lead to misleading inferences about average costs, if premiums are used as output proxy. Furthermore, premiums ignore investment performance. Following Yuengert (1993) and Berger *et al.* (2000) we use 'incurred benefits plus additions to provisions' as measure of insurance output. Insurance provisions of firm *s* in year *t* develop as follows (ignoring non-recurrent items):[12]

$$\text{Provisions}_{s,t+1} = \text{Provisions}_{st} + \text{Net premiums}_{st} - \text{Cost}_{st} - \text{Profit BT}_{st} + \text{Net inv. inc}_{st} - \text{Benefits}_{st} \qquad (4)$$

Costs are operating costs and profits are before taxes. Additions to provisions ($\text{Provisions}_{s,t+1}$ – Provisions_{st}) plus 'incurred benefits' are equal to:

$$\text{Net premiums}_{st} - \text{Costs}_{st} - \text{Profit BT}_{st} + \text{Net inv. inc}_{st} \qquad (5)$$

This term can be split into (i) the prime price of the insurance policy, that is, new production (Net premiums – Costs – Profit BT), which represents services to (new) clients, and (ii) investment services, that is, the annual return on the invested funds (Net investment income), which describes services to existing clients. Note that the key function of the insurance, risk bearing or risk pooling is a service provided 'free' by the insurer, which does not count towards costs: losses and gains on life policies cancel out for the insurance firm, unless the decease pattern of their clients deviates from the used death tables, for instance where longevity risk has been underestimated.

Eq. (5) shows that premiums have been corrected for the profit and cost margins, so that potentially distorting systematic differences in costs and profits across large and small firms are excluded. On the other hand, costs may reflect deadweight (*e.g.* scale inefficiency), but they are also related to administrative and communicational services (providing policies and advice, *et cetera*). Similarly, profits may consist of excess profits, but also of a risk premium for stockholders bearing unexpected risk. We will use premiums as an alternative output measure, with and without 'net investment income', as a robustness test and for comparison with the literature. Cummins and Weiss (2012) recommend splitting insurance output into individual and group policies and, next, into annuities and life insurances,[13] as each of these four categories have their own properties. We have followed this

[12] This is based on Thiele's differential equation (Jørgensen, 1913). This formula holds as long as the discount rate to calculate future liabilities is constant, as until recently has been the case in the Netherlands. When an insurer changes the discount rate, or another basis principle of provisioning, Eq. (4) is extended with a 'non-recurrent item'.

[13] In our data set this second split is replaced by a split into unit-linked policies and policies in euro.

alternative in one of the robustness tests. Alternatively, we have also applied the TCF model of Eq. (1) to each of these four categories separately.

Insurance output, and its components benefits and investment income, are expressed in (euro) amounts and ignore the granularity of the output: each policy is a contract that provides a service to a client. Note that when an insurer, with a given output, has more or, on average, smaller policies, this implies providing more client services, which go hand in hand with higher operating costs. We include this granularity dimension of output by adding the number of policies as a separate, second output measure.

4.1.2. The translog cost model

Besides output terms, the translog cost model (1) contains control variables, which have impact on operational costs and help to refine scale economies measurement. The model we will estimate reads as follows:

$$\ln OC_{st} = \beta_0 + \sum_{i=1}^{N} \beta_{Yi} \tilde{y}_{ist} + \frac{1}{2} \sum_{i=1}^{N} \sum_{k=1}^{N} \beta_{Yik} \tilde{y}_{ist} \tilde{y}_{kst} + \gamma_1 Stock_{st} + \gamma_2 Ac/GP_{st} + \gamma_3 CPP_{st}$$
$$+\gamma_4 LSP/GP_{st} + \gamma_5 ULP/GP_{st} + \gamma_6 HHI_t + \gamma_7 Wage_t + \gamma_8 Time_t + \varepsilon_{st} \qquad (6)$$

$\tilde{y}$ denotes the logarithm of an output component in deviation of its average. An important issue, discussed extensively in the literature, is the effect of the organizational form on performance (*cf.* Cummins, Weiss and Zi, 1999). A prominent hypothesis in this context is the ‘expense preference hypothesis of organizational form’, which is derived from agency theory. This hypothesis predicts that mutual insurers will have higher costs than stock-based insurers, because the stock market imposes a more effective mechanism for corporate control and reduces excessive consumption of perquisites by managers and possible deviation from profit maximization principles (Mester, 1989). We examine the effect of organizational form on scale efficiency by simply adding a dummy variable (*Stock*, which is 1 for stock firms) to allow for different cost levels between stock and mutual companies.

Operating costs consist of managerial costs and acquisition costs. Insurers have different distribution strategies, which cause huge variation in the acquisition costs margin across insurance firms. We express the acquisition costs as share of gross premiums (Acq_{st}/GP_{st}). Similarly, we expect that insurers with relatively more collective policy premiums (CPP_{st}/GP_{st}) and those with lump sum premiums (LSP_{st}/GP_{st}) have, on average, lower costs. The share of unit-linked premiums (ULP_{st}/GP_{st}) may also affect the cost level. The final three control variables describe the entire life insurance market, which show variation over time only. The Herfindahl-Hirschman Index[14] (HHI), based on premiums, measures concentration of insurance firms, and increases over time. Concentration may be the consequence of strong competition (negative coefficient) or may enable tacit collusion so that less pressure exists to cut cost (positive effect). The real wage level is an input price (*Wage*). As the firm-specific 'wage rate' is not observed, the general wage index of the financial industry, deflated by the cost-of-living price index, may at least pick up cost changes over time. Finally, 'time' presents technical progress (*Time*) and is expected to have a negative effect on costs.

4.2. The PCS model of competition

The PCS model assumes that more efficient firms (that is, firms with lower marginal costs) will gain higher market shares or profits, and that this effect will be stronger with competition. In order to support this intuitive market characteristic, Boone develops a broad set of theoretical models (see Boone, 2000, 2001, 2004 and 2008, Boone *et al*., 2004, and CPB, 2000). We use one of Boone's theoretical models to explain the PCS model. Following Boone *et al*. (2004), and replacing 'firms' by 'insurers', we consider an insurance industry where each insurer s produces one product q_s (or portfolio of insurance products), which faces a demand curve of the form:

$$p\ (q_s,\ q_{r\neq s}) = a - b\ q_s - d \sum_{r\neq s} q_r \tag{7}$$

and has constant marginal costs mc_s. This insurer maximizes profits $\pi_s = (p_s - mc_s)\ q_s$ by choosing the optimal output level q_s. We assume that $a > mc_s$ and $0 < d \leq b$. The first-order condition for a Cournot-Nash equilibrium can then be written as:

$$a - 2\ b\ q_s - d \sum_{r\neq s} q_r - mc_s = 0 \tag{8}$$

Where N insurers produce positive output levels, we can solve the N first-order conditions (8), yielding:

$$q_s\ (c_s) = [(2\ b/d - 1)\ a - (2\ b/d + N - 1)\ mc_{is} + \sum_r mc_r]/[(2\ b + d\ (N - 1))(2\ b/d - 1)] \tag{9}$$

We define profits π_s as variable profits excluding entry costs ε. Hence in equilibrium, an insurer enters the insurance industry if, and only if $\pi_s \geq \varepsilon$. Note that Eq. (9) provides a relationship between output and marginal costs. It follows from $\pi_s = (p_s - mc_s)\ q_s$ that profits depend on marginal costs in a quadratic way. Competition in this market increases as the (portfolios of) services produced by the various insurers become closer substitutes, that is, as d increases (with d kept below b). Furthermore, competition increases when entry costs ε decline. Boone *et al*. (2004) prove that market shares of more

[14] HHI is defined as $\sum_{s=1}^{n} ms_s^2$ where ms_s represents the market share of firm s.

efficient insurers (with lower marginal costs *mc*) increase both under regimes of stronger substitution and amid lower entry costs.

4.2.1. The empirical PCS model

Eq. (9) supports the following regression model for market share, defined as $ms_s = q_s / \sum_r q_r$:

$$ms_{st} = \alpha + \beta_t\, mc_{st} + \sum_{t=1,..,(T-1)} \gamma_t\, d_t + u_{st} \qquad (10)$$

where α, β_t and γ_t are parameters, ms_{st} denotes the market share of insurer *s* in year *t*, mc_{st} stands for the marginal costs of the respective insurer, d_t is a time dummy and u_{st} an error term. The parameter of interest, β_t, is expected to have a negative sign, because relatively efficient insurers will gain higher market shares. Eq. (10) may also be specified in log-linear terms in order to deal with heteroskedasticity, which is mainly an empirical issue, that can be investigated with the Box-Cox test (Liu and Hay, 1997, p. 608). Moreover, this specification implies that β_t is an elasticity, which facilitates its interpretation, particularly across industries or countries.[15] We will refer to β_t as the PCS indicator in year *t*. Boone shows that where differences in performance in terms of market shares are increasingly determined by marginal cost differences, this indicates increased competition. The PCS indicator requires data on fairly homogeneous products.

Marginal costs are not observed but can be derived, using the translog cost function, from Eq. (3), as we will do. An alternative, namely average costs, would ignore the distinction between fixed and variable costs, but appears to be a quite useful approximation in practice (Bikker and Van Leuvensteijn, 2008). The competition coefficient β is negative in the case of effective competition and ranges from 0 (no competition) to $-\infty$ (extreme competition). Increases or decreases, in absolute terms, in β, for instance, over time, can be interpreted. With due reservation, the life insurance β can be compared across industrial sectors or across countries. Note that an efficient insurer can use its cost advantage to gain a higher market share through setting the output price below the market price (fitting well with Eq. (10)), to gain a higher profit margin through maintaining its market price, or to pass-through a portion of its efficiency-gains to its customers. In all these cases market shares in Eq. (10) can be replaced by profit, which is the product of profit margin and market share. Finally, efficiency-gains can also translate into innovation attractiveness, improved design or quality.

[15] The few existing empirical studies based on the PCS model have all used a log linear relationship. See, for example, Bikker and van Leuvensteijn (2008).

5. The structure of the Dutch life insurance industry

We explain the structure of the Dutch life insurance industry using the key data presented in Table 5.1. Total assets entrusted to life insurers, expressed in prices of 2010, continuously increase over time, but at a gradually lower pace. The volume of premiums did increase around the turn of the century, but fell slightly during the crisis years. The latter indicates a decline in the sale of new life policies. When total assets are shown as a percentage of GDP, we observe that since 2002, life insurance funds do no longer grow in line with GDP. Other developments such as heavily increased competition from banks and declining trust caused by the investment policy scandal, also contributed to the fall in new production. Net investment income is a major source of funding for future benefits, exceeding premiums by a factor 2 to 3. Investments can be split into the part related to unit-linked policies, where policyholders bear the risk, and the part related to policies with benefits expressed in euro, where

Table 5.1. Key data of the Dutch life insurance market over time (averages)

	1995-1998	**1999-2002**	**2003-2006**	**2007-2010**
	In billions of euro, at 2010 prices			
Total assets (abbreviated to ta)	232.7	289.3	304.4	325.3
Insurance provisions (ip)	177.0	223.3	244.6	264.5
Of which: for own account	126.8	144.6	151.7	158.2
Gross premiums (gp)	22.1	28.4	27.4	25.3
	Percentages			
Total assets/GDP	52.1	56.4	55.8	55.2
Net premiums/gp [a]	94.0	96.4	97.3	96.0
Net investment income/gp	68.7	40.6	61.5	42.6
Net investment income/ta	6.5	4.0	5.5	3.3
Benefit payments/gp	44.0	60.7	71.1	89.9
Addition to ir/gp	72.7	41.0	49.0	27.4
Insurance output/gp [b]	116.7	101.7	120.1	117.2
Profits before taxes/gp	13.1	12.7	13.3	13.0
Equity/ta	11.4	9.7	8.8	7.1
Operating costs/gp [c]	13.1	12.7	13.3	13.0
Marginal costs/gp	12.3	12.0	12.9	12.9
Acquisition costs/gp	6.9	6.0	6.5	6.4
Management costs/gp	6.5	6.8	7.0	7.3
Reinsurance costs/gp	6.0	3.6	2.7	4.0
Collective premiums/gp	31.0	31.7	31.8	34.9
Lump-sum premiums/gp	40.1	45.4	46.5	44.1
Unit-linked premiums/gp	30.2	44.5	43.3	40.7
Insurance provisions endowments/(ip)	63.6	64.0	64.0	64.7
Stock firms/all firms	88.7	90.5	90.0	91.7
	Natural units			
Concentration index (HHI)	7.9	8.0	10.0	14.9
Number of firms	101.5	94.3	77.8	60.0
Number of policies, in millions	34.7	39.2	39.8	38.5

[a] The ratio 'net premiums/gp' is calculated as the sum of net premiums of all firms divided by the gross premiums of all firms. This is equal to the weighted average of individual observations of this ratio, weighted by the firm size in terms of gross premiums. This holds also, *mutatis mutandis*, for the other ratios; [b] Insurance output is the sum of Benefit payments and Addition to insurance provisions; [c] Operating costs is the sum of marginal and acquisition costs minus 'provisions and profit sharing from reinsurers'. For our calculations, we ignored the latter item which is zero or negligible for most insurers but substantial for some.

insurers bear the risk.[16] Net investment income depends, of course, on market conditions. During the dotcom bubble crash (2002), these returns were – on an aggregated level – almost zero, and during the credit crisis (2008), they were even firmly negative. The steady decline in nominal interest rates over time is also important here. Benefit payments increase strongly over time, reflecting increasing numbers of maturing policies. 'Addition to the insurance provisions' makes clear that new production and extension of the coverage of existing policies still exceeds benefit payments albeit by declining amounts. This is in line with the slowdown in growth of the insurance provisions presented in the top of Table 5.1. As a result, insurance output, the sum of incurred benefits and additions to the provisions, is fairly constant over time.

Profits are remarkably stable when presented as four-year averages, but fluctuate strongly within these periods with a close to zero profit in 2002 and a strong loss in 2008. Note that the profit on a life policy becomes available gradually, year after year over the life span of the contract, so that actual profit reflect the result over all the existing policies rather than the expected profit on the new production. In the (hypothetical) case of declining excess profit on new production due to increasing competition, this would show up only gradually in the financial reporting figures, over a long time span. The equity ratio declined due, in part, to stock market losses, reflecting that insurers nowadays no longer have the large buffers of earlier years. Operating costs as a percentage of premiums are stable over time at 13%. Remarkably, increased competition with banks and the criticism on the high hidden costs have so far not led to a reduction in costs.

Management costs tend to rise over time. Note that the profit and cost margins together take away at least one quarter of the new savings under life policies. On top of this, there are the investment costs, varying from one-third to twice the operating costs. During 2008, the investment costs peaked at € 8 billion against investment losses of € 5 billion (in percentages of gross premiums: 30.4% against, for instance, 3.3% in 2005); the full sample average is 8%. We have doubts as to whether investment costs and losses are always separated correctly. We do not analyse investment costs in the same way as operating costs: where operating costs are a pure overhead, investment costs may lead to higher expected returns.[17]

An important feature of the market structure is the number of firms. According to the literature, the larger the number of firms (and the lower the entry/exit barriers in the market), the more competitive the life insurance market will be. The number of insurers increased from 95 in 1995 to 105 in 1998,

[16] Investments and investment returns in Table 5.1 are based on the sum of both types of investments.
[17] Complex assets classes generally require higher research and risk management costs but have higher expected returns. Although scale economies may be expected in investment activities, large insurers may typically move earlier to complex assets classes. This behaviour have been observed for pension funds, see De Dreu and Bikker (forthcoming).

but afterwards declined strongly to 48 in 2010. The decline to 48 insurers is the net effect of 32 new entries and 80 exits, indicating fairly strong dynamics in this market, particularly in the earlier years. The concentration index HHI, based on premiums, raises from 7 in 1998 to 21 in 2010. On the HHI scale running from 0 to 10,000, the index values remain low (and the number of insurers is still high), particularly compared to the banking market, explaining that consolidation on this market is still very low. Competition may come under threat when concentration increases, because – as the theory explains – tacit collusion can then be achieved more easily. But the decrease may also be due to mergers and acquisitions under competitive pressure. Given the existence of substantial unused scale economies (Bikker and Van Leuvensteijn, 2008), it is remarkable that operating costs do not decline over time as a consequence of up scaling. The number of mutual firms declines more rapidly than the number of stock firms. In the bottom part of Table 5.1, we split the life policy market into subdivisions (collective versus individual policies, lump-sum versus periodical payments, unit-linked policies versus fixed benefits in euro policies, and endowment policies versus annuities), based on their shares in premiums and insurance provisions; we do not observe substantial changes over time. An exception is the unit-linked market which increases during the earlier years, but falls back more recently.

5.1 The life submarkets

Table 5.2 presents key data on two submarkets, policies in euro (or fixed-benefit policies) and unit-linked policies, each split further into individual and collective policies, for two subperiods: 1995-2002 and 2003-2010. Expressed in either premiums or insurance output, we observe that fixed-benefit policies take up 60% of the life market, while unit-linked products have a share of 40%. Similarly, the figures show that the market share of individual policies, at 60%, is larger than that of collective contracts. The investment income on unit-linked policies of individuals is rather low, particularly in the initial period. Benefit payouts are relatively large for individual fixed-benefit policies and relative small for individual unit-linked policies, reflecting, on average, more mature individual fixed-benefit policies. While the profit margin on unit-linked policies is very minor, the profit margin on fixed-benefit policies is relatively high, during the first period especially on collective policies, and duringthe second subperiod particularly on individual policies Operating cost margins are much lower for collective policies than for individual ones, reflecting another element of scale economies. This also holds for costs expressed as a percentage of premiums (lower in the subperiod panels of Table 5.2), but much more strongly for unit-linked business than for fixed-benefit policies.

Information about the structure of the market is given by the number of insurers. It is clear that not all insurers are active on all submarkets. Individuals hold mainly endowment policies, often linked to mortgage loans, while collective contracts relate often to pension schemes provided by employers.

Table 5.2. Key data of the Dutch life insurance submarkets over time (averages)

	Fixed-benefit policies		Unit-linked policies	
	Individually	*Collective*	*Individually*	*Collective*
	1995-2002			
	Shares of submarkets in percentages			
Premiums, gross (gp)	43.9	16.9	24.2	15.1
Net investment income	44.2	32.9	2.4	20.5
Benefit payouts	47.3	23.4	6.8	22.5
Addition to insurance provisions	39.9	20.6	23.6	15.9
Insurance output [a]	42.9	21.8	16.6	18.6
Profit	49.4	44.8	1.1	4.7
Operating costs	51.3	13.0	30.7	5.0
Insured future endowment benefits	62.7	10.1	23.9	3.3
Insured future annuity benefits	12.5	50.4	1.1	36.0
	Shares of lump sums in percentages			
Premiums, gross	50	47	31	52
	Annual average			
Operating costs as % of gp	15	10	16	4
Number of firms	88	43	65	20
HHI	10.2	35.5	7.8	104.3
	2003-2010			
	Shares of submarkets in percentages			
Premiums, gross (gp)	41.7	15.7	24.5	18.1
Net investment income	39.5	26.4	14.6	19.6
Benefit payouts	51.0	18.9	14.1	16.1
Addition to insurance provisions	23.0	20.1	30.9	26.0
Insurance output [a]	40.1	19.4	20.6	19.9
Profit	66.0	25.5	5.3	3.2
Operating costs	42.0	14.7	33.2	10.0
Insured future endowment benefits	55.3	8.6	30.5	5.7
Insured future annuity benefits	13.1	48.2	1.4	37.2
	Shares of lump sums in percentages			
Premiums, gross	59	55	17	48
	Annual average			
Operating costs as % of gp	13	12	18	7
Number of firms	60	27	46	18
HHI	15.6	68.2	17.1	108.0

[a] Insurance output is the sum of Benefit payouts and Addition to insurance provisions.

This holds for fixed-benefit policies, but even more strongly for unit-linked policies.[18] Concentration is much stronger on collective contract markets and strongest in the unit-linked part of that market. Large contracts often concern pension benefits for employees of a company and are negotiated between experts at both ends of the table, in sharp contrast to individual contracts with uninformed private persons. Furthermore, they are typically renegotiated every five year. Relatively few insurers are active on this more demanding submarket of collective contracts and much more competition is expected here.

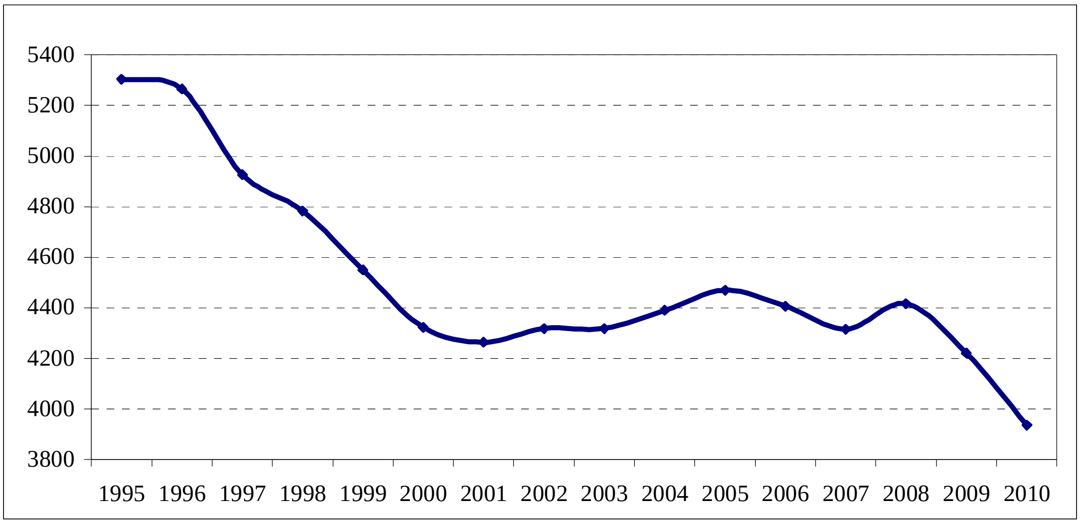
[18] Table 5.2 presents shares of submarkets. Note that the volumes of the annuity markets (in terms of 'insured future benefits') are calculated using the following rule of thumb: multiplying annual rents by 10 as a proxy for the average length of periodical benefits.

Fig. 5.1 presents a weighed HHI of individual insurers' market shares on submarkets which describes the focus of insurers on specialization over time (as in Bikker and Gorter, 2011).

$$\mathrm{HHI}_{\mathrm{t}}^{\mathrm{w}} = \sum_{i=s}^{S} w_{st} \sum_{j=1}^{J} ms_{stj}^{2} \tag{11}$$

with ms_{stj} the premiums of submarket *j* as a share of all gross premiums of insurer *s* (where *ms* refers to market share), while the weight w_{st} is the total premium of insurer *s* as share of all premiums in the life market, all in year *t*. The graph shows a decline in specialization over time, particularly in the earlier years and most recently. Apparently, life insurers tend to operate increasingly on all submarkets, which suggests that scope economies are more important than specialization benefits. This outcome contrasts with the greater focus and specialization Bikker and Gorter (2011) observe on the Dutch non-life market.

Fig. 5.1. Development in life insurance specialisation over time (HHI$^{\mathrm{w}}$)



6. Estimation results on scale inefficiencies

We estimate Eq. (6) using data over 1995-2010 to obtain a measure of the economies of scale (EoS) present in the Dutch life insurance industry. The first row of Table 6.1 presents results for insurance output (I.O.) as output size measure. The coefficient of insurance output, the cost elasticity, is 0.817, so that the average EoS effect is 18.3%, implying that a (small) increase in size would save almost one-fifth of the costs on the additional production. The positive coefficient of the squared insurance output term indicates that the EoS effect is concave, meaning that the effect is largest for the smallest insurance firms and decreases for larger firms, where at some point scale economies swift over to diseconomies in scale, that is, where the cost elasticity becomes larger than 1. We plot this in Fig. 6.1, where the cost elasticity has been calculated for 10 size classes (based on premiums), each with a number of around 120 annual observations of insurance firms.[19] Note that the graph would look fully concave if the insurance firms were not allocated to size classes, but were expressed in (the logarithm) of output size itself. Under constant returns to scale (CRS) the cost elasticity would be 1 and the coefficient of the squared term would be 0. A Wald test on this null hypothesis makes clear that the EoS effect is significantly different from zero, as in the CRS case (see last row in Table 6.1).

The coefficients of the other variables of Eq. (6) implying the following. Stock firms have significantly higher cost than mutual firms. This is in contrast to the 'expense preference hypothesis of organizational form', which predicts that mutuals will have higher costs than stocks, because the stock market imposes a more effective mechanism for corporate control and reduces excessive consumption of perquisites by managers and possible deviation from profit maximization principles. Our outcome is in line with most of the literature, see Section 3.

Table 6.1. Estimates of the translog cost function for Dutch life insurers (1995–2010)

	Uncensored data set				**Cost/premium rate censored at 75%**			
Output measures:	*Insurance output*		*I.O. & policies*		*Insurance output*		*I.O. & policies*	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Insurance output (I.O.), in logs[a]	0.817	67.4	0.474	26.5	0.848	78.7	0.580	33.5
Id. in squares[a]	0.023	8.6	0.003	0.5	0.028	9.6	0.008	1.2
Number of policies, in logs[b]			0.419	23.7			0.318	18.8
Id. in squares[b]			-0.014	-2.0			-0.021	-3.2
Cross term[a,b]			0.022	1.9			0.028	2.3
Stock firms	0.718	7.4	0.721	9.0	0.521	6.1	0.562	7.5
% of collective policy premiums	-0.620	-6.5	-0.222	-2.8	-0.326	-4.1	-0.114	-1.6
% of lump-sum premiums	-0.444	-5.3	0.520	6.5	-0.512	-7.3	0.246	3.4
% of unit-linked premiums	-0.390	-5.4	0.200	3.1	-0.202	-3.3	0.173	3.0
% of acquisition costs	2.604	19.0	1.937	16.8	6.381	32.5	4.737	25.0
Concentration (HHI)	-0.022	-1.7	0.001	0.1	-0.028	-2.6	-0.007	-0.7
Wage rate, in logs	1.961	1.8	0.443	0.5	1.384	1.5	0.400	0.5
Time trend	-0.012	-0.7	-0.016	-1.0	-0.005	-0.3	-0.011	-0.8
Intercept	1.951	0.6	6.419	2.2	3.872	1.3	6.785	2.7
Number of observations	1216		1196		1156		1137	
R^2, adjusted	85.5		90.4		88.9		91.7	
EoS spread [c]	41.0		21.7		43.9		25.8	
F test on CRS	235.1		54.6		200.1		60.4	

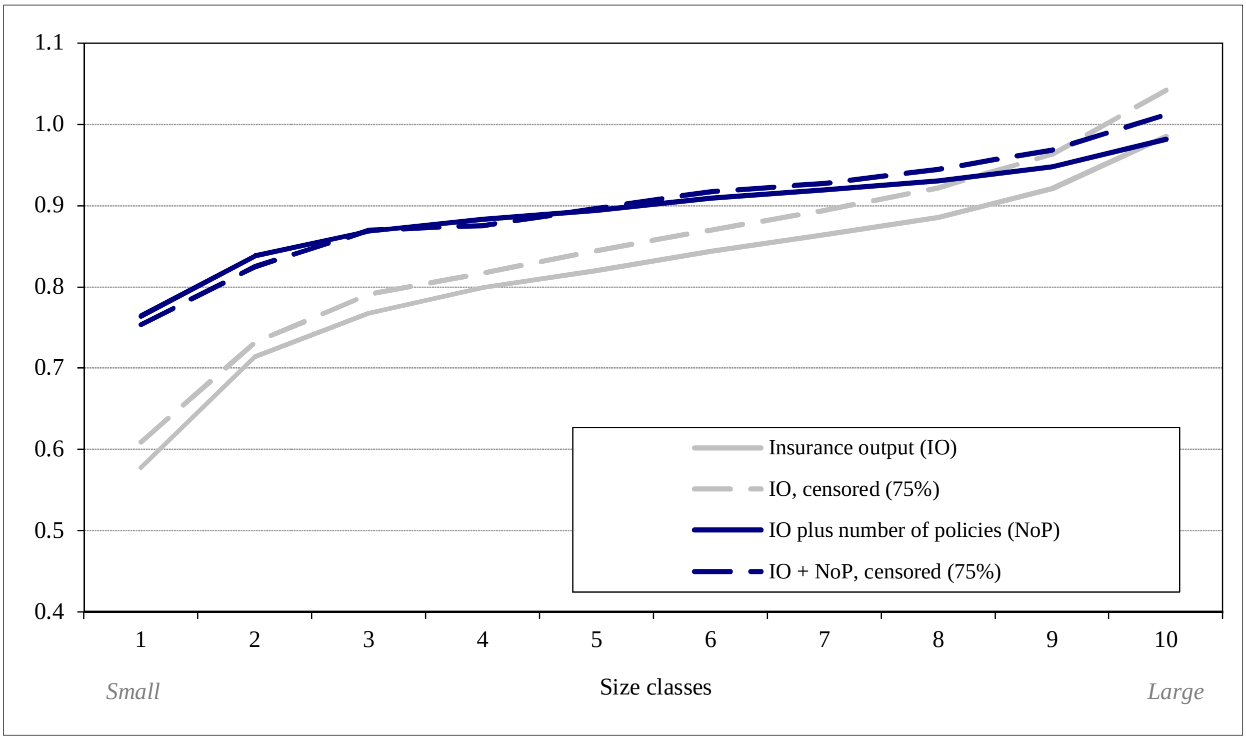
[a] Expressed as the deviation from the average of (logs of) insurance output across all insurer-year combinations, see Eq. (6), allowing for easier interpretation of the coefficients; [b]. Similarly for policies instead of insurance output; [c] EoS spread: difference in EoS between smallest en largest size class where 10 size classes are considered.

The coefficients of the other variables of Eq. (6) implying the following. Stock firms have significantly higher cost than mutual firms. This is in contrast to the 'expense preference hypothesis of

[19] Insurers in class 1 have premiums below 1.5 million euro. Maximum premiums in the other classes are, in million euro: 8, 19, 36, 59, 87, 135, 241, 581 and 4378.

organizational form', which predicts that mutuals will have higher costs than stocks, because the stock market imposes a more effective mechanism for corporate control and reduces excessive consumption of perquisites by managers and possible deviation from profit maximization principles. Our outcome is in line with most of the literature, see Section 3.

Fig. 6.1. Cost elasticities across ten insurer-size classes



Note: Based on the estimates of the four models in Table 6.1.

The acquisition cost share has a strong and significant impact on operational costs as well. Acquisition cost varies strongly across insurers, depending on their distribution model, while managerial costs are more stable. This explains the sensitivity of operational costs for the acquisition cost share. Various types of life insurance business go with lower cost than others: collective contracts are more cost efficient than individual ones, and lump sum policies have lower costs than periodical payment policies, both as expected. Unit-linked policies, with investment risk at the cost of policyholders, are more cost effective than policies in euro, with investment risk at the cost of the insurer. So far, all coefficients are highly significant. Concentration, measured with the HHI, has a negative effect on costs at the 10% level of significance. As we already control for scale economies, the HHI may reflect improved (X-) efficiency, or else its downward trend may pick up the effect of technical progress over time. Wage rate has a positive effect on cost, as expected, albeit at the 10% level of significance only. This coefficient reflects the impact of wage changes over time, as we do not have firm-specific wage information. The time trend coefficient, representing cost-saving technical progress over time, is negative but, again, not statistically significant.

The second column of Table 6.1 shows the estimates of an alternative model with two output variables. Apart from insurance output, which reflects the firm size in money terms, we also have the number of policies, which reflects the insurer's size in terms of administrative (and acquisition) activities, so that they complement each other. The estimates make clear that the average scale economies level – one minus the sum of the two linear output coefficients, see Eq. (2) – is at 0.107 somewhat smaller than in the case of the single-output measure, but still significantly different from CRS, see the test in the last row of the table. Fig. 6.1 confirms that we have again concavity: larger EoS for smaller banks and *vice versa*. The coefficients of number of stock firms and the share of acquisition costs are fairly unchanged, but the impact of type of life insurance policies differs from before, now we not only control for the monetary value of the output but also for the number of policies. Where the annual premium amount per policy is relatively high – as holds for collective contracts, lump-sum policies and, likely, unit-linked contracts – the respective coefficients shift downwards when controlled for the number of policies: the relatively lower costs are now also attributed to a relatively lower number of policies. The goodness of fit for models including number of policies is significantly higher, suggesting that operating costs are better explained by this model specification.

The data were selected by deleting negative or zero values for the key variables; operating costs, premiums, benefits, insurance provisions and number of policies.[20] We notice also that the cost/premium ratio has a non-representative high value for some insurer-year observations. This could point to insurance firms in their start-up or winding-down stages. In a sensitivity analysis, we delete all observations where this ratio is above 75%, a condition which holds for 5% of the sample. The advantage of such censoring is that we exclude 'non- representative' firm-year combinations, at although the risk of omitting observations with (large) positive error terms (reflecting high cost inefficiency, all else being equal). This risk would increase the lower we set the cost-premium threshold. Table 6.1, right-hand panel, and Fig. 6.1 indicate that scale economies are somewhat lower after the data censoring, as might be expected after deletion of large positive errors. As the results lead to the same conclusions, however, including concavity and rejection of CRS, our approach appears to be robust against censoring.

Table 6.2 builds on Eq. (6) but investigates alternative output measures, where the first two rows correspond to Table 6.1. The third row refers to a TCF model with – as output measure – the two components of insurance output taken separately instead of added up. This more general specification allows for different output effects on operating costs for benefits and additions, both in terms of EoS

[20] This concerns small numbers, *e.g.* 6 observations of gross premiums out of the total sample of 1334 observations have been excluded.

and concavity of scale economies. In the literature, premiums are often used as volume measures, although critics argue that this measure includes the profit margin. Premium income is also combined with another output indicators, such as number of policies, total assets and net investment income, where the latter two reflect services provided through the entrusted funds. Furthermore, each of these three output indicators also represents the output volume separately. Finally, we apply the model to stock-based insurers only, with the same results. The cost elasticities, and hence the scale economies, do not depend heavily on the chosen output measure and remain quite stable.

Starting with the uncensored results, we find that all model specifications have a scale economy that is significantly different from CRS. Ten out of twelve output measures give concave scale economies, two yield a linear relationship, indicating constant scale economies across size classes. It is clear that the approach is robust against the choice of output measure. A general outcome is that models based

Table 6.2. Cost elasticity estimates for various life insurance output measures

	Uncensored					**Censored at cost/ premium ratio of 75%**				
Output measure	*CE* [a]	*FF*	*# of obs.*	R^2	*Test*	*CE*	*FF*	*# of obs.*	R^2	*Test*
Insurance output	0.817	Ca	1216	85.5	234.8	0.848	Ca	1156	88.8	200.0
Idem & # of policies	0.893	Ca	1196	90.5	54.6	0.898	Ca	1137	91.8	60.4
Benefits & additions to provisions separately	0.803	Ca	1091	86.5	162.6	0.824	Ca	1054	90.6	215.6
Net premiums	0.847	Ca	1273	90.6	536.8	0.927	Ca	1168	93.7	79.9
Gross premiums	0.887	Ca	1293	90.0	290.8	0.940	Li	1220	93.3	24.1
Idem & # of policies	0.927	Ca	1268	91.4	46.8	0.957	Li	1199	93.9	11.3
Idem & total assets	0.919	Ca	1293	91.1	47.7	0.940	Li	1220	93.6	21.1
Idem & net inv. income	0.904	Ca	1167	92.8	97.9	0.937	Li	1098	93.6	22.8
Total assets	0.870	Li	1293	85.5	61.2	0.885	Li	1220	89.7	64.3
Net investment income	0.772	Ca	1167	84.2	233.6	0.771	Ca	1098	84.2	228.4
Number of policies	0.845	Li	1268	83.0	69.4	0.840	Cv	1199	82.2	60.0
Insurance output, stocks	0.818	Ca	1108	81.2	216.5	0.856	Ca	1068	87.0	208.4

[a] CE is short for cost elasticity, FF stands for functional form, Ca for concave and Li for linear, R^2 refers to the goodness of fit, adjusted for the used number of degrees of freedom, and Test refers to the Wald test on the CRS hypothesis.

on a combination of two indicators have a consistently higher cost elasticity (after summing) than single indicator models. The more volatile variable 'net investment income' has a lower cost elasticity, probably due to the errors-in-variable bias, which may indicate lower suitability. Continuing with the censoring cases, we observe as before that the cost elasticities generally are larger. Still, scale economies are again significantly different from CRS in all model specifications. Note that the concavity reduces under censoring, because – apart from non representative firm-year combinations – the smaller and relatively more scale inefficient insurers are typically excluded. This may affect the reliability of the censoring results. Cummins and Weiss (2012) recommend splitting insurance output into individual and group policies and, next, into annuities and life insurances,[21] as each of these four categories have their own properties. We have followed this suggestion. The outcome is closely in line

[21] In our data set this second split is replaced by a split into unit-linked policies and policies in euro.

with the variant without splitting (the first line in Table 6.2), but the number of observations is much smaller as not all firms operate on all four product markets.

6.1. Estimation results over time

Table 6.3 presents cost elasticity effects over time. For comparison, the first two columns give the full-period estimates of Table 6.1. We find that the cost elasticity is lower in both periods up to 2002 than in the two later periods, 0.82 versus 0.87, on average. This implies that the average scale economies declined over time from 0.18 to 0.13. In a test we find that the hypothesis of 'constant scale economies over time' is rejected at a confidence level of 95%. In Section 5 we observed that concentration in the life insurance sector did increase over time. It is plausible to expect that after a wave of concentration, unused scale economies decrease. On the other hand, other developments may have shifted the optimal scale to a higher level, by increasing fixed costs, thereby causing new unused scale economies. Examples of such developments may be tougher regulatory requirements, the expansion of risk management possibilities and needs, and the steady growth of profitable ICT investments. Apparently, the first effect, concentration, dominated possible optimal scale shifting developments, so that unused scale economies fell over time. Concavity in the scale economies and rejection of CRS are observed in each of the four subperiods.[22]

Table 6.3. Estimates of the translog cost function for life insurers split into four subperiods

	Entire sample		**1995-1998**		**1999-2002**		**2003-2006**		**2007-2010**	
	Coeff.	*t-value*	*Coeff.*	*t-value*	*Coeff.*	*t-value*	*Coeff.*	*t-value*	*Coeff.*	*t-value*
Insurance output, in logs[a]	0.818	67.4	0.830	41.8	0.802	37.1	0.868	41.1	0.872	38.8
Id. in squares[a]	0.023	8.6	0.032	7.3	0.017	3.8	0.034	7.6	0.019	3.2
Stock firms	0.717	7.4	0.802	5.8	0.798	4.7	0.784	4.3	-0.055	-0.3
% of collective policy prem.	-0.620	-6.5	-0.641	-4.3	-0.551	-3.2	-0.419	-2.5	-0.370	-2.2
% of lump-sum premiums	-0.444	-5.3	-0.773	-5.1	-1.015	-7.0	-0.550	-3.8	-0.192	-1.3
% of unit-fund premiums	-0.387	-5.4	-0.358	-2.9	0.202	1.6	-0.519	-4.2	-0.442	-3.2
Concentration (HHI)	-0.021	-1.7	0.254	0.9	0.063	0.4	0.454	1.8	-0.241	-3.1
Wage rate	1.930	1.8	7.481	0.8	-1.909	-0.5	2.309	0.8	21.986	3.0
Time trend	-0.012	-0.7	0.124	0.7	0.242	1.9	-0.253	-2.1	0.504	2.4
% of acquisition costs	2.604	19.0	3.434	11.8	1.232	7.5	6.435	18.4	5.865	16.1
Intercept	2.044	0.6	-18.878	-0.6	12.585	1.0	-1.880	-0.2	-70.191	-2.7
Number of observations	1,216		372		344		289		211	
R^2, adjusted	85.5		87.7		87.6		90.8		91.4	
EoS spread [a]	41.0		55.1		31.2		61.8		33.8	
F test on CRS	234.8		123.3		73.7		78.7		26.4	

[a] See footnotes below Table 6.1.

Going into the detail of the control variable coefficients, we see for some variables a certain variation in the coefficient, *e.g.* for stock firms in the last subperiod. We attribute the latter to the very small and declining number of mutual firms (to 5, on average, against 12 in the first subperiod) in this period,

[22] The critical value of the CRS test statistic for the last column in Table 6.3, F(2, 211), at the 1% significance level, is 4.71.

which may distort the comparison with the stock firms which, by the way also declined in number. As a robustness analysis we repeat the estimates of Table 6.3 for the 75% censoring case and the alternative model which also includes the number of policies as measure of output.[23] In the censoring case, scale economies decline by one-third over time, similarly to the uncensored case, while in the model with 'number of policies' included, the EoS remain unchanged. CRS is rejected significantly in all cases.

6.2. Estimation results per product type

We estimate Eq. (6) also for the four product types: the split of all polices *either* into unit-linked policies versus policies expressed in euro (left-hand panel of Table 6.4), *or* into individual versus collective policies (right-hand side). In line with expectations, the unused scale economies are by far the smallest for collective policies (11%) where concentration is higher than elsewhere and the policies themselves are, of course, already on a larger scale. Other characteristics are similar as before, whether or not we apply censoring. When we split the policies further (first into unit-linked policies and policies expressed in euro and each group further into individual and collective policies), we find the smallest EoS of 10% for Euro-Collective against 20% for the other 3 combinations.[24]

Table 6.4 Estimates of the translog cost function for life insurers split into four product types

	Unit-linked		**Euro**		**Individual**		**Collective**	
	Coeff.	*t-value*	*Coeff.*	*t-value*	*Coeff.*	*t-value*	*Coeff.*	*t-value*
Insurance output, in logs[a]	0.78	36.0	0.79	57.7	0.82	58.9	0.89	45.1
Id. in squares[a]	0.03	4.4	0.02	5.6	0.02	7.3	0.02	4.9
Stock	0.94	3.0	0.62	6.2	0.64	6.1	0.46	2.3
Share lump-sum policies	-0.26	-1.6	-1.21	-11.7	-0.69	-6.9	-1.35	-6.5
HHI	-0.09	-4.7	0.00	0.0	-0.02	-1.5	-0.03	-1.8
Wages, in logs	-0.13	-0.1	1.26	1.0	-0.13	-0.1	-0.98	-0.5
Time	0.09	2.3	0.02	0.9	0.02	0.9	0.13	3.5
Constant	8.22	1.2	3.85	0.9	8.78	2.1	10.41	1.5
Number of observations	740		1130		1180		517	
R^2, adjusted	69.2		81.4		81.4		84.3	
EoS spread[a]	0.39		0.32		0.40		0.35	
F-test on CRS	79.0		188.7		178.6		34.0	

[a] See footnotes below Table 6.1.

7. Estimates of the PCS indicator of competition

The PCS indicator measures competition. The more strongly market shares or profits of life insurance firms are determined by their marginal costs, the stronger competition on that market is. In the literature, market shares or profits are used as the dependent variable in the PCS model. Eq. (10) in Section 4.2.1 regards market shares (ms_{st}), which in our application are based on gross premiums. We

[23] Not shown here, available on request from the author.
[24] Not shown here, available on request from the author.

also apply an alternative model based on profits. Market shares tend to change rather gradually over time. One reason is that not all premiums come from new production.[25] We have two instruments to deal with such gradual adjustment of market shares to marginal costs. First, we use the partial adjustment model of Nerlove (1958) by including the lagged endogenous variable, the so-called Koyk lag. Second, following Hay and Liu (1997), Boone (2004) and Creusen *et al.* (2006), we also introduce so-called fixed effects (or dummy variables d_s) for insurance firms, so that firm specific characteristics such as scale are wiped out.[26] Thus adapted, Eq. (10) appears as:

$$\ln(ms_{st}) = \alpha \ln(ms_{s,t-1}) + \beta_t \ln(mc_{st}) + \sum_k \gamma_k X_{stk} + \sum_s \delta_s d_s + u_{it} \qquad (12)$$

The coefficient (1-α) determines the speed of adjustment. The parameter β_t is the elasticity which reflects competition over time: the more negative beta, the stronger the underlying competition. Following Bikker and Van Leuvensteijn (2008), we include control variables (X_{stk}). These are dummy variables which indicate for each of the four considered product market whether the respective insurer is active on that market or not. Considered product markets are collective, unit-linked, lump-sum and endowment policies. Evidently, an insurer's (overall) market share increases with the number of submarkets where the firm is active. Note that we express *ms* and *mc* in logarithms. Actually, we apply the Box-Cox test for each model specification to find whether linear or log-linear is most appropriate. In all cases, the linear functional form is rejected in favour of the log-linear one.

Table 7.1 presents the estimates of the PCS indicator β for market shares and profits for both marginal costs and average costs. Marginal costs are based on Eq. (6) and the estimates in the first column of Table 6.1. Average costs have the disadvantage of including fixed costs but do not suffer from possible model specification or estimation errors. Here, we assume that $\beta_t = \beta$ for all t, which provides an indication of competitive pressure over the entire sample period (1995-2010) and allows a quick comparison across model specifications. For the development of competitive pressure over time, we refer to Fig. 7.1, which shows the estimated annual β_t, based on the most appropriate specification. We focus first on market shares using marginal costs in a model with a lagged endogenous variable, the first panel on the left in Table 7.1. Estimating with OLS, we observe that the market shares are determined dominantly by previous market shares – the lagged endogenous coefficient is roughly 0.95 – and only slightly by contemporary marginal costs (with coefficient -0.07, significantly different from zero), which points to (some) competitive pressure. Inclusion of fixed effects (FE; our preferred approach) for 124 insurance firms has a heavy impact on the estimates, indicating that market shares depend strongly on (unknown) firm-specific characteristics, and underlining the limited direct role

[25] New production consists of lump-sum premiums, and part of periodical premiums, together around 55% of total premiums.
[26] We also include time or annual effects; they appear to be negligible.

Table 7.1. PCS model for the Dutch life industry

		With lagged dependent		Without lagged dependent	
Market shares		*FE*	*OLS*	*FE*	*OLS*
Marginal costs	S.-T.	-0.374***	-0.071***		
	L.-T.	(-0.92***)	(-1.51***)	-0.749***	-1.335***
Lagged MS		0.594***	0.953***		
Average costs	S.-T.	-0.292***	-0.093***		
	L.-T.	(-0.75***)	(-1.60***)	-0.652***	-1.346***
Lagged MS		0.613***	0.948***		
Profits					
Marginal costs	S.-T.	-0.066	-0.057		
	L.-T.	(-0.10)	(-0.51)	-0.034	-0.509***
Lagged profits		0.334***	0.889***		
Average costs	S.-T.	-0.076	-0.074**		
	L.-T.	(-0.11)	(-0.65**)	-0.009	-0.614***
Lagged profits		0.322***	0.887***		

Note: MC means marginal costs, AV average costs and MS market shares. S.-T. stands for short term en L.-T. for long term. Calculated values are between brackets. One, two or three asterisks indicate significance levels of respectively, 10%, 5% and 1%. Coefficients of control variables (dummies for presence on product markets) have not been shown.

of efficiency.[27] With fixed effects included, the lagged endogenous coefficient α falls roughly from 0.95 to 0.60, so that the speed of adjustment increases from 5% to 40%, indicating that gradual adjustment remains important. At the same time, the (short-term) PCS indicator of competition is, in absolute terms, much larger at -0.37. Its long-term (L.-T.) effect, calculated as $\beta/(1-\alpha)$, is more substantial still, at-0.92. Although all the results are highly significant, they should be considered with great caution, as the market shares in terms of premiums are not determined by new production only but also by existing policies, so that changes may be underestimated (see footnote 19).

Replacing marginal cost by average costs yields similar results as regards sign, level and significance, but may be inferior as proxies are used for marginal costs instead of the true – albeit estimated – values. This outcome makes clear that the PCS model may be applied without the marginal cost calculations based on the cumbersome estimation of the TCF model. The upper right-hand panel of Table 7.1 presents the direct long-term effects of marginal costs, from Eq. (6) without the lagged dependent variable. The estimation results of the PCS indicator are fairly similar, although in absolute terms, the coefficients are slightly larger. An important conclusion is that the indicator invariably carries a negative sign, corresponding with competition, and is always statistically significant. We checked that these results do not depend on the control variables: exclusion would not change the outcomes.

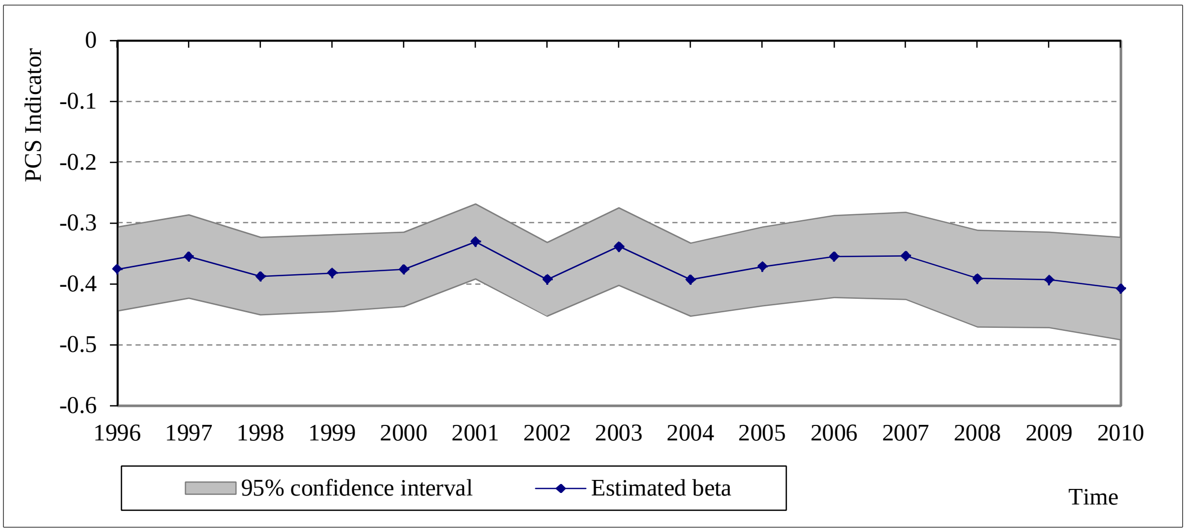
Finally, we consider the profit model instead of the market share model. Again, the PCS indicator carries the negative sign in all specifications, pointing to competitive pressure, although in the FE model it is not statistically significant and, in absolute terms, attains lower levels. Profits, too, adjust

[27] The Wu-Hausman test rejects the random effect model.

only slowly to marginal costs, if at a slightly less sluggish pace than market shares. Firms with high profits tend to be able to maintain that favourable position. This is, of course, partly due to bookkeeping, where profits per policy become available gradually over time. Note also that the within goodness of fit of the FE profit model is low compared to the in-between R^2 (not shown), confirming that profitability differs substantially across insurers, but changes little over time, indicating absence of strong cost-profit dynamics.

Fig. 7.1 shows the development of the PCS indicator over time, using the FE market share model, but estimating beta for each year separately, that is, dropping the $\beta_t = \beta$ restriction. We see that the short-term effect of marginal costs, β_t, hovers around its average value of -0.35, which corresponds with a long-term effect of around -1. An important conclusion is that β is negative in each year and significantly different from zero at the 1% level, pointing to permanent competitive pressure. The confidence interval increases over time, probably due to the 50% reduction in the number of life insurance firms. If anything, the beta is slightly more negative during 2008-2010, indicating slightly more competition.

Fig. 7.1. Effect of marginal costs on market shares of Dutch life insurers over time



Note: This graph presents estimates of β_t over time using the FE market share model, based on marginal costs (see the first rows of the middle columns in Table 7.1).

A remarkable result is that competitive pressure did not change more. Particularly, we do not see a substantial increase in competition in more recent years where the public focus on (hidden) costs has increased dramatically. Apparently, the clients of insurance firms have not responded to the investment policy scandal with greater awareness of costs. A caveat is, again, the fact that premiums do not reflect only new policies (around 55%), but also existing ones, so that market share changes are underestimated.

In order to assess whether our estimates for the PCS indicator are high or low, we compare them with estimates for other industries. Creusen *et al.* (2006) estimated the PCS model based on profits for the Dutch manufacturing and service industries and found elasticities between average variable costs and profits of around, respectively, -5.7 and -2.5. Our long-term 'profit' PCS indicator of the life insurance industry ranges from -0.1 to -0.6, indicating far weaker competitive pressure. Van Leuvensteijn *et al.* (2011, 2012) estimate the 'market share' indicator for the banking sectors in the US, Japan and the larger EU countries. The average long-run value for the Netherlands is -2.5, in absolute terms substantially above that of the life insurance sector, which ranges from -0.8 to -1.6. Bikker and Popescu (2012) finds PCS indicator estimates for the Dutch non-life insurers ranging from -3.1 to -3.8, which also points to more competitive pressure than on the life insurance market. As noted above, differences in accounting practices for profits and losses may impair the cross-sector comparison.

Table 7.2 presents competition estimates for submarkets of the Dutch life industry. We take market shares as dependent variable, marginal costs as explanatory variables, and include the lagged endogenous variable. Of course, the product market control variables disappear. Note that the life insurance market is split *either* into unit-linked policies and policies in euro (or fixed-benefit policies), *or* into individual and collective policies. On the unit-linked policy market, we do not find any evidence of competitive pressure (see upper-left hand panel). These policies are complex and private persons are likely to depend on advice from intermediaries. If intermediaries are sensitive to the relatively high commissions paid by insurance companies for selling this type of product, competition will be impaired (CPB, 2005, pages 55-82; Gorter, 2012). The PCS model does not indicate any competitive pressure on this market. The collective policy markets is more competitive, given the (highly) significant coefficients of marginal costs for both estimation approaches (see lower right-hand panel). This is in line with expectations, as collective contracts are negotiated with experts at both ends of the table. Competition takes an intermediate position in the remaining two submarkets, policies in

Table 7.2. PCS model for submarkets of the Dutch life insurance industry (1995-2010)

		Unit-linked policies		**Policies in euro**	
		FE	*OLS*	*FE*	*OLS*
Marginal costs	S.-T.	0.006	-0.036	-0.121***	-0.011
	L.-T.	(0.02)	(-0.55)	(-0.54***)	(-0.54)
Lagged MS		0.622***	0.934***	0.774***	0.980***
		Individual policies		**Collective policies**	
		FE	*OLS*	*FE*	*OLS*
Marginal costs	S.-T.	-0.047**	-0.006	-0.172***	-0.050**
	L.-T.	(-0.18**)	(-0.35)	(-0.44***)	(-2.08*)
Lagged MS		0.737***	0.983***	0.613***	0.976***

Note: MS means market shares. S.-T. stands for short-term en L.-T. for long-term. One, two or three asterisks indicate significance levels of respectively, 10%, 5% and 1%.

euro and individual policies, where the marginal cost coefficient is significantly negative, but for FE only, and – in absolute terms – less than in the collective policies market.

8. Conclusions

Efficiency and competition on the life insurance sector are important for companies and households to keep prices low and innovation and quality high. This paper investigates efficiency and competition on this market with – given the large unfavourable changes in economic, financial and institutional conditions for life insurers – special attention to developments over time. We use unused scale economies as an indirect measure of competition and find that for 1995-2010 life insurers have, on average, significant unused scale economies of 10% to 20%. These economies decrease significantly with the size of the insurer: they are twice as large for the 10% smallest firms and are close to zero for the 10% largest ones. Such unused economies of scale cannot exist under strong competition and suggest that further consolidation for the smaller life insurers would be cost efficient.

When we split the sample into two subperiods, we observe less unused economies of scale in more recent years (13% versus 18%). This is in line with the consolidation of the last 15 years: the number of life insurers fell from 100 to 50. In this light it is remarkable that the operating costs as a percentage of gross premiums remained the same: cost efficiency did not improve. Another structural change over time is the tendency for life insurers to operate increasingly on all product submarkets, indicating that scope economies are more important than specialization benefits.

The PCS indicator observes competitive pressure directly. We find a statistical significant effect of marginal costs on market shares, indicating that competitive pressure exists: efficient insurers are rewarded with increasing market shares. But this competitive pressure is weak, compared to the indicator values in other industrial sectors including banks and non-life insurers. Changes in the indicator over time are limited but, if anything, they point to further weakening. Competition is relatively stronger on the collective policies market and, according to the indicator, weaker or absent on the individual policies market. These results should be considered with great caution as the market shares in premiums are not only determined by premiums of the new production but also by premiums over existing policies, so that changes in market shares (and hence, competition) may be underestimated.

References

Bain, J.S., 1951, Relation of profit ratio to industry concentration: American manufacturing 1936-1940, *Quarterly Journal of Economics* 65, 293-324.

Berger, A.N., J.D. Cummins, M.A. Weiss, H. Zi, 2000, Conglomeration versus strategic focus: evidence from the insurance industry, *Journal of Financial Intermediation* 9, 323-362.

Bikker, J.A., M. van Leuvensteijn, 2008, Competition and efficiency in the Dutch life insurance industry, *Applied Economics* 40: 2063-2084.
Bikker, J.A., J.W.B. Bos, 2008, *Bank Performance: A Theoretical and Empirical Framework for the Analysis of Profitability, Competition, and Efficiency,* London: Routledge.
Bikker, J.A., J.K. Gorter, 2011, Restructuring of the Dutch nonlife insurance industry: consolidation, organizational form, and focus, *Journal of Risk and Insurance* 78, 163-184.
Bikker. J.A., A. Popescu, 2012, Efficiency and competition in the Dutch non-life insurance industry: effects of the 2006 health care reform, *DNB Working paper*, De Nederlandsche Bank, forthcoming.
Boone, J., 2000, Competition, *CEPR Discussion Paper Series* No. 2636.
Boone, J., 2004, A New Way to Measure Competition, *CEPR Discussion Paper Series No. 4330*, March.
Boone J., R. Griffith, R. Harrison, 2004, *Measuring Competition*, presented at the Encore Meeting 2004 'Measuring competition', the Hague.
Boone, J., 2008, A new way to measure competition, *Economic Journal* 118, 1245-1261.
Christensen, L.R., D.W. Jorgenson, L.J. Lau, 1973, Transcendental Logarithmic Production Frontiers, *Review of Economics and Statistics* 55, 28-45.
CPB, 2005, Competition in markets for life insurance, *CPB Document 96*, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.
Creusen, H., B. Minne, H. Van der Wiel, 2006, Competition in the Netherlands: an analysis of the period 1993–2001, *CPB Document No. 136*, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.
Cummins, D., H. Zi, 1998, Measuring economic efficiency of the US life insurance industry: econometric and mathematical programming techniques, *Journal of Productivity Analysis* 10, 131-152.
Cummins, J.D., M. Rubio-Misas, H. Zi, 2004, The effect of organizational structure on efficiency: evidence from the Spanish insurance industry" *Journal of Banking and Finance* 28, 3113-3150.
Cummins, J.D., M. Rubio-Misas, 2006, Deregulation, consolidation and efficiency: evidence from the Spanish insurance industry, *Journal of Money, Credit, and Banking* 38, 323-355.
Cummins, J.D., M.A. Weiss, X. Xie, H. Zi, 2010, Economies of scope in financial services: a DEA efficiency analysis of the US insurance industry, *Journal of Banking and Finance* 34, 1525-1539.
Cummins, J.D., M.A. Weiss, 2012, Analyzing firm performance in the insurance industry,using frontier efficiency and productivity methods, in: Georges Dionne (ed.), *Handbook of Insurance Economics*, Boston: Kluwer Academic Publishers.
Eling, M., M. Luhnen, 2010, Frontier efficiency methodologies to measure performance in the insurance industry: overview, systematization, and recent developments, *The Geneva Papers: Issues and Practices* 35, 217-265.
Erhemjamts, O., J.T. Leverty, 2010, The demise of the mutual organizational form: an investigation of the life insurance industry, *Journal of Money, Credit and Banking* 42, 1011-1036.
Fecher, F., S. Perelman, P. Pestieau, (1991), Scale economies and performance in the French insurance industry, *The Geneva Papers on Risk and Insurance* 16, 315–326.
Fecher, F., D. Kessler, S. Perelman, P. Pestieau, 1993, Productive Performance of the French Insurance Industry, *Journal of Productivity Analysis* 4, 77-93.
Fenn, P., D. Vencappa, S. Diacon, P. Klumpes, C. O'Brien, 2008, Market structure and the efficiency of European insurance companies: a stochastic frontier analysis, *Journal of Banking and Finance* 32,,86-100.
Fukuyama, H., 1997, Investigating productive efficiency and productivity changes of Japanese life insurance companies, *Pacific-Basin Finance Journal* 5, 481-509.
Gardner, L., M. Grace, 1997, Cost Differences Between Mutual and Stock Life Insurance Companies, *Working paper*, Center for Risk Management and Insurance Research, College of Business Administration, Georgia State University, Atlanta, GA.
Goldberg, L., A. Rai, 1996, The structure-performance relationship for European banking, *Journal of Banking and Finance* 21, 589-611.

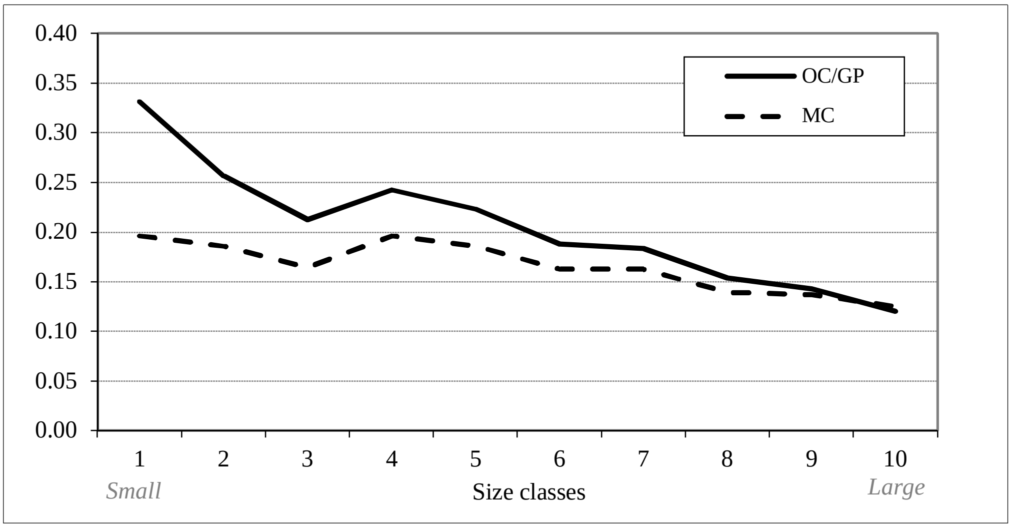
Gorter, J.K., 2012, The effect of commission bans on the quality of financial advice, *DNB Working paper* nr 350, De Nederlandsche Bank, Amsterdam.
Grace, M.F., S.G Timme, 1992, An examination of cost economies in the US life insurance industry, *Journal of Risk and Insurance* 59, 72–103.
Hanoch, G., 1975, The elasticity of scale and the shape of average costs, *American Economic Review* 65, 492-497.
Hay D.A., G.S. Liu, 1997, The efficiency of firms: What difference does competition make?, *The Economic Journal* 107, 597-617.
Jeng, V., G.C. Lai, M.J. McNamara, 2007, Efficiency and demutualization: evidence from the U.S. life insurance industry in the 1980s and 1990s, *Journal of Risk and Insurance* 74, 683-711.
Jørgensen, N.R., 1913, *Grundzüge einer Theorie der Lebensversicherung*, Jena: G. Fischer, 408 pages.
Kasman, A., E. Turgutlu, 2009, Cost Efficiency and Scale Economies in the Turkish Insurance Industry, *Applied Economics* 41, 3151-3159.
Klumpes, P.J.M., 2004, Performance benchmarking in financial services: evidence from the UK life insurance industry, *Journal of Business* 77, 257-273.
Leuvensteijn, M. van, J.A. Bikker, A.A.R.J.M. van Rixtel, C. Kok Sørensen, 2011, A new approach to measuring competition in the loan markets of the euro area, *Applied Economics* 43, 3155-3167.
Leuvensteijn, M. van, C. Kok Sørensen, J.A. Bikker, A.A.R.J.M. van Rixtel, 2012, Impact of bank competition on the interest rate pass-through in the euro area, *Applied Economics* (forthcoming)
Mahlberg, B., T. Url, 2010, Single market effects on productivity in the German insurance industry, *Journal of Banking and Finance* 34, 1540-1548.
Mason, E.S., 1939. Price and Production Policies of Large-Scale Enterprise, *American Economic Review* 29 (1, Supplement), 61-74.
Mester, L.J., 1989, Testing for expense preference behavior: mutual versus stock savings and loans, *The RAND Journal of Economics* 20, 483-498.
Nerlove, M., 1958, *Distributed lags and demand analysis for agricultural and other commodities,* Agricultural Research Service, US Dept. of Agriculture, Washington, D.C., 121 pages.
Pottier, S.W., 2011, Life insurer efficiency and state regulation: evidence of optimal firm behavior, *Journal of Regulatory Economics* 39, 169-193.
Shaffer, S., 1998, Functional forms and declining average costs, *Journal of Financial Services Research* 14, 91-115.
Smirlock, M., 1985, evidence on the (non-)relationship between concentration and profitability in banking, *Journal of Money, Credit and Banking* 17, 69-83.
Yuengert, A., 1993, The measurement of efficiency in life insurance: estimates of a mixed normal-gamma error model, *Journal of Banking and Finance* 17, 483-496.
White, H., 1980, Using least squares to approximate unknown regression functions, *International Economic Review* 21, 149-170.

Appendix: Marginal costs versus average costs

Fig. A.1 shows, for ten size classes, estimated marginal costs versus observed average costs, both expressed as share of gross premiums. Estimates are based on the TCF model of Table 6.1 with ‘insurance output’ as output measure and with ‘cost/premium rate censored at 75%’. The latter guarantees that extreme values are excluded. The estimated average operational costs (OC) are higher than marginal costs (MC), in line with expectations, as fixed costs are included in average OC and not in MC. The graph shows clearly that fixed costs play a smaller role, the larger the life insurance firm is. Additionally, we see that marginal cost decrease also with firm size.

Fig. A.2 presents for the same ten size classes the spread in average costs per unit. The graph clearly shows the left-hand leg of the theoretical U shape, be it with some irregularities. Most remarkably is the huge spread in operational costs within each size class, particularly in the small and medium-sized life insurers, which reflects that the sample is hugely heterogeneous. It is difficult to distinguish

Fig. A.1. Marginal costs versus average costs of Dutch life insurers



Note: Based on the TCF model of Table 6.1 with 'insurance output' as output measure and with 'cost/premium rate censored at 75%'.

inefficiency from model and measurement errors, for instance, omitted (unobserved) special characteristics of insurers, their products or their markets. Given the huge spread in operational costs it is obvious to assume that inefficiency is also a major component.

Fig. A.2. Average costs distribution of Dutch life insurers across size classes