S. Kellner and G. Frank Mathewson
*University of Toronto**

Entry, Size Distribution, Scale, and Scope Economies in the Life Insurance Industry

This article estimates the production parameters for life insurance firms. To do so, we first build a firm and industry model which yields an equilibrium distribution of firm size predicated on the firm's maximizing profits and segmenting markets in the face of variability in the degree of information held by consumers on the attributes of life insurance policies issued by different firms. We estimate directly the first-order conditions defining the optimal outputs for each firm. Our results indicate the absence of any natural monopoly effects and are consistent with the hypothesis of zero profits in the industry.

I. Introduction

Existing empirical studies in the life insurance industry in the United States and Canada find evidence of economies of scale in at least some aspects of the production of life insurance, with consequent implications for the efficiency of life insurance markets. However, these empirical results are inconsistent with the institutional facts in the industry. First, overall economies of scale (Houston and Simon 1970) are at odds with the evidence that over a substantial period of time, the life insurance market in Canada and the United States has supported a large and growing number of independent firms of vastly different sizes.[1] Second, if economies of scale account for the differences in firm size, with small firms ex-

* We wish to thank Jack Carr, Mel Fuss, and Herb Mohring for helpful comments and suggestions.

1. For life insurance firms, we may identify four product lines as ordinary (individual) and group insurance and ordinary and group annuities. Over the period 1961–76 in Canada, the number of firms selling ordinary insurance increased by 40%, group insurance by 45%, for ordinary annuities by 37%, and group annuities by 27%. Further, summary concentration statistics (4, 7, 16, 30 firm concentration ratios and Herfindahl indices) calculated by product line reveal decreasing levels of concentration through time. Vastly different firm sizes prevail throughout the period (e.g., the smallest firm's assets are approximately .01 times the largest firm's assets). Stigler's (1958) survivorship principle emphasizes a natural selection argument for firm sustainability. Continuing observable economies of scale are inconsistent with increasing numbers of firms of varying sizes in an industry over a relatively long period of time.

(*Journal of Business,* 1983, vol. 56, no. 1)

0021-9398/83/5601-0006$01.50

isting only temporarily, then prices for the large firms should be lower than the small firms. Casual evidence indicates that, on the contrary, price and firm size are positively correlated.[2] Third, economies of scale in some part of the life insurance production process—for example, head-office functions (Geehan 1977)—are inconsistent with the lack of any move toward specialization in that activity, especially in a growing market for the product (as argued by Stigler [1951]).

We argue in this paper that there is an alternative explanation of the institutional facts in this industry, embedded in a competitive industry model. This model is characterized by identifiable differences in prepurchase product information among consumers, caused by: variation in their search costs, market segmentation by life insurance firms, open entry into the industry, efficient life insurance and annuity prices equal to marginal costs, efficient but diverse firm sizes (with larger firms characterized by higher prices and selling expenditures and with long-run plant scale *never* characterized by increasing returns), and competitive long-run rates of return to all firms regardless of size. Conditional on this competitive specification for firms in this industry, our principal contribution is to estimate a set of first-order conditions for profit-maximizing output decisions by life insurance firms to discover that (1) as we predict, we cannot reject the hypothesis of nonincreasing returns to plant scale in this industry; (2) production functions for life insurance are not Cobb-Douglas; (3) pairwise economies of scope for joint outputs in this industry have disappeared by the late 1970s; and (4) on average, firms in this industry earn competitive rates of return.

All of these results are consistent with our competitive model. The conclusion that the market works has even greater force in the United States, where there is a renewed questioning of the exemption of life insurance firms from federal antitrust law (see, e.g., U.S. Department of Justice 1977). (U.S. life insurance firms are currently exempt from federal antitrust law because they are state regulated.)

Of more general interest is our explanation of the diversity of firm size. It is important for the model to generate an explanation of firm size not only for consistency with industry history but to justify our use of cross-sectional data. Our model is estimated at several points using data for the worldwide operations of a set of Canadian life insurance firms (Superintendent of Insurance for Canada 1961, 1966, 1971, 1976).

2. Price comparisons require a simple life insurance contract to keep the product standardized across companies. Such a contract is yearly renewable nonparticipating term insurance issued to males at age 25, renewable at constant rates to age 70, convertible to whole life insurance with no further evidence of insurability. For example, in a sample of six firms selling this product in Canada (McLeod 1978), the rank correlation coefficient between price (annual premium) and firm size (measured as total premiums collected by the firm) is .94 with a probability of .04 that the two measures are not positively related. (For this product and sample, the maximum price is 37% higher than the minimum price.)

An explanation for the diversity of firm size circumvents Friedman's (1955) objection that in an industry with identical technologies, all firms in equilibrium should be alike so that firm differences represent temporary disequilibria rendering cross-section estimation invalid.

At the heart of our explanation of the diversity of firm size is our contention that consumers of life insurance products hold varying stocks of prepurchase information about the product (i.e., have varying price elasticities of demand for death benefits or annuity size in contracts underwritten by different firms) and that life insurance firms segment the retail market to specialize and to offer consumers coverage and sales effort–advertising (service) packages that are appropriate for their level of information. Consumers hold varying levels of information on product substitutes because of variability in their marginal costs of search across competing underwriting firms. Market segmentation by firms is feasible, as each contract sale is a negotiated bilateral exchange in which we assume that agents estimate accurately consumer price elasticities for coverage, because consumers reveal extensive personal and financial information to agents to facilitate risk rating and policy recommendations.[3] Consumers who hold higher price elasticities for coverage from any single firm are aware of the attributes of insurance or annuities and other substitute products and, therefore, require less advertising and direct selling effort. In the limit, knowledgeable consumers buy from firms that incur no selling or advertising expenditures (direct writers).

There are large numbers of competing underwriting firms and open entry into the industry. This has two consequences for our model. First, while premium prices for each firm depend on items of "quality" such as advertising, sales effort, and coverage, there is no reason to assume that the price for each contract sold by the firm (total premiums per contract net of the actuarial costs of the contract) should depend on the number of contracts underwritten by each firm. Therefore, firms are price takers and prices per contract should equal the marginal costs of each contract. Second, all firms should earn zero expected profits in long-run equilibrium. If entry improves the productivity of consumer search, then entry by stimulating search should increase consumer awareness of alternative underwriting sources and increase each consumer's price elasticity of coverage from a particular firm. In turn, this should reduce the price of each contract in the conventional competitive fashion to yield efficient firm size. If we order firms according to the prepurchase knowledge or price elasticity of coverage of their con-

3. This consumer prepurchase search model is not formalized or tested here. In a separate paper, Mathewson (1983) tests predictions generated by this model on the distribution of life insurance prices. In general, the results offer support to the market segmentation price-quality consumer search model.

sumers, then the model predicts that in long-run equilibrium, break-even output levels, prices, advertising, and direct selling expenses per contract for each respective firm are positively correlated and increase as consumer price elasticities for coverage fall.

Therefore, the equilibrium distribution of firm size is determined by the underlying (common) production and marketing technologies and the density of consumers holding various price elasticities of coverage, that is, the variance of consumer search costs. The model predicts that the smaller the distribution of consumer search costs (the more alike consumers are in terms of ignorance), the tighter the distribution of firm size. A lowering of high search costs results in a shift in market share toward low-priced, low-advertising small firms and away from high-priced, high-advertising firms. These predictions are similar to those developed in a more formal and general model of consumer sampling by Lynk (1981).

Section II sets forth the details of insurance or annuity sales in a one-output (for simplicity) model of a life insurance firm together with a characterization of industry equilibrium. Section III sets out and estimates directly the analogous set of first-order conditions for a multiple-output firm that determines the profit-maximizing output levels in each insurance or annuity activity. We report our tests on production structure and on long-run industry equilibrium. Finally, Section IV contains our summary comments and conclusions.

II. A Model of a Life Insurance Firm and Industry

Our model of a life insurance firm and industry is conditional upon six assumptions:

1. Independent of their corporate structure (joint stock or mutual), life insurance firms seek to maximize expected profits.

2. Consumers of life insurance firms respond differently to individual sales effort and advertising; further, firms segment the retail market according to this characteristic and possess some price-setting powers in negotiating the size of each contract.

3. Save for some joint set-up costs, life insurance firms are characterized by two separable technologies, marketing and production.

4. The risk of death for each consumer is uncorrelated across consumers; further, all firms completely diversify away any insurance risk either by writing a sufficiently large number of contracts and/or by exchanging risk through a complete and perfect reinsurance market.

5. Capital markets for the investment of each firm's portfolios are competitive and in equilibrium, in the sense that life insurance firms neither outperform nor underperform other financial intermediaries in the construction of optimal investment portfolios.

6. Entry into the life insurance industry is unlimited.

Each of these assumptions deserves some comment. We recognize that mutual ownership (firms are mutualized when they purchase their own shares) constitutes a nontransferable form of ownership that is typically less concentrated than a joint stock company. Therefore, as ownership control is more costly, there appears to be a greater potential than usual for managers in these organizations to pursue non-profit-maximizing objectives. However, in the absence of any forces that generate rents (e.g., natural monopoly tendencies) that may be dissipated in this fashion and any differences in tax treatment, the presence of joint stock firms, supplying substitute products to the market at profit-maximizing prices, should discipline mutual firms to pursue the same profit objective.

Our second assumption is that consumers of life insurance products differ in their response to both the direct sales effort of agents and advertising of life insurance firms. Further, agents can readily identify consumers according to this characteristic and know that consumers have downward sloping demand curves for insurance coverage. The basis for this assumption lies in an argument about consumer search and information. We begin by asserting that initially uninformed consumers are capable of prepurchase search across both product types and selling agents. As long as consumers have varying marginal opportunity costs to search, they hold varying levels of knowledge on the value of these substitute products supplied in the marketplace. Each contract sale is a negotiated bilateral exchange. Consumers reveal extensive personal and financial information to agents to facilitate risk rating and policy recommendations. This information permits agents to sort consumers by their knowledge of product and firm alternatives, that is, their individual price elasticities for coverage.[4] It then pays to offer different sales effort, coverage, and price packages to consumers according to this price elasticity. Increased knowledge about brand alternatives should yield increased individual price elasticities for coverage by any single firm, that is, reduced price-setting powers on coverage for individual sellers. For simplicity, we assume that each firm services only one segment of the retail market. By this, we mean that each firm sells a homogeneous commodity to consumers characterized by a common knowledge of life insurance alternatives and, consequently, a common price elasticity of demand for death benefits.

Production and marketing technologies that are separable except for joint set-up costs lead us to asssume that, except for these set-up costs,

4. In describing the functions and powers of the life insurance agent, Mowbray, Blanchard, and Williams (1969) point out that the agent armed with a rate book specifying rates, surrender values, and policy features as well as sample policies and application blanks is expected to screen the field for his employer, operating with explicit instructions on potential clients, "canvassing only the types of person covered by his instructions" (p. 396).

no economies of scope exist between producing and selling insurance and other annuity contracts. Given the vertical relation between the activities of producing and marketing and advertising policies within the firm, this is a reasonable assumption.

Uncorrelated insurance risk and complete reinsurance markets mean that all insurance firms, independent of their size, can successfully diversify away any underwriting risk. Therefore, smaller firms are not at an underwriting disadvantage to larger firms simply because of size. This is a reasonable assumption given the existence of active reinsurance markets and the size of most insurance firms.[5] This assumption eases considerably the specification burden in our model, for it means that, without loss of generality, we may consider a "typical" consumer's risk of death within each information class of consumers.

Our similar assumption of equal competitive returns on investment portfolios across institutions means that, in equilibrium, there are no special advantages due to size (and, therefore, ability to diversify away investment risk) accruing either to insurance firms as a group or to large insurance firms within the group.

We assume that there are no capital market barriers to entry into the industry. Freedom of entry is the standard condition that firms enter the most profitable segment of the industry until, in equilibrium, expected profits are zero everywhere.

While our empirical investigation recognizes several potential outputs for each firm, for reasons of diagrammatic simplicity and ease of economic interpretation, without loss of generality, we assume here that each firm sells a single product, a one-period nonparticipating insurance policy. We proceed to specify a corresponding single-output firm model and describe the industry equilibrium. At the empirical stage, we permit firms to have multiple outputs and multiple sales efforts corresponding to each output to facilitate tests on the sustainability of product bundles.

First, for each firm, we derive an expression for the revenue per policy underwritten and retained (i.e., net of reinsurance ceded) net of actuarial costs. Typical consumers face with probability $\Gamma(0 < \Gamma \leq 1)$ the event that they die during the period, an event that represents in our model a loss of income for their family. With family utility functions concave in income, consumers purchase life insurance contracts that pay death benefits to their heirs. We define the inverse demand function for death benefits for a consumer of a typical firm as: $P = P(\mu, Z; \Gamma)$, where P is the premium price of insurance conventionally

5. In 1976, e.g., the average Canadian life insurance firm serviced 760,000 insurance and annuity contracts. If the probability of collecting on any contract is on average .2 and the value of the contract is \$25,000, then the coefficient of variation for payouts per contract for the average firm is .0023, a number which indicates that any underwriting risk is low relative to the expected payout per contract.

measured per dollar of death benefits per contract, μ is the size of the contract in dollars of death benefits ($\partial P/\partial\mu < 0$), and Z measures the firm's goodwill ($\partial P/\partial Z > 0$). (As well, $\partial P/\partial\Gamma > 0$.) By assuming that $\partial^2 P/\partial\mu\partial Z > 0$, we guarantee that increases in goodwill increase the demand elasticity of the individual consumer for a given contract size (i.e., if ϵ measures demand elasticity, then $\partial\epsilon/\partial Z(\bar{\mu}) > 0$).[6]

Premium revenue from each policy may earn additional investment revenue. If we define the portion of premium revenues invested in an investment portfolio as λ, the expected rate of return as r, and the opportunity cost per dollar of premium revenue to the policyholder of this investment portfolio as ρ (ρ depends on the nondiversified riskiness of the investment portfolio and the capital structure of the company), then the expected premium revenue from the sale and retention of one contract may be written as (defined to be net of any reinsurance ceded, given assumption 4) $[P(\mu, Z; \Gamma)(1 + \lambda)(r - \rho) - \Gamma]\mu$. Invoking our capital markets assumption (assumption 5) means that in equilibrium $r = \rho$. Therefore, we may define the solution to the firm's optimal contract size problem for that market segment it chooses to enter as:

$$\Phi(Z) \equiv \max_{\mu}[P(\mu, Z; \Gamma) - \Gamma]\mu, \tag{1}$$

where we assume that $\Phi(Z)$ is concave (Γ, which has no further interest, is suppressed.)

Next, we specify sales and production technologies together with corresponding costs. Firm goodwill with each consumer (Z) is created by variable expenditures on sales effort and advertising. We define N as the hours of sales effort by an agent to sell each contract. Although salesmen are paid on commission that is a proportion of gross premium revenues, the absence of barriers to entry into the ranks of agents means that in equilibrium salesmen earn a competitive hourly wage of w. Total dollars of advertising by each firm are measured by A, and β is a parameter that reflects the response of each consumer group to advertising and sales effort. This response is conditional on the prepurchase level of information held by each consumer on the product's attributes and alternatives. Consumers are ordered so that higher β's mean an increase in the marginal productivity of both salesmen's efforts and advertising. Therefore, the marketing technology is $Z \equiv Z(N, A; \beta)$ where Z is concave with $\partial Z/\partial N$, $\partial Z/\partial A > 0$ and $\partial^2 Z/\partial N\partial\beta$, $\partial^2 A/\partial A\partial\beta > 0$.

The output of each firm is the number of policies written and retained (i.e., again net of reinsurance ceded) by each firm and defined as m.

6. This assumption on the impact of firm-specific goodwill on individual price elasticities for consumers is consistent with Nelson's argument that in the aggregate, advertising informs consumers of alternative products increasing their price elasticities and thus increasing the degree of competition in the market (Ferguson 1974, pp. 31–33).

Price per contract, $\Phi(Z)$, is independent of m due to the levels of competition faced by each firm. Defining output for firms in terms of contracts retained after reinsurance is sensible as we expect the non-sales and advertising expenses of the firm to vary with this measure of the physical volume of output independent of the size (death benefits) of each contract. Therefore, the costs of production assuming the standard cost minimizing selection of inputs, are written as $G(m, w)$, where the inclusion of w serves to remind us that head-office production costs depend on factor prices. The average cost curve $G(m, w)/m$ is assumed to exhibit the traditional U shape. Under our assumptions and given (1) and our marketing technology, the expected profit problem for a typical life insurance firm may be written as:

$$\max_{m,N,A} \quad E\Pi \equiv m[\xi(N, A; \beta) - wN] - G(m, w) - A, \tag{2}$$

where $\xi(N, A; \beta) \equiv \Phi[Z(N, A; \beta)]$ and ξ is concave.

We proceed to describe an equilibrium in this model. A firm equilibrium is characterized by the usual marginal conditions:

$$m\colon \xi(N, A; \beta) = G_m + wN \tag{3}$$

$$N\colon \xi_N \leqslant w \tag{4}$$

$$A\colon m\xi_A \leqslant 1. \tag{5}$$

Equation (3) tells us that profit-maximizing contract levels occur when marginal revenue from the sales of an additional contract equals marginal head-office expenses plus marginal commissions to salesmen to sell that contract. It is the multiproduct analogue to (3) that forms part of our estimated system. Equations (4) and (5) define optimal levels for N and A that appear in (3). The inequality signs refer to direct writers of insurance. Equation (4) tells us that profit-maximizing sales effort per contract occurs where the increase in the value of goodwill from a salesman's efforts just equal the competitive wage rate. Sales agents' efforts are on a per contract basis as there is no word-of-mouth creation of goodwill in this world. Equation (5) tells us that profit-maximizing firms should advertise until the marginal value product of advertising (over all consumers) just equals \$1.00 (the marginal cost).

An industry equilibrium occurs when $E\Pi = 0$ or when $\xi(N, A; \beta) = [G(m, w) + A]/m + wN$. In this case, (3) is replaced by the usual result:

$$(G_m + wN) = \{[G(m, w) + A]/m + wN\}. \tag{6}$$

At the industry level, we assume an inelastic number of consumers, all buying life insurance contracts. The questions for each consumer are (i) how much coverage and (ii) at what price and level of information. Entry squeezes prices in two directions. First, more firms mean that consumers are spread more thinly across firms, reducing price. Second, more firms, by improving the productivity of search, should stimulate consumer search. This, in turn, should lower contract prices by increasing consumer price elasticities for coverage. In an industry equilibrium, returns to firms are competitive, the standard result. As (6) holds only when $(G_m + wN) \geqslant [G(m, w)/m + wN]$ in this model, long-run firm sustainability and free entry mean that we *never* expect to observe any economies of scale in the production of life insurance contracts in the long run. This is consistent with our entry observations and concentration measures for the industry.[7]

Our empirical work directly estimates a multiproduct analogue of (3) and a specification of total head-office production costs conditional on optimal levels of advertising and sales effort given by multiproduct analogues of (4) and (5). We then test the consistency of our estimates with an industry equilibrium. Before we turn our attention to this task, we investigate the nature of an equilibrium distribution of firm size in this model. As we have argued, an explanation of the distribution of firm size is important both to explain the institutional facts of the industry and to justify the use of cross-sectional estimation.

In our model, for an industry equilibrium (described by [4], [5], and [6]), conventional comparative statics (using routine, if tedious, mathematics and therefore omitted) yield the predictions that $\partial m/\partial\beta$, $\partial A/\partial\beta$, $\partial N/\partial\beta > 0$. Therefore, larger firms have larger advertising expenditures and larger direct sales expenses per contract sold and higher prices. Figure 1 illustrates the nature of the firm long-run equilibrium in this industry.

In figure 1, for consumers of type β^*, who receive N^* hours of direct agent effort per contract and A^* dollars of advertising expenditures, a long-run equilibrium for the firm exists with sales of m^* contracts, that is, (N^*, A^*, M^*) simultaneously satisfy (3), (4), (5), and (6) for consumers of type β^* (m^* is a zero-profit output, as the average cost curve with advertising and sales expenditures set at the optimal levels corresponding to m^* is minimized at m^*).[8]

In industry equilibrium, there is a distribution of firms that is demand determined. Corresponding to each value of β, there is an industry equilibrium along the corresponding $(G_m + wN)$ marginal cost curve. The number of firms in each segment of the retail market is determined

7. Potential exists for increasing returns in the production of goodwill. Data at our disposal do not afford a test of this possibility. In regulatory regimes such as the United States where life insurance has always been distributed through independent agents (in Canada, regulatory rules have historically constrained life insurance agents to work for only one company), the large number of agencies suggests the absence of any economies of scale in sales distribution.

8. The nature of an industry equilibrium here is similar to the monopolistically competitive equilibrium described in Demsetz (1959, 1968) and Rosse (1978), only simpler because of the price-taker assumption for firm output (m).

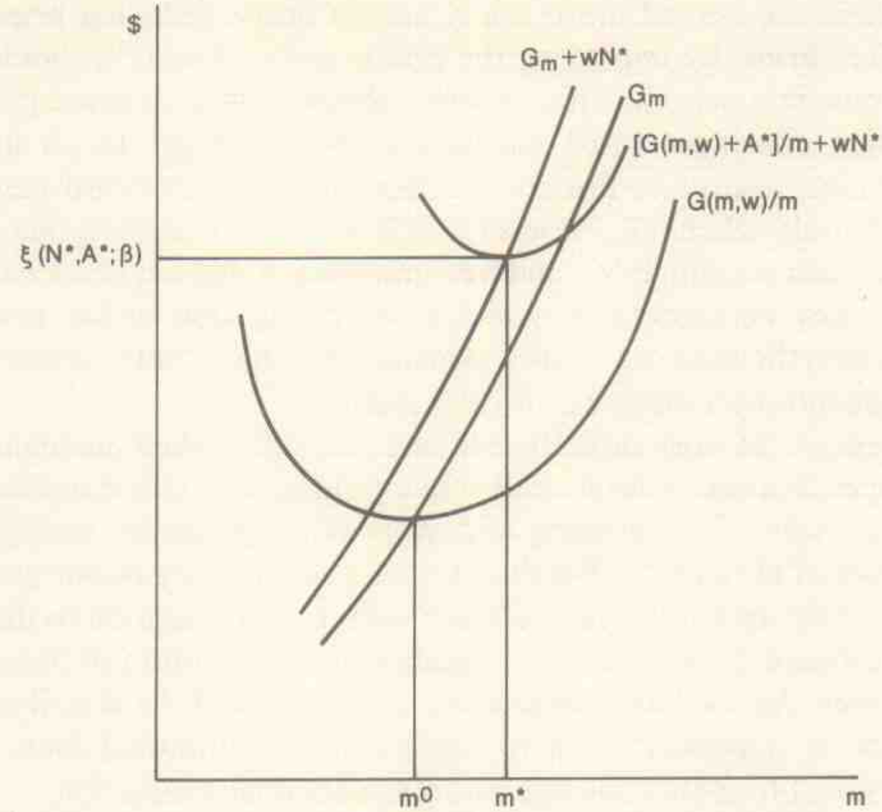

FIG. 1.—Industry equilibrium for life insurance firms

by technological considerations and the density of consumers at any level of β. Because the efforts of agents to sell contracts is on a per contract basis, increases in wN in response to higher β's increase each firm's marginal cost but in the one-output model leave efficient plant scale unchanged. However, changes in A in response to higher β's are similar to increments in fixed costs and, therefore, increase efficient plant scale.

In principle, these equilibrium conditions are similar to competitive equilibrium conditions. For a firm to exist in any part of the retail market, consumers must be sufficiently dense in number to support that firm's earning a normal rate of return. Starting from a market where consumers are sufficiently knowledgeable that there are no sales expenditures ($\xi_N < w$) and no advertising expenditures ($m\xi_A < 1$) and firm output is illustrated in figure 1 as m^0 (these firms are the direct writers), providing consumers are sufficiently dense in number, firm-size and goodwill expenditures increase as β increases (for less knowledgeable consumers). As figure 1 illustrates, in industry equilibrium, life insurance firms always produce on that portion of their marginal cost curve ($G_m + wN$) that lies above their corresponding average cost curve, $G(m, w)/m + wN$. Therefore, in an industry equilibrium, life insurance firms are never characterized by increasing returns to scale in production. Life insurance firms are in fact multiproduct firms that produce insurance and annuities sold to individuals and groups. The principal contribution of this paper is to use a multiproduct analogue of the profit-maximizing marginal conditions for the firm to estimate directly the production characteristics for life insurance and then to test the consistency of these parameter estimates with an industry equilibrium.

III. Empirical Estimation of the Life Insurance Firm Model

For empirical purposes, we recognize four potential contract types that firms write in each year: ordinary life insurance policies written for individual consumers (m_1); group life insurance policies written for groups of consumers (m_2); ordinary annuities sold to individual consumers (m_3); and group annuities sold to groups of consumers (m_4). We measure m_2 and m_4 as the total number of individuals in groups serviced by each company, in each year, thus ignoring any substitution between numbers of groups and group size. Our estimation proceeds directly from the corresponding set of first-order conditions for a multiple-output firm. Conditions corresponding to (4) and (5) generate optimal sales effort by salesmen and advertising. While we exclude word-of-mouth goodwill creation across consumers purchasing a particular line of insurance, we can permit information spillovers across lines of insurance in a multiple-output setting. It is reasonable to expect life insurance firms that accumulate brand names in life insurance products, for example, should be able to trade on that reputation in annuity markets. In this case, our specification for the marketing technology producing goodwill in activity i may be written as:

$$Z_i = Z^i(\mathbf{N}, A; \beta), \tag{7}$$

where $\mathbf{N} = (N_1, N_2, N_3, N_4)$ and N_i represents the hours of sales effort per contract line i. Firms then allocate sales effort to each activity to capture not only the direct increase in revenue in that line but also the indirect effect of an improvement in the firm's general brand name that increases revenues from other lines of activity. A similar condition holds for advertising.

With the inclusion of these indirect effects, the marginal conditions determining sales effort and advertising, analogous to (4) and (5) become, respectively,

$$\sum_j m_j \, \xi^j_{N_i} = w m_i, \tag{8}$$

$$\sum_j m_j \, \xi^j_A = 1, \tag{9}$$

where $\xi^j_{N_i} \equiv \partial \xi^j / \partial N_i$, $\xi^j_A \equiv \partial \xi^j / \partial A$, and ξ^i is the net expected revenue per contract line j; $i, j = 1, \ldots, 4$.

If we use asterisks to denote optimal solution values for decision variables, then, for these multiple-output lines, the optimal number of contracts by line, conditional upon the optimal production of goodwill, given by (8) and (9), is described by:

$$\xi^i(\mathbf{N}^*, A^*; \beta) = \partial G/\partial m_i + wN_i^* \qquad i = 1, \ldots, 4, \tag{10}$$

where $G \equiv G(m, w)$, $\mathbf{m}$ is a vector of contract outputs, and $\mathbf{N}^*$ is a vector of optimal sales efforts.

Due to the multiproduct nature of production and sales, (8), (9), and (10) are more complex than before. However, the same general characteristics still prevail for an industry equilibrium.

As (10) is the multiproduct analogue of (3), economic interpretations are identical. One slight manipulation in (10) is required for empirical purposes. Upon multiplication by $m_i^*/G(\mathbf{m}^*, w)$ to yield an elasticity formulation, (10) may be evaluated at the optimal values (m^*, A^*, N^*) as:

$$m_i^*\xi^i(\mathbf{N}^*, A^*; \beta)/G(\mathbf{m}^*, w) = m_i^*/G(\mathbf{m}^*, w) \cdot \partial G/\partial m_i + wm_i^*N_i^*/G(\mathbf{m}^*, w), \tag{11}$$
$$i = 1, \ldots, 4.$$

Equation (11) and the production cost function $G(\mathbf{m}^*, w)$ with an appropriate functional form and error specification, lend themselves to estimation using Zellner's (1962) multivariate regression technique of "seemingly unrelated estimation." Our sample consists of the set of Canadian-owned life insurance firms selling in Canada for 1961, 1966, 1971, and 1976. For each of these years, the data are taken from the consolidated balance sheets and income statements for the worldwide operations of these firms (Superintendent of Insurance for Canada 1961, 1966, 1971, 1976). The sample consists exclusively of Canadian-owned firms, as the data available on foreign-owned firms selling in Canada are solely for their Canadian branch plant operations and therefore, exclude head-office expenses. The data do not measure factor prices (w) for each firm. However, as life insurance firms compete in competitive factor markets, it is reasonable to assume that factor prices are constant across firms in each of our cross-section samples and may be safely ignored. However, factor prices do vary through time if only for reasons of inflation, and such variation cannot be ignored. Consequently, we estimate the model separately for each of the 4 years of cross-section data.

One potential problem that arises from the direct estimation of our system of equations is possible simultaneous equation bias due to the endogenous nature of $\mathbf{m}^*$ and $\mathbf{N}^*$, variables appearing on the right-hand side of (11), and the total cost function. In this cross-section sample, there are no exogenous final demand or factor supply variables available to facilitate any simultaneous equation methodology for avoiding potential simultaneous equation bias.[9]

Insurance and annuity contracts extend on average beyond one year. However, our model is atemporal. With some care, each variable in our static model has a measured counterpart in the data. Expected net premium revenues, that is, total premium revenues net of expected payouts and reinsurance, are required on the left-hand side of (11). The multioutput analogues to the firm's optimal death benefits problem of contract size (given by [1] for the single-output firm) yield, by definition,

$$m_i^* \xi^i(\mathbf{N}^*, A^*; \beta) \equiv m_i^*[P^i(\mu_i^*, Z_i^*; \Gamma^i) - \Gamma^i]\mu_i^*, \qquad (12)$$

where $Z_i^* \equiv Z^i(\mathbf{N}^*, A^*; \beta)$.

Contracts sold by each underwriting firm are measured as the contracts in force for the firm in each sample period. In this case, total premium revenues by line net of reinsurance ceded, $m_i^* P^i(\mu_i^*, Z_i^*; \Gamma^i)\mu_i^*$, required in (12) for each firm are directly measurable. For each life insurance or annuity contract, regulation requires the firm to establish minimum policy reserves, calculated as the present value of future claims minus the present value of future premiums using prescribed discount rates and mortality tables. These reserves appear as liabilities on the firm's balance sheet in anticipation of future payouts over the life of the contract.[10] Aggregating over all contracts in force in any line for each sample period, we measure the expected payouts associated with any revenue stream ($m_i^*\Gamma^i\mu_i^*$) as the difference in these regulatory–constrained reserve liabilities for each line between the sample year and the previous year.

As agent-related expenses (commissions, agent convention expenses, agent pension contributions, etc.) are aggregated to provide estimates of $wm_i^* N_i^*$ for each line of activity, G^* includes all other head-office expenses except advertising.[11] To the extent that these measurements reflect arbitrary accounting convention, such as the calculation of depreciation on fixed assets, such accounting expenses in-

9. In a simultaneous equation model of property-liability insurance firms, Halpern and Mathewson (1975) were able to construct firm-specific exogenous demographic variables to facilitate a two-stage estimation procedure. Such a procedure was made possible by information on the geographical distribution of firm sales. No information is available on the distribution of the sales of the life insurance firms in this sample.

10. The accounting procedures used in the Canadian industry are described in Canadian Institute of Chartered Accountants (1973, pp. 102–14). Policy reserve liabilities record the present value of future claims minus the present value of future net premiums. The choice of mortality tables and discount rates obviously affects these calculations.

11. Advertising appears explicitly as a decision variable, determined by (9), which through the marketing technology affects the left-hand side of (11). However, advertising expenditures are not required explicitly for purposes of estimation.

corporate measurement bias. There is every reason to believe that such bias is minimized in these cost figures. Most expenses incurred by life insurance firms reflect factor purchases through market-determined prices. The principal fixed asset for life insurance firms is their head-office building. For the regulatory accounting statements, life insurance firms are required to impute a rental rate to their use of head-office space equal to either the rental rate charged to other firms renting space in these buildings or the rental rate for equivalent commercial space in neighboring sites. Therefore, G^* is relatively free of the usual accounting biases due to arbitrary depreciation rules.

Not all firms in our sample produce in all four lines. Therefore, we must select a functional form for $G(\mathbf{m}^*, w)$ which permits outputs to be zero by some firms in some lines. For any study of economies of scope, inclusion of such firms is important. We use a translog approximation to our cost function with a Box-Cox metric on the output lines to facilitate zero outputs in some lines by some firms. Consequently, following Caves, Christensen, and Tretheway (1980), under our condition of constant factor prices for all firms in each of our cross samples, $G(\mathbf{m}^*, w)$ may be approximated as:

$$\begin{aligned} \ln G = \alpha_0 &+ \sum_i \alpha_i (m_i^{*\lambda} - 1)/\lambda \\ &+ \tfrac{1}{2} \sum_i \sum_j \delta_{ij} [(m_i^{*\lambda} - 1)/\lambda] \cdot [(m_j^{*\lambda} - 1)/\lambda]. \end{aligned} \tag{13}$$

Substituting the marginal cost relation corresponding to (13) into (11), the first-order conditions for profit-maximizing contract levels,[12] plus (13) itself, all conditional on profit-maximizing levels of sales effort and advertising, yields a five-equation nonlinear estimation problem for each of the 4 sample years amenable to multivariate regression techniques. Here, we use Zellner's (1960) "seemingly unrelated" estimation technique which permits symmetry constraints on parameters across equations ($\delta_{ij} = \delta_{ji}$) and captures any mutual correlations on the regression disturbances (and therefore, is known to be more efficient than ordinary least squares). Our system of equations is linear conditional on λ. We proceed by specifying a set of values for λ, estimating the remainder of the parameters conditional on λ, and selecting those parameter estimates that correspond to that λ which maximizes the log of the likelihood function for the system.

Parameter estimates together with their standard errors are reported in table 1. While these estimates yield no direct economic interpretation, it is worth noting that the values of λ that maximize the likelihood

12. The first-order conditions, which together with (13) constitute the estimated equations, are: $m_i^* \xi^i(\mathbf{N}^*, A^*\beta)/G(\mathbf{m}^*, w) = m_i^{*\lambda}[\alpha_i + \Sigma_j \delta_{ij}(m_j^{*\lambda} - 1)/\lambda] + w m_i^* N_i^*/G(\mathbf{m}^*, w)$, $i = 1, \ldots, 4$.

TABLE 1 Estimates of Cost Function Parameters for a Sample of Canadian-owned Life Insurance Firms

	Estimates			
Parameter	1961	1966	1971	1976
α_0	.108E + 02*	.122E + 02*	.117E + 02*	.647E + 01*
	(.279)	(.133)	(.226)	(.104E + 01)
α_1	.252E − 01*	.445E − 02*	.301E − 01*	.595*
	(.754E − 02)	(.455E − 03)	(.295E − 02)	(.968E − 01)
δ_{11}	.943E − 04	−.177E − 05*	−.848E − 04*	.231E − 02
	(.130E − 03)	(.443E − 06)	(.208E − 04)	(.344E − 02)
δ_{12}	−.583E − 04	.171E − 06	.269E − 04*	−.113E − 01*
	(.682E − 04)	(.260E − 06)	(.112E − 04)	(.444E − 02)
δ_{13}	−.914E − 05	−.919E − 06*	−.128E − 04	.637E − 02*
	(.348E − 04)	(.358E − 06)	(.860E − 05)	(.280E − 02)
δ_{14}	−.115E − 03*	−.173E − 05*	−.463E − 04*	.397E − 02
	(.218E − 04)	(.439E − 06)	(.616E − 05)	(.273E − 02)
α_2	.793E − 02	−.181E − 03	−.208E − 02	.175*
	(.526E − 02)	(.298E − 03)	(.153E − 02)	(.531E − 01)
δ_{22}	.757E − 05	.286E − 06	.282E − 05	.280E − 02
	(.538E − 04)	(.226E − 06)	(.649E − 05)	(.170E − 02)
δ_{23}	−.687E − 04*	−.408E − 06	−.298E − 04*	.114E − 03
	(.240E − 04)	(.326E − 06)	(.482E − 05)	(.367E − 03)
δ_{24}	.671E − 05	−.898E − 06*	−.319E − 04*	.631E − 03
	(.143E − 04)	(.256E − 06)	(.931E − 05)	(.376E − 03)
α_3	.368E − 02*	−.222E − 03	−.445E − 02*	.214E − 01
	(.159E − 02)	(.194E − 03)	(.987E − 03)	(.418E − 01)
δ_{33}	.798E − 03*	.108E − 04*	.224E − 03*	−.164E − 01*
	.508E − 04	(.211E − 05)	(.313E − 04)	(.018E − 02)
δ_{34}	−.188E − 03*	.281E − 05	.540E − 04*	.846E − 02*
	.291E − 04	(.175E − 05)	(.243E − 04)	(.832E − 03)
α_4	.146E − 02	−.107E − 03	.654E − 02	−.120*
	.157E − 02	(.246E − 03)	(.685E − 03)	(.322E − 01)
δ_{44}	.384E − 02*	.974E − 05*	.210E − 03*	−.101E − 02*
	(.566E − 04)	(.210E − 05)	(.315E − 04)	(.414E − 03)
λ	.30	.50	.35	.025

NOTE.—Standard errors in parentheses under coefficient estimate.
* Variables are significant at the 5% level, two-tailed.

function are positive.[13] Before we use these parameter estimates to analyze the nature of underlying production in this industry, we may easily test whether or not the underlying structure is Cobb-Douglas, a structure frequently imputed to insurance production (e.g., Cummins and Vanderhei [1979, p. 713] specify Cobb-Douglas production functions for property-liability insurance). Table 2 reports the results of log likelihood tests where rejection of the Cobb-Douglas structure appears possible in all 4 years. This suggests that previous work that imputes a Cobb-Douglas production to insurance firms may be incorrect.

Table 3 contains the summary statistics for our tests of the structure

13. Positive finite λ's mean that the transformed variables converge to a finite value as m_i goes to zero; i.e., $\lim\,(m_i^{\lambda} - 1)/\lambda = -1/\lambda,\ m_i \rightarrow 0$.

TABLE 2 Test of the Production Structure

Year and Structure	Log Likelihood	Test Statistic $(-2\ln L_1/L_0)$	*df*	Critical Value (5%)
1976:				
Joint cost function (maintained hypothesis)	−465.75	...	...	...
Cobb-Douglas	−517.14	102.78	10	20.5
1971:				
Joint cost function (maintained hypothesis)	−262.29	...	...	...
Cobb-Douglas	−338.33	152.08	10	20.5
1966:				
Joint cost function (maintained hypothesis)	−172.20	...	...	...
Cobb-Douglas	−202.29	60.18	10	20.5
1961:				
Joint cost function (maintained hypothesis)	−287.36	...	...	...
Cobb-Douglas	−341.24	107.76	10	20.5

of production. Overall scale elasticity may be measured as (Panzar and Willig 1977) $SE \equiv 1/(\Sigma_i \partial \ln G/\partial \ln m_i)$. Point estimates together with estimated standard errors are reported in table 3 for an "average" firm, that is, a firm whose output in each line is the average over all firms in the industry. The results for 1966 and 1971 indicate that the hypothesis of scale economies for the average firm may be rejected safely. For 1961 and 1976, only weaker statements are possible. Although the 1976 results indicate economies of scale as a point estimate, rejection of nonincreasing returns to scale is possible only with significant risk of error.

Multioutput production requires consideration of the optimal output packages, economies of scope. Panzar and Willig (1979) show that a sufficient condition for a twice differentiable multiproduct cost function to exhibit economies of scope is that $\partial^2 G/\partial m_i \partial m_j < 0$. In the translog structure, using the Box-Cox transform on outputs, approximate tests for this condition, identical to those developed by Denny and Fuss (1977), are that $\alpha_i \cdot \alpha_j + \delta_{ij} < 0$, where the data are scaled so that $m_i = 1$ for all i at the point of approximation. This requires division of each output variable by its respective sample mean. Table 3 indicates that there are pairwise scope economies between certain lines in certain years, for example, ordinary insurance and group annuities in 1961, 1966, and 1971. Other scope economies for which we could not test with our data could arise from production jointness in marketing, that is, brand-name effects. There are also positive estimates of some cross-partial derivatives which indicate that some joint production may be uneconomic, for example, ordinary and group annuities in 1971 and ordinary insurance and group annuities in 1976. In the presence of fixed

TABLE 3 Estimates of Profitability, Scale Economies, and Jointness Economies Evaluated at the Mean for a Selected Sample of Canadian-owned Firms

Parameter	1961	1966	1971	1976
Scale economies	.98	.83	.83	1.09
	(.64E − 01)	(.37E − 01)	(.63E − 01)	(.86E − 01)
Hypothesis	SE ≥ 1	SE ≥ 1	SE ≥ 1	SE ≤ 1
May be rejected with a probability of error of approximately	.38	Less than .0005	Less than .0005	.22
Pairwise jointness economies:				
J_{12}[a]	−.12	.85E − 01	.19	.51E − 01
	(.18)	(.14)	(.14)	(.51E − 01)
J_{13}	.14*	.20E − 01*	−.58E − 02	.42E − 01
	(.26E − 01)	(.94E − 02)	(.12E − 01)	(.27E − 01)
J_{14}	−.51E − 01*	−.93E − 01*	−.70E − 01*	.26E − 01
	(.13E − 01)	(.26E − 01)	(.16E − 01)	(.10E − 01)
J_{23}	−.43E − 01*	−.17E − 01	−.57E − 01*	.42E − 02
	(.24E − 01)	(.15E − 01)	(.97E − 02)	(.45E − 02)
J_{24}	.36E − 02	−.72E − 01*	−.79E − 01*	.36E − 02
	(.12E − 01)	(.21E − 01)	(.23E − 01)	(.20E − 02)
	−.36E − 01*	.24E − 01	.27E − 01*	.14E − 01
J_{34}	(.63E − 02)	(.15E − 01)	(.12E − 01)	(.21E − 02)

NOTE.—Standard errors are reported in parentheses under each coefficient estimate.

[a] $J_{ij} \equiv \alpha_i \cdot \alpha_j + \delta_{ij}$ where 1 ≡ ordinary insurance; 2 ≡ group insurance; 3 ≡ ordinary annuities; 4 ≡ group annuities.

* Significantly different from zero at the 5% level, two-tailed.

costs, positive cross partials of the cost function are necessary but not sufficient for uneconomic multiproduct production. If demands for the products of the industry were truly independent, with uneconomic joint production, then specialized firms would drive multiproduct firms from the market. This is not necessarily the case where there are some demand interdependencies, for example, where goodwill of the firm crosses product lines. Table 3 does not reveal any consistent pattern of scope economies through time for the average firm.

One fact does emerge for the average firm. By 1976, any significant economies of scope had disappeared. If this is the case, we would expect new firms who should exhibit the greater flexibility on the choice of product lines not to produce in all four lines. Of the 24 new foreign and domestic firms that entered the Canadian market between 1971 and 1976, one entered producing in all four lines, the others producing in fewer lines. At best, such evidence is very casual. While our estimates of the sign of $\partial^2 G/\partial m_i \partial m_j$ are for the average firm, there is no guarantee that the cost surface is identical in shape for other feasible output levels. However, it is possible to conduct these tests for any output levels simply by the appropriate scaling of the output variables.

Our final inquiry is whether or not our firm model is consistent with a zero-profit industry equilibrium. At any moment, expected profits for firms in our model may not be zero due to lags in entry and/or exit responses by firms or transitory efforts that may be firm specific (e.g., errors by firms in identifying one or more optimal decision margins specified in the model) or industry specific (e.g., an unanticipated influenza epidemic that dramatically alters mortality experience). We observe a large distribution of firm size in this industry. We have already demonstrated that this observation is consistent with an industry equilibrium as an equilibrium distribution of firm size in our model, conditional upon a distribution of consumer ignorance (i.e., β's). Zero expected profits plus profit maximization for each firm in the industry yield the following condition as the multioutput analogue of (6):

$$\sum_i \partial \ln G/\partial \ln m_i = 1 + A^*/G^* \equiv \tilde{\eta}, \tag{14}$$

where A^* represents optimal advertising expenses and G^* is total production costs at optimal output levels measured for each firm.

An alternative estimate of $\Sigma_i \partial \ln G/\partial \ln m_i$ may be derived from our estimated cost function. We denote this measure as $\hat{\eta}$. Our test of the consistency of our cost function parameter estimates with zero profits is to test whether the calculated mean of $d \equiv \hat{\eta} - \tilde{\eta}$ is equal to zero. This tests whether on average the industry is in equilibrium.[14] Table 4

14. We would also be interested to know whether each firm is in equilibrium. While we can calculate the sample variance of $\hat{\eta}$, we do not know the underlying distribution for A^*/G^* (or $\tilde{\eta}$), so that significance tests are not possible.

TABLE 4 **Tests of the Consistency of the Cost Function Parameter Estimates with an Industry Equilibrium**

Year	$\hat{\eta}$	$\bar{\eta}$	Test Statistic	Critical Value (5%)
1961	1.04	.85	−.58	2.0
1966	1.03	.56	−.84	2.0
1971	1.05	.96	−1.70	2.0
1976	1.02	.78	−.53	2.0

reports the *t*-statistic appropriate for the null hypothesis that $\bar{d} = 0$. This test reveals that we cannot reject the null hypothesis that our production parameter estimates based on a firm equilibrium are on average consistent with the industry equilibrium generated by open entry into our model for each of our sample years.

Summary and Conclusions

This paper builds a model of profit-maximizing life insurance firms in an industry with open entry. Market segmentation based on the response of consumers to firm expenditures on goodwill is sufficient to yield an equilibrium distribution of firm size. Based on the marginal conditions for profit maximization, we estimate the production parameters for life insurance firms. There is no evidence of any natural monopoly tendencies in the production of life insurance. While this result is at odds with the empirical results of other authors, it is consistent with the institutional evolution of the industry in the United States and Canada over the past 20 or so years. We find evidence of some economies of scope at points during the period for some product lines evaluated at the size of an average firm. Scope economies may exist for other product combinations for firms of a different size and may also result from the presence of fixed costs independent of the bundle of products produced. The absence of scope economies does not rule out the existence of multiproduct firms due to demand interdependencies such as the firm's brand name. Our production parameter estimates are consistent with a zero-profit industry equilibrium. If the industry is in equilibrium at each point of time over the sample period, then observed entry is a response to the growth in demand for the product rather than a competitive response to excess profits.

Our results are a strong statement about certain elements of the operation of the competitive marketplace in one industry. With respective insurance prices defined as net premium (premiums net of expected payouts) in each line, price appears to equal marginal cost. The market is sufficiently large that there are no advantages to size from any fixed costs. Firm size for our sample is consistent with zero profits. In this sense, the market works in this industry.

References

Benston, G. 1972. Economies of scale of financial institutions. *Journal of Money, Credit and Banking* 4 (2): 312–41.

Canadian Institute of Chartered Accountants. 1973. *Financial Reporting for Life Insurance Companies*. Toronto: Canadian Institute of Chartered Accountants.

Caves, D. W.; Christensen, L. R.; and Tretheway, M. W. 1980. Flexible cost functions for multiproduct firms. *Review of Economics and Statistics* 62 (3): 477–81.

Cummins, J. D., and Vanderhei, J. 1979. A note on the relative efficiency of property-liability insurance distributions systems. *Bell Journal of Economics* 10 (2): 709–19.

Demsetz, H. 1959. The nature of equilibrium in monopolist competition. *Journal of Political Economy* 67 (1): 21–30.

Demsetz, H. 1968. Do competition and monopolistic competition differ? *Journal of Political Economy* 76 (1): 146–48.

Denny, M., and Fuss, M. 1977. The use of approximation analysis to test for separability and the existence of consistent aggregates. *American Economic Review* 67 (3): 404–18.

Ferguson, J. 1974. *Advertising and Competition: Theory and Measurement Fact*. Cambridge, Mass.: Ballinger.

Friedman, M. 1955. Comment. In *Conference on Business Concentration and Price Policy*. Princeton, N.J.: Princeton University Press.

Geehan, R. 1977. Returns to scale in the life insurance industry. *Bell Journal of Economics* 8 (2): 497–514.

Halpern, P. J., and Mathewson, G. F. 1975. Economies of scale in financial institutions: a general model applied to insurance. *Journal of Monetary Economics* 1 (2): 203–20.

Houston, D. B., and Simon, R. M. 1970. Economies of scale in financial institutions: a study of life insurance. *Econometrica* 38 (6): 856–64.

Lynk, W. J. 1981. Information, advertising, and the structure of the market. *Journal of Business* 54 (2): 271–304.

McLeod, W. E. 1978. *Shopper's Guide to Canadian Life Insurance Prices*. Ottawa: Consumer's Association of Canada.

Mathewson, G. F., with Todd, J. 1982. *Information, Entry, and Regulation in Markets for Life Insurance*. Toronto: University of Toronto Press.

Mathewson, G. F. 1983. Information, search and price variability of individual life insurance contracts. *Journal of Industrial Economics*, in press.

Mowbray, A.; Blanchard, R.; and Williams, C. A., Jr. 1969. *Insurance: Its Theory and Practice in the U.S.* New York: McGraw-Hill.

Panzar, J. C., and Willig, R. D. 1977. Economies of scale in multi-output production. *Quarterly Journal of Economics* 91 (3): 481–94.

Panzar, J. C., and Willig, R. D. 1979. Economies of scope, product specific economies of scale and the multi-product competitive firm. Bell Laboratories Economics Discussion Paper no. 152.

Rosse, J. 1978. Lecture notes: monopolistic competition, excess capacity, and inefficiency. Studies in Industry Economics 74. Stanford, Calif.: Stanford University.

Stigler, G. 1951. The division of labor is limited by the extent of the market. *Journal of Political Economy* 59 (3): 185–93.

Stigler, G. 1958. The economies of scale. *Journal of Law and Economics* 1:54–71.

Superintendent of Insurance for Canada. Various years. Report of the Superintendent of Insurance. Vol. 3. *Annual Statements—Life Insurance Companies and Fraternal Benefit Societies*. Ottawa: Queen's Printer of Canada.

U.S. Department of Justice. 1977. *The Pricing and Marketing of Insurance*. Stock no. 027-000-00477-6. Washington, D.C.: Government Printing Office, January.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57:348–68.