

基于时间序列分析的寿险需求模型分析

白 丽 李海刚

(上海交通大学安泰经济与管理学院管理科学与工程系, 上海 200052)

摘 要 分别运用最小二乘法、时间序列分析对寿险市场进行了实证研究。与已有的研究相比,通过建立多元线性模型和改进的回归模型,说明国内生产总值的增长是寿险需求增长的根本原因,教育水平的提高和赡养(抚养)率对寿险需求影响显著;同时利用单位根检验指出由于缴纳保费存在分期缴纳的现象,导致保费收入有时间滞后效应。

关键词 寿险需求 最小二乘法 时间序列模型 单位根检验

中图法分类号 F840.6; **文献标志码** A

我国人寿保险业务在最近几年取得了巨大发展,寿险保费收入在最近十年中一直占据保险市场的大半壁江山,并一直保持高速增长。卓志^[1]等人撰文分析了我国寿险需求的影响因素。但是他们的研究是将年度数据视作独立样本点,没有考虑时间序列的平稳性,因而可能导致回归结果有差异。本文用 Eviews 软件^[2]分别运用最小二乘法、时间序列分析的理论来进行对寿险需求模型的分析进行改进,最后对解释变量序列和被解释变量序列做单位根检验,对模型进行第二次改进。对影响我国寿险需求的因素进行了比较深入的分析。在本文中,一共使用了1990年至2005年共计16年的全国寿险保费收入年度。

1 寿险需求影响因素的选择

通过相关资料和已有的研究结果^[3,4],影响寿险需求的量化因素表现在两个方面:一是内生因素,如收入变量;二是外生因素,一般包括:经济的结构、发展水平、国民收入水平等。本文本来考虑了以下

几个因素:GDP、小孩抚养率、老人赡养率、预计通货膨胀率、存款利率、实际利率、教育水平、城镇人口占总人口的比例、社会保障、人口死亡率、金融发展水平和其它影响因素,其它因素主要是制度方面的因素:如政策、法律等等,但是这些因素中有些数据难以获得,并且其它因素难以量化。因此,尽管在实证研究中可以分析的影响因素很多,但由于数据不易获得和难以量化,本文将最主要的影响寿险的因素进行保留,概括起来可以分成两大类:人口因素、经济及金融因素。人口因素主要指年龄、期望寿命、教育程度和赡养(抚养)率等与人口结构相关的因素;经济及金融因素主要包括收入、财富和价格水平等与经济有关的因素:

1.1 国内生产总值(GDP)

理论研究表明,寿险需求与收入密切相关。GDP作为反映一国国民收入的一项宏观指标,自然成为影响寿险需求的重要因素,国内生产总值与寿险需求正相关。本文的数据均出自中华人民共和国国家统计局网站^[5]。

1.2 小孩赡养率、老人抚养率

小孩赡养率、老人抚养率简称赡养(抚养)率,也称负担少儿系数和负担老年系数,是指一国14岁以下少儿人口和65岁以上老年人口数目之和比15岁至65岁的人口数,该比率反映劳动年龄人口对被

2009年11月9日收到 国家自然科学基金(70571052)资助
第一作者简介:白 丽(1982—),女,上海交通大学安泰经济与管理学院硕士生,研究方向:新产品开发和知识管理。E-mail:bellyhome@gmail.com。

赡养人口的负担情况。随着人均寿命的提高,以及计划生育的实施,独生子女家庭增加,我国人口逐渐向老龄化发展。我国是一个具有“尊老爱幼”传统的国家,人口的老龄化和独生子女的增加可能会影响寿险需求。赡养(抚养)率数据全部来自于中华人民共和国国家统计局网站^[5]。

1.3 利率

人寿保险作为一项金融产品,利率的影响不容忽视。利率有名义利率、实际利率之分。一般经济理论认为实际利率而非名义利率影响寿险需求。名义利率选择银行一年期整存整取利率,真实利率为当年名义利率减当年通货膨胀率,当年通货膨胀率可以用历年居民消费价格指数 CPI 来衡量。本文的名义利率、居民消费价格指数均来自于中华人民共和国国家统计局网站^[5]。

1.4 预期通货膨胀率

人寿保险一般都具有长期性,未来的通货膨胀将影响保单的价值,使人寿保险的吸引力下降。因

此,预期通货膨胀率与寿险需求负相关。预期通货膨胀率采用过去四年平均通货膨胀率,当年通货膨胀率可以用历年居民消费价格指数 CPI 来衡量,该数据出自均来自于中华人民共和国国家统计局网站^[5]。

1.5 教育水平

受教育程度较高的人群,其收入水平相对较高,同时这类人群厌恶风险的程度也相对较高,因而教育水平的高低与寿险需求正相关。可以用普通高校毕业生人数来衡量,该数据出自均来自于中华人民共和国国家统计局网站^[5]。

2 模型的建立

2.1 样本数据

本文中寿险保费收入系寿险、人身意外伤害险、健康险三者总和。从中华人民共和国国家统计局网站上收集了 1990—2005 年寿险保费收入年度数据(见表 1)。

表 1 1982—2005 年相关因素数据

年份	国内生产总值 d/亿元	居民价格 指数	预期通货膨胀 率 i/%	名义利率 /%	实际利率 Y/%	抚养小孩率 c/%	赡养老人率 o/%	教育水平 e/个	保费收入 m/亿元
1990	18 667.8	3.1	10.6	8.64	(1.91)	45.73	10.84	649 440.0	59.97
1991	21 781.5	3.4	8.2	7.56	(0.61)	46.30	10.69	64 6537.0	82.70
1992	26 923.5	6.4	7.7	7.56	(0.16)	45.03	10.58	62 9692.0	143.53
1993	35 333.9	14.7	6.9	10.08	3.18	43.87	10.13	599 214.0	196.73
1994	48 197.9	24.1	12.2	10.08	(2.07)	42.03	10.43	665 047.0	202.79
1995	60 793.7	17.1	15.6	10.08	(5.50)	40.16	10.06	836 877.0	196.47
1996	71 176.6	8.3	16.1	9.18	(6.87)	38.49	10.32	878 652.0	301.73
1997	78 973.0	2.8	13.1	5.67	(7.41)	36.74	10.35	875 539.0	633.64
1998	84 402.3	-0.8	6.8	4.59	(2.26)	35.61	10.89	877 077.0	767.40
1999	89 677.1	-1.4	2.2	2.25	0.03	34.99	11.15	902 270.0	872.10
2000	99 214.6	0.4	0.3	2.25	2.00	22.89	6.96	1 008 567.0	997.47
2001	109 655.2	0.7	-0.3	2.25	2.52	32.71	10.15	1 104 109.0	1 423.96
2002	120 332.7	-0.8	-0.3	1.98	2.26	30.17	11.57	1 418 141.0	2 274.84
2003	135 822.8	1.2	0.4	1.98	1.61	28.58	11.96	198 8091.0	3011.00
2004	159 878.3	3.9	1.3	2.25	1.00	26.76	11.87	2 541 777.0	3 194.00
2005	183 084.8	1.8	1.5	2.25	0.72	27.39	12.71	3 257 684.0	3 649.00

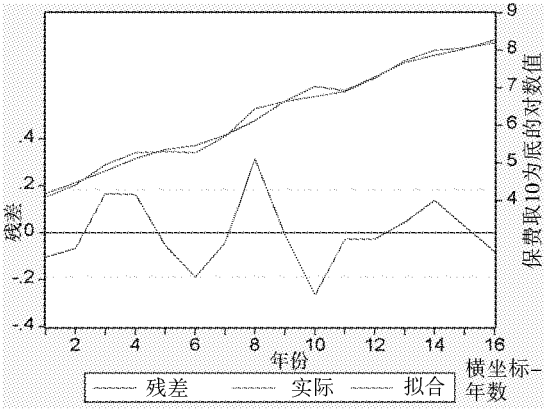
数据来源:中华人民共和国国家统计局网站

2.2 预设模型和最小二乘法估计

最小二乘法的模型为:

$$\lg m = \beta_0 + \beta_1 \lg d + \beta_2 i + \beta_3 r + \beta_4 \lg e + \beta_5 c + \beta_6 o + \varepsilon_t$$
其中,lgm 表示对保费收入取以 10 为底的对数,lgd

为对国内生产总值(亿元)取以10为底的对数, lge 为对教育水平毕业人数取以10为底的对数。 ε_t 表示残差,数据采用了(1990—2005)年共16年的数据。



蓝色线—残差,红色线—实际曲线,绿色线—拟合曲线

图1 最小二乘法拟合曲线

表2 最小二乘法输出结果

变量	系数	标准差	t 统计量	显著性水平检验
β_0	-6.136 654	7.808 576	-0.785 886	0.4521 **
lgd	1.287 180	0.256 542	5.017 429	0.000 7 **
i	-0.062 018	0.039 565	-1.567 503	0.151 4 **
r	-0.015 732	0.047 912	-0.328 346	0.750 2 **
lge	-0.153 067	0.427 700	-0.357 883	0.728 7 **
c	-0.031 072	0.049 180	-0.631 814	0.543 2 **
o	0.163 126	0.129 405	0.129 405	0.239 2 **
拟合优度	0.988 544	调和拟合优度 0.980 906		
杜宾-瓦特森检验	1.725 929	F 统计值 129.433 6		

说明:1. 表中数据都是数值;2. β_0 为常数项, lgd 和 lge 表示分别对变量 d 和 e 取以10为底的对数;3. *, **, *** 分别表示10%, 5%和1%的显著性水平。

图1的拟合曲线与实际曲线拟合的效果不是很好,残差曲线波动较大,在表2的输出结果中,除了 lgd 通过了变量显著性水平检验,其余的解释变量均没有通过显著性水平检验,说明,模型还需要进一步改进。

3 时间序列分析

3.1 单位根检验

分别对被解释变量和解释变量做单位根检验(Augmented Dickey-Fuller Test,简称ADF检验),包

括加入时间常数,一阶差分和时间趋势项^[4]。得出的结论如下。

对于保费收入变量 lgd 的单位根检验,原假设 $H_0: \rho = 1$,在5%和10%的临界值均大于ADF统计值-7.231 9(如表3所示),因此拒绝原假设,认为保费收入不存在单位根;对于保费收入是否存在时间趋势项,原假设 $H_0: \beta = 0$,在5%临界值为2.85小于 t 统计值7.019 4,因此拒绝原假设,即保费收入存在时间趋势项。对此,经济上的解释可以理解

表3 保费收入变量 lgd 的单位根检验		
ADF 检验统计	显著性水平检验	t 检验量
0.000 3	-7.231 938	0.000 3
1% 水平		-4.800 080
5% 水平		-3.791 172
10% 水平		-3.342 256

说明:临界值代表在不同的显著性水平下,当ADF值大于临界值时,相应的变量序列有单位根

表4 保费的时间趋势检验				
变量	系数	标准差	t 统计量	显著性水平检验
$lgm(-1)$	1.383 588	0.191 178	-7.231 938	0.000 0
$D(lgm(-1))$	0.905 561	0.157 481	5.750 292	0.000 2
Cons	5.505 654	0.725 419	7.589 617	0.000 0
$TREND(lgm)$	0.385 635	0.054 938	7.019 405	0.000 0
拟合优度	0.852 453	调和拟合优度 0.808 189		
杜宾-瓦特森检验	2.215660	F 统计值 19.25834		

说明:1. $lgm(-1)$ 表示 lgm 的一阶时间滞后项, D 表示变量的一阶差分, $TREND$ 表示检验保费收入是否具有时间趋势项,Cons表示常数项;2. *, **, *** 分别表示10%, 5%和1%的显著性水平。

(2)按上述方法,分别对变量 lgd, i, r, lge, c, o 做单位根检验,并加入时间常数、一阶差分和时间趋势,将结果整理成表格如表5。

表5 单位根检验结果				
变量	ADF 检验	临界值	时间趋势项检验	临界值
lgd	-5.731 5	-4.800 0 ***	4.876 3	2.85 **
i	-5.921 8	-4.992 2 ***	-6.243 1	2.85 **
r	-7.432 7	-4.992 2 ***	6.747 7	2.85 **

续表

变量	ADF 检验	临界值	时间趋势项检验	临界值
lge	-0.966 9	-3.342 2 *	1.696 4	2.85 **
c	-4.290 8	-3.759 7 ** ; -4.7282 *	-3.940 7	2.85 **
o	-2.617 5	-3.324 9 ***	1.265 8	2.85 **

由表 5 中数据得知,变量 lge 和 o 也没有通过单位根(10%)的检验,即接受了原假设,因此 lgm、lge、o 都存在单位根,需要进行一阶差分或二阶差分。在 Eviews 软件中分别对这三个序列分别进行时滞为 3 和 5 单位根检验,发觉它们的一阶差分都拒绝原假设,即这三个序列都应该用一阶差分进行回归。

对于时间趋势项检验,lgm、lge 和 r 都大于临界值 2.85,拒绝原假设,即模型包含时间趋势项。

3.2 时间序列模型

改写原模型:

$$\lg m = \beta_0 + \beta_1 D(\lg d) + \beta_2 c + \beta_3 D(o) + \beta_4 r + \beta_5 i + \beta_6 D(\lg e) + \beta_7 \lg m_{t-1} + \beta_8 \lg d_{t-1} + \beta_9 r_{t-1} + \varepsilon_t$$

对上述模型做最小二乘估计后,重新获得估计结果。

表 6 第一次改进模型后的回归输出结果

变量	系数	标准差	t 统计量	显著性水平检验
β_0	-12.842 89	2.353 954	-5.455 881	0.005 5 **
$D(\lg d)$	8.364 120	0.903 800	9.254 392	0.000 8 **
c	0.021 021	0.011 138	1.887 400	0.132 1 **
$D(o)$	0.086 309	0.015 756	5.477 720	0.005 4 **
r	-0.253 131	0.028 220	-8.969 998	0.000 9 **
i	-0.202 398	0.018 262	-11.082 76	0.000 4 **
$D(\lg e)$	2.416 731	0.553 456	4.366 619	0.012 0 **
$\lg m(-1)$	0.009 333	0.117 309	0.079 562	0.940 4 **
$\lg d(-1)$	1.616 408	0.238 582	6.775 509	0.002 5 **
$r(-1)$	-.086 890	0.017 109	-5.078 696	0.002 5 **
拟合优度	0.999 235		调和拟合优度	0.997 515
杜宾-瓦特森检验	3.633 977		F 统计值	580.775 0

由表 6 可见,拟合优度得到了改善,几个解释变量如 lgd 的一阶差分项 $D(\lg d)$,真实利率项 r,预期通货膨胀率项 i 等的显著性检验也得到了改进,但保费收入变量 lgm 的一阶时间滞后项 $\lg m(-1)$ 和变量 c 的显著性检验还是不够理想,尝试去掉一阶时间滞后项 $\lg m(-1)$,得到拟合曲线和输出结果如

图 2。

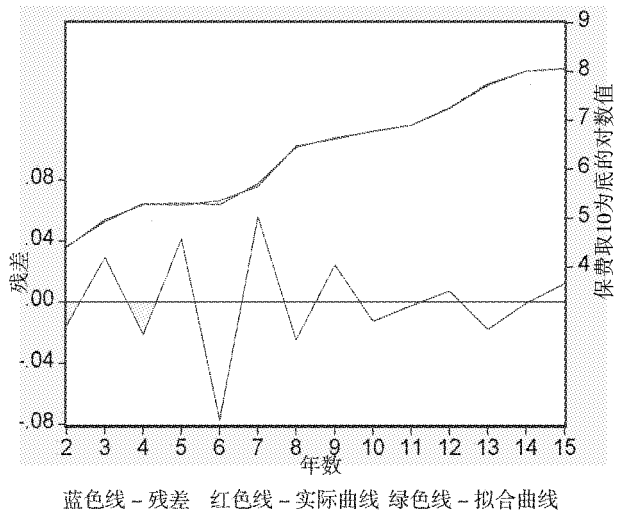


图 2 拟合结果

此时为了验证之前去掉一阶时间滞后项 $\lg m(-1)$ 项是否合理,对该项进行冗余变量 (Redundant Variables) 检验,得到 F 统计值为 0.006 33,接受原假设,认为一阶时间滞后项 $\lg m(-1)$ 项的确是冗余变量。

表 7 第二次改进模型后的回归输出结果

变量	系数	标准差	t 统计量	显著性水平检验
β_0	-12.947 64	1.746 711	-7.412 585	0.000 7 **
$D(\lg d)$	8.409 844	0.624 397	13.468 74	0.000 0 **
c	0.020 841	0.009 761	2.135 201	0.085 8 *
$D(o)$	0.086 196	0.014 046	6.136 691	0.001 7 **
r	-0.253 658	0.024 556	-10.329 71	0.000 1 **
i	-0.203 044	0.014 644	-13.865 59	0.000 0 **
$D(\lg e)$	2.431 166	0.468 043	5.194 324	0.003 5 **
$\lg d(-1)$	1.631 169	0.134 277	12.147 81	0.000 0 **
$r(-1)$	-.087 261	0.014 732	-5.923 154	0.002 0 **
拟合优度	0.999 234		调和拟合优度	0.997 515
杜宾-瓦特森检验	3.633 977		F 统计值	815.423 4

由表 7 第二次改进的输出结果可以看到,基本上每个解释变量都通过了显著性检验,变量 r 也通过了 10% 的显著性检验。因此模型最终定为:

$$\lg m = \beta_0 + \beta_1 D(\lg d) + \beta_2 c + \beta_3 D(o) + \beta_4 r + \beta_5 i + \beta_6 D(\lg e) + \beta_7 \lg d_{t-1} + \beta_8 r_{t-1} + \varepsilon_t$$

4 结论

通过对比最小二乘法模型、时间序列以及改进的模型,我们得到如下的结论:

(1) 国内生产总值的增长是寿险保费收入增长的主要原因。国内生产总值的增长带来国民收入的增加,随着收入的增加,人们对人寿保险的需求也增加。本文中逐步改进的模型都是 GDP 系数最大,有力的说明了我国寿险业高速增长是建立在 GDP 增长的基础上。

(2) 前期寿险保费收入对当期保费收入影响不是很大,从第一次的改进模型到第二个的改进模型过程中,一方面说明了虽然保费收入有延续性,长期寿险保单有部分分期缴纳,但是另一方面也说明了保费收入有滞后效应对寿险需求影响不是很大。

(3) 我国不断下降的少年儿童赡养(抚养)率,对寿险需求构成了显著的影响,因为随着我国独生子女政策的实施,小孩在家庭中的地位越来越重要。父母为未成年子女购买保险促进了寿险的增长,并且是正相关的。我国是一个具有“尊老爱幼”传统的国家,并且现在的子女都对赡养父母越来越重视,并且由于生活水平的提高,收入提高,子女为老人买寿险的需求大大增强,加上保险公司不断开发了针对老年人的寿险产品品种,子女、或老人本身购买寿

险产品也越来越方便,因而社会的老龄化对寿险需求的有了显著影响。

(4) 大学生在校人数的增加,受教育的水平提高对寿险需求产生显著的影响。

(5) 通过参阅相关资料和有关的研究证明^[7]: 实际利率(而非名义利率)应该对寿险需求有正面的影响,实际利率增加寿险需求增加。但本文中是负相关,这可能是模型还存在一定的偏误,或者数据量不够的缘故。因为在实际生活中,由于寿险产品具有储蓄性质,实际利率增加时,人们对储蓄的需求增加,因而对寿险的需求也相应增加。

参 考 文 献

- 1 卓 志. 人寿保险的经济分析引论. 北京: 中国金融出版社, 2001
- 2 张文璋. Eviews 在计量经济学中的应用. <http://www.pinggu.org/BBS/thread-267697-1-1.html>, 2005
- 3 贾罗 R A, 马斯科西莫 V, 津巴 W T. 金融经济学手册. 吴文峰, 仲黎明, 冯 芸, 译. 上海: 上海人民出版社, 2007
- 4 (美)威廉. H. 格林. 计量经济分析. 第五版. 费剑平, 译. 北京: 中国人民大学出版社, 2007
- 5 中华人民共和国国家统计局. 年度数据: 金融和保险. <http://www.stats.gov.cn/tjsj/ndsj/>, 1990—2005
- 6 吕振通, 张凌云. SPSS 统计分析与运用. 北京: 机械工业出版社, 2009
- 7 江生忠. 2003 年中国保险业发展报告. 天津: 南开大学出版社, 2003: 46—47

Empirical Study on Life Insurance Demand Model Based on Time Series Analysis

BAI Li, LI Hai-gang

(Shanghai Jiaoong University, Antai College of Economics & Management, Shanghai 200052, P. R. China)

[Abstract] The Ordinary Least Square (OLS) and time series analysis on the life insurance demand model are empirically applied. Compared with the existing research, a multiple linear regression model and an improved one are created, which indicates that the growth of gross domestic product is main contributors to the development of life insurance. Also, improving education environment and supporting the old ratio and raising the young ratio affect life insurance demand significantly. Meanwhile, there is an apparent time lag effect on premium income by ADF test as result of installment payment of premiums.

[Key words] life insurance demand ordinary least square time series model ADF test