Economics Letters 124 (2014) 520–522

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

# Is there adverse selection in the life insurance market? Evidence from a representative sample of purchasers

Timothy Harris [1], Aaron Yelowitz *

*Department of Economics, 335 Gatton Business & Economics Building, University of Kentucky, Lexington, KY, 40506-0034, USA*

## HIGHLIGHTS

- We use a representative sample of life insurance purchasers from the SIPP.
- We find that 62% of life insurance purchasers are under age 50.
- Individuals who have higher mortality are no more likely to hold life insurance.
- This analysis finds no compelling evidence for adverse selection for life insurance.
- Underwriting and institutional features may mitigate adverse selection.

## ARTICLE INFO

*Article history:*
Received 12 May 2014
Received in revised form
26 July 2014
Accepted 26 July 2014
Available online 4 August 2014

*JEL classification:*
D82

*Keywords:*
Adverse selection
Advantageous selection
Life insurance

## ABSTRACT

This paper examines asymmetric information in the life insurance market using data that link life insurance holdings with death records for a representative sample of purchasers. This analysis finds no compelling evidence for adverse selection in a broad age cohort.



## 1. Introduction

Recent work exploring asymmetric information in insurance markets has emphasized both the roles of adverse and advantageous selection (Einav and Finkelstein, 2011). The life insurance market is unique for high take-up and low reliance on government involvement or mandates relative to markets for health, auto, and long-term care insurance, as well as annuities. Several recent published studies – all using the Health and Retirement Survey (HRS) – have reached differing conclusions about the degree or existence of adverse selection in the life insurance market (Cawley and Philipson—hereafter CP, 1999; He, 2009, 2011).[2] Although this data sheds light on the degree of adverse selection, its sample consists of a cohort aged 51–61 in 1992, somewhat older than the typical life insurance holder. He (2009) notes that "evidence for or against asymmetric information among this cohort may not be representative of what one may find in other cohorts". Furthermore, He (2011) stated that individuals older than middle fifties have "passed their peak need for life insurance". We examine adverse selection using a much broader age range contained in the Survey of Income and Program Participation (SIPP). When the SIPP sample is restricted to the age cohort examined in the HRS, the analysis finds similar results to previous work. However, when we examine broader age groups there is no compelling evidence of adverse selection.

## 2. Data

The data come from the 1990 and 1991 panels of the SIPP. This nationally-representative longitudinal sample is constructed through individual interviews in four-month intervals known as "waves". The 1990 and 1991 samples follow individuals for eight

* Corresponding author. Tel.: +1 859 257 7634; fax: +1 859 323 1920.
*E-mail addresses:* tim.harris@uky.edu (T. Harris), aaron@uky.edu (A. Yelowitz).
[1] Tel.: +1 801 710 9910.

[2] Hedengren and Stratmann (2012) find evidence of advantageous selection using a restricted-use version of the SIPP. However, they do not address the concerns of He (2009, 2011).

http://dx.doi.org/10.1016/j.econlet.2014.07.029
0165-1765/

**Table 1**
Summary statistics.

| | Full sample | Age 25–50 | Age 51–61 | Age 62+ |
|---|---|---|---|---|
| Age | 52.3 | 36.9 | 56.1 | 71.9 |
| Male | 49.3% | 50.1% | 51.8% | 47.2% |
| White (non-Hispanic) | 77.7% | 73.0% | 78.8% | 83.6% |
| Mortality (Died by 1996) | 7.3% | 0.5% | 3.2% | 18.5% |
| Has term life insurance? | 19.7% | 21.0% | 22.9% | 16.5% |
| Observations | 17,600 | 8725 | 2550 | 6325 |

Note: Data from 1990 and 1991 SIPP panels.

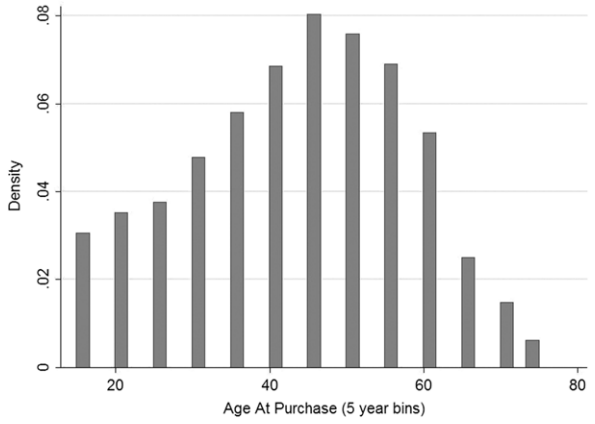


**Fig. 1.** New life insurance purchases from the 2008 SIPP panel.

waves. Each wave contains responses regarding income, labor force activity, and participation in government assistance programs. In addition to the "core" monthly questions, the survey covers less-frequently asked subjects in "topical modules". The wealth topical modules contain detailed information on assets and liabilities (including individual life insurance holdings) and are asked twice per panel. Another topical module asks about health conditions such as cancer, stroke, and high blood pressure, along with an indicator for self-reported disability or illness.

The key motivation for using the SIPP data from 1990 and 1991 comes from the unique availability of mortality records from the Social Security Administration's Master Beneficiary Record (MBR) matched to the public data. The MBR reports deaths through 1996 with a high degree of accuracy. Table 1 contains summary statistics for the applicable sample and subsamples. Although the death record window – as long as six years after the introduction of the SIPP panel – is relatively short, He (2009) notes that "the period during which buyers are most likely to take advantage of their private information is 4–6 years before death".

Following CP, our primary focus is on individual term life insurance holdings. We abstract away from group and whole life insurance markets due to the confounding factors of group rating and investment mechanisms.

The SIPP has an important advantage because it samples all relevant age groups that purchase life insurance. Fig. 1 shows the distribution of new purchasers using the SIPP panel from 2008.[3] It shows that a majority (62%) of life insurance purchasers are age 50 and under.[4] Therefore, a more thorough analysis of adverse selection in life insurance markets can be ascertained using SIPP data.

## 3. Empirical setup

Following the setup of CP, the following logit model is estimated to gauge the extent or existence of adverse selection.

$$Prob\,(individual\_term_i = 1) = \Lambda(\alpha_0 + \alpha_1\ mortality_i + \alpha_2 X_i). \quad (1)$$

In this model, $individual\_term_i$ is a latent indicator variable for having any individual term life insurance and $mortality_i$ is an indicator variable for mortality from 1990 to 1996.[5] $X_i$ is a vector of covariates which includes demographic, financial, health and bequest variables. The SIPP data lacks some important health variables including smoking status, which are used to price insurance premia. $\Lambda(\cdot)$ is the logistic cumulative distribution function. In order to reject the null hypothesis that market participants possess symmetric information, $\alpha_1$ must be greater than zero.

He (2009) questioned the validity of the CP findings using a sample of existing life insurance holders. She argued that there was survivorship bias associated with high-risk individuals dying and dropping out of the sample. Therefore, she looked at only those individuals who purchased life insurance. Following her general setup, we run regressions where $newbuyer_i$ is unity for someone who did not have life insurance in the initial survey and then reported having life insurance the following year.[6]

$$Prob\,(newbuyer_i = 1) = \Lambda(\beta_0 + \beta_1\ mortality_i + \beta_2 X_i). \quad (2)$$

In later work, He (2011) hypothesized that there is dynamic adverse selection in the life insurance market from individuals with better health letting their policies lapse. Following her work, Eq. (3) sets up a regression to look at how actual mortality correlates with lapses in contracts.

$$Prob\,(lapse_i = 1) = \Lambda(\gamma_0 + \gamma_1\ mortality_i + \gamma_2 X_i). \quad (3)$$

The analysis undertaken here differs from He (2011) in that she used voluntary cancellation as the dependent variable whereas the subsequent analysis uses a combination of voluntary or end-of-policy lapses as the dependent variable.[7]

## 4. Results

Table 2 reports the findings from each regression under four specifications, each of which successively includes more covariates. The baseline regression contains the following covariates: mortality, age, indicator variables for male, white (non-Hispanic), and highest level of education. Finance adds income, net worth, employment status, industry, and census region. Bequest adds indicators for married and children. Health augments the regression

[3] The 2008 panel has questions regarding life insurance in topical modules 4, 7, and 10 that represent interviews in September to December of 2009, 2010, and 2011. A new purchaser is defined as an individual who did not have life insurance in the initial survey and then reported having life insurance the following year.

[4] Industry surveys indicate that two-thirds of first-time life insurance buyers are under the age of 40. See http://www.slideshare.net/MarionGuthrie/first-time-life-insurance-buyers (accessed 4/8/2014).

[5] Observations with deaths prior to the sample year were omitted do to "probable mismatches" as described in the SIPP data dictionaries.

[6] SIPP topical modules 4 and 7 in the 1990 and 1991 panels both contain information about life insurance. Questions regarding term versus whole and employer versus individual are staggered between waves 4 and 7. Consequently, when looking at new buyers (and lapses) the dependent variable is restricted to be either new buyer of term or new buyer of non-employer provided life insurance.

[7] *Lapse_i* is unity if an individual had life insurance in wave 4 and did not have life insurance in wave 7.

**Table 2**
Logit results.

| $mortality_i$ coefficient | Life insurance coverage | | | New purchases | | | Lapsation | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) Age 51–61 | (2) Age 25–50 | (3) Age 62+ | (4) Age 51–61 | (5) Age 25–50 | (6) Age 62+ | (7) Age 51–61 | (8) Age 25–50 | (9) Age 62+ |
| Baseline | −0.484 | −0.600 | −0.101 | 0.267 | −0.275 | −0.269 | 0.571 | 0.922 | 0.510 |
| | (0.317) | (0.491) | (0.098) | (0.396) | (0.551) | (0.144) | (0.400) | (0.518) | (0.115) |
| | [−0.070] | [0.071] | [−0.012] | [0.044] | [−0.040] | [−0.034] | [0.088] | [0.183] | [0.072] |
| Baseline + finance | −0.430 | −0.352 | −0.073 | 0.308 | −0.046 | −0.272 | 0.659 | 0.842 | 0.496 |
| | (0.325) | (0.496) | (0.099) | (0.413) | (0.558) | (0.145) | (0.408) | (0.514) | (0.117) |
| | [−0.062] | [−0.044] | [−0.009] | [0.049] | [−0.007] | [−0.034] | [0.099] | [0.158] | [0.068] |
| Baseline + finance + bequest | −0.407 | −0.049 | −0.067 | 0.337 | −0.034 | −0.262 | 0.663 | 0.626 | 0.493 |
| | (0.329) | (0.500) | (0.099) | (0.416) | (0.559) | (0.146) | (0.412) | (0.515) | (0.117) |
| | [−0.058] | [−0.006] | [−0.008] | [0.054] | [−0.005] | [−0.033] | [0.097] | [0.111] | [0.067] |
| Base + finance + bequest + health | −0.413 | 0.172 | −0.047 | 0.367 | 0.200 | −0.242 | 0.607 | 0.543 | 0.502 |
| | (0.337) | (0.510) | (0.100) | (0.433) | (0.567) | (0.147) | (0.417) | (0.543) | (0.118) |
| | [−0.057] | [0.023] | [−0.006] | [0.059] | [0.031] | [−0.030] | [0.086] | [0.094] | [0.068] |
| Observations | 2546 | 9039 | 6321 | 1090 | 5341 | 2689 | 1460 | 4056 | 3609 |

Notes: standard errors in parentheses. Marginal effects in brackets.

with health insurance, self-reported disability, family medical expenses, and the following work limiting illnesses: cancer, heart trouble, high blood pressure, kidney problems, lung or respiratory trouble, paralysis, and stroke.

Column (1) replicates the findings of CP using the same age range as contained in the HRS. In addition, the SIPP data for 1990 and 1991 captures roughly the same cohort as those sampled in the 1992–1994 HRS. The results reported in column (1) fall within the upper and lower bound given by CP and are consistent in sign with the results of He (2009). Columns (2) and (3) report coefficients for other age ranges that are not statistically different from zero and in all but one case are negative. This provides evidence against adverse selection in all relevant cohorts of life insurance purchasers.

Columns (4)–(6) contain the regression results of Eq. (2) including a sample of all individuals who did not have life insurance at the time of the initial survey. The dependent variable, *newbuyer*, is one if an individual reported having any type of life insurance in a later period.[8] Column (4) again uses the same age group as the HRS sample and has a positive, statistically insignificant coefficient consistent in sign of with the findings of He (2009). However, columns (5) and (6) representing other ages indicate that actual mortality has a negative or insignificant effect on life insurance holdings. This finding provides evidence that the age group observed in the HRS is not fully representative of life insurance purchasers and that adverse selection is not prevalent across all groups.

The last test for adverse selection looks at those individuals who lapse their life insurance policies. Columns (7)–(9) contain the results from Eq. (3). Here the positive coefficient on *mortality* indicates that individuals who die are more likely to lapse their policies. In this analysis, there is no evidence for adverse selection and even some evidence in favor of advantageous selection in the oldest group of life insurance holders.

## 5. Conclusion

Using data on a representative sample of life insurance purchasers, we find no significant evidence of adverse selection. In virtually all specifications, those who have higher mortality are no more likely to hold life insurance. Although the empirical findings are consistent with the concept of advantageous selection, it is important to recognize the importance of underwriting in the life insurance market. All existing empirical analyses examine life insurance holdings, not applications. Insurers ask extensive questions and require medical exams prior to approval of an application. These institutional features suggest caution before claiming that applicants are advantageously selected; rather the underwriting process potentially screens out high-risk applicants who would otherwise obtain life insurance.

## References

Cawley, J., Philipson, T., 1999. An empirical examination of information barriers to trade in insurance. Amer. Econ. Rev. 89 (4), 827–846.
Einav, L., Finkelstein, A., 2011. Selection in insurance markets: theory and empirics in pictures. J. Econ. Perspect. 25 (1), 115–138.
He, D., 2009. The life insurance market: symmetric information revisited. J. Public Econ. 93 (10), 1090–1097.
He, D., 2011. Is there dynamic adverse selection in the life insurance market? Econom. Lett. 112, 113–115.
Hedengren, D., Stratmann, T., 2012. Adverse vs. advantageous selection in life insurance markets. Working Paper, Accessed from http://ssrn.com/abstract=2194376 (May 1, 2014).

[8] Similar results were produced looking at individuals who purchased only term life insurance as well as individuals who purchased only non-group rated life insurance. Due to data limitations, both cannot be observed simultaneously.