

Is There Dynamic Adverse Selection in the Life Insurance Market?

Daifeng He*

March 7, 2011

College of William and Mary

Abstract: This paper finds evidence of dynamic adverse selection in the life insurance market. Lower-risk individuals are more likely to cancel a policy, and to cancel one of greater face value conditional on cancellation, than are individuals with higher mortality risk.

Keywords: dynamic adverse selection, reclassification risk, mortality risk, lapse, life insurance.

JEL classification: D8

* Economics Department, College of William and Mary, Williamsburg, VA 23187, U.S.A; Phone: 757-221- 4645; Fax: 757-221-1175; dhe@wm.edu.

1. Introduction

In long-term insurance, customers who cannot commit to the contractual relationship may have an incentive to voluntarily cancel the contract (or “lapse”) if favorable information about their risk arrives over time. Such dynamic adverse selection would undermine the quality of the risk pool and boost premiums, impairing consumer welfare. Moreover, dynamic selection may explain why long-term contracts are missing in many insurance markets, subjecting customers to substantial reclassification risk (Cochrane 1995).¹ The presence of this dynamic selection has, however, rarely been empirically explored. One of the few exceptions is Finkelstein et al.'s (2005) examination of lapse behavior and *ex-post* nursing home utilization in the long-term-care insurance market.

Using data from the Health and Retirement Study (HRS), I examine whether dynamic selection is present in the life insurance market. I find that individuals with lower mortality risk are more likely to lapse a contract than are those with higher mortality risk; and that conditional on lapsation, lower-risk individuals appear to lapse policies of greater face value than do higher-risk individuals, although the estimate in the latter case is imprecise perhaps due to the small sample size. These results suggest that despite substantial front-loading in the life insurance market evidenced by Hendel and Lizzeri (2003)², consumers still engage in risk-based dynamic selection. This selection renders the remaining risk pool significantly sicker than otherwise. Because premiums must rise to cover the cost of this deteriorated risk pool, the market fails to provide full insurance against reclassification risk. My results echo Finkelstein et al. (2005), who provide evidence of substantial dynamic selection in long-term-care insurance.³

Interestingly however, once I control for observable information upon which insurers could have priced (but in practice typically do not), the above results disappear. Instead, I find that, after controlling for observable information revealed after underwriting, lower-risk individuals are not more likely to lapse, or to lapse a larger amount, than are otherwise statistically equivalent higher-risk individuals. This suggests that risk-based dynamic selection could have been eliminated by pricing on the same observable information. Doing so would, of

¹ For example, long-term contracts do not exist in health or auto insurance markets.

² Front-loading refers to the underwriting practice under which premiums are higher than actuarially fair in early periods and lower than actuarially fair in later periods. Hendel and Lizzeri (2003) argue that liquidity constraints are important factors in explaining incomplete front-loading.

³ Konezka and Luo (forthcoming), however, find that other factors such as financial considerations, rather than a reassessment of the health risk, explain most of the lapsation in long-term care insurance.

course, undermine the very purpose of providing insurance against reclassification risk in dynamic contracts.

2. Data and Empirical Implementation

I use the Health and Retirement Study (HRS) dataset. The HRS is a nationally representative longitudinal dataset containing rich information on demographics, health status, insurance ownership, and family dynamics. I use the 1996 - 2004 waves of the initial HRS cohort, members of which were born between 1931 and 1941.⁴ The cohort was interviewed every two years.

Table 1 presents summary statistics of the main variables. Eight percent of the sample reported cancelling life coverage in the 1996 or 1998 wave;⁵ six percent reported *voluntary* cancellation (or “lapse”) during the same period.⁶ Cancelled coverage had an average face value of about \$73,000. Voluntarily cancelled coverage had a slightly higher face value of \$77,000. By 2004, nine percent of the sample had died.

To examine the presence of dynamic selection on the extensive margin, I estimate the following Probit model:

$$\Pr(\text{lapse}_i = 1) = \Phi(\beta_0 + \beta_1 \text{mortality}_i) \quad (1)$$

where *lapse* is an indicator set to unity if the individual reported having voluntarily cancelled coverage in either the 1996 or 1998 wave, and zero otherwise. $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. *mortality* indicates whether the individual had died by 2004. It proxies the respondent’s mortality risk. Coefficient β_1 is the parameter of interest, measuring the correlation between mortality risk and the lapse decision. A significant and negative β_1 indicates the presence of dynamic adverse selection.

I also examine the lapse-mortality correlation on the intensive margin. To do so, I estimate the following model for those who reported a voluntary contract cancellation:

$$\text{lapse_amount}_i = \alpha_0 + \alpha_1 \text{mortality}_i + \varepsilon_i \quad (2)$$

⁴ 1996 was the first wave in which HRS began asking about life insurance lapsation behavior.

⁵ Because lapsation is relatively rare, I combine two waves together to obtain enough lapsation observations.

⁶ If a respondent reported cancelling or lapsing a contract, HRS asked whether the respondent herself or the provider, employer, or another individual cancelled the coverage.

Here, *lapse_amount* is the face value of a voluntarily cancelled contract and *mortality* is defined as in model (1). If dynamic selection is present on the intensive margin, α_1 is significantly negative. That is, conditional on lapsing a contract, those with lower mortality risk lapse a larger face value.

For comparison, I also estimate the following variants of models (1) and (2):

$$\Pr(\text{lapse}_i = 1) = \Phi(\beta_0 + \beta_1 \text{mortality}_i + \mathbf{X}_i \mathbf{B}) \quad (3)$$

$$\text{lapse_amount}_i = \alpha_0 + \alpha_1 \text{mortality}_i + \mathbf{X}_i \mathbf{\Pi} + \varepsilon_i \quad (4)$$

where *lapse*, *lapse_amount*, and *mortality* are defined the same as in models (1) and (2). \mathbf{X} is the vector of 1998 levels of the risk-classification variables which life insurers commonly consider when setting insurability and premiums at underwriting (thus, “pricing” variables).⁷

Specifically, \mathbf{X} includes the respondent’s: (a) age, gender, and smoking and drinking status (whether she has ever smoked, whether she smokes now, and whether she drinks now); (b) health status and medical history (whether she has been diagnosed with diabetes, high blood pressure, cancer, heart disease, arthritis, lung disease, or stroke; whether she has had a hospital stay in the previous 12 months; and whether her BMI classifies her as overweight, obese, or of a healthy weight);⁸ and (c) family history (whether her father or mother died before age 60).⁹

Note that the post-underwriting variables \mathbf{X} can be considered observable in the sense that insurers could have observed them in 1998 and employed those observations to adjust the insureds’ risk classifications and premium.¹⁰ In practice, however, these factors typically are not used to determine premiums after the initial underwriting. By controlling for the observable \mathbf{X} in models (3) and (4), I therefore can examine the counterfactual case in which insurers update their pricing variables and revised insurability or premiums accordingly. As discussed in Cardon and Hendel (2001, footnote 15), controlling for all information available to insurers, regardless of whether the information is actually priced upon, reveals the intrinsic nature of the market.

⁷ For the importance of controlling for risk classification when testing for *static* adverse selection, see Chiappori and Salanié (2003) and Finkelstein and McGarry (2006); for life insurance underwriting practices, see Cummins et al. (1983), McGill’s Life Insurance (2000), and He (2009).

⁸ In the BMI, 18.5-24.9 is healthy weight, 25-30 is overweight, and above 30 is obese.

⁹ These variables are crude indicators of unfavorable family medical history. HRS data do not, however, provide information about the cause of parents’ deaths.

¹⁰ Insurers had collected that information at initial underwriting. It therefore is difficult to imagine they would be unable to update it had they wished to do so.

Again, β_j and α_j are the coefficients of interest. I will compare their estimates to those obtained from models (1) and (2).¹¹

3. Results

Table 2 reports the results. Columns (1) and (2)'s dependent variable indicates a lapse, while columns (3) and (4)'s is the face value of a lapsed contract, measured in thousands of dollars. In column (1), the estimated marginal effect of *mortality* is negative and statistically significant, indicating that individuals with low mortality risk are more likely to lapse than are individuals with higher risk. Indeed, the point estimate of -0.02 indicates that those who died within the sample period were 33 percent less likely to have dropped life coverage than were those who survived beyond. Column (3)'s estimate of *mortality*'s effect in the lapse-amount regression is also negative, indicating that, conditional on lapsation, those who died early lapsed a smaller amount. Perhaps because of the small sample size, this estimate is not statistically significant. Together, however, these estimates suggest that dynamic selection does occur in the life insurance market.

When the 1998 post-underwriting observables are controlled for in column (2), the estimated *mortality* effect drops significantly relative to the column (1) estimate and is no longer statistically significant. This suggests that insurers could have avoided the risk-based selection had they re-set their premiums on the basis of the updated information. Column (4)'s estimate actually turns positive when controlling for these observables, similarly suggesting that dynamic selection would have been absent if insurers had priced on the updated observables. Note that consistent with the dynamic selection evidenced by columns (1) and (3), some life-style and health controls are significantly negative, indicating that those with less-desirable life styles or health conditions – for example smokers and cancer patients in column (2), and smokers and arthritis or stroke patients in column (4) – are less likely to cancel a contract.

Why do insurers not update premiums on the observables, even though by doing so they would have avoided the erosion of their risk pool by dynamic adverse selection? The answer lies in the tradeoff between insuring against reclassification risk and overcoming static adverse selection. If insurers had used the 1998 observables to update their pricing of previously written

¹¹ The author thanks the anonymous referee for his/her insightful comments, which led to this control-vs-no-control comparison.

contracts, the long-term life insurance market would have collapsed to a spot market in which no risk reclassification insurance would be provided. I find nevertheless that when insurers forgo pricing upon the post-underwriting observables for the sake of providing reclassification risk insurance, the market still fails to provide full reclassification risk insurance in the absence of consumer commitment mechanisms such as full front-loading.

4. Conclusion

This paper finds evidence of dynamic adverse selection in life insurance, the largest private individual insurance market. Using the HRS dataset, I find that individuals with lower mortality risk are more likely than higher-risk individuals to lapse a contract and to lapse a greater contract face value. This is consistent with the fact that, beyond their middle 50s, the sample individuals have passed their peak need for life insurance and thus likely will renew the contracts only if the odds are in their favor. Interestingly, such risk-based selection could have been avoided by pricing on the updated observables after initial underwriting. Insurers refrain from doing so precisely to reinforce dynamic insurance.

Acknowledgement:

I am grateful to an anonymous referee, Sebastian Galiani, Steve Buccola, Bart Hamilton, Peter McHenry, Jennifer Mellor, Alvin Murphy, and Bruce Petersen for their helpful comments and suggestions. Any errors are my own.

References:

- Chiappori, P., Salanié, B., 2003. Testing contract theory: a survey of some recent work. In: Dewatripont, M., Hansen, L., Turnovsky, P. (Eds.), *Advances in Economics and Econometrics - Theory and Applications*, Eighth World Congress, Econometric Society Monographs. Cambridge University Press, Cambridge, pp. 115-149.
- Cardon, J., Hendel, I., 2001. Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey. *The RAND Journal of Economics*, 32(3), 408-427.
- Hendel, I., Lizzeri, A., 2003. The role of commitment in dynamic contracts: evidence from life insurance. *Quarterly Journal of Economics*, 118(1), 299-327.
- Cochrane, J., 1995. Time-consistent health insurance. *Journal of Political Economy* 103(3), 445-473.
- Cummins, J. D., Smith, B.D., Smith, R.N., Vanderhei, J.L., 1983. *Risk classification in life insurance*. Kluwer-Nijhoff Publishing.
- Finkelstein, A., McGarry, K., Sufi, A., 2005. Dynamic inefficiencies in insurance markets: evidence from Long-Term Care Insurance. *American Economic Review Papers and Proceedings*, 95 (2), 224-228
- Finkelstein, A., McGarry, K., 2006. Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review* 96(4), 938-958.
- He, D., 2009. The life insurance market: symmetric information revisited. *Journal of Public Economics* 93(10), 1090-1097.
- Konetzka, R. T. and Luo, Y., forthcoming. Explaining lapse in long-term care insurance markets. *Health Economics*, in press.
- McGill's Life Insurance, 2000, third ed. Edward E. Graves, editor. Bryn Mawr, American College.

Table 1 Sample Summary Statistics

variable	definition	mean	std. dev
life	whether reported owning life insurance	0.74	0.44
general termination	whether reported having a life insurance contract cancelled in wave 1996 or 1998 (either the respondent's or another's decision)	0.08	0.026
voluntary termination	whether reported <i>voluntarily</i> cancelling a life insurance contract in wave 1996 or 1998 (the respondent's own decision)	0.06	0.24
general termination amount	face value of terminated contracts (\$000)	72.71	132.89
voluntary termination amount	face value of <i>voluntarily</i> terminated contracts (\$000)	76.63	145.08
mortality	whether dead by wave 2004	0.09	0.29

Note: Summary statistics are, unless noted, based on the 1998 wave of the HRS cohort. They are weighted with HRS individual sampling weights.

Table 2 Termination Decision and Mortality Risk

VARIABLES	(1)	(2)	(3)	(4)
	dep var= <i>lapse</i>		dep var= <i>lapse_amount</i> (\$1000s)	
mortality	-0.020** (0.010)	-0.012 (0.011)	-14.605 (22.704)	12.572 (21.048)
male		0.017** (0.007)		42.567** (18.565)
smoke_ever		0.002 (0.007)		-37.854* (22.411)
smoke_now		-0.020*** (0.007)		-18.552 (19.581)
drink		0.003 (0.006)		38.163** (18.012)
diabetes		0.002 (0.010)		1.991 (37.012)
HBP		0.002 (0.007)		29.547 (23.944)
cancer		-0.020** (0.009)		28.538 (36.819)
heart		-0.007 (0.009)		-4.307 (18.950)
arthritis		0.000 (0.006)		-35.393* (20.538)
lung		0.003 (0.013)		11.097 (31.406)
stroke		0.003 (0.017)		-38.816* (20.014)
hospital_stay		0.008 (0.009)		-8.737 (24.094)
healthyweight		0.067 (0.067)		-238.345*** (47.349)
overweight		0.049 (0.052)		-286.375*** (43.990)
obese		0.046 (0.062)		-300.144*** (55.365)
history_father		0.002 (0.008)		17.648 (25.005)
history_mother		0.005 (0.010)		35.089 (31.573)
Observations	6399	6231	341	333
R-squared	.	.	0.001	0.167
Pseudo R2	0.00123	0.0251	.	.

Notes: Columns (2) and (4) regressions also control for age dummies and a constant, the coefficients of which are suppressed. All estimations are weighted with HRS individual sampling weights. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1