Received: 10 August 2020 | Revised: 26 October 2020 | Accepted: 1 November 2020

DOI: 10.1111/jori.12332

ORIGINAL ARTICLE

The Journal of Risk and Insurance

Why do life insurance policyholders lapse? The roles of income, health, and bequest motive shocks

Hanming Fang[1,2,3] | **Edward Kung**[4]

[1]Department of Economics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2]School of Entrepreneurship and Management, ShanghaiTech University, Shanghai, China

[3]The National Bureau of Economic Research, Cambridge, Massachusetts, USA

[4]Department of Economics, California State University at Northridge, Northridge, California, USA

Correspondence
Hanming Fang, Department of Economics, University of Pennsylvania, 133 S. 36th St., Philadelphia, PA 19104, USA.
Email: hanming.fang@econ.upenn.edu

Funding information
National Science Foundation, Grant/Award Number: SES-0844845

Abstract

We present and empirically implement a dynamic discrete choice model of life insurance decisions to assess the importance of various factors in explaining life insurance lapsation. We estimate a model using information on life insurance holdings from the Health and Retirement Study. Counterfactual simulations using the estimates of our model suggest that a large fraction of life insurance lapsations are driven by idiosyncratic shocks, uncorrelated with health, income, and bequest motives, particularly when policyholders are relatively young. As the remaining policyholders get older, however, the role of such independent and identically distributed (i.i.d.) shocks gets smaller, and more of their lapsation is driven by income, health, or bequest motive shocks. As anticipated, income and health shocks are relatively more important than bequest motive shocks in explaining lapsation when policyholders are young, with bequest motive shocks playing a more important role as we age.

KEYWORDS

life insurance lapsations, sequential Monte Carlo method

JEL CLASSIFICATION

G22, L11

1 | INTRODUCTION

The life insurance market is large and important. According to Life Insurance Marketing and Research Association International (LIMRA International), 78% of American families owned some type of life insurance in 2004. By the end of 2008, the total number of individual life insurance policies in force in the United States stood at about 156 million; and the total individual policy face amount in force reached over 10 trillion dollars (see American Council of Life Insurers, 2009, pp. 63–74).

1.1 | Life insurance market

There are two main types of traditional individual life insurance products, term life insurance and whole life insurance.[1] A term life insurance policy covers a person for a specific duration at a fixed or variable premium for each year. If the person dies during the coverage period, the life insurance company pays the face amount of the policy to his/her beneficiaries, provided that the premium payment has never lapsed. The most popular type of term life insurance has a fixed premium during the coverage period and is called level term life insurance. A whole life insurance policy, on the other hand, covers a person's entire life, usually at a fixed premium. In the United States at year-end 2008, 54% of all life insurance policies in force were Term Life insurance. Of the new individual life insurance policies purchased in 2008, 43%, or 4 million policies, were term insurance, totaling $1.3 trillion, or 73%, of the individual life face amount issued (see American Council of Life Insurers, 2009, pp. 63–74). Besides the difference in the period of coverage, term and whole life insurance policies also differ in the amount of cash surrender value (CSV) received if the policyholder surrenders the policy to the insurance company before the end of the coverage period. For term life insurance, the CSV is zero; for whole life insurance, the CSV is typically positive and prespecified to depend on the length of time that the policyholder has owned the policy. One important feature of the CSV on whole life policies relevant to our discussions below is that by government regulation, CSVs do *not* depend on the health status of the policyholder when surrendering the policy.[2]

1.2 | Lapsation

Lapsation is an important phenomenon in life insurance markets. Both LIMRA and the Society of Actuaries consider a policy to lapse if its premium is not paid by the end of a specified time (often called the grace period).[3] According to LIMRA (2009, p. 11), the life insurance industry calculates the annualized lapsation rate as follows:

---

[1]The whole life insurance has several variations such as universal life (UL), variable life (VL), and variable-universal life (VUL). Universal life allows flexible premiums subject to certain minimums and maximums. For variable life, the death benefit varies with the performance of a portfolio of investments chosen by the policyholder. Variable-universal life combines the flexible premium options of UL with the varied investment option of VL (see Gilbert & Schultz, 1994).

[2]The life insurance industry typically thinks of the CSV from the whole life insurance as a form of tax-advantaged investment instrument (see Gilbert & Schultz, 1994).

[3]This implies that if a policyholder surrenders his/her policy for cash surrender value, it is also considered as a lapsation.

$$\text{Annualized policy lapse rate} = 100 \times \frac{\text{Number of policies lapsed during the year}}{\text{Number of policies exposed to lapse during the year}}.$$

The number of policies exposed to lapse is based on the length of time the policy is exposed to the risk of lapsation during the year. Termination of policies due to death, maturity, or conversion is not included in the number of policies lapsing and contributes to the exposure for only the fraction of the policy year they were in force. Table 1 provides the lapsation rates of individual life insurance policies, calculated according to the above formula, both according to face amount and the number of policies for the period of 1998–2008. Of course, the lapsation rates also differ significantly by the age of the policies. For example, LIMRA (2009, p. 18) showed that the lapsation rates are about 2%–4% per year for policies that have been in force for more than 11 years in 2004–2005.

1.3 | Reasons for lapsation have important welfare implications

Our interest in the empirical question of why life insurance policyholders lapse their policies is motivated by recent research on the effects of lapsation in life insurance and life settlements markets. It is well known that life insurance pricing is supported by lapsation but recent work has highlighted the fact that the reasons for lapsation have important welfare consequences. For example, Gottlieb and Smetters (2020) note that the efficiency implications of lapsation depend on whether policyholders lapse due to forgetfulness or income shocks, and also on whether these shocks are anticipated or unanticipated. Daily et al. (2008) and Fang and Kung (2010b, 2020) also showed that the efficiency implications of the growing secondary market for life insurance depends crucially on whether lapses are driven by loss in bequest motive or other factors. They showed that if policyholders' lapsation is driven only by the loss of bequest motives, then consumer welfare is unambiguously lower with a secondary market than without, but if lapsation is driven by income or liquidity shocks, then a life settlement market may potentially improve consumer welfare.[4]

To understand why the reasons for lapsation matters, it is important to remember that life insurance premiums are front-loaded, meaning in the early part of the policy period, the premium payments exceed the actuarially fair value of the risk insured, but in the later part of the policy period, the premium payments are lower than the actuarially fair value. As a result, policyholders who lapse after holding the policy for some time give up value, which the life insurance company pockets as a profit. Due to competition, these so-called lapsation profits are factored into the pricing of the life insurance policy to start with (Gilbert & Schultz, 1994), and so policyholders who lapse end up cross-subsidizing policyholders who do not. The welfare implications of this cross-subsidization depend on the marginal utility of income of lapsers relative to entire pool (since everyone benefits from lower pricing), and it therefore depends on the reasons for lapsation. If lapsation happens mainly for idiosyncratic reasons, such as a loss in

[4]Related, Fang and Wu (2020) showed that when policyholders are overconfident about the strength of their bequest motive at the time of purchasing their life insurance policy, they will underestimate their probability of lapsation, and end up being exploited by the life insurance by purchasing "too much" risk reclassification insurance. A life settlements market can potentially improve consumer welfare by imposing a limit on the extent to which primary insurers can exploit overconfident consumers.

TABLE 1 Lapstion rates of individual life insurance policies, calculated by face amount and by number of policies: 1998–2008

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
By face amount	8.3	8.2	9.4	7.7	8.6	7.6	7.0	6.6	6.3	6.4	7.6
By number of policies	6.7	7.1	7.1	7.6	9.6	6.9	7.0	6.9	6.9	6.6	7.9

Source: American Council of Life Insurers (2009).

bequest motive, and if lapsers have relatively high income relative to the pool (including, e.g., young parents buying life insurance) then the cross-subsidization is welfare-improving and a reduction in the ability of life insurance companies to benefit from lapsation profits, such as due to competition from secondary markets, reducing the welfare benefits of the cross-subsidization. On the other hand, if lapsation primarily happens due to income and health shocks which raise the marginal utility of income of lapsers (e.g., if someone lapses to pay for health costs), then a secondary market may be welfare improving, especially if the shocks are unanticipated.

Our paper complements recent work on the determinants of lapsation by Fier and Liebenberg (2013), Cole and Fier (2020), and Gottlieb and Smetters (2020). A common finding is that lapsation is related to income shocks. In a survey of the universe of TIAA customers, Gottlieb and Smetters (2020) found that two theories account for the majority of lapses: policyholders often forget to pay their premiums and many policyholders underestimate their future need for money. Our paper contributes to this literature a methodology to use observational data and account for serially correlated unobservables to *quantify* the importance of the different factors driving lapsation patterns over the tenure of the life insurance policy.

1.4 | What do we do in this paper?

For this purpose, we present and empirically implement a dynamic discrete choice model of life insurance decisions. The model is "semistructural" and is designed to bypass data limitations where researchers only observe whether an individual has made a new life insurance decision (i.e., purchased a new policy, or added to/changed an existing policy) but do not observe what the actual policy choice is nor the choice set from which the new policy is selected. We empirically implement the model using the limited life insurance holding information from the Health and Retirement Study (HRS) data. An important feature of our model is the incorporation of serially correlated unobservable state variables. In our empirical analysis, we provide ample evidence that such serially correlated unobservable state variables are necessary to explain some key features in the data.

Methodologically, we deal with serially correlated unobserved state variables using posterior distributions of the unobservables simulated using Sequential Monte Carlo (SMC) methods.[5] Relative to the few existing papers in the economics literature that have used similar

[5]See also Norets (2009) which develops a Bayesian Markov Chain Monte Carlo procedure for inference in dynamic discrete choice models with serially correlated unobserved state variables. Kasahara and Shimotsu (2009) and Hu and Shum (2012) present the identification results for dynamic discrete choice models with serially correlated unobservable state variables.

SMC methods, our paper is, to the best of our knowledge, the first to incorporate multi-dimensional serially correlated unobserved state variables. To give the three unobservable state variables in our empirical model their desired interpretations as unobserved income, health, and bequest motive shocks respectively, we propose two channels through which we can anchor these unobservables to their related observable counterparts. We also discuss how the additional unobservable state variables significantly improve our model fit.

Our estimates for the model with serially correlated unobservable state variables are sensible and yield implications about individuals' life insurance decisions consistent with the both intuition and existing empirical results. In a series of counterfactual simulations reported in Table 10, we find that a large fraction of life insurance lapsations are driven by idiosyncratic shocks which are uncorrelated with health, income, and bequest motives, particularly when policyholders are relatively young. But as the remaining policyholders get older, the role of such independent and identically distributed (i.i.d.) shocks gets smaller, and more of their lapsations are driven either by income, health or bequest motive shocks. Income and health shocks are relatively more important than bequest motive shocks in explaining lapsation when policyholders are young, but as they age, the bequest motive shocks play a more important role. We discuss the implications of these findings on the effects of life settlement markets on consumer welfare.

The remainder of the paper is structured as follows. In Section 2 we describe the data set used in our empirical analysis and describe how we constructed key variables, and we also provide the descriptive statistics. In Section 3 we present the empirical model of life insurance decisions. In Section 5 we estimate the dynamic model with serially correlated unobserved state variables, describe the SMC method to account for them in estimation, and provide the estimation results.[6] In Section 6 we report our counterfactual experiments using the model with unobservables. In Section 8 we conclude.

2 | DATA

We use data from the HRS. The HRS is a nationally representative longitudinal survey of older Americans which began in 1992 and has been conducted every 2 years thereafter.[7] The HRS is particularly well suited for our study for two reasons. First, the HRS contains rich information about income, health, family structure, and life insurance ownership. If family structure can be interpreted as a measure of bequest motive, then we have all the key factors motivating our analysis. Second, the HRS respondents are generally quite old: between 50 and 70 years of age in their first interview. As we described in the introduction, the life settlements industry typically targets policyholders in this age range or older, so it is precisely the lapsation behavior of this group that we are most interested in.

Our original sample consists of 4512 male respondents who were successfully interviewed in both the 1994 and 1996 HRS waves, and who were between the ages of 50 and 70 in 1996. We chose 1996 as the period to begin decision modeling because the 1996 wave is *the first time* the HRS began to ask questions about whether or not the respondent lapsed any life insurance

[6]In the online appendix, we also present the estimates from a dynamic model without serially correlated unobservable state variables, and show via simulations that the dynamic model without serially correlated unobservable state variables fails to replicate some important features of the data.

[7]See http://hrsonline.isr.umich.edu/concord, for the survey instruments used in all the waves of HRS.

policies and whether or not the respondent obtained any new life insurance policies since the last interview. As we will explain below, these questions are used prominently in the construction of the key decision variable used in our structural model. We use only respondents who were also interviewed in 1994 so we can know whether or not they owned life insurance in 1994.

We follow these respondents until 2006. Any respondent who missed an interview for any reason other than death between 1996 and 2006 was dropped from the sample. Any respondent with a missing value on life insurance ownership any time during this period was also dropped. This leaves us a sample of 3567 males. We also dropped 243 individuals who never reported owning life insurance in any wave of HRS data. Our final analysis sample thus consists of 3324 in wave 1996 and the survivors among them in subsequent waves, 3195 in wave 1998, 3022 in wave 2000, 2854 in wave 2002, 2717 in wave 2004 and 2558 in wave 2006. Table 2 describes how we come to our final estimation sample.

2.1 | Construction of variables related to life insurance decisions

Here we describe the questions in HRS we use to construct the life-insurance related variables.

- For whether or not an individual owned life insurance in the current wave, we use the individual's response to the following HRS survey question, which is asked in all waves: [Q1] "Do you currently have any life insurance?"
- For whether or not an individual obtained a policy since the previous wave, we use the individual's response to the following HRS question, which is asked all waves starting in 1996: [Q2] "Since (previous wave interview month-year) have you obtained any new life insurance policies?" If the respondent answers "yes," we consider him to have obtained a new policy.
- For whether or not an individual lapsed a policy since last wave, we use the individual's response to the following HRS question, which is asked all waves starting in 1996: [Q3] "Since (previous wave interview month-year) have you allowed any life insurance policies to lapse or have any been cancelled?" We also use the response to another survey question, which is also asked all waves starting in 1996: [Q4] "Was this lapse or cancellation something you chose to do, or was it done by the provider, your employer, or someone else?" If the respondent answers "yes" to the first question *and* answers "my decision" to the second question, we consider him to have lapsed a policy.

TABLE 2 Sample selection criterion and sample size

Selection criterion	Sample size
All individuals who responded to both 1994 and 1996 HRS interviews	17,354
Males who were aged between 50 and 70 in 1996 (wave 3)	4512
Did not having any missing interviews from 1994 to 2006	3696
Did not have any missing values for reported life insurance ownership status from 1994 to 2006	3567
Reported owning life insurance at least once from 1996 to 2006	3324

Note: The selection criteria are cumulative.

In the notation of the model we will present in the Section 3, we construct an individual's period-t (or wave-t) decisions as follows:

- For the individual who reported not having life insurance in the previous wave ($d_{t-1} = 0$), we let $d_t = 0$ if the individual reports not having life insurance this wave; and $d_t = 1$ if the individual reports having life insurance this wave ("yes" to Q1). Because the individual does not own life insurance in period $t-1$ but does in period t, we interpret that he chose the optimal policy in period t given his state variables at t.
- For the individual who reported having life insurance in the previous wave ($d_{t-1} \geq 1$), we let $d_t = 0$ if the individual reports not having life insurance this wave ("no" to Q1); and we let $d_t = 1$ (i.e., the individual re-optimizes his life insurance) if the individual reports having life insurance this wave ("yes" to Q1) *and* he obtained new life insurance policy ("yes" to Q2), OR if the individual answered "yes" to Q1, reported lapsing (i.e., answered "yes" to Q3) and reported that lapse was his own decision (answered "my own decision" to Q4). Note that under this construction, we have interpreted the "lapsing or obtaining" of any policies as an indication that the respondent reoptimized his life insurance coverage. Finally, we let $d_t = 2$ (i.e., he kept his previous life insurance policy unchanged) if the individual reports having life insurance this wave ("yes" to Q1) AND the individual reported no to obtaining new policy ("no" to Q2) AND the individual did not lapse any existing policy (either reported "no" to Q3 or reported "yes" to Q3 but did not report "my decision" to Q4).

2.2 | Information about the details of life insurance holdings in the HRS data

HRS also has questions regarding the face amount and premium payments for life insurance policies. However, there are several problems with incorporating these variables into our empirical analysis. First, the questions differ across waves. In the 1994 wave, questions were asked regarding the total face amount and premiums for term life policies; but for whole life policies only total face amount was collected.[8] In the 1996 and 1998 waves, information about lapsation and face amounts are available, but not premiums. From 2000 on, the HRS asked about the combined face value for all policies, combined face value for whole life policies, and the combined premium payments for whole life policies. Note that information about the premiums for term life policies were not collected from 2000 on. Second, there is a very high incidence of missing data regarding life insurance premiums and face amounts. In our selected sample, 40.3% of the respondents have at least one instance of missing face amount in waves when they reported owning life insurance. The incidence of missing values in premium payments is even higher. Third, even for those who reported face amount and premium payments for their life insurance policies, we do not know the choice set they faced when purchasing their policies.

For these reasons, we decide to only model the individuals' life insurance decisions regarding whether to reoptimize, lapse or maintain an existing policy, and only use the observed information about the above decisions in estimating the model.

[8]The questions in 1994 wave related to premium and face amount are: [W6768]. About how much do you pay for (this term insurance/these term insurance policies) each month or year? [W6769]. Was that per month, year, or what? [W6770]. What is the current face value of all the term insurance policies that you have? [W6773]. What is the current face value of (this [whole life] policy/these [whole life] policies?)

2.3 | Descriptive statistics

2.3.1 | Patterns of life insurance coverage and its transitions

Table 3 provides the life insurance coverage and patterns of transition between coverage and no coverage in the HRS data. Panel A shows that among the 3324 live respondents in 1996, 88.1% are covered by life insurance; among the 3195 who survived to the 1998 wave, 85.7% owned life insurance, and so forth. Over the waves, the life insurance coverage rates among the live respondents seem to exhibit a declining trend, with the coverage rate among the 2558 who survived to the 2006 wave being about 78.6%.

TABLE 3 Life insurance coverage and transition patterns in HRS: 1996–2006

	Wave					
	1996	**1998**	**2000**	**2002**	**2004**	**2006**
Panel A: Life insurance coverage status						
Currently covered by life insurance	2927	2739	2524	2313	2187	2011
	88.1%	85.7%	83.5%	81.0%	80.5%	78.6%
No life insurance coverage	397	456	498	541	530	547
	11.9%	14.3%	16.5%	19.0%	19.5%	21.4%
Total live respondents	3324	3195	3022	2854	2717	2558
Panel B: Life insurance coverage status conditional on no coverage in previous wave						
Life insurance coverage this wave	243	125	130	150	163	123
	47.5%	33.4%	31.9%	33.7%	32.7%	25.6%
No life insurance coverage this wave	269	249	277	295	336	357
	52.5%	66.6%	68.1%	66.3%	67.3%	74.4%
Total live respondents with no coverage last wave	512	374	407	445	499	480
Panel C: Life insurance coverage status conditional on coverage in previous wave						
Life insurance coverage this wave	2684	2614	2394	2163	2024	1888
	95.4%	92.7%	91.5%	89.8%	91.3%	90.9%
No life insurance coverage this wave	128	207	221	246	194	190
	4.6%	7.3%	8.5%	10.2%	8.7%	9.1%
Total respondents with coverage last wave	2812	2821	2615	2409	2218	2078
Panel D: Whether changed terms of coverage conditional on coverage in both current and previous waves						
Did not change terms of coverage	2430	2395	2233	2034	1881	1769
	90.5%	91.6%	93.3%	94.0%	92.9%	93.7%
Changed terms of coverage	254	219	161	129	143	119
	9.5%	8.4%	6.7%	6.0%	7.1%	6.3%
Total live respondents with coverage in both waves	2684	2614	2394	2163	2024	1888

Note: Panel A shows life insurance coverage status for live respondents in each wave. Panel B shows life insurance coverage status for respondents who reported no coverage in previous wave. Panel C shows life insurance coverage status for respondents who reported coverage in previous wave. Panel D shows the fraction of respondents who reported changing the terms of their life insurance coverage, conditional on being covered in both the current and previous waves.

Panels B and C show, however, that there is substantial transition between coverage and no coverage. Panel B shows that among the 512 individuals who did not have life insurance coverage in 1994, almost a half (47.5%) obtained coverage in 1996; in later waves between 25.6% and 33.7% of individuals without life insurance in the previous wave ended up with coverage in the next wave. Panel C shows that there is also substantial lapsation among life insurance policyholders. In our data, between-wave lapsation rates range from 4.6% to 10.2%. Considering that our sample is relatively old and the tenures of holding life insurance policies in the HRS sample are also typically longer, these lapsation rates are in line with the industry lapsation rates reported in the introduction (see Table 1).

Panel D shows that even among those individuals who own life insurance in both the previous wave and the current wave, a substantial fraction has changed the terms of their coverage, or in the words of our model, reoptimized. Between 6.0% and 9.5% of the sample who have insurance coverage in adjacent waves reported changing the terms of their coverages.

2.4 | Summary statistics of state variables

Table 4 summarizes the key state variables for the sample used in our empirical analysis. It shows that the average age of the live respondents in our sample is 61.1. The means of household income in our sample are quite stable around $62,000 to $66,000; and similarly the means of log household incomes are stable around 10.58 to 10.73. The next eight rows report the mean of the incidence of health conditions, including high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychological problems, and arthritis. It shows clear signs of health deterioration for the surviving samples over the years. The sum of the above eight health conditions steadily increases from 1.37 in 1996 to 2.34 in 2006. Finally, the marital status of the surviving sample seems to be quite stable, with the fraction married being in the range of 83%–85%.

Tables A1 and A2 in the appendix summarize the mean and *SD* of the state variables by the life insurance coverage status. There does not seem to be much of a difference in ages between those with and without life insurance coverage, but the mean log household income is significantly higher for those with life insurance than those without and life insurance policyholders are much more likely to be married than those without. We also find that those with life insurance tend to be healthier than those without life insurance.

2.5 | Reduced-form determinants of the life insurance decisions

Table 5 presents the coefficient estimates of a Logit regression on the probability of purchasing life insurance among those who did not have coverage in the previous wave. It shows that the richer, younger, healthier, and married individuals are more likely to purchase life insurance coverage than the poorer, older, unhealthier, and widowed or unmarried individuals. Table 6 presents the estimates of a multinomial Logit regression for the probability of lapsing, changing coverage, or maintaining the previous coverage. The omitted category is lapsing all coverage. The estimates show that richer individuals are more likely to either maintain the current coverage or change existing coverage than to lapse all coverage; individuals who experienced negative income shocks are more likely to

TABLE 4 Summary statistics of state variables

Variable description	Wave 1996	1998	2000	2002	2004	2006
Age of respondent	61.102 (4.3535)	63.035 (4.3441)	64.925 (4.3070)	66.852 (4.2861)	68.789 (4.2858)	70.676 (4.2512)
Household income ($1000s)	62.632 (73.846)	61.978 (73.512)	65.187 (87.389)	63.246 (76.716)	66.278 (87.937)	66.549 (80.344)
Log household income	10.582 (1.3013)	10.587 (1.2086)	10.623 (1.2051)	10.611 (1.2045)	10.688 (1.0377)	10.727 (0.9150)
Whether ever diagnosed with high blood pressure	0.4025 (0.4905)	0.4372 (0.4961)	0.4765 (0.4995)	0.5172 (0.4998)	0.5683 (0.4954)	0.6197 (0.4856)
Whether ever diagnosed with diabetes	0.1439 (0.3510)	0.1594 (0.3661)	0.1755 (0.3805)	0.2062 (0.4046)	0.2273 (0.4191)	0.2523 (0.4344)
Whether ever diagnosed with cancer	0.0599 (0.2373)	0.0811 (0.2731)	0.1007 (0.3009)	0.1287 (0.3349)	0.1591 (0.3659)	0.1874 (0.3903)
Whether ever diagnosed with lung disease	0.0710 (0.2569)	0.0777 (0.2677)	0.0791 (0.2700)	0.0880 (0.2834)	0.1083 (0.3108)	0.1170 (0.3215)
Whether ever diagnosed with heart disease	0.1902 (0.3926)	0.2145 (0.4106)	0.2351 (0.4241)	0.2672 (0.4426)	0.3050 (0.4605)	0.3435 (0.4750)
Whether ever diagnosed with stroke	0.0494 (0.2167)	0.0579 (0.2337)	0.0652 (0.2470)	0.0712 (0.2572)	0.0829 (0.2757)	0.0986 (0.2982)
Whether ever diagnosed with psychological problem	0.0599 (0.2373)	0.0714 (0.2575)	0.0788 (0.2695)	0.0873 (0.2823)	0.0932 (0.2907)	0.1002 (0.3003)
Whether ever diagnosed with arthritis	0.3904 (0.4879)	0.4460 (0.4972)	0.4831 (0.4998)	0.5312 (0.4991)	0.5790 (0.4938)	0.6205 (0.4854)
Sum of above conditions	1.3672 (1.2409)	1.5453 (1.2912)	1.6940 (1.3103)	1.8969 (1.3372)	2.1230 (1.3941)	2.3392 (1.4204)
Whether married	0.8504 (0.3567)	0.8431 (0.3638)	0.8444 (0.3626)	0.8447 (0.3623)	0.8438 (0.3631)	0.8326 (0.3734)
Whether has children	0.9419 (0.2340)	0.9402 (0.2372)	0.9394 (0.2386)	0.9386 (0.2400)	0.9366 (0.2436)	0.9335 (0.2492)
Age of youngest child	24.832 (10.059)	26.813 (10.020)	28.764 (9.959)	30.892 (9.839)	32.933 (9.779)	34.894 (9.718)
# of live respondents	3324	3195	3022	2854	2717	2558

Note: SDs are in parenthesis.

lapse all coverage; individuals who are either divorced or widowed are more likely to lapse all coverage; finally, individuals who have experienced an increase in the number of health conditions are somewhat more likely to lapse all coverage, though the effect is not statistically significant.

TABLE 5 Reduced-form logit regression on the probability of buying life insurance, conditional on having no life insurance in the previous wave

Variable	Coefficient	SE
Constant	0.9224	0.6039
Age	−0.0407***	0.0084
Logincome	0.0929***	0.0315
Number of health conditions	−0.0810***	0.0297
Married	0.0831	0.1007
Has children	0.2704	0.1713
Age of youngest child	0.0006	0.0043
Observations	2707	
Log-likelihood	−1713.4	

Note: *, **, and *** respectively represent significance at 10%, 5%, and 1% levels.

TABLE 6 Reduced-form multinomial logit regression on the probability of lapsing, changing coverage, or maintaining coverage, conditional on owning life insurance in the previous wave

	Change existing coverage		Maintain existing coverage	
Variable	Coefficient	SE	Coefficient	SE
Constant	−0.1119	0.9182	1.6361***	0.6333
Age	−0.0789***	0.0099	−0.0480***	0.0069
Logincome	0.4542***	0.0578	0.2522***	0.0396
Number of health conditions	−0.0493	0.0391	−0.0411	0.0267
Married	0.3358**	0.1372	0.3092***	0.0908
Has children	−0.0912	0.2027	0.0464	0.1366
Age of youngest child	0.0097*	0.0051	0.0075**	0.0034
ΔAge	−0.0956	0.0623	0.2067***	0.0439
$(\Delta\text{Age})^2$	0.0090*	0.0048	−0.0082**	0.0034
ΔLogincome	−0.1406***	0.0457	−0.0386	0.0305
$(\Delta\text{Logincome})^2$	0.0174***	0.0053	0.0102***	0.0037
ΔConditions	0.1268	0.1362	0.0159	0.0877
$(\Delta\text{Conditions})^2$	−0.0843	0.0514	−0.0178	0.0285
ΔMarried	0.3218	0.1998	−0.0954	0.1391
Observations	14,951			
Log-likelihood	−7565.6			

Note: Conditional on owning life insurance, the three choices are: (a). to lapse all coverage; (b). to change the existing coverage; and (c). to maintain the existing coverage. The base outcome is set to choice (a). For any variable x, Δx is the difference between the current value of x and the value of x which occurred during the last period in which the respondent changed his coverage. *, **, and *** represent significance at 10%, 5%, and 1% levels, respectively.

3 | AN EMPIRICAL MODEL OF LIFE INSURANCE DECISIONS

In this section, we present a dynamic discrete choice model of how individuals make life insurance decisions, which we will later empirically implement. Our model is simple, yet rich enough to capture the dynamic intuition behind the life insurance models of Hendel and Lizzeri (2003) and Fang and Kung (2010a).

Time is discrete and indexed by $t = 1, 2, \ldots$ In the beginning of each period t, an individual i either has or does not have an existing life insurance policy. If the individual enters period t without an existing policy, then he chooses between not owning life insurance ($d_{it} = 0$) or optimally purchasing a new policy ($d_{it} = 1$). If the individual enters period t with an existing policy, then, besides the above two choices, he can additionally choose to keep his existing policies ($d_{it} = 2$). If an individual who has life insurance in period $t - 1$ decides not to own life insurance in period t, we interpret it as lapsation of coverage. As we describe in Section 2, the choice $d_{it} = 1$ for an individual who previously owns at least one policy is interpreted more broadly: an individual is considered to have reoptimized his existing policy if he reported purchasing a new policy or choosing to lapse one of the existing policies (while he continues to hold at least one other policy). The key interpretation for choice $d_{it} = 1$ is that the individual re-optimized his life insurance holdings.

3.1 | Flow payoffs from choices

Now we describe an individual's payoffs from each of these choices. First, let $x_{it} \in \mathcal{X}$ denote the vector of *observable* state variables of individual i in period t, and let $z_{it} \in \mathcal{Z}$ denote the vector of *unobservable* state variables.[9] These characteristics include variables that affect the individual's preference for or cost of owning life insurance, such as income, health and bequest motives. We normalize the utility from not owning life insurance (i.e., $d_{it} = 0$) to 0; that is,

$$u_0(x_{it}, z_{it}) = 0 \text{ for all } (x_{it}, z_{it}) \in \mathcal{X} \times \mathcal{Z}. \tag{1}$$

The utility from optimally purchasing a new policy in state (x_{it}, z_{it}), that is, $d_{it} = 1$, regardless of whether he previously owned a life insurance policy, is assumed to be:

$$u_1(x_{it}, z_{it}) + \varepsilon_{1it}, \tag{2}$$

where ε_{1it} is an idiosyncratic choice specific shock, drawn from a Type-I extreme value distribution. In our empirical analysis, we will specify $u_1(x_{it}, z_{it})$ as a flexible polynomial of x_{it} and z_{it}.

Now we consider the flow utility for an individual i entering period t with an existing policy which was last re-optimized at period $\hat{t}$. That is, let $\hat{t} = \sup\{s \mid s < t, d_{is} = 1\}$. Let $(\hat{x}_{it}, \hat{z}_{it}) = (x_{i\hat{t}}, z_{i\hat{t}})$ denote the state vector that i was in when he last reoptimized his life

[9] We present the model here assuming the presence of both the observed and unobserved state variables. In the Online Appendix, we also present estimates of a model with only observed state variables; in that case, we simply ignore the unobserved state vector z_{it}.

insurance. We assume that the flow utility individual i obtains from continuing the existing policies purchased when his state vector was $(\hat{x}_{it}, \hat{z}_{it})$ is given by

$$u_2(x_{it}, z_{it}, \hat{x}_{it}, \hat{z}_{it}, \varepsilon_{2it}) = u_1(x_{it}, z_{it}) - c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it})) + \varepsilon_{2it}, \tag{3}$$

where $c(\cdot,\cdot)$ can be considered as a *suboptimality penalty function*, which may also include adjustment costs (see discussion below in Section 4), that depends on the "distance" between the current state (x_{it}, z_{it}) and the state in which the existing policy was purchased $(\hat{x}_{it}, \hat{z}_{it})$. The adjustment cost can be positive or negative, depending on the factors that have changed. For example, if the individual was married when he purchased the existing policy but is not married now, then, all other things equal, the adjustment cost is likely to be negative; he would have less incentive to keep the existing policy. On the other hand, if the individual's health has deteriorated substantially, then obtaining a new policy could be prohibitively costly, in which case the adjustment cost is likely to be positive; he would have more incentive to keep the existing policy which was purchased during a healthier state.

To summarize, we model the life insurance choice as the decision to either: (1). hold no life insurance; (2). purchase a new insurance policy that is optimal for the current state; or (3). continue with an existing policy. By decomposing the ownership decision into continuation versus reoptimization, our model is able to capture the intuition that an individual who has suffered a negative shock to a factor that positively affects life insurance ownership (such as income or bequest motive) may still be likely to keep his insurance if the policy was initially purchased a long time ago during a better health state.

Moreover, the decomposition of the ownership decision allows us to examine two separate motives for lapsation: lapsation because the individual no longer needs *any* life insurance, and lapsation because the policyholder's personal situation, that is, (x_{it}, z_{it}), has changed such that new coverage terms are required.

3.2 | Parametric assumptions on u_1 and c functions

In our empirical implementation of the model, we let the observed state vector x_{it} include age, log household income, sum of the number of health conditions, marital status, an indicator for whether the individual has children, and the age of the youngest child. We let the unobserved state vector z_{it} include z_{1it}, z_{2it}, and z_{3it} which respectively represent the unobserved components of income, health, and bequest motives. In Section 5 below, we will describe how we anchor these unobservables to their intended interpretations and how we use sequential Monte Carlo methods to simulate their posterior distributions.

The primitives of our model are thus given by the utility function of optimally purchasing life insurance u_1, and the suboptimality adjustment function c. In our empirical analysis we adopt the following parametric specifications for $u_1(x_{it}, z_{it})$ and $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$:

$$\begin{aligned} u_1(x_{it}, z_{it}) = {} & \theta_0 + \theta_1 \mathrm{AGE}_{it} + \theta_2(\mathrm{LOGINCOME}_{it} + z_{1it}) + \theta_3(\mathrm{CONDITIONS}_{it} + z_{1it}) \\ & + \theta_4(\mathrm{MARRIED}_{it} + z_{1it}) + \theta_5 \mathrm{AGE}_{it}^2 + \theta_6(\mathrm{LOGINCOME}_{it} + z_{1it})^2 \\ & + \theta_7(\mathrm{CONDITIONS}_{it} + z_{2it})^2 + \theta_8(\mathrm{MARRIED}_{it} + z_{3it})^2 + \theta_9 \mathrm{HAS\ CHILDREN}_i \\ & + \theta_{10} \mathrm{HAS\ CHILDREN}_i \times \mathrm{AGE\ OF\ YOUNGEST\ CHILD}_{it}; \end{aligned} \tag{4}$$

$$\begin{aligned} c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it})) = \theta_{11} &+ \theta_{12}(\mathrm{AGE}_{it} - \widehat{\mathrm{AGE}}_{it}) + \theta_{13}(\mathrm{AGE}_{it} - \widehat{\mathrm{AGE}}_{it})^2 \\ &+ \theta_{14}(\mathrm{LOGINCOME}_{it} + z_{1it} - \widehat{\mathrm{LOGINCOME}}_{it} - \hat{z}_{1it}) \\ &+ \theta_{15}(\mathrm{LOGINCOME}_{it} + z_{1it} - \widehat{\mathrm{LOGINCOME}}_{it} - \hat{z}_{1it})^2 \\ &+ \theta_{16}(\mathrm{CONDITIONS}_{it} + z_{2it} - \widehat{\mathrm{CONDITIONS}}_{it} - \hat{z}_{2it}) \\ &+ \theta_{17}(\mathrm{CONDITIONS}_{it} + z_{2it} - \widehat{\mathrm{CONDITIONS}}_{it} - \hat{z}_{2it})^2 \\ &+ \theta_{18}(\mathrm{MARRIED}_{it} + z_{3it} - \widehat{\mathrm{MARRIED}}_{it} - \hat{z}_{3it}) + \theta_{19} \\ &(\mathrm{MARRIED}_{it} + z_{3it} - \widehat{\mathrm{MARRIED}}_{it} - \hat{z}_{3it})^2. \end{aligned} \tag{5}$$

In Section 4 below, we will provide an interpretation of the above-specified $c(\cdot,\cdot)$ function as a suboptimality penalty function.

3.3 | Transitions of the state variables

The state variables that an individual must keep track of depend on whether the individual is currently a policyholder. If he currently does not own a policy, his state variable is simply his current state vector (x_{it}, z_{it}); if he currently owns a policy, then his state variables include both his *current state vector* (x_{it}, z_{it}) and the state vector $(\hat{x}_{it}, \hat{z}_{it})$ at which he purchased the policy he currently owns.

In our empirical analysis, we assume that the current state vectors (x_{it}, z_{it}) evolve exogenously (i.e., not affected by the individual's decision) according to a joint distribution given by

$$(x_{it+1}, z_{it+1}) \sim f(x_{it+1}, z_{it+1} | x_{it}, z_{it}).$$

In particular, for the observed state vector x_{it}, that includes age, log household income, sum of the number of health conditions, and their respective squares, marital status, whether the individual has children, and the age of the youngest child, we estimate their evolutions directly from the data. For the unobserved state vector z_{it}, we will use sequential Monte Carlo methods to simulate its evolution (see Section 5.2 below for details).

The evolution of the state vector $(\hat{x}_{it}, \hat{z}_{it})$ is endogenous, and it is given as follows. If the individual does not own life insurance at period t, which we denote by setting $(\hat{x}_{it}, \hat{z}_{it}) = \varnothing$, then

$$[((\hat{x}_{it+1}, \hat{z}_{it+1}) \mid (\hat{x}_{it}, \hat{z}_{it}) = \varnothing] = \begin{Bmatrix} (x_{it}, z_{it}) \text{ if } d_{it} = 1, \\ \varnothing \text{ if } d_{it} = 0, \end{Bmatrix} \tag{6}$$

where $\varnothing$ denotes that the individual remains with no life insurance. If the individual owns life insurance at period t purchased at state $(\hat{x}_{it}, \hat{z}_{it})$, then

$$[((\hat{x}_{it+1}, \hat{z}_{it+1}) \mid (\hat{x}_{it}, \hat{z}_{it}) \neq \varnothing] = \begin{Bmatrix} \varnothing \text{ if } d_{it} = 0, \\ (x_{it}, z_{it}) \text{ if } d_{it} = 1, \\ (\hat{x}_{it}, \hat{z}_{it}) \text{ if } d_{it} = 2. \end{Bmatrix} \tag{7}$$

4 | DISCUSSION

4.1 | Dynamic discrete choice model without the knowledge of the choice and choice set

As we mentioned in Section 2, we do not have complete information about the exact life insurance policies owned by the individuals, and for those whose life insurance policies we do know about, we do not know their choice sets. However, we do know whether an individual has reoptimized his life insurance policy holdings (i.e., purchased a new policy or lapsed an existing one), or has dropped all coverage.

In fact the data restrictions we face are fairly typical for many applications.[10] For example, in the study of housing the market, it is possible that all we observe is whether a family moved to a new house, remained in the same house, or decided to rent; but we may not observe the characteristics (including the purchase price) of the new house the family moved into, or the characteristics of the house/apartment the family rented; and most likely, we are not able to observe the set of houses or apartments the family has considered purchasing or renting (see, e.g., Kung, 2012).

Our formulation provides an *indirect utility approach* to deal with such data limitations. Suppose that when individual i's state vector is (x_{it}, z_{it}), he has a choice set $\mathcal{L}(x_{it}, z_{it})$ which includes all the possible life insurance policies that he could choose from. Note that $\mathcal{L}(x_{it}, z_{it})$ depends on i's state vector (x_{it}, z_{it}), which captures the notion that life insurance premium and face amount typically depend on at least some of the characteristics of the insured. Let $\ell \in \mathcal{L}(x_{it}, z_{it})$ denote one such available policy. Let $u^*(\ell; x_{it}, z_{it})$ denote individual *i's primitive* flow utility from purchasing policy ℓ. If he were to choose to own a life insurance policy, his choice of the life insurance contract from his available choice set will be determined by the solution to the following problem:

$$V(x_{it}, z_{it}) = \max_{\ell \in \mathcal{L}(x_{it}, z_{it})} \{u^*(\ell; x_{it}, z_{it}) + \varepsilon_{\ell it} + \beta \mathrm{E}[V(x_{it+1}, z_{it+1}) | \ell, x_{it}, z_{it}]\}. \tag{8}$$

Let $\ell^*(x_{it}, z_{it})$ denote the solution. Then the flow utility $u_1(x_{it}, z_{it})$ we specified in (2) can be interpreted as the *indirect* flow utility, that is,

$$u_1(x_{it}, z_{it}) = u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it}). \tag{9}$$

It should be pointed out that, under the above indirect flow utility interpretation of $u_1(x_{it}, z_{it})$, in order for the error term ε_{1it} in (2) to be distributed as i.i.d. extreme value as assumed, we make the assumption that $\varepsilon_{\ell it}$ in (8) does *not* vary across $\ell \in \mathcal{L}(x_{it}, z_{it})$. We therefore interpret ε_{1it} as a shock to the cost of reoptimizing rather than a shock that varies across contracts.

[10] McFadden (1978) and Fox (2007) studied problems where the researcher only observes the choices of decision-makers from a subset of choices. McFadden (1978) showed that in a class of discrete-choice models where choice specific error terms have a block additive generalized extreme value (GEV) distributions, the standard maximum-likelihood estimator remains consistent. Fox (2007) proposed using semiparametric multinomial maximum-score estimator when estimation uses data on a subset of the choices available to agents in the data-generating process, thus relaxing the distributional assumptions on the error term required for McFadden (1978).

4.2 | Interpretations of the sub-optimality penalty function $c(\cdot)$

The suboptimality penalty function $c(\cdot)$ we introduced in (3) permits the interpretation that changing an existing life insurance policy may incur adjustment costs. To see this, consider an individual whose current state vector is (x_{it}, z_{it}) and owns a life insurance policy he purchased at $\hat{t}$ when his state vector was $(x_{i\hat{t}}, z_{i\hat{t}})$. Suppose that he decides to change (lapse or modify) his current policy and reoptimize, but there is an adjustment cost of $\kappa > 0$ for changing. Thus, the flow utility from lapsing into no coverage for this individual will be

$$u_0(x_{it}, z_{it}) = 0.$$

The flow utility from reoptimizing, using the notation from (9), will be

$$u_1(x_{it}, z_{it}) = u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it}) - \kappa.$$

And the flow utility from keeping the existing policy is

$$\begin{aligned} u_2(x_{it}, z_{it}, \hat{x}_{it}, \hat{z}_{it}) = u^*(\ell^*(\hat{x}_{it}, \hat{z}_{it}); x_{it}, z_{it}) &= \overbrace{u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it}) - \kappa}^{u_1(x_{it}, z_{it})} \\ &\quad - \left[\overbrace{u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it}) - u^*(\ell^*(\hat{x}_{it}, \hat{z}_{it}); x_{it}, z_{it})}^{\text{sub-optimality penalty}} - \kappa \right] \\ &= u_1(x_{it}, z_{it}) - c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it})) \end{aligned} \tag{10}$$

where

$$c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it})) \equiv [u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it}) - u^*(\ell^*(\hat{x}_{it}, \hat{z}_{it}); x_{it}, z_{it})] - \kappa.$$

Note that in the above expression for $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$, the term in the square bracket is the difference between the flow utility the individual could have received by purchasing the optimal policy for his current state (i.e., $\ell^*(x_{it}, z_{it})$), and the flow utility he receives from his existing policy, $\ell^*(\hat{x}_{it}, \hat{z}_{it})$, that he purchased when his state was $(\hat{x}_{it}, \hat{z}_{it})$. The suboptimality penalty therefore measures the utility loss from holding a policy $\ell^*(\hat{x}_{it}, \hat{z}_{it})$ that was optimal in state $(\hat{x}_{it}, \hat{z}_{it})$, but suboptimal when state vector is (x_{it}, z_{it}). But by not reoptimizing, the individual saves the adjustment cost κ. Given the presence of adjustment cost κ, we would expect that an existing policyholder will hold on to his policy until the sub-optimality penalty $\left[u^*(\ell^*(x_{i\hat{t}}, z_{i\hat{t}}); x_{it}, z_{it}) - u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it})\right]$ exceeds the adjustment cost κ, if we ignore decisions driven by i.i.d. preference shocks ε_{1it} and ε_{2it}.

It is clear from the above discussion that, in this formulation, we can also allow the adjustmentt cost κ to be made a function of (x_{it}, z_{it}), though we will not be able to separate the sub-optimality penalty $\left[u^*(\ell^*(x_{i\hat{t}}, z_{i\hat{t}}); x_{it}, z_{it}) - u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it})\right]$ from $\kappa(x_{it}, z_{it})$.[11]

[11]If the adjustment cost κ is incurred both when the individual lapses into no coverage, and when he reoptimizes, that is, if $u_0(x_{it}, z_{it}) = -\kappa$, and $u_1(x_{it}, z_{it}) = u^*(\ell^*(x_{it}, z_{it}); x_{it}, z_{it}) - \kappa$, then we can allow κ to depend on both (x_{it}, z_{it}) and $(x_{i\hat{t}}, z_{i\hat{t}})$.

It is also worth pointing out that our parametric specifications of $u_1(x_{it}, z_{it})$ and $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$, as given in (4) and (5), respectively, are consistent with the above interpretations of the suboptimality penalty function.[12]

4.3 | Limitations of the "indirect flow utility" approach

In this paper, we adopt the "indirect flow utility" approach to deal with the lack of information regarding individuals' actual choices of life insurance policies and their relevant choice set. This is useful for our purpose of understanding why policyholders lapse their coverage (as we will demonstrate later), but it comes with a limitation. The indirect flow utilities $u_1(x_{it}, z_{it})$ and $u_2(x_{it}, z_{it}, \hat{x}_{it}, \hat{z}_{it})$, defined in (9) and (10), respectively, are derived only under the *existing life insurance market structure*. As a result, the estimated indirect flow utility functions are *not* primitives that are invariant to counterfactual policy changes that may affect the equilibrium of the life insurance market. Of course, this limitation is also present in other dynamic discrete choice models where the flow utility functions can have the interpretation as a reduced-form, indirect utility function of a more detailed choice problem.[13]

Another limitation is that this approach allows for decisions that are possibly inconsistent with traditional state-dependent preferences over money. For example, it is possible that not all combinations of premiums and face amounts can be rationalized in a fully rational model, and violations of stochastic dominance are not ruled out. While this "black box" approach suits our purpose of predicting and fitting the basic life insurance ownership patterns we observe in our data, it is not designed to fit data with more detailed life insurance contract information if such data were available.[14]

Finally, our approach makes no assumptions about the nature of the suboptimality penalty function. These costs could be "real" (e.g., due to transaction costs or real changes to circumstance) or "psychological" (e.g., inertia, confusion). While the nature of the penalty function matters little for predicting choices, it may matter for the welfare implications. In our discussion of welfare implications, we assume that the suboptimality costs are due to real changes in circumstances and that policyholders behave rationally, but it is worth keeping in mind that the welfare implications may be different if these costs are psychological.

5 | ESTIMATES FROM A DYNAMIC DISCRETE CHOICE MODEL WITH UNOBSERVABLE STATE VARIABLES

In this section, we present our estimation and simulation results for the dynamic structural model of life insurance decisions presented in Section 3. As described in Section 3 the flow utilities are given by Equations (1)–(3). We start by describing how to estimate the dynamic

[12]Due to the ages of the individuals in our estimation sample, we have practically no changes in the number of children. Thust the difference between AGE OF YOUNGEST CHILD$_{it}$ and AGE OF YOUNGEST CHILD$_{i\hat{t}}$ is essentially the same as the difference between AGE$_{it}$ and AGE$_{i\hat{t}}$.

[13]For example, in many I.O. papers a reduced-form flow profit function is assumed. Presumably, the profit function is not invariant to changes in the market structure.

[14]We thank an anonymous referee for pointing out this detail to us.

discrete choice problem arising from these utilities. To simplify notation, let $w_{it} = (x_{it}, z_{it})$ denote the full vector of state variables, both observed and unobserved, and let $\hat{w}_{it} = (\hat{x}_{it}, \hat{z}_{it})$ be the state variables when i last reoptimized. Let $P(w_{it}|w_{it-1})$ be the transition distribution function of the state variables. As mentioned in Section 3, we assume that the evolution of the state variables w_{it} does not depend on the life insurance choices, though the evolution of the "hatted" variables, $\hat{w}_{it}$, does depend on the choices.

At period t, let $V_{0t}(w_{it})$ be the present value from choosing $d_{it} = 0$ (no life insurance); let $V_{1t}(w_{it})$ be the present value from choosing $d_{it} = 1$ (reoptimize); and let $V_{2t}(w_{it}, \hat{w}_{it})$ be the present value of choosing $d_{it} = 2$ (keep existing policy), for those who owned policies previously purchased at state $\hat{w}_{it}$. To derive the choice-specific value functions, it is useful to first derive the inclusive continuation values from being in a given state vector. Let $V_t(w_{it}, \hat{w}_{it})$ denote the period-t inclusive value for being in state w_{it} and having an existing policy purchased when the state was $\hat{w}_{it}$. Let $W_t(w_{it})$ denote the period-t inclusive value for being in state w_{it} and not having any existing life insurance. Under the assumption of additively separable choice specific shocks drawn from i.i.d. Type-1 extreme value distributions, and using $G(\cdot)$ to denote the joint distribution of the random vector $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})$, $V_t(w_{it}, \hat{w}_{it})$ and $W_t(w_{it})$ can be, following Rust (1994), expressed as follows:

$$\begin{aligned} V_t(w_{it}, \hat{w}_{it}) &= \int \max\{V_{0t}(w_{it}), V_{1t}(w_{it}) + \varepsilon_{1t}, V_{2t}(w_{it}, \hat{w}_{it}) + \varepsilon_{2t}\} dG(\varepsilon_t) \\ &= \log\{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})] + \exp[V_{2t}(w_{it}, \hat{w}_{it})]\} + 0.57722, \end{aligned} \tag{11}$$

$$\begin{aligned} W_t(w_{it}) &= \int \max\{V_{0t}(w_{it}), V_{1t}(w_{it}) + \varepsilon_{1t}\} dG(\varepsilon_t) \\ &= \log\{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})]\} + 0.57722, \end{aligned} \tag{12}$$

where 0.57722 is Euler's constant. Then, the choice-specific value functions can be written as follows:

$$V_{0t}(w_{it}) = \beta \int W_{t+1}(w_{it+1}) dP(w_{it+1}|w_{it}), \tag{13}$$

$$V_{1t}(w_{it}) = u_1(w_{it}) + \beta \int V_{t+1}(w_{it+1}, w_{it}) dP(w_{it+1}|w_{it}), \tag{14}$$

$$V_{2t}(w_{it}, \hat{w}_{it}) = u_2(w_{it}, \hat{w}_{it}) + \beta \int V_{t+1}(w_{it+1}, \hat{w}_{it}) dP(w_{it+1}|w_{it}). \tag{15}$$

In estimation, we assume that $\beta = 0.81$ (for an implied annual discount factor of 0.9) and we assume that age 80 is the final decision period, so that $V_{0,80}(w_{it}) = 0$, $V_{1,80}(w_{it}) = u_1(w_{it})$, and $V_{2,80}(w_{it}, \hat{w}_{it}) = u_2(w_{it}, \hat{w}_{it})$. Under these assumptions, we can solve for the choice-specific value functions at each age using backward recursion.

The choice probabilities at each period t are then given as follows. For individuals without life insurance in the beginning of period t, their choice probabilities for $d_{it} \in \{0, 1\}$ are given by:

$$\Pr\{d_{it} = 0 | w_{it}, \hat{w}_{it} = \varnothing\} = \frac{\exp[V_{0t}(w_{it})]}{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})]},$$

$$\Pr\{d_{it} = 1 | w_{it}, \hat{w}_{it} = \varnothing\} = \frac{\exp[V_{1t}(w_{it})]}{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})]}.$$

For individuals who own life insurance in the beginning of period t, which are purchased in previous waves when state vector is $\hat{w}_{it}$, their choice probabilities for $d_{it} \in \{0, 1, 2\}$ are given by:

$$\Pr\{d_{it} = 0 | w_{it}, \hat{w}_{it} \neq \varnothing\} = \frac{\exp[V_{0t}(w_{it})]}{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})] + \exp[V_{2t}(w_{it}, \hat{w}_{it})]},$$

$$\Pr\{d_{it} = 1 | w_{it}, \hat{w}_{it} \neq \varnothing\} = \frac{\exp[V_{1t}(w_{it})]}{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})] + \exp[V_{2t}(w_{it}, \hat{w}_{it})]},$$

$$\Pr\{d_{it} = 2 | w_{it}, \hat{w}_{it} \neq \varnothing\} = \frac{\exp[V_{2t}(w_{it}, \hat{w}_{it})]}{\exp[V_{0t}(w_{it})] + \exp[V_{1t}(w_{it})] + \exp[V_{2t}(w_{it}, \hat{w}_{it})]}.$$

If all the state variables are observed, then the model is easily estimated by maximum-likelihood. The transition distribution function $P(w_{it} | w_{it-1})$ can be estimated directly from the data, the value functions can be computed by backward recursion, and the conditional choice probabilities can be calculated at any value of the parameters. The parameters that maximize the likelihood of the data can then be estimated.

The difficulty comes when there are unobserved state variables, z_{it}. We now turn to describing how we add three unobserved state variables: z_1, z_2, and z_3 that are meant to represent the serially correlated unobservable components of income, health, and bequest motive. Below we first describe how we anchor the interpretations of these unobserved state variables, then we describe how we use sequential Monte Carlo methods to simulate their conditional distributions.

5.1 | Anchoring the unobserved state variables

In this specification, we would like to give the unobserved state variable z_1 the interpretation of an *unobserved liquidity (or income)* shock, and normalize its unit to the same as log income, z_2 the interpretation of an *unobserved health shock* that is normalized to the units of health conditions, and finally z_3 the interpretation as an *unobserved component of bequest motive* that is normalized to the units of marital status.

We assume that the initial distribution in 1994 (which we set to be $t = 0$) for each of these unobserved variables is degenerate and given by:[15]

$$z_{1i0} = \theta_{20} h_{i0}, \tag{16}$$

$$z_{2i0} = \theta_{21} h_{i0}, \tag{17}$$

$$z_{3i0} = \theta_{22} h_{i0}, \tag{18}$$

where h_{i0} is an indicator dummy for whether the individual reported owning life insurance in 1994. To anchor z_1, z_2, and z_3 to have the desired interpretation given above, we expect, but do not restrict, that the coefficients θ_{20}, θ_{21} and θ_{22} to be of certain signs. For example, because income is a positive

[15] The assumption that the initial distribution of the unobservable state variables $\mathbf{z}_0$ is degenerate is for computational simplicity.

factor for life insurance ownership, we expect that the sign of θ_{20} to be positive, so that individuals who owned life insurance in the initial period also have higher z_1. This anchors the interpretation of z_1 as an unobserved component of income. Similarly, because being married is also a positive factor for life insurance ownership, we also expect the sign of θ_{22} to be positive, so that individuals who owned life insurance in the initial period also have higher values of unobserved bequest motive z_3.

The second channel that anchors the unobserved state variables to having the desired interpretation is incorporated in our specifications for $u_1(\cdot)$ and $c(\cdot)$, as formulated in (4) and (5). Note that we restricted z_{1i} to entering both $u_1(\cdot)$ and $c(\cdot)$ in the same way as LOGINCOME, z_{2i} the same way as CONDITIONS, and z_{3i} the same way as MARRIED. These restrictions, together with the sequential Monte Carlo method (described in the next section below) we use to simulate the distributions of $\mathbf{z}_t \equiv (z_{1t}, z_{2t}, z_{3t})$, ensures that the unobserved variables $\mathbf{z}_t$ have the desired interpretations.

If we know the distributions of the unobservable state vectors $(\mathbf{z}_t, \hat{\mathbf{z}}_t)$, solving this model would be straightforward. Given a vector of parameters $\Theta = (\theta_0, ..., \theta_{19})$, we can compute the value functions at each age through backward recursion. The difficulty of handling unobserved state variables comes during estimation, because we have to integrate over the unobservables when computing the likelihood. We now turn our attention to this problem.

5.2 | Using sequential Monte Carlo (SMC) method to simulate the distributions of the unobserved state variables

We use a SMC method to simulate the distributions of the unobservable state vectors.[16] SMC is a set of simulation-based methods which provides a convenient and attractive approach to computing the posterior distributions of non-Gaussian, nonlinear, and high dimensional random variables.[17] A thorough discussion of the method, from both the theoretical and the practical perspectives, is available in Doucet et al. (2001). The SMC method has been widely used in fields such as speech recognition, biology, physics, and so forth. Despite the obvious potential importance of serially correlated unobservable state variables, there are few applications of SMC in the economics literature. Fernandez-Villaverde and Rubio-Ramirez (2007) used SMC for estimating macroeconomic dynamic stochastic general equilibrium models with serially correlated unobservable state variables using a likelihood approach. Blevins (2016) proposed the use of SMC to allow for serially correlated unobservable state variables in estimating dynamic single agent models and dynamic games. Hong et al. (2018) also discusses the method in an application to the pharmaceuticals industry. All of the papers allow for a single serially correlated unobservable state variable. In our application, we believe that there might be important serially correlated unobservable components for each of the three potential sources of lapsation, shocks to income, health, and bequest motives.

Now we provide a detailed discussion about the SMC algorithm. For a given individual, we observe the sequence of choices $\{d_t\}_{t=0}^T$, observed state variables $\{x_t, \hat{x}_t\}_{t=0}^T$, and whether the individual had life insurance in 1994 h_0. The data set is thus $\{d_t, x_t, \hat{x}_t, h_0\}_{t=0}^T$ (we have dropped the i subscript for convenience). Let $p(d_{0:T} | x_{0:T}, \hat{x}_{0:T}, h_0)$ denote the conditional likelihood of the observed data. We can write:

[16]SMC algorithms are also called bootstrap filters, particle filters, and sequential importance samplers with resampling.

[17]SMC for nonlinear, non-Gaussian models is the analog of Kalman filter for linear, Gaussian models. Gordon et al. (1993) is the seminal paper that proposed this algorithm, which they refer to as the *bootstrap filter*.

$$p(d_{0:T}|x_{0:T}, \hat{x}_{0:T}, h_0) = p(d_0|x_0, \hat{x}_0, h_0)\prod_{t=1}^{T} p(d_t|d_{t-1}, x_t, \hat{x}_t) \tag{19}$$

Because the initial distribution of $\mathbf{z}_0$ is degenerate and depends only on h_0, we can write:

$$p(d_0|x_0, \hat{x}_0, h_0) = p(d_0|x_0, \hat{x}_0, \mathbf{z}_0).$$

Assuming we have solved for the value functions in a first stage, we should be able to compute $p(d_0|x_0, \hat{x}_0, \mathbf{z}_0)$.

Now, for each $t > 0$ we can write:

$$p(d_t|d_{t-1}, x_t, \hat{x}_t) = \int p(d_t|d_{t-1}, x_t, \hat{x}_t, \mathbf{z}_t, \hat{\mathbf{z}}_t)p(\mathbf{z}_t, \hat{\mathbf{z}}_t|d_{t-1})d\mathbf{z}_t d\hat{\mathbf{z}}_t. \tag{20}$$

We know how to compute $p(d_t|d_{t-1}, x_t, \hat{x}_t, \mathbf{z}_t, \hat{\mathbf{z}}_t)$ for a given set of parameter values $\boldsymbol{\theta}$. What we need is a method to draw from $p(\mathbf{z}_t, \hat{\mathbf{z}}_t|d_{t-1})$; and we use the SMC method for this purpose.

SMC is a recursive algorithm that begins by drawing a swarm of particles approximating the initial distribution of the hidden state. The initial swarm is then used to draw a swarm for the next period, and this swarm is then filtered according to sequential importance weights.

The unobservables transition according to the following equations:

$$z_{1it} = \theta_{23} z_{1it-1} + \theta_{24}\epsilon_{z_1 t}, \tag{21}$$

$$z_{2it} = \theta_{25} z_{2it-1} + \theta_{26}\epsilon_{z_2 t}, \tag{22}$$

$$z_{3it} = \theta_{27} z_{3it-1} + \theta_{28}\epsilon_{z_3 t}, \tag{23}$$

where $\epsilon_{z_1 t}$, $\epsilon_{z_2 t}$, and $\epsilon_{z_3 t}$ are i.i.d. random variables with standard normal distributions $\mathcal{N}(0, 1)$. The transition distribution of the observed state variables, $P(x_{it}|x_{it-1})$ is estimated directly from the data in a first step.[18]

The filtered particles are then used to draw another swarm for the next period, and so on. In the following notation, we will absorb $\hat{\mathbf{z}}$ into $\mathbf{z}$ and use $\mathbf{z}$ to denote any unobserved variable, including the "hatted" $\mathbf{z}$'s.

The method proceeds as follows:

0. Set $t = 0$, draw a swarm of particles $\{\mathbf{z}_0^{(r)}\}_{r=1}^{R}$ from the initial distribution $p(\mathbf{z}_0)$. This distribution must be parametrically assumed, with potentially unknown parameters. In our case it is assumed to be degenerate as described by (16)–(18). Set $t = 1$.
1. For $t > 0$, use $\{z_{t-1}^{(r)}\}_{r=1}^{R}$ to draw a new swarm $\{\tilde{\mathbf{z}}_t^{(r)}\}_{r=1}^{R}$ from the distribution $p(\mathbf{z}_t|\mathbf{z}_{t-1}, d_{t-1})$. This distribution is known because we have imposed a parametric specification on it. The previous period's choice, d_{t-1}, is required because that determines how the "hatted" $\mathbf{z}$'s evolve. The swarm of particles $\{\tilde{\mathbf{z}}_{t,r}\}_{r=0}^{R}$ now approximates the distribution $p(\mathbf{z}_t|d_{t-1})$.

[18] Thus, we do not allow the transitions of the observed state variables to depend on the realization of the unobserved state variables. This is only for simplicity and is not a limitation of the SMC algorithm.

2. For each $r = 1, \ldots R$, compute $w_t^{(r)} = p(d_t | d_{t-1}, x_t, \hat{x}_t, \tilde{\mathbf{z}}_{t,r})$. The vector $\{w_t^{(r)}\}_{r=1}^{R}$ is known as the vector of importance weights. We can now approximate the integral in (20) by $\frac{1}{R}\sum_{r=1}^{R} w_t^{(r)}$.
3. Draw a new swarm of particles $\{\mathbf{z}_t^{(r)}\}_{r=1}^{R}$ by drawing with replacement from $\{\tilde{\mathbf{z}}_t^{(r)}\}_{r=1}^{R}$. Use the normalized importance weights as sampling probabilities.
4. Set $t = t + 1$ and go to step 1.

Figure 1 presents a graphical representation of the SMC algorithm. The SMC starts at time $t - 1$ with an unweighted measure $\{\tilde{\mathbf{z}}_{t-1}^{(r)}, R^{-1}\}$, which provides an approximation of $p(\mathbf{z}_{t-1} | d_{1:t-2})$. For each particle, the importance weight is computed using the information about the observed choice d_{t-1}. The importance weight is given by the model's prediction of the likelihood $p(d_{t-1} | \tilde{\mathbf{z}}_{t-1}^{(r)})$ of observing d_{t-1} when the particle is $\tilde{\mathbf{z}}_{t-1}^{(r)}$, properly renormalized. This results in the weighted measure $\{\tilde{\mathbf{z}}_{t-1}^{(r)}, \tilde{w}_{t-1}^{(r)}\}$, which yields an approximation of $p(\mathbf{z}_{t-1} | d_{1:t-1})$. Subsequently, the resampling with replacement (or the selection) step selects only the fittest particles to obtain the unweighted measure $\{\mathbf{z}_{t-1}^{(r)}, R^{-1}\}$, which is still an approximation of $p(\mathbf{z}_{t-1} | d_{1:t-1})$. Finally, the prediction step draws new varieties of particles from the parametric process $p(\mathbf{z}_t | \mathbf{z}_{t-1})$, resulting in the measure $\{\tilde{\mathbf{z}}_t^{(r)}, R^{-1}\}$, which is an approximation of $p(\mathbf{z}_t | d_{1:t-1})$. The measures $\{\tilde{\mathbf{z}}_{t-1}^{(r)}, R^{-1}\}$ and $\{\tilde{\mathbf{z}}_t^{(r)}, R^{-1}\}$ are the posterior distributions of the unobservables we use in the numerical integration of the choice probabilities (20).

At each iteration, we compute the per period probability of observing the data given by Equation (20). Because this is done iteratively, starting from $t = 0$, we can eventually work our way up to $t = T$ and compute the entire likelihood given by (19). Repeating this process for each individual for the data gives us the entire likelihood of the data. We can then estimate the parameters via maximum-likelihood.[19]

We use simulated maximum-likelihood to estimate the parameters of $u_1(\cdot)$, $c(\cdot)$, the initial values of z_1, z_2, and z_3, namely θ_{20}–θ_{22} as described in (16)–(18), as well as their AR(1) autocorrelation coefficients and variance terms θ_{23}–θ_{28} as described in (21)–(23). We compute the standard errors using a bootstrap procedure. In each iteration of the procedure, a new random seed is used to create a bootstrapped sample of individuals from the original roster. The structural parameters are then re-estimated using this bootstrapped sample. A total of 50 bootstrapped samples were used. For each structural parameter, the standard error is calculated as the *SD* of the estimates from the 50 bootstrapped samples.[20]

5.3 | Estimation results

Table 7a,b present the estimation results. Panel A shows the estimated coefficients for $u_1(x_{it}, z_{it})$ as specified in (4). The estimates suggests that younger, higher income, healthier, and married

[19]We employed 64 particles in each swarm to integrate out the $\mathbf{z}$'s when computing the conditional choice probabilities. We use 40 particles when computing the expected future value term. One evaluation of the likelihood takes about 2 min on an 8-core, 2.5 GHz, 64 bit AMD computer while using all eight CPUs, and the entire estimation routine took about 4 days when starting from an initial guess of all zeros.

[20]See Olsson and Rydén (2008) for a discussion about the asymptotic performance of approximate maximum-likelihood estimators for state space models obtained via sequential Monte Carlo methods. It provides criteria for how to increase the number of particles and the resolution of the grid to produce estimates that are consistent and asymptotically normal.

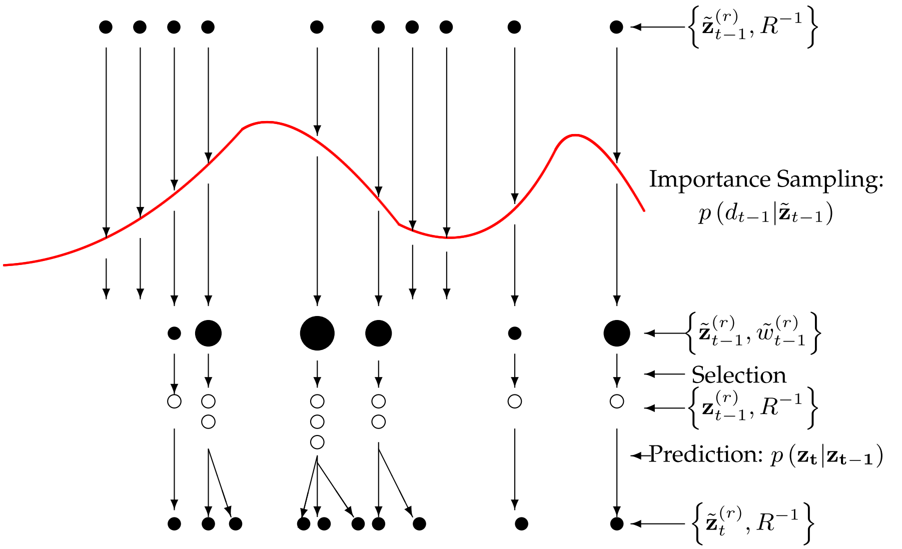

FIGURE 1 A graphical representation of the sequential Monte Carlo algorithm. *Source*: Adapted from Doucet et al. (2001, Chapter 1, p. 12) [Color figure can be viewed at wileyonlinelibrary.com]

individuals and individuals with children have higher utility from purchasing new life insurance (and correspondingly, the opposite are more likely to lapse).[21]

Panel B in Table 7a presents the estimated coefficients for $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$ as specified in (5). The estimates suggest that as ΔAge increases, the suboptimality cost increases, meaning policyholders are more likely to reoptimize as age increases. As income increases, the suboptimality also increases, meaning policyholders are more likely to reoptimize when their incomes increase, but less likely when their incomes go down. The estimates also suggest that individuals who experience negative health shocks have a lower suboptimality penalty—meaning they are more likely to keep their existing policy. This result is consistent with the empirical findings in He (2011) and the theoretical predictions of Hendel and Lizzeri (2003). Finally, the estimates suggest that policyholders who get married are more likely to reoptimize.

Panel C in Table 7b presents the estimated relationship, as specified in (16)–(18), between the initial distributions of the three unobservables z_{10}, z_{20}, and z_{30} and the indicator of whether the individuals owned life insurance in 1994. The positive estimates for coefficients θ_{20} in (16) and θ_{22} in (18) indicate that those who owned life insurance policies in 1994 tend to have higher values of unobservable income and bequest motives; on the other hand, the negative estimate of coefficient θ_{21} in (17) indicates that the policyholders in 1994 tends to be healthier. This result is consistent with the findings in Cawley and Philipson (1999), and most likely reflects survivorship bias as explained in He (2009) and He (2011).

Panel D in Table 7b presents the estimates of the coefficients of the autoregressive processes described in (21)–(23). The estimates for coefficients θ_{23}, θ_{25}, and θ_{27} are all positive and significant (both economically and statistically), suggesting that the unobservable income, health and bequest

[21]Although the coefficient on $(\text{MARRIED} + z_3)$ and $(\text{MARRIED} + z_3)^2$ nearly cancel each other out, the marginal effect of MARRIED is still positive because z_3 is positive on average.

TABLE 7a Estimation results from dynamic model with serially correlated unobservables

	Estimate	*SE*
Panel A: Coefficients for $u_1(x_{it}, z_{it})$		
Constant (θ_0)	−2.4438***	0.0071
Age (θ_1)	−0.0134***	0.0001
Logincome $+z_1(\theta_2)$	0.0139***	0.0007
Conditions $+z_2(\theta_3)$	−0.0358***	0.0018
Married $+z_3(\theta_4)$	−1.1704***	0.0105
Age$^2(\theta_5)$	0.0001***	0.0000
$(\text{Logincome}+z_1)^2(\theta_6)$	0.0070***	0.0001
$(\text{Conditions}+z_2)^2(\theta_7)$	0.0014***	0.0001
$(\text{Married}+z_3)^2(\theta_8)$	1.1663***	0.0092
Has children (θ_9)	0.0374***	0.0022
Has children × Age of youngest child (θ_{10})	0.0046***	0.0002
Panel B: Coefficients for $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$		
Constant (θ_{11})	1.3024***	0.0077
$\Delta\text{Age}(\theta_{12})$	0.1865***	0.0007
$(\Delta\text{Age})^2(\theta_{13})$	−0.0074***	0.0000
$\Delta(\text{Logincome}+z_1)(\theta_{14})$	0.0103***	0.0006
$(\Delta(\text{Logincome}+z_1))^2(\theta_{15})$	−0.0003***	0.0000
$\Delta(\text{Conditions}+z_2)(\theta_{16})$	−0.0443***	0.0014
$(\Delta(\text{Conditions}+z_2))^2(\theta_{17})$	0.0060***	0.0002
$\Delta(\text{Married}+z_3)(\theta_{18})$	0.0136***	0.0010
$(\Delta(\text{Married}+z_3))^2(\theta_{19})$	0.0284***	0.0025

Note: The specifications for $u_1(x_{it}, z_{it})$ and $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$ are given in (4) and (5) respectively. For any variable x, Δx is the difference between the current value of x and $\hat{x}$, which is the value of x at the time when the respondent changed his coverage. The annual discount factor β is set at 0.9.

*, **, and *** represent significance at 10%, 5%, and 1% levels, respectively.

motives shocks are rather persistent, though there are sizeable variations in the unobservables, particularly the unobserved income.

5.4 | Model fit

Table 8 presents an assessment of the performance of the dynamic model with serially correlated unobservable state variables. We report in Panel A of Table 8 the comparisons between the simulated model predictions regarding aggregate choice probabilities by wave and those in the data. Our simulation is able to replicate the increase in the fraction of individuals without life insurance coverage from 12.12% in 1996 to 22.86% in 2006, an increase that almost matches what is in the actual data (from 11.92% in 1996 to 21.36% in 2006).[22]

[22]This is a significant improvement in model fit over an estimated alternative model without serially correlated unobservables, which we report in the online appendix.

TABLE 7b Estimation results from dynamic model with serially correlated unobservables

	Estimate	*SE*
Panel C: Initial distribution of unoversvables		
z_1: whether covered in 1994 (θ_{20})	2.7331**	0.1482
z_2: whether covered in 1994 (θ_{21})	−10.3789***	0.2273
z_3: whether covered in 1994 (θ_{22})	0.4779***	0.0358
Panel D: Transition distribution of unobservables		
z_1: autocorrelation (θ_{23})	0.6957***	0.0158
z_2: autocorrelation (θ_{25})	0.8765***	0.0143
z_3: autocorrelation (θ_{27})	0.4997***	0.0655
z_1: *SD* (θ_{24})	0.2623***	0.0101
z_2: *SD* (θ_{26})	0.0012***	0.0001
z_3: *SD* (θ_{28})	0.0793***	0.0050
Log-likelihood	−9,164.13	

Note: The specifications for $u_1(x_{it}, z_{it})$ and $c((x_{it}, z_{it}), (\hat{x}_{it}, \hat{z}_{it}))$ are given in (4) and (5), respectively. For any variable x, Δx is the difference between the current value of x and $\hat{x}$, which is the value of x at the time when the respondent changed his coverage. The annual discount factor β is set at 0.9.

*, **, and *** represent significance at 10%, 5% and 1% levels, respectively.

TABLE 8 Model fit for dynamic model with serially correlated unobservable state variables

	Wave					
	1996	**1998**	**2000**	**2002**	**2004**	**2006**
Panel A: Aggregate choice probabilities by wave						
Actual data						
No life insurance coverage	0.1192	0.1421	0.1642	0.1889	0.1948	0.2136
Covered, but changed or bought new coverage	0.1493	0.1077	0.0963	0.0978	0.1123	0.0942
Covered, and kept existing coverage	0.7314	0.7500	0.7394	0.7131	0.6928	0.6920
Simulation using dynamic model with serially correlated unobservables						
No life insurance coverage	0.1212	0.1423	0.1608	0.1789	0.2005	0.2286
Covered, but changed or bought new coverage	0.1698	0.1142	0.1046	0.1029	0.1023	0.1032
Covered, and kept existing coverage	0.7089	0.7434	0.7344	0.7180	0.6971	0.6681
Panel B: Cumulative outcomes for 1994 policyholders						
Actual data						
Lapsed to no life insurance	0.0455	0.0899	0.1358	0.1792	0.2094	0.2339
Changed coverage amount	0.0903	0.1472	0.1845	0.2080	0.2283	0.2489
Kept 1994 coverage	0.8641	0.7315	0.6137	0.5202	0.4491	0.3812
Policyholder died	0.0000	0.0312	0.0657	0.0924	0.1130	0.1358
Simulation using dynamic model with serially correlated unobservables						
Lapsed to no life insurance	0.0459	0.0913	0.1317	0.1674	0.2003	0.2343
Changed coverage amount	0.1165	0.1798	0.2183	0.2445	0.2647	0.2816
Kept 1994 coverage	0.8375	0.6981	0.5876	0.4967	0.4229	0.3522
Policyholder died	0.0000	0.0306	0.0622	0.0913	0.1119	0.1318

In Panel B of Table 8, we report the comparison between the simulated model predictions and the data of the cumulative decision outcomes by wave for individuals who owned life insurance in 1994. By incorporating serially correlated unobservable state variables, we are able to capture the pattern of a steadily increasing cumulative fraction of 1994 policyholders lapsing to no insurance in the actual data. In the data, this cumulative fraction went from 4.55% in 1996 to 23.39% by 2006; in our simulation, it goes from 4.59% in 1996 to 23.43% in 2006.

6 | COUNTERFACTUAL SIMULATIONS

In this section, we report the results from a large number of counterfactual simulations to address two important questions. The first set of counterfactual simulations highlights the importance of serially correlated unobserved state variables in explaining the patterns of life insurance decisions observed in the data. The second set of counterfactual simulations attempts to disentangle the contributions of income, health and bequest motives shocks, both observed and unobserved, in explaining the observed lapsations.

It is useful to emphasize at the outset the nature of our counterfactual simulations. When we remove the shocks, we are assuming that the market environment faced by consumers remains unchanged from when all shocks are present. That is, our counterfactual simulation does not allow for the market to re-equilibrate to respond to the fact that there are now fewer shocks. In particular, we must assume that the choice set of life insurance contracts that each individual faces in a given state does not change. Thus our counterfactual simulations are a partial equilibrium exercise.

6.1 | The importance of serially correlated unobserved state variables

In this section, we report a series of counterfactual simulations to demonstrate the importance of including serially correlated unobservable state variables. Panel A of Table 9 is identical to the bottom subpanel of Panel B in Table 8 and it reports the model's predictions about the cumulative outcomes for 1994 policyholders.

In Panel B of Table 9, we report the predictions of the model using the coefficient estimates as reported in Panel A and B of Table 7, but under the counterfactual assumption that the *unobservable* state variables did not change over time. It shows that without the shocks to the unobserved state variables, the model is unable to match the sharply increasing cumulative fraction of 1994 policyholders that lapse to no life insurance, and the model also over-predicts by a large margin the cumulative fraction of 1994 policyholders who kept their 1994 coverage.

In Panel C of Table 9, we report the predictions of the model using the coefficient estimates as reported in Table 7, but under the counterfactual assumption that the *observable* state variables stayed the same as their values in 1994, except for age, while keeping the changes in the unobserved state variables. Surprisingly, assuming away the changes in the observable state variables *barely changes* the model's predictions about the cumulative outcomes for 1994 policyholders.

In Panel D, we report the predictions of the model using the coefficient estimates of the model as reported in Panel A and B of Table 7, but under the counterfactual assumption that *neither* the unobservable state variables *nor* the observed state variables (except for age) changes over time. Only i.i.d. choice specific shocks are retained in these simulations. The

TABLE 9 Counterfactual simulations using the estimates of the dynamic model with serially correlated unobservables: Cumulative outcomes for 1994 policyholders

	Wave					
Outcome	**1996**	**1998**	**2000**	**2002**	**2004**	**2006**
Panel A: All shocks included						
Lapsed to no life insurance	0.0459	0.0913	0.1317	0.1674	0.2003	0.2343
Changed coverage amount	0.1165	0.1798	0.2183	0.2445	0.2647	0.2816
Kept 1994 coverage	0.8375	0.6981	0.5876	0.4967	0.4229	0.3522
Policyholder died	0.0000	0.0306	0.0622	0.0913	0.1119	0.1318
Panel B: No shocks to the unobserved state variables						
Lapsed to no life insurance	0.0433	0.0651	0.0788	0.0889	0.0976	0.1071
Changed coverage amount	0.1171	0.1773	0.2150	0.2425	0.2666	0.2899
Kept 1994 coverage	0.8395	0.7267	0.6419	0.5719	0.5150	0.4571
Policyholder died	0.0000	0.0307	0.0641	0.0965	0.1207	0.1456
Panel C: No shocks to the observable state variables except for age						
Lapsed to no life insurance	0.0428	0.0832	0.1186	0.1502	0.1809	0.2137
Changed coverage amount	0.1149	0.1768	0.2150	0.2416	0.2632	0.2819
Kept 1994 coverage	0.8421	0.7091	0.6033	0.5153	0.4419	0.3697
Policyholder died	0.0000	0.0307	0.0629	0.0927	0.1138	0.1345
Panel D: Only i.i.d. choice specific shocks						
Lapsed to no life insurance	0.0353	0.0520	0.0622	0.0696	0.0761	0.0832
Changed coverage amount	0.1150	0.1728	0.2079	0.2330	0.2554	0.2778
Kept 1994 coverage	0.8496	0.7440	0.6644	0.5985	0.5445	0.4887
Policyholder died	0.0000	0.0310	0.0653	0.0987	0.1238	0.1501

predictions in Panel D are very similar to Panel B where only changes in unobservable state variables are eliminated.

The counterfactual simulations in Table 9 thus provide very strong evidence for the importance of serially correlated unobservable state variables in explaining the choice patterns in the data. Qualitatively, it plays a much more important role than the variations in the observable state variables in capturing the key features in the data.

6.2 | Disentangling the contribution of income, health and bequest motive shocks to lapsations

In this section, we present a series of counterfactual simulations aimed at disentangling the contributions of income, health and bequest shocks, including both observed and unobserved components, to the lapsation of life insurance policies observed in the data. We present our results in four panels in Table 10. There are two sub-panels in Panels A–C. Let us first discuss Panel A, which illustrates the contribution of income shocks to the lapsation of 1994 policyholders. In the shaded subpanel, we use as baseline the model's prediction of the cumulative lapsation rates of 1994 policyholders when only i.i.d. choice specific shocks are present (the first row), and examine how the addition of income shocks increases the model's predicted lapsation

TABLE 10 Disentangling the contributions of income, health and bequest motive shocks to the lapsations of 1994 policyholders

	Wave					
	1996	**1998**	**2000**	**2002**	**2004**	**2006**
Panel A: The Role of Income Shocks						
[1] i.i.d. choice specific shocks only	0.0353	0.0520	0.0622	0.0696	0.0761	0.0832
[2] i.i.d. and income shocks only	0.0384	0.0613	0.0772	0.0897	0.1011	0.1134
[3] Incremental contribution of income shocks (%)	6.75	10.19	11.39	12.01	12.48	12.89
[4] All but income shocks	0.0415	0.0740	0.0988	0.1186	0.1365	0.1555
[5] All shocks included	0.0459	0.0913	0.1317	0.1674	0.2003	0.2343
[6] Incremental contribution of income shocks (%)	9.59	18.95	24.98	29.15	31.85	33.63
Panel B: The Role of Health Shocks						
[1] i.i.d. choice specific shocks only	0.0353	0.0520	0.0622	0.0696	0.0761	0.0832
[2] i.i.d. and health shocks only	0.0409	0.0642	0.0793	0.0905	0.1002	0.1101
[3] Incremental contribution of health shocks (%)	12.20	13.36	12.98	12.49	12.03	11.48
[4] All but health shocks	0.0386	0.0701	0.0957	0.1178	0.1386	0.1617
[5] All shocks included	0.0459	0.0913	0.1317	0.1674	0.2003	0.2343
[6] Incremental contribution of health shocks (%)	15.90	23.22	27.33	29.63	30.80	30.99
Panel C: The Role of Bequest Motive Shocks						
[1] i.i.d. choice specific shocks only	0.0353	0.0520	0.0622	0.0696	0.0761	0.0832
[2] i.i.d. and bequest motive shocks only	0.0352	0.0579	0.0739	0.0866	0.0980	0.1113
[3] Incremental contribution of bequest shocks (%)	−0.22	6.46	8.88	10.16	10.93	11.99
[4] All but bequest shocks	0.0444	0.0759	0.0997	0.1187	0.1357	0.1526
[5] All shocks included	0.0459	0.0913	0.1317	0.1674	0.2003	0.2343
[6] Incremental contribution of bequest shocks (%)	3.27	16.87	24.30	29.09	32.25	34.87
Panel D: Contributions of i.i.d. Choice Specific Shocks						
[1] Lower bound (%)	71.24	40.96	23.39	12.13	5.09	0.51
[2] Upper bound (%)	81.26	69.99	66.74	65.35	64.55	63.64

rates (the second row).[23] Notice that the addition of income shocks to the i.i.d. choice specific shocks leads to more lapsation. The incremental contribution of income shocks accounts for about 6.75% of the total lapsation predicted by the model when all shocks are included in 1996.[24] The incremental contributions of income shocks over time are reported in the third row. It reveals that the importance of income shocks are increasing over time in explaining lapsation. By 2006, income shocks alone were able to explain about 12.89% of the predicted lapsation.

The bottom, unshaded, subpanel in Panel A uses a different baseline. The baseline is instead the model's prediction of lapsation rates when *all but* income shocks are included

[23]Note that the first row numbers in Panel A of Table 10 are identical to the numbers in the first row of Panel D of Table 9.

[24]That is, $(0.0384 - 0.0353)/0.0459 \approx 6.75\%$, where 0.0459 is lapsation rates predicted by the model when all shocks are included (reported in the fifth row of the table, as well as Panel B of Table 8). The other percentages are calculated analogously.

(reported in the fourth row of Panel A). This baseline prediction is compared to the predicted lapsation rates when all shocks are included (fifth row of Panel A). The difference is attributed to the incremental contribution of income shocks (sixth row of Panel A). Using this baseline, we see that the contribution of income shocks to lapsation is also increasing over time, from 9.59% in 1996 to 33.63% in 2006.[25]

Panel B of Table 10 carries analogous calculations to illustrate the contribution of health shocks to the lapsations of 1994 life insurance policyholders. The shaded subpanel shows that if we use the predicted lapsation rates with only i.i.d. choice specific shocks as the baseline, the incremental contribution from adding health shocks using this baseline is more or less stable over time, staying at about 12% throughout the years. If we use the predicted lapsation rates when all but health shocks are included as the baseline, the incremental contribution from adding health shocks goes from 15.90% in 1996 to 31% in 2006.[26]

Panel C of Table 10 shows the contribution of bequest motive shocks to life insurance lapsation. As in Panels A and B, the top shaded subpanel calculates the incremental contribution of bequest motive shocks using the predicted lapsation with only i.i.d. choice specific shocks as the baseline, and the bottom subpanel uses the predicted lapsation with all shocks except for bequest shocks as the baseline. We find that the importance of bequest motives in explaining lapsation increases over time. In 1996 the bequest motive shocks explain between −0.22% and 3.27% of the lapsation; but by 2006, it explains between 11.99% and 34.87% of the lapsation.

Panel D of Table 10 bounds the contributions of i.i.d. choice specific shocks in explaining the lapsations. The lower bounds are calculated as the residuals after subtracting the upper bound contributions from income, health and bequest motive shocks.[27] Panel D reveals that lapsation of life insurance policies are largely driven by i.i.d. choice specific shocks for younger individuals, but for surviving policyholders an ever larger fraction of lapsations is explained by either income, or health, or bequest motive shocks.[28] By 2006, between more than 1/3 to almost 100% of the lapsations are driven by one of these shocks.

To summarize, the simulations reported in Table 10 indicate that when individuals are young, most of the life insurance policy lapses are driven by i.i.d. choice specific shocks, and the rest is explained, in descending order of importance, by health shocks and income shocks; the bequest motive shocks only account for a very minor fraction of lapsation. However, as policyholders get older, the importance of the i.i.d. choice specific shocks declines dramatically, and the three shocks eventually account for about the same fraction

[25]There are two other possible counterfactual baselines that we do not report. We could have used “i.i.d. choice specific shock and health shocks” as baseline and contrast it with “i.i.d. choice specific shock, health shocks and income shocks” (which is the same as “all but bequest motive shocks”). Alternatively, we could have used “i.i.d. choice specific shock and bequest motive shocks” as the baseline and contrast it with “i.i.d. choice specific shock, bequeest motive shocks and income shocks” (which is the same as “all but health shocks”). Note that the information needed to carry out these calculations is presented in other rows in Table 10. For space reasons, we do not present these calculations separately.

[26]It is worthwhile pointing out that the discrepancy in the results for different baselines is due to the nonlinearities in our model.

[27]For example, we obtain 71.24% lower bound number for year 1996 from 1 − 9.59% − 15.90% − 3.27%, where 9.59%, 15.90%, and 3.27% are respectively the upperbound contributions of income, health, and bequest motive shocks reported in Panels A to C. The other bounds are calculated analogously.

[28]The large role for i.i.d. shocks that we find is consistent with Gottlieb and Smetters (2020), who found that forgetting to pay premiums is a significant driver of lapsation, since forgetfulness can be subsumed into our i.i.d. shocks to lapsation.

of lapsation (ranging from about 12% to around 30%). If anything, the bequest motive shocks are more important in explaining lapsation among older individuals than the income and health shocks.

6.3 | Policy implications

Our findings on the relative contributions of income, health, bequest motive, and i.i.d. choice shocks to the lapsation rates of 1994 policyholders have implications regarding the possible consumer welfare effects of emerging life settlement markets. As we mentioned in the Introduction, theoretical studies by Daily et al. (2008) and Fang and Kung (2010a, 2010b, 2020) show that the reasons for lapsation will determine whether life settlement markets can improve consumer welfare. Specifically, Fang and Kung (2010b) show that a key determinant for whether consumers may benefit from life settlement markets is whether lapsation is driven by factors that are positively correlated with the marginal utility of income. Thus, to the extent we found in Section 5.3 that decreases in log income lead to a higher probability of lapsation, and decreases in log income certainly lead to higher marginal utility of income, the fraction of lapsation that can be attributed to changes in income, both observed and unobserved, should be a potential source of welfare gain for consumers when life settlement markets are introduced. On the other hand, lapsation driven by i.i.d. choice specific shocks are not positively correlated to the marginal utility of income, and thus lapsation driven by i.i.d. choice specific shocks can lead to a reduction in consumer welfare. Analogously, to the extent that health shocks and bequest motive shocks are orthogonal to income shocks, and are thus not necessarily positively correlated to the marginal utility of income, we suspect that the fractions of lapsation attributable to these two shocks are likely sources for consumer welfare reduction when life settlement is introduced.

Our finding that i.i.d. choice specific shocks explain the bulk of the policy lapsation when individuals are relatively young (when they are in their early 60 s) suggests that life settlement is likely to lead to a welfare loss for relatively young policyholders; but may lead to welfare gains for older policyholders in their early seventies, as changes in income become a more important source for lapsation. We should emphasize, however, that these implications from our analysis are only suggestive; a more definite study of the welfare effect of the life settlement market would require that we estimate a fully structural model of the behavior of both the consumers who choose life insurance policies and the life insurance companies who offer such policies.

7 | DISCUSSION

We now discuss two important issues. The first issue is about the identification of the three components of the serially correlated unobservable state variables intended to capture income, health and bequest motive shocks. To the best of our knowledge, this is the first paper that allows for more than one unobservable state variable. So a natural question is whether the distributions of such unobservable state variables can be separately identified. This is obviously an important question to be addressed in future research. For now, we would first like to emphasize that in this paper, we tried to anchor the interpretation of these three shocks by restricting that each has the same effect on behavior as their respective observable counterparts (see Section 5.1).

TABLE 11 Log-likelihoods of various specifications of unobservable state variables

Specification	Serially correlated unobservables? Income?	Health?	Bequest motive?	Log-likelihood
All	Yes	Yes	Yes	−9164.13
1	No	Yes	Yes	−9246.65
2	Yes	No	Yes	−9233.80
3	Yes	Yes	No	−9219.66
4	No	No	Yes	−9274.13
5	No	Yes	No	−9288.69
6	Yes	No	No	−9285.87
None	No	No	No	−9338.59

To give further evidence that patterns in the data justify the inclusion of multiple dimensions of serially correlated unobserved state variables, we also estimated a series of alternative models where we include only a subset of the three shocks. In the same spirit of Heckman and Singer (1984) for the case of unobserved types, we ask whether the inclusion of additional unobserved state variables increases the log-likelihood of the estimated model. In Table 11, we report the log-likelihood of the estimated models with various specifications of the unobservable shocks. In particular, the specification labeled "All" corresponds to the model estimated in Section 5 where we include all three unobservable state variables, and the specification labeled "None" corresponds to a model estimated without any unobservable state variables. In specifications labeled 1–6, various combinations of the three unobservable state variables are included in the estimation. From the last column in Table 11, we can see that the inclusion of the additional unobservable state variables significantly increases the log-likelihood of the models. For example, in specifications 4–6, we estimated models with only one of the unobservable state variables respectively. The log-likelihoods of these models improve over specification "None." Similarly, in specifications 1–3, we estimated models with two of the three unobservable state variables; and again, the log-likelihoods of these models improve over specifications with only one of the unobservables. Finally, the log-likelihood of the model with all three shocks is higher than specification 1–3. The results in Table 11 show that the data indeed seems to be more consistent with a model using all three serially correlated unobserved state variables.[29]

The second issue is regarding our finding that the importance of the unobserved state variables in explaining lapsation increases over time. The concern is whether this is a mechanical result due to the way we simulate the unobservable state variables using SMC. In particular, recall that the initial distribution of the unobservable state variables is assumed to have smaller support than in later periods (see Sections 5.1 and 5.2). While this is a possibility, we would like to make two counter-arguments. First, even though the unobservable state variables in the earlier periods have smaller support (in fact, just one point support in the initial period), these points in the support were chosen to best fit the data; thus there is no a priori reason that the unobservable state variables in the early periods should have less impact just because they have a small support.

[29]We also conducted formal likelihood ratio tests of the specifications 1-6 and "None" against specification "All." The tests are all in favor of specification "All."

Second, if the simulated unobservable state variables were pure noise that the individuals do not take into account, then the unobservables' importance in explaining the observed lapsations should not have changed over time at all, that is, the importance should be about zero in all periods. Thus the fact that the importance of the unobservable state variables were found to be increasing over time is an indication that these simulated unobservable state variables are capturing something informative.

8 | CONCLUSION

In this paper, we empirically investigate the contributions of income, health, and bequest motive shocks to life insurance lapsation. We present a dynamic discrete choice model of life insurance decisions allowing for serially correlated unobservable state variables. The model is designed to deal with the data reality where researchers only observe whether an individual has made a new life insurance decision (i.e., purchased a new policy, or added to/changed an existing policy) but do not observe the actual policy choice or the choice set from which the new policy is selected. The semistructural dynamic discrete choice model allows us to bypass these data limitations. We empirically implement the model using the limited life insurance holding information from the HRS data.

We deal with serially correlated unobserved state variables using posterior distributions of the unobservables simulated from SMC methods. Relative to the few existing papers in the economics literature that have used similar SMC methods, our paper is the first to incorporate multi-dimensional serially correlated unobserved state variables. To give the three unobservable state variables in our empirical model their desired interpretations as unobserved income, health, and bequest motive shocks, this paper proposes two channels through which we can anchor these unobservables to their related observable variables.

Our estimates for the model with serially correlated unobservable state variables are sensible and yield implications about individuals' life insurance decisions consistent with the both intuition and existing empirical results. In a series of counterfactual simulations reported in Table 10, we find that a large fraction of life insurance lapsation is driven by i.i.d. choice specific shocks, particularly when policyholders are relatively young. But as the remaining policyholders get older, the role of such i.i.d. shocks gets less important, and more of their lapsation is driven either by income, health or bequest motive shocks. Income and health shocks are relatively more important than bequest motive shocks in explaining lapsation when policyholders are young, but as they age, the bequest motive shocks play a more important role. We also show that in the model with unobserved state variables, the contribution of the shocks to unobservables is much larger than the contribution of the shocks to observed state variables (Table 9).

Our empirical findings have important implications regarding the effect of the life settlement industry on consumer welfare. As shown in theoretical analysis in Daily et al. (2008), Fang and Kung (2010a, 2010b, 2020) and Fang and Wu (2020), the theoretical predictions about the effect of life settlement on consumer welfare crucially depend on why life insurance policyholders lapse their policies. If bequest motive shocks are the reason for lapsation, then the life settlement industry is shown to reduce consumer welfare in equilibrium; but if income shocks are the reason for lapsation, then life settlements may increase consumer welfare. To the extent that we find both income shocks and bequest motive shocks play important roles in explaining life insurance lapsations, particularly among the elderly population targeted by the life settlement industry, our research suggests that the effect of life settlement on consumer welfare is ambiguous. Unfortunately, our "semistructural," partial equilibrium model of life insurance decisions (which is necessitated by data limitations) is not suitable for a quantitative general equilibrium evaluation of the welfare impacts of introducing a life settlement market. This is an important, but challenging, area for future research.

ACKNOWLEDGMENTS

An earlier version of the paper appeared as NBER Working Paper No. 17899 (March 2012). We have received helpful comments and suggestions from two anonymous referees, the Editor, Steven Berry, Xu Cheng, Daniel Gottlieb, Han Hong, Aprajit Mahajan, Panle Jia Barwick, Alessandro Lizzeri, Jim Poterba, Kent Smetters, Xun Tang, Ken Wolpin, Motohiro Yogo and seminar/conference participants at New York University, Cowles Foundation Summer Conference in Structural Microeconomics (2010), AEA/CEANA (2011) and SED (2012) for many helpful comments, questions and suggestions. Fang would also like to gratefully acknowledge the generous financial support from the National Science Foundation through Grant SES-0844845. All remaining errors are our own.

REFERENCES

American Council of Life Insurers. (2009). Life Insurers Fact Book.

Blevins, J. R. (2016). Sequential Monte Carlo methods for estimating dynamic microeconomic models. *Journal of Applied Econometrics*, *31*, 773–804.

Cawley, J., & Philipson, T. (1999). An empirical examination of information barriers to trade in insurance. *American Economic Review*, *89*, 827–846.

Cole, C. R., & Fier, S. G. (2020). An examination of life insurance policy surrender and loan activity. *Journal of Risk and Insurance*. (forthcoming). https://doi.org/10.1111/jori.12329

Daily, G., Hendel, I., & Lizzeri, A. (2008). Does the secondary life insurance market threaten dynamic insurance? *American Economic Review Papers and Proceedings*, *98*(2), 151–156.

Doucet, A., de Freitas, N., & Gordon, N. (2001). *Sequential Monte Carlo methods in practice*. Springer.

Fang, H. & Kung, E. (2010a). *How does the life settlements market affect the primary market for life insurance?* (NBER Working paper 15761).

Fang, H. & Kung, E. (2010b). *The welfare effect of life settlement market: the case of income shocks*. Working Paper, University of Pennsylvania and Duke University.

Fang, H., & Kung, E. (2020). Life insurance and life settlements: the case for health-contingent cash surrender values. *Journal of Risk and Insurance*, *87*(1), 7–39.

Fang, H., & Wu, Z. (2020). Life insurance and life settlement markets with overconfident policyholders. *Journal of Economic Theory*, *189*, 105093.

Fernandez-Villaverde, J., & Rubio-Ramirez, J. F. (2007). Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies*, *2007*(74), 1059–1087.

Fier, S. G., & Liebenberg, A. P. (2013). Life insurance lapse behavior. *North American Actuarial Journal*, *17*(2), 153–167.

Fox, J. T. (2007). Semiparametric estimation of multinomial discrete-choice models using a subset of choices. *Rand Journal of Economics*, *38*(4), 1002–1019.

Gilbert, J., & Schultz, E. (1994). *Consumer reports life insurance handbook*. Consumer Reports Books.

Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEEE Proceedings of Radar and Signal Processing*, *140*(2), 107–113.

Gottlieb, D., & Smetters, K. (2020). Lapse-based insurance. *American Economic Review*. (forthcoming).

He, D. (2009). The life insurance market: Asymmetric information revisited. *Journal of Public Economics*, *93*, 1093–1097.

He, D. (2011). Is dynamic selection present in the life insurance market? *Economics Letters*, *112*(1), 113–115.

Heckman, J., & Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, *52*(2), 271–320.

Hendel, I., & Lizzeri, A. (2003). The role of commitment in dynamic contracts: evidence from life insurance. *Quarterly Journal of Economics*, *118*(1), 299–327.

Hong, H., Gallant, R. A., & Khwaja, A. (2018). A bayesian approach to estimation of dynamic models with small and large number of heterogeneous players and latent serially correlated states. *Journal of Econometrics*, *203*(1), 19–32.

Hu, Y., & Shum, M. (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics*, *171*(1), 32–44.

Kasahara, H., & Shimotsu, K. (2009). Nonparametric identification of finite mixture models of dynamic discrete choice. *Econometrica*, *77*(1), 135–175.

Kung, E. (2012). *Efficient mortgage design in an equilibrium model of housing and mortgage markets*. Working paper, Duke University.

Life Insurance Marketing and Research Association International (LIMRA). (2009). *U.S. Individual Life Insurance Persistency: A Joint Study Sponsored by LIMRA and the Society of Actuaries. A 2009 Report*. Windsor, CT, USA.

McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, & J. Weibull (Eds.), *Spatial interaction theory and planning models* (pp. 75–96). North Holland.

Norets, A. (2009). Inference in dynamic discrete choice models with serially correlated unobserved state variables. *Econometrica*, *77*(5), 1665–1682.

Olsson, J., & Rydén, T. (2008). Asymptotic properties of particle filter-based maximum likelihood estimators for state space models. *Stochastic Processes and their Applications*, *118*(4), 649–680.

Rust, J. (1994). Structural estimation of Markov decision processes? In R. Engle, & D. McFadden (Eds.), *Handbook of econometrics* (pp. 3082–3139). North Holland.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Fang H, Kung E. Why do life insurance policyholders lapse? The roles of income, health, and bequest motive shocks. *J Risk Insur*. 2020;1–34. https://doi.org/10.1111/jori.12332