Insurance: Mathematics and Economics 71 (2016) 50–62

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

Longevity risk and retirement income tax efficiency: A location spending rate puzzle

Huaxiong Huang [a], Moshe A. Milevsky [b,*]

[a] Department of Mathematics and Statistics, York University and the Fields Institute, Toronto, Canada

[b] Schulich School of Business, York University and the IFID Centre, Toronto, Canada

ARTICLE INFO

Article history:
Received April 2016
Received in revised form
August 2016
Accepted 11 August 2016
Available online 24 August 2016

Keywords:
Pension planning
Retirement planning
Lifecycle model
Longevity risk
Optimal consumption

ABSTRACT

In this paper we model and solve a retirement consumption problem with differentially taxed accounts, parameterized by longevity risk aversion. The work is motivated by some observations on how Canadians de-accumulate financial wealth during retirement — which seem rather puzzling. While the Modigliani lifecycle model can justify a variety of (pre-tax) de-accumulation or draw down rates depending on risk preferences, the existence of asymmetric taxes implies that certain financial accounts should be depleted faster than others. Our analysis of data from the *Survey of Financial Security* indicates that Canadian retirees maintain approximately two-thirds of their financial wealth in tax-sheltered accounts and a third in taxable accounts *regardless of age*. The ratio of taxable to tax-sheltered wealth increases slightly or remains relatively constant depending on household income which is not what one would expect from the lifecycle model. Indeed, using our model we cannot locate a plausible tax function that justifies a constant "account ratio" regardless of age. For example under flat rates taxable accounts should be depleted well before tax-sheltered accounts are ever touched. The account ratio should go to zero quite rapidly in the absence of government mandated withdrawals. We also demonstrate that under progressive income taxes withdrawals are made from both accounts but at different rates depending on account size, pension income and longevity risk preferences. Again, the "account ratio" should eventually decline. We postulate that this sort of behavior is likely due to irrational considerations linked to mental accounting, etc. It remains to be seen whether this will persist over time and under a more careful analysis of Canadian cohorts or if retirees in other countries exhibit the same behavior.

Let me tell you how it will be.
There's one for you, nineteen for me.
Cause I'm the taxman. Yeah, I'm the taxman.
Should five per cent appear too small,
Be thankful I don't take it all.

[Beatles, The Taxman 1966]

1. Introduction and motivation

The classic Modigliani and Brumberg (1954) lifecycle hypothesis predicts that young households accumulate wealth during their working years and then slowly deplete it and spend-down during their retirement years. The precise rate at which a household's net-worth is accumulated or de-accumulated over the lifecycle depends on a large number of subjective (personal preference) and objective (financial economic) parameters, creating a wide range of potentially observable behavior.

Ceteris paribus, impatient households accumulate less wealth and/or spend it down faster. Likewise, equally patient households facing uncertainty about how long they will live might decide to draw down financial assets at a slower rate if they are *longevity-risk* averse. Those who are more tolerant of this risk – for example willing to reduce their standard of living in the unlikely event they become centenarians – might decide to draw down their wealth at a faster rate, thus mimicking an impatient household. The lifecycle model allows for quite a bit of flexibility, especially when bequest motives are included in the mix. Nevertheless, longevity risk and the household's attitude to this uncertainty plays a big part in the optimal de-accumulation rate, as per the original work by Yaari (1965) — with or without annuities.

In support of this, Fig. 1 provides a sense of the heterogeneity of retirement wealth depletion rates between the ages of 65 and 80, based on data from Statistics Canada's Survey of Financial Security (SFS, 2012). More about this survey will come later, but for now we

* Correspondence to: 4700 Keele Street, Toronto, Ontario, Canada, M3J 1P3.
E-mail addresses: hhuang@mathstat.yorku.ca (H. Huang), milevsky@yorku.ca (M.A. Milevsky).

http://dx.doi.org/10.1016/j.insmatheco.2016.08.002
0167-6687/

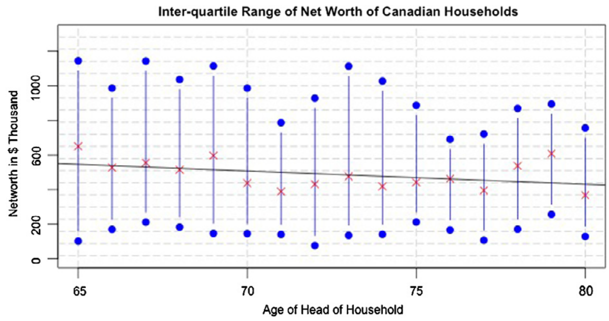

Fig. 1. (Most recent wave.) 50% of financial wealth is within given range. Sum of the market value of 10 financial categories plus 5 non-financial categories minus the value of 7 debt categories. Sample weighted. Compiled by authors.
Source: SFS 2012.

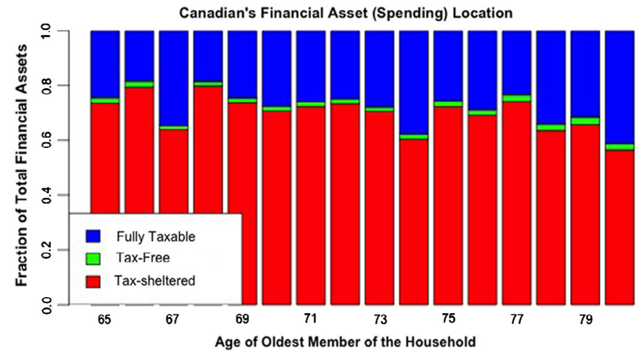

Fig. 2. Ratio of market values for 10 financial categories (only). The fully-taxable is what we call the B-account, and the tax-sheltered is the A-account. We do not address the third (tax-free) C-account in this analysis.
Source: SFS 2012.

present a simple regression.

$$w_i = \alpha_0 + \alpha_1 x_i + e_i.$$

With age x_i as the independent variable, net-worth w_i declines by approximately $\alpha_1 = -\$18{,}000$ per year for the 3,179 households (observations) at or above the age of 65 in the SFS2012 dataset. And although this (negative) coefficient is significant at the 99.9% level of significance, the adjusted R^2 is a mere 0.0095, which leaves much unexplained variation in wealth depletion or draw down rates. Not shown in Fig. 1 are the financial assets (a.k.a. investment accounts), which are obviously a subset of net-worth, which decline by approximately 4% per year above the age of 65. Of course, both these estimates are based on a cross-section as opposed to an actual time series for a given household, which is something we will return to later.

Intuitively though, households with (very) strong bequest motives might rationally decide to spend less than the interest, dividends and capital gains they earn on their financial assets. Others might continue to accumulate wealth for the same legacy reasons. Again the heterogeneity in observed behavior is justifiable and fits within the standard lifecycle framework and is not the focus of our attention. The Modigliani lifecycle 'tent' is wide enough to accommodate many observed spending or drawdown practices.

However, in contrast to the rather broad range of possible **pre-tax** depletion or draw-down rates during retirement – see for example the widely cited work by Blake et al. (2003) – once a household decides how much they want to spend from their financial accounts or assets, the exact source or location is rather limited — if they are behaving tax optimally. In other words, whether a retiree would like to (only) spend 1% of their financial accounts in a given year, or a more aggressive 10%, the various tax-accounts from where the cash-flow is sourced should not depend on subjective and personal factors. It should be driven by tax efficiency considerations alone.

Here is the bottom line, certain *accounts* should be depleted faster than others. And yet, this does not seem to be the case in practice (using our Canadian data.) The same SFS2012 dataset indicates a puzzling phenomenon regarding the pattern by which Canadians de-accumulate and draw down wealth during their retirement years. On average Canadians maintain approximately two thirds of their investable wealth in tax-sheltered accounts regardless of age. The relative wealth ratio does not appear to decline with the age of the household and is remarkably (and statistically significant) constant for higher net-worth households. Fig. 2 provides a visual illustration of aggregate assets by age. We offer a more detailed statistical analysis in Section 6.

So much for the motivating empirical evidence. To show that this behavior is inconsistent with rationality, we formulate and solve a lifecycle optimization problem using techniques from the Calculus of Variation (CV). In particular, we extend the standard no-tax (lifecycle) model to one in which there are a variety of tax accounts with different tax treatments. We investigate the optimal tax-adjusted depletion rates and specifically the ratio of taxable investment wealth (what we label F_t^B) to total investment wealth ($F_t = F_t^A + F_t^B$) over the entire retirement horizon. We label this $R_t^B := F_t^B / F_t$ and discuss its properties for a variety of tax rate specifications.

Indeed, we cannot locate a plausible marginal-tax rate function (denoted by $\xi(z)$ in the paper) that justifies the ratios we observe in the data. For example, under flat marginal tax rates, the heavily taxed B-accounts should be completely depleted well before tax-sheltered A-accounts are ever touched. In other words, the ratio R_t^B should go to zero quite rapidly. And, under progressive marginal income taxes withdrawals are made from both accounts at the same time, but at different rates depending on account size, pre-existing pension income and longevity risk aversion. Eventually the ratio should decline, even when the model is expanded to include required minimum distributions (RMD) and other frictions.

In sum, we are hard-pressed to find a rational explanation for maintaining an increasing and/or constant tax-account ratio. And, while our empirical result is only suggestive, it certainly suggests the need for further and careful research. At the very least, this paper offers some normative guidance on how it *should* be done.

1.1. Methodological contributions of this study

The resulting nonlinear coupling between the two types of wealth accounts (F_t^A, F_t^B) poses interesting technical challenges for solving the optimal retirement depletion problem and is really the main technical contribution of this paper. The simplest approach would be to use brute numerical force (i.e. the Bellman equation) directly, based on the dynamic optimization principle. While this method is conceptually simple and easy to implement it is computationally intense and extremely slow. In this paper we use a traditional method based on the Calculus of Variations (CV). The main technical challenge then, is to develop an efficient method that can handle the nonlinearity of the tax-function as well as the various constraints, effectively. We illustrate and apply the algorithm using a continuous and monotonic approximation to the tax rate function.

One rarely-highlighted result in (this corner of) lifecycle theory, is that in the presence of a minimal pension income and lifetime uncertainty, it is optimal to deplete all liquid wealth prior to the maximum length of life. The so-called wealth depletion time (WDT) and the corresponding initial withdrawal rate (IWR) are both unknown variables which have to be solved simultaneously in any lifecycle model with pension income. This subtle fact was