

Mortality Swaps and Tax Arbitrage in the Canadian Insurance and Annuity Markets

Narat Charupat and Moshe A. Milevsky*

*Charupat is an Assistant Professor of Finance at the DeGroote School of Business, McMaster University. Milevsky is an Associate Professor of Finance at the Schulich School of Business, York University, and the Director of the Individual Finance and Insurance Decisions (IFID) Centre in Toronto, Canada. The Financial support from the Life Underwriters Association of Canada (Charupat), the York University Research Authority and the SSHRC (Milevsky) are gratefully acknowledged. The authors would like to thank Glenn Daily and Clarence Kwan for their helpful suggestions, as well as Lowell Aronoff (CANNEX), John Hitchcock (SunLife), Jon Archer (RBC) and Zale Newman (PanFinancial) for providing data on insurance and annuity quotes. Any errors are the authors' responsibilities.

Abstract

We analyze a tax arbitrage opportunity that results from engaging in two seemingly counter-intuitive transactions in the Canadian insurance market. Specifically, if an individual acquires a fixed immediate life annuity and then uses the periodic annuity income to fund a term-to-life insurance policy, these two transactions, which we refer to as a "mortality swap", will generate a payoff pattern that is risk-free. In other words, a mortality swap replicates a risk-free security, albeit one with a stochastic liquidation date. We show theoretically that the rate of return on a mortality swap is equal to the risk-free rate on a *before-tax* basis, but exceeds it on an *after-tax* basis. This is confirmed by the results of our empirical test which uses observed annuity and insurance quotes that already reflected adverse-selection and transaction costs. We also observe that the older an individual is and/or the higher his/her marginal tax rate is, the more he/she stands to gain from this tax arbitrage.

This advantageous investment opportunity exists because of the arguably lenient method that Canadian authorities use to tax annuity income. We provide two major reasons why this method leads to an arbitrage opportunity. We then compare this method to that under the U.S. tax rules, and show that the U.S. method renders tax arbitrage very unlikely.

The demand for life insurance and annuities is usually attributed to risk aversion, consumption smoothing, and the desire for household protection. Our findings provide an arbitrage-based reason for their demand. As a natural by-product, our research contains policy implications for the optimal taxation of annuities and insurance policies.

1. Introduction

The essence of arbitrage-free pricing is that assets with identical payoff patterns should have the same price. The equality of prices and, by extension, rates of return holds true if markets are sufficiently free of frictions. In such markets, investors will be indifferent among the choices of assets with similar payoffs.

In practice, however, market frictions such as transaction costs, income taxes, and asymmetry of information can significantly distort investment decisions by causing one investment alternative to be less costly - thus providing a higher rate of return - than others. Previous research has investigated this possibility, both theoretically and empirically. (See, for example, Jarrow and O'Hara (1989) and Kamara and Miller (1995))

In this paper, we document a return discrepancy between the Canadian insurance and fixed-income markets. We show that by engaging in seemingly counter-intuitive transactions involving two insurance products, one can create a risk-free portfolio whose after-tax return is greater than that of available risk-free securities. The two insurance products in questions are (i) a standard term-to-100 life insurance policy; and (ii) a single-premium fixed immediate life annuity with no guarantee period.

Consider an individual who invests \$100,000 in a fixed immediate life annuity, and then uses part of the periodic income from the annuity to pay the premium on a life insurance policy whose death benefit is also \$100,000. This 'back-to-back' transaction, which we shall henceforth refer to as a "mortality swap", will create a constant periodic flow of income and will return the original \$100,000 upon the death of the policy owner. This payoff pattern is similar to that of a risk-free investment such as a bank deposit whose principal is redeemed (by the individual's estate) at the time of death.¹

In a frictionless market, the individual would be indifferent between a mortality swap and a bank deposit. However, we show that in the presence of Canadian personal income taxes, a mortality swap yields a considerably higher *after-tax* rate of return than the risk-free rate. This discrepancy is due to the (arguably) lenient methods by which Canadian authorities tax "prescribed" annuity income.² The benefit of this lenient taxation cannot be captured by

¹ The payoff of a mortality swap is risk-free in the sense that there is no variation in it, given that the insurance company does not default. In the same sense, a bank deposit is risk-free except for the bank's default risk. Therefore, our comparison of the two is valid, especially for insurance companies and banks of the highest credit rating.

² In Canada, if an individual purchases a life annuity with assets from his/her registered retirement savings plan, the annuity income will be wholly taxable. On the other hand, if the annuity is purchased with assets outside a registered plan, only a portion of its income will be taxable. How that portion is determined depends on whether or not these "non-registered" life annuities qualify for the "prescribed" status. If they

investing in an annuity by itself because of the mortality risk of its returns. However, by combining an annuity and an insurance policy into a mortality swap, the mortality risk of the annuity return is offset by the mortality risk of the insurance return, and the benefit of the lenient tax treatment is captured.

Under the standard arbitrage argument, one can therefore create "arbitrage" profits by borrowing money to fund a mortality swap, thus creating a portfolio with zero initial cost and a positive payoff with probability one. However, this profitable opportunity is not an arbitrage opportunity in the traditional sense. This is because it is not scalable since individuals are limited by the amount of insurance that they can purchase. In addition, it is not entirely riskless as tax laws might change without providing a "grandparent" clause. Therefore, a more appropriate way to look at this strategy is from the point of view of an investor who is deciding between a risk-free security and a mortality swap. Based on this point of view, we define "arbitrage" for our context as a strategy where a mortality swap provides a superior return to the risk-free rate. It is under this weaker definition that we will use the term "arbitrage" in the remaining of the paper.

Our research sheds light on the interaction between the insurance and financial markets *vis a vis* the proper taxation of these products. While the practice of back-to-back transactions has been around for a while (mainly among wealthy individuals) there has not been a rigorous study of the magnitude of, or the underlying reasons for, its benefits.³

From a micro-economic perspective, our study provides a tax-driven explanation for the demand for insurance and annuities. This explanation differs from, but does not contradict, the classical life-cycle smoothing (risk-aversion) argument proposed in Yaari (1965), Fischer (1973), Campbell (1980), Karni and Zilcha (1986) and Lewis (1989). In addition, while life insurance is extensively studied and widely understood (see the classic book by Black and Skipper (1999), immediate life annuities are still a growing area (see the work by Poterba (1997), Mitchel, Poterba, Warshawsky and Brown (2000) and Brown, Mitchel, Poterba, Warshawsky (1999), and Milevsky (2001) for some possible explanations). This paper contributes to the annuity literature and also links it to the concept of arbitrage.

Although our main analysis is done under the Canadian tax rules, the results have implications on other tax jurisdictions as well. For example, under the present U.S. tax regulations, life annuities are taxed using a slightly different scheme that results in a higher effective tax rate. We explain why tax arbitrage is unlikely under this scheme. Recently,

do, then the taxable portion remains constant throughout the annuitant's life. In practice, most life annuities in Canada qualify for the prescribed status.

³ Back-to-back transactions are sometimes referred to in the industry as "insured annuities". In practice, insured annuities may not be exactly in the same form or involve the same insurance policies and annuities as we describe.

however, Brown, Mitchel, Poterba, and Warshawsky (1999) proposed an alternative scheme that is similar to the Canadian scheme, while preserving the higher effective tax rate. We also discuss the possibility of tax arbitrage under this proposed scheme.

The paper is organized as follows. In section 2, we analyze a mortality swap in a frictionless market. In section 3, we introduce the Canadian personal income taxation into our model, and identify the reasons for the existence of an arbitrage opportunity. We then present a numerical analysis using annuity and insurance quotes from Canadian insurance companies. In addition, we discuss the effect of the U.S. personal income taxation. Then, in section 4, we examine the effects of other market frictions such as adverse-selection costs and transaction costs that may prevent individuals from taking full advantage of this arbitrage opportunity. Section 5 concludes the paper. Proofs and description of the data used in our numerical analysis are contained in the appendices.

2. Frictionless Markets

We start our analysis of mortality swaps by showing that a tax arbitrage opportunity does not exist in a "frictionless" market, which we define as one with no taxes, transaction costs or adverse-selection concerns. In addition, we assume a flat term structure of interest rates at the (continuously compounded) per-annum rate of r . The real-world non-flat term structure of interest rates will not change the main results of our analysis. Rather, it merely complicates the framing of the arbitrage transaction. In other words, the basic intuition for the existence of arbitrage is independent of the current term structure.

2.1. Life Annuities in Frictionless Markets

Consider an immediate life annuity that guarantees a lifetime payment of \$1 per annum (in continuous time), starting immediately. Under the assumed market structure, the actuarially fair price of this annuity for an individual aged x , is given by:

$$a_x(r) = \int_0^{\infty} e^{-rs} {}_s p_x ds, \quad (1)$$

where e^{-rs} is the discount factor applicable to a \$1 payment at time s , and ${}_s p_x$ is the conditional probability that an x -year-old individual will survive for another s years (i.e., to age $x + s$).⁴ The pricing relationship in equation (1) is indeed the *actuarial principle of equivalence*, which equates the present value of the expected benefits to the initial price.⁵

As a corollary, a one-dollar investment in a life annuity at age x will entitle the annuitant to lifetime annual payments in the amount of $1/a_x(r)$ per year.

Clearly, there is a mortality risk associated with the effective rate of return that an individual will earn on an annuity. The longer he/she lives, the more payments he/she will receive and thus the higher is the effective return.

2.2. Life Insurance in Frictionless Markets

Consider a life insurance policy that pays \$1 at the time of death and requires a constant premium to be paid every year (in continuous time). One can think of this policy as a term-to-life insurance policy. In a world without adverse selection, insurance companies use the same mortality table to price both annuities and insurance policies. Given this assumption, the actuarially fair annual premium for an individual aged x is given by:

$$i_x(r) = \frac{\int_0^\infty e^{-rs} {}_s p_x \lambda_{x+s} ds}{\int_0^\infty e^{-rs} {}_s p_x ds} = \frac{\int_0^\infty e^{-rs} {}_s p_x \lambda_{x+s} ds}{a_x(r)}, \quad (2)$$

where $\lambda_{x+s} > 0$ is the instantaneous force of mortality, and the second equality comes from equation (1).⁶

As in the case of annuities, there is a mortality risk associated with the effective rate of return on an insurance policy. The longer a policy holder lives, the more premium payments he/she makes and thus the lower the effective return.

2.3. Mortality Swaps in Frictionless Markets

⁴ As a special case, if lifetime is exponentially distributed, then the annuity price, $a_x(r)$, is equal to $1/(\lambda + r)$ where λ is the (constant) instantaneous force of mortality. We note, however, that exponential distributions are not consistent with empirical mortality patterns. We mention it here because it has a bearing on the argument that we will develop later in the paper.

⁵ See Bowers et. al. (1986) for a description of the underlying assumptions and the law of large numbers used in actuarial pricing.

⁶ Again, as a special case, if lifetime is exponentially distributed, then equation (2) reduces to $i_x(r) = \lambda$.

Suppose an x -year-old individual uses \$1 to purchase a life annuity. According to the above corollary, his/her lifetime annuity income is constant at $\$1/a_x(r)$ per year. Suppose the individual then buys a life insurance policy whose death benefit is also \$1. The premium for this policy is constant at $\$i_x(r)$ per year. By constructing this mortality swap, the individual has created a cash flow stream of $\$1/a_x(r) - i_x(r)$ per year as long as he/she is alive, and \$1 at the time of death. This cash flow pattern is similar to that of a \$1 investment in a risk-free instrument such as a bank deposit that pays interest at a percentage rate of $1/a_x(r) - i_x(r)$ every year, and whose principal is redeemed at par (by the individual's estate) at the time of death. Therefore, by holding both an annuity and an insurance policy, the individual removes the mortality risk in the effective rate of return on the combined position.

In a frictionless market, the rate of return on a mortality swap is equal to the risk-free rate; i.e.,

$$\frac{1}{a_x(r)} - i_x(r) = r. \quad (3)$$

Equation (3) is in fact a well-known actuarial identity whose proof can be found in Bowers et al. (1986).⁷ We state this formally as a theorem.

Theorem 1 In a frictionless market - i.e. no taxes, transaction costs or adverse selection - with a flat term structure of interest rates, a mortality swap will not admit arbitrage. In other words, an investment in a mortality swap has the same lifetime and death payoffs as those of a risk-free investment such as a bank deposit.

Proof As stated.

In the special case of exponential mortality, the yearly annuity income on a \$1 investment is $\$(\lambda + r)$, and the insurance premium on a \$1 policy is $\$\lambda$, resulting in a cash flow stream of $\$r$ per year from a mortality swap.

3. Markets with Personal Income Taxes

⁷ The proof by Bowers et al. was done in discrete time. A continuous-time proof is available from the authors upon request.

In this section, we relax the frictionless assumptions by introducing personal income taxation into our framework. Specifically, we assume that individuals are subject to the Canadian income tax rules. We want to examine the effect of the rules on the payoffs of annuities and insurance policies. We will then compare a mortality swap to a risk-free investment. Our main result is that when income taxation on annuities and insurance policies is not implemented properly, a tax arbitrage opportunity will exist. We then discuss the U.S. tax rules and the possibility of tax arbitrage under them.

3.1. Life Annuities in the Canadian Markets

The Canadian *Income Tax Act* (hereafter referred to as the "*Act*") considers income from a prescribed life annuity as consisting of two parts. The first part is the return *of* capital (principal), and is, therefore, not taxed. The second part is considered interest (or return *on* capital) and is fully taxable. The percentage of each annuity payment that will be taxed (hereafter referred to as the "taxable portion", and denoted by ρ) is specified by the *Act* based on the price of the annuity, the income from the annuity, and the age of the purchaser, x (which determines his/her expected remaining lifetime), according to the following formula:⁸

$$\rho_x = 1 - \frac{1}{E[T_x^g] \frac{1}{a_x(r)}} = 1 - \frac{a_x(r)}{E[T_x^g]}, \quad (4)$$

where $1/a_x(r)$ is, as before, the yearly annuity income that an x -year-old individual will receive from a \$1 investment. The central variable $E[T_x^g]$ is the individual's expected remaining lifetime based on the mortality table specified by the *Act*, which currently is *The Society of Actuaries 1971 Individual Annuity Mortality (IAM) Table*. The second term on the right-hand side of either equality represents the portion of each annuity income that is considered to be the return *of* capital. Once determined, the taxable portion, ρ_x , will remain the same throughout the annuitant's life, even if he/she lives beyond his/her expected remaining lifetime, $E[T_x^g]$ years.

The after-tax annual income from a \$1 annuity investment is therefore constant at:

$$\frac{1}{a_x(r)} - \frac{\tau \rho_x}{a_x(r)},$$

⁸ See Reg. 300 (1.1) of the Canadian *Income Tax Act* (2000). The U.S. Internal Revenue Services also use this formula. However, the implementation is different. We will discuss the U.S. case later in the paper.

where τ is the annuitant's marginal tax rate which is assumed to be constant throughout his/her remaining lifetime.⁹

There are three things worth noting on the effect of taxes on annuity income, all of which have implications on the possibility of tax arbitrage. First, the taxable portion, ρ_x , remains constant throughout the annuitant's life. This means that annuity income is never fully taxed, even after the principal is considered to be completely returned. Secondly, the choice of the significantly outdated 1971 individual annuitant mortality (IAM) table causes $E[T_x^g]$ to be shorter than present life expectancies, which are used by insurance companies to price annuities. This causes the taxable portion of annuity income to be lower than it should be. Finally, given the annuity price, $a_x(r)$, the value of ρ_x depends only on the expected value or the first moment of the distribution of future lifetime, $E[T_x^g]$, and not on any higher moments. These three issues will be discussed in details shortly.

3.2. Life Insurance in the Canadian Markets

Under the Canadian tax rules, life insurance policies that are issued after December 1, 1982 are classified into two classes depending on their tax treatment - exempt and non-exempt.¹⁰ The distinction is based on the extent of the investment (or savings) component of the policies. The lower this component is, the greater is the chance that the policies will be exempt. The policy holder is required by the *Act* to have his/her policy tested on each of its anniversaries. For policies that pass this test, taxation is deferred until disposition occurs. Since the death of policy holders does not constitute disposition, death benefits from exempt policies are not taxable.

Generally, term-life policies (including the one used in our construction of a mortality swap) qualify as exempt policies. Therefore, their death benefits are usually not taxable. As for their premiums, they are not tax-deductible. Hence, there is no effect of taxes on the amounts of premium and death benefit. The annual premium of $\$i_x(r)$ results in an after-tax payment of \$1 at the time of death.

3.3. Mortality Swaps in the Canadian Markets

⁹ As will be shown, the existence (although not the magnitude) of the tax arbitrage is independent of personal income tax rates. Therefore, this assumption will not affect our results, but will facilitate the exposition.

¹⁰ See Regulation 306 of the Canadian *Income Tax Act*. For a practitioner's discussion of this rule, see, for example, Lengvari and Joshua (1995).

We again compare a \$1 investment in a mortality swap to a \$1 bank deposit. The after-tax payoffs from the two investment choices to an individual who is currently x -year-old are:

Investment's After-Tax Payoff	State of Nature	
	While Alive	At Death
Mortality Swap	$\$ \frac{1}{a_x(r)} - \$ \frac{\tau \rho_x}{a_x(r)} - \$ i_x(r)$ per year	\$1
Bank Deposit	$\$ (1 - \tau) r$ per year	\$1

Due to the constant nature of its after-tax payoff, this mortality swap can still be thought of as a synthetic bank deposit on an after-tax basis. Because the costs and the final payoffs under both investment alternatives are the same, tax arbitrage is possible if the mortality swap provides higher intermediate after-tax payoffs than that of the bank deposit; i.e.,

$$\frac{1}{a_x(r)} - \frac{\tau \rho_x}{a_x(r)} - i_x(r) > (1 - \tau)r. \quad (5)$$

Intuitively, since their before-tax payoffs are the same (as per equation (3)), tax arbitrage will occur if the tax payment under the swap is lower than under the bank deposit; i.e.,

$$\frac{\tau \rho_x}{a_x(r)} < \tau,$$

which leads to the arbitrage condition:

$$\rho < r a_x(r) \quad (6)$$

Therefore, in the current framework, the existence of tax arbitrage depends on the interaction among the taxable portion, the interest rate, and the price of the annuity. On the other hand, it does not depend on the individual's marginal tax rate, τ , as long as τ is greater than zero. However, from (5), the higher the individual's marginal tax rate, τ , the more he/she stands to gain from this mortality swap.

Before we discuss the reasons why this tax arbitrage does exist in practice, we first present the results of our numerical analysis of the real-life annuity and insurance quotes.

3.3.1. A Numerical Study

Table 1 reports annual insurance premiums and annuity income for female individuals of various ages. The insurance premiums are based on a term-to-100 life insurance policy with a death benefit of \$100,000, while the annuity income is based on a \$100,000 investment in an immediate life annuity.¹¹ Based on the swap's after-tax payoffs for various marginal tax rates, we calculate "grossed-up returns". These are the equivalent pre-tax rates of return that an individual would have to obtain on a fixed-income investment of the same risk, in order to end up with the same after-tax returns.

For example, consider a 60-year-old female who is in good health and whose marginal tax rate is 50%.¹² She would pay \$1,820 per annum (in monthly installments) under the insurance policy. The average of the best five annuity quotes based on a \$100,000 purchase would provide a 60-year-old female with \$7,397 per annum (in monthly installments), of which \$3,272 is taxable. At the 50% marginal tax rate, this would result in an annual after-tax annuity income of \$5,761, and a net cash flow from the swap of \$3,941 per annum for life. Therefore, the after-tax rate of return on the swap is 3.94% p.a. Based on her 50% tax bracket, this translates to an equivalent pre-tax rate of return of 7.88% p.a. This equivalent rate was higher than to the yield on a bank deposit or a guaranteed investment certificate (GIC) which, at the time we obtained the quotes (July 1, 1999), was around 5% p.a.¹³ As a result, this individual would have an arbitrage opportunity.

The existence of this arbitrage opportunity can also be ascertained by checking whether the arbitrage condition in inequality (6) holds. In this case, a_{60} , which is the annuity price per \$1 of yearly income, is equal to $\$100,000/7,397 = \13.519 . The taxable portion, ρ_{60} , is equal to $\$3,272/7,397 = 0.44234$, while the risk-free rate is 5% p.a. Hence, ρ_{60} is less than ra_{60} , and tax arbitrage exists.

Table 1
Annual Insurance Premiums, Annuity Income and Equivalent Pre-Tax Rates of Return on a Mortality Swap for Females of Various Ages and Marginal Tax Rates

¹¹ The insurance and annuity quotes are as of July 1, 1999. Please see Appendix A for a detailed discussion about the quotes.

¹² In Canada, 50% is approximately the highest combined federal and provincial marginal tax rate. In the province of Ontario, it takes effect beyond an annual taxable income of \$74,000 (as of September 2000).

¹³ Note that in this paper, we assume a flat deterministic term structure. In practice, the rates for bank deposits or GICs can change from year to year. As a result, a more appropriate rate to be compared to in our framework may be the yield of a corporate (i.e., bank's) bond whose maturity is approximately the same as the individual's estimated remaining lifetime. Even with this new benchmark, which was around 6% to 6.5% as of July 1, 1999, the swap's 7.88% return still far exceeded it.

Age	Annual Insurance Premium	Annual Annuity Payment	Annual Taxable Income	Equivalent Pre-Tax Return		
				$\tau = 25\%$	$\tau = 40\%$	$\tau = 50\%$
50	\$ 890	\$ 6,612	\$ 3,603	6.43%	7.13%	7.84%
55	\$ 1,276	\$ 6,917	\$ 3,421	6.38%	7.12%	7.86%
60	\$ 1,820	\$ 7,397	\$ 3,272	6.35%	7.11%	7.88%
65	\$ 2,566	\$ 8,112	\$ 3,137	6.35%	7.15%	7.96%
70	\$ 3,566	\$ 9,357	\$ 3,136	6.68%	7.56%	8.45%
75	\$ 4,919	\$ 10,812	\$ 2,709	6.95%	8.02%	9.08%
80	\$ 6,654	\$ 13,162	\$ 2,145	7.96%	9.42%	10.87%

In order to get a measure of the sensitivity of the above results to the insurance premium, we perturb the numbers and then calculate the equivalent rates of return again. For example, suppose the 60-year-old female does not qualify for the "cheap insurance", and has to pay 50% more (\$2,730 instead of \$1,820) for the term-to-100 policy. The grossed-up return from the mortality swap would be 6.06% (in contrast to 7.88%). In the other direction, if the life insurance price were reduced by 25% (\$1,365 instead of \$1,820), the grossed-up return from the mortality swap would increase to 8.79%. Clearly, the state of health (i.e., the ability to pass a medical examination) will be crucial to the possibility and benefit of tax arbitrage.

A few stylized facts are evident from Table 1. First, all the grossed-up rates of return are higher than the current deposit rates (or bond yields). Secondly, the higher the individual's marginal tax rate, the higher is the grossed-up return from a mortality swap. Therefore, everything else being equal, a higher tax rate implies (i) a higher likelihood that an individual will have an arbitrage opportunity from a mortality swap ; and (ii) a larger magnitude of gain from it. Thirdly, the grossed-up rates of return on the mortality swap (roughly) increase with age. We conjecture that this is most likely driven by the insurance loads. It appears that the older an individual is, the more likely is the arbitrage opportunity and the more gain he/she can obtain from it. In fact, for certain combinations of age and tax rate, the swap's returns are approximately twice as high as the risk-free rate.¹⁴

3.3.2. The Reasons for Tax Arbitrage

¹⁴ In addition, the swap's returns are higher if, for example, a couple create it by using a last-to-die insurance policy (i.e., a policy issued on two lives under which premiums are required, and death benefits are not paid, until both of them have died) and a joint-and-100%-survivor annuity (i.e., an annuity issued on two lives under which payments continue in whole until both have died).

The arbitrage condition in inequality (6) suggests that tax arbitrage will exist if the taxable portion is less than the product of the interest rate and the annuity price; i.e.,

$$\rho_x < ra_x(r)$$

For a given level of interest rate, r , the higher the annuity price, $a_x(r)$, the more likely is tax arbitrage. This may seem counter-intuitive considering the appearance of inequality (5). However, we know from Theorem 1 that the difference between $1/a_x(r)$ and $i_x(r)$ is always equal to r , regardless of the value of $a_x(r)$.¹⁵ That is, the effect of an increase (decrease) in the annuity price is offset by a decrease (increase) in the insurance premium. Therefore, the only effect of a higher $a_x(r)$ is on the term $\tau\rho_x/a_x(r)$, which can be shown to be a decreasing function of $a_x(r)$. Intuitively, an increase in the annuity price is beneficial since it reduces the tax payments.

As a result, given a level of r , the discussion on the possibility of tax arbitrage can be done in terms of the annuity price, $a_x(r)$. For this purpose, we rewrite inequality (6) by using the definition of ρ_x as follows:

$$a_x(r) > \frac{1}{\frac{1}{E[T_x^g]} + r}. \quad (7)$$

In this form, the condition explicitly states that tax arbitrage will occur when the annuity price is higher than a certain threshold. We note that this threshold,

$$\frac{1}{\frac{1}{E[T_x^g]} + r},$$

can be thought of as the annuity price under an exponential mortality distribution with a life expectancy of $E[T_x^g]$ years. This follows from the fact that under exponential distributions, $a_x(r) = 1/(\lambda + r)$ and λ is the reciprocal of life expectancy.

The intuition for this threshold is crucial for our discussion on the reasons why the arbitrage condition such as inequality (7) can be satisfied. As noted earlier, given the annuity price, $a_x(r)$, the value of ρ_x depends only on the expected value or the first moment of the

¹⁵ We assume that return on the money invested by insurance companies is not taxable, which is a close approximation to the case in practice.

distribution of lifetime, $E[T_x^g]$. This is *as if* Canada Customs and Revenue Agency (CCRA) were assuming that human mortality could be completely specified by its first moment. This would be correct if and only if population mortality were exponentially distributed, since one would need only to know the force of mortality, λ , to generate an exponential mortality distribution. It follows by extension that CCRA inadvertently assumes that annuities are priced according to an exponential distribution whose first moment is equal to $E[T_x^g]$ years. In other words, by stipulating $E[T_x^g]$ and the formula for ρ_x , CCRA indirectly specifies a threshold for the price of an annuity. Tax arbitrage will not exist *if and only if* the annuity price is less than or equal to that threshold. We state it as a theorem.

Theorem 2 In a frictionless market with taxation and a flat term structure of interest rates, there is an upper bound for the annuity price for an x -year-old individual, below which a mortality swap will not admit tax arbitrage. That upper bound is the price of the annuity if it is priced based on an exponential distribution with a life expectancy of $E[T_x^g]$ years.

Proof As stated and based on inequality (7) and the properties of an exponential distribution whose first moment is $E[T_x^g]$ years.

Theorem 2 provides an insight into two major reasons why the annuity price quoted by insurance companies can be higher than the threshold and thus tax arbitrage can occur. The first and more obvious reason follows directly from the theorem. It involves the case where CCRA specifies a life expectancy, $E[T_x^g]$, that is too short relative to that assumed by the insurance companies, $E[T_x^c]$. To isolate its effect, we assume for now that insurance companies use an exponential distribution. In this case, the quoted annuity price will be higher than the threshold.

Theorem 3 *First Reason for Tax Arbitrage:* Assume a frictionless market with taxation and a flat term structure of interest rates. If insurance companies use an exponential mortality distribution in their pricing, then tax arbitrage will exist if $E[T_x^g] < E[T_x^c]$.

Proof Under the exponential assumption, annuity prices are an increasing function of life expectancies. Therefore, if $E[T_x^g] < E[T_x^c]$, then the price obtained with a life expectancy of $E[T_x^c]$ will be higher than that obtained with $E[T_x^g]$, which is the upper bound for tax arbitrage.

$E[T_x^g]$ will be smaller than $E[T_x^c]$ if CCRA uses a mortality table that is not up to date and, therefore, does not reflect recent improvements in life expectancies. As stated earlier, for individual annuitants, the *Act* is currently using the 1971 individual annuitant mortality (IAM) table, as compiled by the Society of Actuaries. This 30-year-old IAM table is significantly out of date, as evidenced by Table 2 where we compare it to two more recent IAM tables for 55-year-old individuals. Life expectancies are lower in the 1971 table than in the 1996 table by approximately 3.5 years for females and 4.5 years for males. The probabilities of survival to various ages in the 1971 table are also lower, especially at the tail of the distribution.

Table 2
A Comparison of Probabilities of Survival and Life Expectancies
of a 55-Year-Old Individual under Three Annuity Mortality Tables

Probability of Survival to Age	1971 IAM Table		1983 IAM Table		1996 IAM Table	
	Female	Male	Female	Male	Female	Male
55	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
60	97.6%	95.2%	98.2%	96.6%	98.5%	97.4%
65	93.8%	88.6%	95.6%	91.9%	96.2%	93.7%
70	88.9%	79.9%	91.4%	84.8%	92.6%	88.0%
75	81.2%	68.2%	84.9%	74.2%	89.9%	79.1%
80	68.9%	53.0%	74.5%	59.6%	77.5%	66.3%
85	50.4%	35.3%	58.6%	41.5%	62.8%	49.6%
90	28.1%	18.1%	37.9%	23.4%	42.7%	31.3%
95	10.3%	5.6%	18.1%	10.0%	22.1%	15.4%
100	2.6%	0.7%	5.9%	2.89%	8.2%	5.6%
Life Expectancy at Age 55 (years)	31.76	27.49	34.03	29.70	35.22	31.93

As an example of the existence of tax arbitrage under the premise of Theorem 3, suppose that insurance companies price their annuities and insurance policies based on an exponential distribution with a mean lifetime, $E[T_x^c]$, equal to that specified in the 1996 table. Suppose also that the current risk-free rate is 7% p.a. In this case, from Table 2, $E[T_x^c]$ of a 55-year-old female is 35.22 years, which means λ is 0.0284. Therefore, the price of the annuity that will pay off \$1 per year (in continuous time) to a 55-year-old female is:

$$a_{55}(0.07) = \frac{1}{\lambda + r} = \frac{1}{0.0284 + 0.07} = \$10.1626.$$

On the other hand, the arbitrage threshold for the price of this annuity is calculated based on the 1971 table with $E[T_x^g] = 31.76$ years to be:

$$\frac{1}{\frac{1}{31.76} + 0.07} = \$9.8536.$$

Hence, the annuity price quoted by the insurance companies is higher than the threshold, and tax arbitrage will exist.

Intuitively, by using an outdated annuity table, CCRA is inadvertently enhancing the after-tax payoff of the annuity. This is because the outdated life expectancies will cause ρ_x to be too small, and thus a larger portion of annuity income to be exempted from taxes. Suppose a 55-year-old female individual with a 40% marginal tax rate invests \$1 in the above annuity. The taxable portion of her annuity income based on the 1971 table is:

$$\rho_{55} = 1 - \frac{10.1626}{31.76} = 0.6800,$$

resulting in an after-tax income of:

$$\frac{1 - 0.4(0.6800)}{10.1626} = \$0.0716.$$

On the other hand, if CCRA were to use the 1996 table in specifying $E[T_x^g]$, ρ_{55} would be 0.7115, and the after-tax annual income would be \$0.0704, a difference of \$0.0012 per year (or 0.12% p.a. of the annuity principal, \$1).

The above example shows that tax arbitrage will exist if insurance companies use an exponential distribution in their pricing, and $E[T_x^g] < E[T_x^c]$. However, observed mortality patterns are not exponentially distributed. Hence, insurance companies do not use exponential distributions. Without specifying the exact distribution used, we cannot state with certainty that the annuity price will be higher than the threshold (and thus tax arbitrage will exist). However, we can claim that tax arbitrage will exist whenever insurance companies price their products

using a certain class of mortality distributions. More precisely, we argue that if actuaries use a mortality distribution from this class, tax arbitrage can exist *even without* $E[T_x^g]$ being less than $E[T_x^c]$. This is indeed the second reason why annuity prices can be higher than their thresholds.

To prove this argument, we isolate the effect of a different mortality distribution by assuming for now that $E[T_x^g]$ is the same as $E[T_x^c]$. Under this assumption, Theorem 2 states that if an exponential distribution is used to price annuities, then tax arbitrage will not exist. Consider, however, an alternative mortality distribution for an x -year-old individual with the following properties:¹⁶

- There exists some age y ; $y > x$, such that (a) the probability of survival up to *any* age before y under the distribution is higher than its exponential counterpart; and (b) the probability of survival up to *any* age beyond y under the distribution is lower than its exponential counterpart.

This property partitions the time horizons into two parts according to the difference in the probabilities of survival between the two distributions. We call the class of distributions that satisfy this property the "equal-mean, thinner-tail" (EMTT) distributions.¹⁷ We argue that annuity prices under this class of distributions will be higher than the exponential prices, and thus arbitrage exists. We state this as a theorem.

Theorem 4 *Second Reason for Tax Arbitrage:* Assume a frictionless market with taxation and a level of interest rate that is positive and the same for all terms. Assume also that $E[T_x^g] = E[T_x^c]$. Then, tax arbitrage will exist if insurance companies use a mortality distribution from the "equal-mean, thinner-tail" (EMTT) class in their pricing.

Proof We show in Appendix B that annuity prices under the EMTT class of distributions are higher than the exponential prices, given that $E[T_x^g] = E[T_x^c]$. By inequality (7), this implies a tax-arbitrage opportunity.

¹⁶ By definition, a mortality distribution implies probabilities of survival that are non-increasing with time horizon.

¹⁷ This "equal-mean, thinner-tail" property may appear to be similar to one of the definitions of "greater riskiness" proposed by Rothschild and Stiglitz (1970). In that paper, Rothschild and Stiglitz show that if random variables x and y have the same mean but y has more weight in the tail of its density function than x does, then y is riskier than x in the sense that every person with a concave utility function will prefer x to y . In our context, the difference in the weights in the tails does not equate with the difference in riskiness as the time of death is not the object of choice.

Empirical mortality distributions such as the IAM tables compiled by the Society of Actuaries belong to this EMTT class. In Table 3, we compare the probability that a 55-year-old female will survive to various ages under the 1996 IAM table, to that implied by the exponential distribution with the same life expectancy (i.e., $E[T_{55}] = 35.22$ years, and $\lambda = 1/35.22 = 0.0284$). The IAM numbers satisfy the above property. That is, the probability of survival to various ages is higher than its exponential counterpart up to (approximately) age 90, and lower than the exponential numbers afterwards.

Table 3
Probabilities of Survival of a 55-Year-Old Female
under an Exponential Distribution and under the 1996 IAM Table

Probability of Survival to Age	Exponential Distribution	1996 IAM Table
55	100.0%	100.0%
60	86.8%	98.5%
65	75.3%	96.2%
70	65.3%	92.6%
75	56.7%	86.9%
80	49.2%	77.5%
85	42.7%	62.8%
90	37.0%	42.7%
95	32.1%	22.1%
100	27.9%	8.2%
Life Expectancy at Age 55 (years)	35.22	35.22

In this case, if annuities are priced based on the 1996 IAM table, then $a_x(r)$ will be higher than its exponential price and tax arbitrage will exist. The intuition for this is as follows. Recall from equation (1) that the price of an annuity is a weighted sum of the present values of all future annuity income, where the weights are the probabilities that the buyer will live to receive those payments. For any given time horizon, the 1996 IAM table assigns a different weight, ${}_s p_x$, from that assigned by the exponential distribution. If interest rate were zero, then the two set of weights would not matter since the weighted sum would turn out to be the same and equal to the life expectancy. However, when interest rate is positive, the present value factors cause earlier weights to be more important than later weights. Since earlier weights are

higher under the 1996 IAM table than under the exponential distribution, the weighted sum will be greater under the 1996 IAM table.

As another example, one theoretical distribution that has been recognized as providing a realistic description of the pattern of mortality at adult ages is the Gompertz distribution.¹⁸ In Figure 1, we provide a visual illustration of what is meant by the "equal-mean, thinner-tail" class of distributions. The dotted line represents the probability that a person with a remaining life expectancy of 18.5 years will survive to various points in time under the Gompertz distribution. The solid line represents the probability of survival under the exponential distribution with the same life expectancy. The two lines intersect at the time horizon of 24.08 years, beyond which the probability of survival under the Gompertz distribution is below that of the exponential distribution. In other words, the Gompertz distribution has a "thinner-tail" than that of the exponential distribution.

[Insert Figure 1 here]

As an example of the existence of tax arbitrage under the premise of Theorem 4, we calculated the per-annum percentage *difference* between the after-tax return on a mortality swap, and the after-tax return on a risk-free bank deposit, for individuals of various ages, assuming that Gompertz distributions are used to price annuities and insurance policies. The Gompertz distributions were calibrated to the 1996 IAM numbers for those individuals so that, among other things, their life expectancies are the same as in the 1996 IAM tables. As before, the before-tax risk-free return is assumed to be 7% p.a. and the marginal tax rate is 40%. Table 4 reports the results of our calculations.¹⁹

Table 4
After-Tax Gains from a Mortality Swap as a Percentage
of Investment, Assuming that the Tax Authority Uses Gompertz
Distributions with Life Expectancies Based on the 1996 IAM Table

Age	Male	Female
35	0.666%	0.651%
45	0.722%	0.720%
55	0.763%	0.782%
65	0.773%	0.821%

¹⁸ See, for example, Carriere (1992).

¹⁹ The details of our calculations are available upon request.

75	0.736%	0.813%
85	0.637%	0.731%
95	0.482%	0.563%

Note: The gains are calculated under an assumption that the risk-free rate is 7% p.a. and the marginal tax rate is 40%.

As can be seen, all differences are positive, consistent with Theorem 4 that arbitrage exists in this case. The magnitude of the differences varies, depending on the age of the individuals. The greatest difference occurs when the buyers are either a 65-year-old male or a 65-year-old female. The differences are lower in magnitude on either side of that age.²⁰ We caution the reader that this 'maximum difference' at age 65 may be driven by the mortality function in question, as opposed to the actual taxable portion.

In sum, there are two related reasons why a mortality swap should lead to a tax arbitrage opportunity in the presence of taxation. First, the after-tax payoff from the annuity part of the swap is increased by the fact that an outdated mortality table is specified by CCRA for the purpose of computing the taxable portion. This reason alone would be sufficient to cause tax arbitrage if insurance companies priced their products by an exponential mortality distribution. With a different distribution, this reason will at least increase the likelihood of tax arbitrage. Secondly, and more importantly, the choice of mortality distribution used by insurance companies can cause tax arbitrage, even if CCRA does not use an outdated table. If the employed distribution belongs to the EMTT class, then tax arbitrage will exist. Since it is reasonable to expect actuaries to use a table that is similar in structure to the IAM tables, which themselves belong to the EMTT class, we believe that both reasons for tax arbitrage are applicable in practice.

The numerical examples under both reasons show that the second reason is more important than the first. While it is straightforward to see where the difference in after-tax returns come from under the first reason, the second reason warrants further explanation.

When the employed distribution comes from the EMTT class, the existence of tax arbitrage can be traced back to the manner in which prescribed annuities are taxed. To see this, note that in reality, each payment of annuity income consists of four different portions - (i) interest on own principal; (ii) return of own principal; (iii) a share of principal of annuitants who have already died; and (iv) a share of interest on the principal of annuitants who have already

²⁰ It may appear that there is a 'best' time (i.e., age) to purchase a mortality swap. However, the differences in Table 4 are in per-annum percentages, not dollar amounts. Clearly, if a 35-year-old individual constructs a mortality swap, the percentage difference may be lower than what a 55-year-old will get. However, the (expected) length of time over which the difference will be realized is much longer since a 35-year-old has a longer life expectancy than a 55-year-old does.

died. There are two reasons why the relative weights of these four portions will vary through time. First, since the annuitant's own principal is gradually returned, earlier annuity payments contain a higher portion of interest on it than later payments do. Secondly, as more and more annuitants die, the share of principal and interest of the deceased annuitants also increases with time. Therefore, in reality, taxable income from an annuity will vary through time. The variation depends on interest rate and mortality patterns.

On the other hand, for tax purposes, each prescribed annuity payment is considered to have only two portions - (i) return of own principal and (ii) the rest, which we will refer to as the "taxable income". This implies that the *Act* treats as one portion the interest on own principal and the share of principal and interest of the deceased annuitants. In addition, the *Act* does not take into account the variation in the relative weights of the two portions. Rather, by specifying ρ_x that does not vary through time, the *Act*, in effect, dictates that the total taxable income earned *during the annuitant's life expectancy*, $E[T_x^g]$ years²¹, should be assigned to each tax year within that period on a straight-line basis. To see this, recall that a \$1 investment in an annuity will yield annual before-tax income of $\$1/a_x(r)$, of which $\$\rho_x/a_x(r)$ is taxable. From the definition of ρ_x in equation (4),

$$\frac{\rho_x}{a_x(r)} = \frac{E[T_x^g] \frac{1}{a_x(r)} - 1}{E[T_x^g]}.$$

The denominator on the right-hand side is indeed the total income that the annuitant will receive over his/her life expectancy minus the initial investment. In other words, the numerator is the total *taxable income* to be earned over $E[T_x^g]$ years. It is then divided by the life expectancy to arrive at the taxable income for each of the $E[T_x^g]$ years. Note that this determination has no connection at all to the actual composition of each annuity payment.

Moreover, the rules allow ρ_x to remain the same, even beyond the annuitant's life expectancy. This means that even though the principal should be considered completely returned and the subsequent payments should wholly be considered taxable income, the annuitant will still be taxed on only part of them.

Now, recall from the above discussion that tax arbitrage will occur if the tax payment under a mortality swap is lower than under a bank deposit; i.e.,

²¹ Note that in order to isolate the effect of "equal-mean, thinner-tail" distributions, we continue to assume that $E[T_x^g] = E[T_x^c]$.

$$\frac{\tau\rho_x}{a_x(r)} < \tau r,$$

or

$$\tau\left(r - \frac{\rho_x}{a_x(r)}\right) > 0.$$

Based on the interpretation of $\rho_x/a_x(r)$ as the taxable income, this condition simply states that tax arbitrage will occur if the tax rules assign "too-low" taxable income from the annuity to each tax year. When an exponential distribution is used to price annuities, arbitrage does not occur because the annuity price, $a_x(r)$, is such that $\rho_x/a_x(r)$ is equal to r . That is, the taxable income assigned by the tax rules happens to be appropriate for the mortality swap.²² However, when a distribution from the EMTT class is used, $a_x(r)$ is higher than in the exponential case for a given life expectancy and a given level of interest rate. This causes the assigned taxable income to be too low since, as mentioned earlier, $\rho_x/a_x(r)$ is a decreasing function of $a_x(r)$.

$a_x(r)$ under an EMTT distribution is higher than its exponential counterpart for the following reason. Under an EMTT distribution, the probability of survival during the early years is higher than under an exponential distribution. Therefore, the contributions of deceased annuitants to the annuity pool during the early years will be lower under an EMTT distribution. This means that in these early years, insurance companies have to rely more on interest income in order to meet their payment obligation. For a given level of interest rate, the only way to generate sufficient interest income is to start with a sufficiently high principal. This results in a higher value for $a_x(r)$.

3.4. Mortality Swap in the U.S. Markets

Under the U.S. tax rules, for the vast majority of life insurance policies, their death benefits are not taxable. Also, insurance premiums are not tax deductible. Therefore, the tax treatment for the insurance side of a mortality swap is the same as in the Canadian case above.

On the other hand, for annuities, the U.S. Internal Revenue Service (IRS) specifies that the taxable portion, ρ_x , be increased to one (i.e., 100%) after the annuitant's life expectancy is reached.²³ Under such rules, there will be a decline in after-tax income when the annuitant

²² Note that we do not claim that this assigned amount is appropriate for the annuity when it is held by itself.

²³ To be precise, the IRS specifies an "exclusion" ratio, which is $1 - \rho_x$. Therefore, this means that the exclusion ratio will decline to zero after the life expectancy is reached. See Brown, Mitchell, Poterba and Warshawsky (1999) for a discussion of the U.S. regulations.

exceeds his/her life expectancy. As a result, the expected effective tax rate will be higher. The IRS adopted this practice in 1986. Prior to that time (from 1954 to 1986), the U.S. taxed annuities in the same manner as the current Canadian method (i.e., the exclusion ratio remained constant during the annuitant's life.).

As a result of the difference in the way annuities are taxed in the U.S., a mortality swap no longer replicates a bank deposit on an after-tax basis, and the comparison technique that we use in the Canadian case no longer applies. One can alternatively compare the present values of the two investment choices to determine whether tax arbitrage exists. That determination will, however, depend on one's (subjective) choice of mortality distribution used in the calculation. Also, if tax arbitrage exists, it will do so only in an "expected" sense. It is possible that *a posteriori*, a mortality swap turns out to be worse than a bank deposit. This will happen if the individual *significantly* outlives his/her life expectancy as specified by the tax authorities, $E[T_x^g]$. In other words, under the U.S. rules, a mortality swap cannot completely remove the mortality risk associated with annuities and insurance policies. Therefore, we believe that it is impossible that a sure tax-arbitrage opportunity will exist under the U.S. tax rules.

Recently, Brown, Mitchel, Poterba, and Warshawsky (1999) propose an alternative scheme that is similar to the current Canadian scheme (i.e., with a constant ρ_x), while preserving the higher expected effective tax rate. Their proposed scheme is intended to prevent a significant drop in annuitant's income after they reach their life expectancies. They show that the current U.S. tax rules can result in a decline of as much as 60% in after-tax annuity income in real term. To preserve the higher effective tax rate, they determine ρ_x such that it provides the same expected present value of tax payments as the expected present value under the current rules.²⁴

The proposed scheme will allow a mortality swap to again replicate a bank deposit, and thus completely remove the mortality risk. Nevertheless, because their proposed ρ_x will be at a higher level than in Canada, we believe that tax arbitrage will still be unlikely, though less so than under the current scheme.

There is yet another reason why tax arbitrage from a mortality swap is unlikely in the U.S. market. U.S. courts have in the past issued rulings on "combination plans" into which an insurance company packaged an annuity and an insurance policy in the manner not significantly different from the mortality swap discussed here. Under the rulings, the proceeds from the

²⁴ Therefore, their proposed method is different from the one the U.S. used from 1954 to 1986. Under their method, they first calculate the expected present value of taxes that an annuitant has to pay under the current rules (under which the taxable portion increases to 1 after the annuitant reaches his/her life expectancy). Based on that amount, they then come up with a new ρ_x that does not change through time but yields the same expected present value of taxes. On the contrary, the old U.S. rules, like the current Canadian rules, determined ρ_x based on the annuitant's life expectancy and the annuity's price and payments.

insurance policy would be subject to income tax (in the hands of the beneficiary) to the extent that the proceeds exceed the net premiums that would have to be paid for the insurance policy if it were a stand-alone policy.²⁵ Assuming that the rulings also apply to our case, the after-tax payoffs of a mortality swap will be further reduced, and thus tax arbitrage from it will be even more unlikely.

3.5. Mortality Swaps under Stochastic Interest Rates

Our analysis so far has been done under the assumption of a flat term structure of interest rates. With a non-flat term structure and stochastic interest rates, the comparison between a mortality swap and a risk-free bank deposit has to be done in a different manner. This is because the payoff from the deposit will vary from year to year. As a result, the existence of tax arbitrage should be determined by comparing the expected after-tax present values of the two investments where the payoff of the bank deposit for any future year is based on the one-year forward interest rate for that year.²⁶ While this determination will depend on one's (subjective) choice of mortality distribution used in the calculation, the above numerical analysis suggests that, *a priori* a mortality swap will still be a better investment choice than a bank deposit. This is because the swap's equivalent pre-tax returns (in Table 1) are generally higher than the interest rates for all points in the term structure. Note also that this tax arbitrage exists only in an "expected" sense. It is possible that *a posteriori*, a mortality swap turns out to be worse than a bank deposit. This will happen if a dramatic rise in interest rates occurs shortly after the swap was constructed. A perfect analogy to this is the case of interest-rate swaps whose current values are positive. The current positive values are the result of the current term structure of interest rates. Since interest rates can move in an adverse direction in the future, it is not certain that the positive values will be realized at the maturity of the swaps.

4. The Effects of Market Frictions

²⁵ See *Helvering v. LeGierse*, 312 U.S. 531 (1941) and *Kess v. U.S.*, 26 AFTR 2d 70-5839 (S.D. Ohio 1970).

²⁶ Since the discounting is done using the term structure of after-tax interest rates, the present value of the bank deposit is equal to the initial investment. Therefore, this comparison amounts to determining whether the present value of the payoffs a mortality swap is greater than the initial investment.

Except for income taxation, we have so far assumed away other market frictions. In this section, we investigate the effects of transaction costs and asymmetry of information on the arbitrage opportunities involving mortality swaps.

4.1. Asymmetry of Information and Loading

To protect themselves from the problem of adverse selection, insurance companies use mortality distributions that are different from that of the population as a whole. When pricing an annuity, they assume higher probabilities of survival than the population average. In other words, they assume that their annuitants will live longer than average. On the other hand, they assume lower-than-average probabilities of survival (i.e., their policy holders have a shorter life expectancy than average) when they price an insurance contract. This practice causes both the annuity price and the insurance premium to be higher than they would be had the population-average probabilities been used. Insurance companies must also pay commissions and other expenses, and make a profit. This further raises the price and the premium.

Let l_a represent the proportional loading costs associated with the annuity, expressed as a percentage of the annuity's frictionless price, $a_x(r)$. Similarly, let l_i represent the friction costs associated with the insurance policy, expressed as a percentage of the frictionless premium, $i_x(r)$. Then, the annuity's price in the presence of frictions becomes:

$$(1 + l_a) a_x(r) \quad l_a > 0, \tag{8}$$

while the insurance premium becomes:

$$(1 + l_i) i_x(r) \quad l_i > 0, \tag{9}$$

To see the effect of l_a and l_i on the opportunity for tax arbitrage, consider again an individual who is confronted with a choice between a \$1 investment in a mortality swap and the same amount of investment in a risk-free bank deposit. The after-tax payoffs from the two alternatives are as follows:

Investment's After-Tax Payoff	State of Nature	
	While Alive	At Death

Mortality Swap	$\$ \frac{1 - \tau \rho_x}{(1 + l_a) a_x(r)} - \$(1 + l_i) i_x(r)$ per year	\$1
Bank Deposit	$\$(1 - \tau) r$ per year	\$1

where we assume that there is no cost associated with the bank deposit.

In this case, the tax arbitrage condition becomes:

$$\$ \frac{1 - \tau \rho_x}{(1 + l_a) a_x(r)} - \$(1 + l_i) i_x(r) > (1 - \tau) r, \quad (10)$$

which can be rewritten in terms of ρ_x as:

$$\tau \rho_x < [\tau + l_a - l_i] r a_x(r) - [l_i + l_a + l_i l_a] a_x(r) i_x(r). \quad (11)$$

By comparing inequality (11) to inequality (6),

$$\rho < r a_x(r) \quad \text{or} \quad \tau \rho < \tau r a_x(r),$$

one can see that in the presence of frictions, the taxable portion, ρ_x , has to be smaller than before for arbitrage to exist. In other words, the portion of annuity income exempted from taxes must be greater than before. Mathematically, this is because (i) $[\tau + l_a - l_i]$ is less than τ by the fact that $0 < \tau < 1$ and $l_a > 0$; and (ii) $[l_i + l_a + l_i l_a] a_x(r) i_x(r)$ is positive by the fact that $l_i, l_a, a_x(r)$, and $i_x(r)$ are all positive. Intuitively, this is because \$1 now buys a smaller stream of pre-tax annuity income which then has to be used to pay taxes and finance a higher insurance premium in a mortality swap. As a result, the swap's return is reduced and tax arbitrage might not exist in the presence of frictions.

Unlike the frictionless (with taxes) case, the existence of arbitrage depends on the individual's marginal tax rate. The higher the marginal tax rate, τ , the easier it is for inequality (10) to be satisfied. The intuition for this is that the impact of taxes on the payoffs of a mortality swap is less than on a bank deposit since only a portion of the annuity income is taxable. This suggests that everything else being equal, individuals with high tax brackets have a better chance of benefiting from a mortality swap.

While the logic of Theorems 3 and 4 still applies in this case, the existence of arbitrage is, more practically, an empirical question.

5. Conclusion

In this paper, we document and explain the existence of a tax arbitrage opportunity involving a mortality swap. We label the transaction a mortality swap because the investor takes on mortality risk by acquiring an immediate life annuity, and then swaps it back by purchasing life insurance. We show that on a before-tax basis, investors will be indifferent between a mortality swap and a bank deposit which it replicates. However, in the presence of the Canadian tax rules and under realistic mortality distributions that are used to price annuities and insurance, a mortality swap will provide a higher after-tax rate of return than that of a generic bank deposit. This tax arbitrage opportunity is the result of the way annuity income is taxed in Canada, and to a certain extent, the fact that the inside buildup in a life insurance policy accumulates tax free. Under the Canadian rules, the taxable portion of annuity income is too low and remains constant at that level throughout annuitants' lives. This allows a mortality swap to replicate the payoff of a bank deposit on a before-tax basis, and exceed it on the after-tax basis.

We use observed annuity prices and insurance premiums to estimate the magnitude of the excess return for individuals of various ages and tax rates. The discrepancy in the rates of return indeed exists. The results show that the older the individuals are and/or the higher their tax rates are, the more they stand to gain from this strategy.

We discuss the reasons why this tax arbitrage is less likely under the U.S. tax rules. Under the U.S. rules, the taxable portion increases to 100% after annuitants reach their life expectancies. In this case, a mortality swap can no longer replicate a bank deposit. As a result, the existence of tax arbitrage depends on each individual's subjective probability of survival. We believe that tax arbitrage is unlikely under the U.S. rules because the increase in the taxable portion results in a higher effective tax rate for annuities. This is consistent with the results of Brown et al (1999) which show small differences between before- and after-tax expected values of life annuities in the U.S. market. In addition, past court rulings in the U.S. markets on the tax status of a comparable product called "hedged insurance policies" indicate that the proceeds from the insurance portion of the swap may be taxable. If this is the case, the possibility of tax arbitrage is remote.

Past studies on the effects of income taxes on security prices (see, for example, Dybvig and Ross (1986) and Dammon and Green (1987)) showed that equilibrium will not exist unless the investors are prevented from taking advantage of a price (or return) discrepancy by short-sale restrictions, progressive marginal tax rates, or some other form of friction. In the case of mortality swaps, however, the return discrepancy can persist in the Canadian markets due mainly to two reasons. First, a mortality swap is not scalable since individuals are limited by the amount

of insurance that they can purchase. Secondly, a mortality swap is not entirely riskless as tax laws might change without providing a "grandparent" clause, in which case the swap's after-tax return may turn out to be lower than that of a bank deposit.

References

- 1) K. Black and H. Skipper, *Life and Health Insurance (13th Ed.)*, Prentice Hall, Upper Saddle River, New Jersey, 1999.
- 2) N. Bowers, H. Gerber, J. Hickman, D. Jones, and C. Nesbit, *Actuarial Mathematics*, The Society of Actuaries, 1986.
- 3) J. Brown, O. Mitchell, J. Poterba, and M. Warshawsky, "Taxing Retirement Income: Non-qualified Annuities and Distributions from Qualified Accounts", *National Tax Journal*, 14(3): 219 - 234, 1999.
- 4) R. A. Campbell, "The Demand for Life Insurance: An Application of the Economics of Uncertainty", *Journal of Finance*, 35(5): 1155 - 1172, 1980.
- 5) J. F. Carriere, "Parametric Models for Life Tables", *Transactions of the Society of Actuaries*, XLIV: 77 - 99, 1992.
- 6) R. M. Dammon and R. C. Green, "Tax Arbitrage and the Existence of Equilibrium Prices for Financial Assets", *Journal of Finance*, 42(3): 1143 - 1166, 1987.
- 7) P. H. Dybvig and S. A. Ross, "Tax Clienteles and Asset Pricing", *Journal of Finance*, 41(3): 751 - 761, 1986.
- 8) S. Fischer, "A Life Cycle Model of Life Insurance Purchases", *International Economic Review*, 14(1): 132 - 152, 1973.
- 9) R. A. Jarrow and M. O'Hara, "Primes and Scores: An Essay on Market Imperfections", *Journal of Finance*, 44(5): 1263 - 1288, 1989.
- 10) A. Kamara and T. Miller, "Daily and Intradaily Tests of European Put-Call Parity", *Journal of Financial and Quantitative Analysis*, 30(4): 519 - 539, 1995.
- 11) E. Karni and I. Zilcha, "Uncertain Lifetime, Risk Aversion and Life Insurance", *Scandinavian Actuarial Journal*, 109 - 123, 1985.
- 12) G. F. Lengvari and R. S. Joshua, "Matters of Life and Death", *CA Magazine*, 44 - 46, August 1995.
- 13) F. D. Lewis, "Depedents and the Demand for Life Insurance", *American Economic Review*, 79(3): 452 - 467, 1989.

- 14) M.A. Milevsky, "Optimal Annuitization Policies: Analysis of the Options", *North American Actuarial Journal*, 5(1), 57 - 69, 2001.
- 15) O. S. Mitchell, J. M. Poterba, M. J. Warshawsky, and J. Brown, "New Evidence on the Money's Worth of Individual Annuities", *American Economic Review*, 89(5): 1,299 – 1,318, 1999.
- 16) J. M. Poterba, "The History of Annuities in the United States", *NBER Working Paper*, 1997.
- 17) M. Rothschild and J. E. Stiglitz, "Increasing Risk: I. A Definition", *Journal of Economic Theory*, 2: 225 - 243, 1970.
- 18) M. E. Yaari, "Uncertain Lifetime, Life Insurance and the Theory of the Consumer", *Review of Economic Studies*, 137 - 150, 1965.

Appendix A

We obtained Canadian insurance and annuity quotes as of July 1, 1999 for female individuals who are between 50 and 80 years old, in five-year increments. The annuity quotes were provided by CANNEX Financial Exchanges Limited and the insurance quotes were provided by SunLife and Transamerica.²⁷

The life annuity prices were quoted in terms of the monthly payments that the buyer will receive from the annuities, assuming a \$100,000 purchase on July 1st 1999, with the first annuity payment on August 1, 1999. The annuity quotes were from contracts that have no guarantee periods and are based on a single individual's life.²⁸

The CANNEX system also provides the taxable amount of each annuity payment, which is based on CCRA's method of computing ρ_x . Interestingly, the 10-15 most competitive annuity quotes as of July 1, 1999 showed a wide variation which can range from 1% - 15% of the best rate in the group. This variation could not be attributed to credit quality alone. Casual discussions with pricing actuaries indicated that sometimes insurance companies deliberately propagate non-competitive annuity prices in order to discourage the inflow of funds. For this reason, we decided to use the average of the five best rates in the group as our proxy for annuity prices. In any event, due to the collection and distribution of the quotes by CANNEX, the life annuity market in Canada seems to have been commoditized and is quite transparent. The annuity side of a mortality swap is therefore easy to price.

The same cannot be said about the life insurance market. The myriad of life insurance products (such as term, whole life and universal) do not allow for one standard comparison across companies and rates. Health status complicates matters by creating a collection of rates based on various underwriting criteria. For the purposes of a mortality swap, we chose a "term-to-100" insurance policy, which is a policy that maintains constant level of premiums until age 100. At age 100, the policy is considered paid in full and no further premiums are required. A term-to-100 policy does not have a cash value, and provides coverage for life even if the insured survives beyond age 100. As we mentioned earlier, this product is the closest policy to the insurance side of a mortality swap.

We obtained quotes from SunLife and Transamerica for term-to-100 policies for female individuals of various ages. These quotes were applicable only to those with the highest quality

²⁷ CANNEX is a financial data intermediary that collects life annuity quotes and quotes of other products from all major Canadian insurance companies. CANNEX then provides an on-line report that summarizes the 10-15 most competitive life annuity quotes available for a particular sex, age and investment combination. The annuity quotes are listed together with the names of the issuers and their credit ratings. These quotes are statistically reliable and usually guaranteed for a period of 15-30 days.

²⁸ In other words, the quotes are from contracts that do not allow co-annuitants.

health risk (i.e., preferred rates) and would probably be conditional on the insured passing a medical examination. The insurance premiums were quoted in terms of monthly rates based on a \$100,000 constant death benefit. The policies became effective on July 1, 1999 but would require a 30-day standard approval period. On August 1, 1999, the first (monthly) premium would be due and payable, conditional on passing the medical exams. The first (monthly) annuity income would therefore coincide with the first (monthly) insurance payment.

Since the annuity income was quoted on a monthly basis, we annualized the quotes by multiplying them by 12. The annualized insurance premiums divided by \$100,000 corresponds to $i_x(r)$ in our notation. Likewise, \$100,000 divided by the annual payment from the annuity corresponds to $a_x(r)$.

Appendix B

Proof that annuity prices under the EMTT class of distributions are higher than the exponential prices, given that $E[T_x^g] = E[T_x^c] = E[T]$.

Let $p_e(x, s)$ be the probability that an x -year-old individual will still be alive s years from now under the exponential distribution whose first moment (i.e., life expectancy) is $E[T]$. Also, let $p_a(x, s)$ be the probability that an x -year-old individual will still be alive s years from now under a distribution from the EMTT class *with the same life expectancy*. Based on the equation for annuity prices in (1), we will show that:

$$\int_0^{\infty} e^{-rs} p_a(x, s) ds > \int_0^{\infty} e^{-rs} p_e(x, s) ds.$$

First, we note that life expectancy is simply the integration of the probability of survival over time horizons of zero to infinity. Since life expectancy is the same under both distributions,

$$\int_0^{\infty} p_a(x, s) ds = \int_0^{\infty} p_e(x, s) ds. \tag{12}$$

Now, based on our definition of the EMTT class, any mortality distribution in this class has the following property:

- There exists some age y ; $y > x$, such that (a) the probability of survival up to *any* age before y under the distribution is higher than its exponential counterpart; and (b) the probability of

survival up to *any* age beyond y under the distribution is lower than its exponential counterpart; i.e., $p_a(x, s) > p_e(x, s)$ for all $s \leq y - x$, and $p_a(x, s) < p_e(x, s)$ for all $s > y - x$.

By the additive property of the integration, we can rewrite equation (12) as:

$$\int_0^{y-x} p_a(x, s) ds + \int_{y-x}^{\infty} p_a(x, s) ds = \int_0^{y-x} p_e(x, s) ds + \int_{y-x}^{\infty} p_e(x, s) ds.$$

Equivalently,

$$\int_0^{y-x} p_a(x, s) ds - \int_0^{y-x} p_e(x, s) ds = \int_{y-x}^{\infty} p_e(x, s) ds - \int_{y-x}^{\infty} p_a(x, s) ds,$$

or

$$\int_0^{y-x} [p_a(x, s) - p_e(x, s)] ds = \int_{y-x}^{\infty} [p_e(x, s) - p_a(x, s)] ds.$$

Then, by the fact that e^{-rs} is decreasing in s , we have that:

$$\int_0^{y-x} e^{-rs} [p_a(x, s) - p_e(x, s)] ds > \int_{y-x}^{\infty} e^{-rs} [p_e(x, s) - p_a(x, s)] ds.$$

or

$$\int_0^{y-x} e^{-rs} [p_a(x, s) - p_e(x, s)] ds + \int_{y-x}^{\infty} e^{-rs} [p_a(x, s) - p_e(x, s)] ds > 0,$$

which, by the additive property, can be rewritten as:

$$\int_0^{\infty} e^{-rs} [p_a(x, s) - p_e(x, s)] ds > 0,$$

or

$$\int_0^{\infty} e^{-rs} p_a(x, s) ds > \int_0^{\infty} e^{-rs} p_e(x, s) ds.$$