1994

An Introduction to Individual Disability Income Insurance

Mark J. Chartier
Monarch Life Insurance Company

Journal of Actuarial Practice Vol. 2, No. 1, 1994

An Introduction to Individual Disability Income Insurance

Mark J. Chartier*

Abstract

There are several actuarial software packages purporting to calculate expected benefit cash flows on disability income insurance policies. To the author's knowledge, however, there is no published text that explains how to perform these calculations. This paper is intended to fill this gap in the literature. It describes some of the more common techniques for pricing disability income insurance. Those techniques for which claim costs can be used and those for which the pricing actuary must project cash flows are identified.

Key words and phrases: *benefits, claim costs, cash flows, incidence rate, exposure, survivorship function*

1 Purpose and Product Features

1.1 Purpose

One of the primary purposes of life insurance is to protect a family from the catastrophic financial consequences of a breadwinner's premature death. Death is not the only reason, however, that a breadwinner may be unable to work. For example, what if the breadwinner is a victim of serious illness or injury and is unable to work for a significant period of time? Will his or her dependents be worse off financially than if he or she had died?

*Mark Chartier received his B.Sc. in mathematics from the Massachusetts Institute of Technology in 1980. He served four years in the United States Navy before joining Monarch Life Insurance Company. Afflicted with a congenital inability to pass fellowship exams, he bears the title "Career Associate" as a badge of honor.

Mr. Chartier's address is: Monarch Life Insurance Company, One Monarch Place, Springfield MA 01133, USA.

This other need defines the purpose of disability income insurance: to provide income when illness or injury renders the insured unable to work. The need is all too real and is arguably greater than that for life insurance. For example, according to the 1980 Commissioner's Standard Ordinary Table,[1] a male age 35 has a 0.00211 probability of dying during the next year. According to the 1985 Commissioner's Individual Disability Tables A,[2] the probability that a 35 year old male will suffer a disability lasting at least 90 days at some time during the next year ranges from 0.00164 in the best occupation class to 0.01176 in the worst occupation class.

Most United States residents have two forms of disability income protection provided by social insurance: workers' compensation and Social Security. Although workers' compensation is technically liability insurance for an employer, it effectively provides disability income protection to the employee. There are, however, some important limitations to workers' compensation. An obvious one is that it only protects employees from disabilities that are work-related. Another is that some segments of the population (the self-employed, for example) are not covered.

Social Security provides much broader protection. The *D* in OASDI[3] represents the United States' most comprehensive response to the need to care for the incapacitated. There are, however, serious shortcomings to Social Security including: (i) a restrictive definition of disability that makes it difficult to qualify for benefits; (ii) a claim settlement process that does not determine eligibility for benefits on a consistent basis; and (iii) an absence of promises. Entitlement to Social Security benefits of any kind is a statutory right (not a contractual right) and the terms of coverage can be changed at any time by an act of Congress. No promise is made about the level of benefits, the size or the timing of any cost of living adjustment, or the definition of disability used to determine eligibility for benefits. Only private insurers can make promises by entering into contracts with private individuals. See Rejda (1984, Chapter 2, pp. 19–46) for a more detailed discussion of the shortcomings of Social Security.

The providers of individual disability income insurance form a small segment of the insurance industry. Based on net earned premium fig-

[1] "Report of the Committee to Recommend New Mortality Tables for Valuation." *Transactions of the Society of Actuaries* XXXIII (1981): 617–674.

[2] "Report of the Committee to Recommend New Disability Tables for Valuation." *Transactions of the Society of Actuaries* XXXVII (1985): 449–601.

[3] *Old-Age, Survivors, Disability and Health Insurance* (OASDI) refers to the monthly retirement, disability, and survivor benefits paid under the United States Social Security system.

ures for 1992, more than half of the business is written by only four carriers: Paul Revere, Provident Life and Accident, Northwestern Mutual, and UNUM. (See Conning & Company, 1993.)

Similarly, not many Americans purchase individual disability income insurance. According to Conning & Company (1993), 75 percent of the individual disability income insurance in-force is on the top occupation class, the professional elite consisting mainly of physicians, attorneys, highly paid corporate executives, certified public accountants, actuaries, etc. For example, Soule (1993), reports that 80 percent of dentists, 78 percent of physicians, and 68 percent of lawyers have disability income insurance. The situation with group disability income insurance is slightly better. Contrary to popular belief, however, most Americans do not have group long-term disability income protection. According to Goldman (1990), while there were over 50 million workers covered by short-term disability (benefit period of two years or less) in 1986, the number covered by group long-term disability was small, less than 20 million.

1.2 Policy Features

Below is a list of some of the more common policy features of individual[4] disability income insurance. This is by no means an exhaustive description. The reader should consult Kidwell (1988, Chapter 3) for more detailed information.

1. **Rate Structure:** Premiums can be level (in which case they vary by issue age but not by attained age) or they can increase by attained age. Level premium rate structures are common and involve a significant prefunding of future benefits. This prefunding is quantified in the active life reserve (to be discussed later).

2. **Renewability:** Three categories are common: nonrenewable for stated reasons only, guaranteed renewable, and noncancellable.

   A contract is *nonrenewable for stated reasons only* if the insurer reserves the right to cancel coverage but may do so only for one of the reasons stated in the contract. For example, the contract might specify that the insured must be working full time in order to renew. The premium rate the insured pays is not guaranteed for the life of the contract.

[4]Group disability income coverage is a different topic. This important subject will not be discussed here. See Goldman (1990) for more on group disability income insurance.

A contract is *guaranteed renewable* if the insurer is contractually prohibited from canceling coverage for any reason other than failure to pay premiums. The premium rate the insured pays is not guaranteed. Even when the contract has a level rate structure, the insurer reserves the right to increase the rate if experience is significantly worse than anticipated when the contract originally was priced.

The most generous renewability provision is found in *noncancellable contracts.* The insurer makes timely payment of premiums the only condition for renewal and guarantees the premium rate until policy expiration, typically at age 65. An insured who buys such a contract with a level rate structure at age 25 will pay the same rate for his or her coverage for the next 40 years.

3. **Elimination Period:** The elimination period is the period of time for which the policyholder self-insures. The insured only begins to accrue benefits after the elimination period is completed. The elimination period can be likened to the deductible on a medical expense policy. Let's illustrate the concept by a few examples.

   Suppose the insured has a contract that, should he or she become disabled, will pay an annuity of $100 per month.[5] Furthermore, assume the insured is disabled for 45 days before returning to work:

   (a) If the insured's contract has a zero day elimination period, he or she is paid $150 (one and a half month's benefit) for those 45 days.

   (b) If the insured's contract has a 30 day elimination period, he or she would be paid $50 (half a month's benefit) for the remaining 15 days.

   (c) If the insured's contract has a 60 day elimination period, he or she would accrue no benefit because the disability did not extend beyond the end of the 60 day elimination period.

   This paper distinguishes time on claim and time disabled. A claim does not begin until the elimination period is completed. The length of time on claim plus the length of the elimination period equals the length of time disabled. Because the longer the elimination period, the lower the premium the insured must pay, the choice of an elimination period is important. It should reflect the

[5]One month is assumed to have 30 days

insured's judgment of how long he or she can rely on personal savings. Elimination periods of 30, 60, and 90 days are common. For financially well-endowed individuals, 180 day and 365 day elimination periods are available.

4. **Benefit Period:** This is the maximum length of time an insured can collect benefits for a single claim. Benefit periods can range from a few months to the entire life of the insured. A common benefit period is one that would allow the policyholder to remain on claim until age 65.

5. **Definition of Disability:** Determining whether to pay a life insurance benefit is a relatively straightforward matter. This is not the case for disability income insurance; the potential for policyholder abuse is enormous. It is essential that the policy state as precisely as possible what constitutes a disability. Two definitions are common. The most generous is the regular occupation definition. According to this definition, the insured is disabled if an illness or injury renders her or him unable to perform the substantial and material duties of her or his regular occupation; whether the insured is able to perform the duties of some other occupation is irrelevant.

   An alternative is the reasonable occupation definition under which the insured is disabled if he or she is unable to do the substantial and material duties of any reasonable occupation. A reasonable occupation is one the insured could be expected to perform by virtue of education, training, and work experience. Some insurers modify the definition of a reasonable occupation by asserting that due regard must be given to the insured's earnings before disability began. The intent is to protect the insured from being forced into an occupation in which he or she would suffer a substantial loss of income.

   Let us illustrate the difference between these two definitions with a realistic example. Suppose an insured's regular occupation is that of a surgeon. The insured begins to suffer from an impairment in her or his right wrist, perhaps carpal tunnel syndrome or arthritis. He or she no longer can perform surgery. But by entering general medical practice, the surgeon still can earn a much higher income than that of the average individual. Is the surgeon disabled? By the reasonable occupation definition, the answer is no; the surgeon now can perform the duties of what is for her or him a reasonable occupation. By the regular occupation definition, however, the surgeon is considered to be disabled because he or she cannot perform the duties of his or her regular occupation. Therefore, the surgeon can collect benefits while she or he earns income from a new occupation!

Some policies offer a hybrid of the two definitions. For example, the regular occupation definition may apply only during the first 24 months of a claim. After 24 months, a claimant can continue to collect benefits only if he or she meets the reasonable occupation definition.

6. **Partial or Residual Disability:** The above definitions refer to total disability, the complete inability to perform the substantial and material duties of the insured's regular or some other reasonable occupation. What if the disabled insured can work, but only on a part-time basis? What if she or he can work full-time but can perform only some of the key duties of the job? If a contract only covers total disability, he or she is not eligible for benefits. Some contracts will pay a fraction of the policy's full benefit amount under a partial or residual disability clause. Under a partial disability clause, the benefit is a function of time unable to work. Under a residual disability clause, the benefit is a function of income lost. Contracts providing residual benefits are common and will be discussed later.

7. **Presumptive Disability:** Some contracts will assert that for certain conditions, the policyholder will be presumed disabled and able to collect benefits even if he or she continues to work and suffers no loss of income. Such conditions might include total blindness in both eyes, loss of use of both hands or both feet, total deafness, etc. The presumptive disability provision also might extend the benefit period under such circumstances, e.g., pay benefits for life in case of blindness.

8. **Protection Against Overinsurance:** The insured is not supposed to profit from a disability claim. Therefore, the insurer should not issue so much coverage that when social insurance and any other private insurance is added, the insured's income on claim is higher than the income before disability. There are policy provisions that can be added to a disability contract to provide the insurer with added protection against overpayment.

   One example is a coordination with social insurance clause. Under such a clause, the amount the insured is paid by his or her private insurer can be reduced dollar for dollar by what he or she receives from social insurance. If a contract has a benefit amount of $1000 per month and a claimant receives $400 a month from OASDI, then he or she will receive only $600 per month from private insurance. Other variations are possible. The formula for coordination of benefits need not be a dollar-for-dollar offset. It may be a percentage reduction in the event social insurance is received.

The above provision applies only to social insurance. It does not take into account other private insurance policies the insured may own. A more comprehensive provision is the relation of earnings to insurance clause, an optional provision under the NAIC's[6] Uniform Individual Accident and Sickness Policy Provision Law. For a discussion of the Uniform Law, see O'Grady (1988, Appendix 2). Such a clause allows the benefit the insurer would pay to be reduced proportionately. Suppose that the insured's income before disability is only 90 percent of the sum of all of his or her disability income insurance benefits. If a contract contained the relation of earnings to insurance clause, the insurer would pay only 90 percent of the policy's benefit amount. Neither of the above provisions typically is found in contracts sold to the industry's target market, the self-employed professional; intense competition has pushed the industry to produce very generous contracts.

9. **Waiver of Premiums:** The typical waiver of premium provision specifies that once a disabled insured has satisfied a certain waiting period (typically longer than the elimination period), premiums will be waived for the duration of disability and any premium paid after commencement of disability will be reimbursed.

10. **Riders:** Various optional benefits can be purchased to supplement the basic contract. A rider can be purchased to adjust benefits for cost of living increases while the insured is on claim. The insured can pay for the right to purchase additional coverage in the future without evidence of medical insurability. The insured could purchase a social insurance contingency rider; if he or she applies for Social Security and is turned down, he or she can collect benefits under this rider. Halpern (1979) discusses how to price such a rider.

[6]The National Association of Insurance Commissioners (NAIC) is an association consisting of state insurance commissioners. The NAIC drafts model laws and recommends their adoption by state legislatures. The NAIC has no legal authority to force states to enact its recommendations.

1.3 Risk Variables

Claim experience is affected by many variables including the following: age, gender, elimination period, benefit period, relationship of benefit amount to income, occupation class, and state (geographical location) of residence. The effects of some of these variables are quantified in the rate manual. For example, premium rates increase with issue age and benefit period, and they decrease with elimination period. Some insurers use a single rate manual for the entire nation, others have surcharges in high risk states and discounts in low risk states. Women, especially those in the child-bearing ages, tend to be charged a higher rate than men, a practice justified by higher claim incidence rates. Occupation class is a crucial risk variable in disability income insurance. As a life insurance risk, a carpenter may be no different from an office worker; however, this is not the case in disability income insurance. Disability insurers therefore group occupations into broad risk classes. The class charged the lowest premium rate normally contains highly skilled professionals: physicians, lawyers, accountants, etc. The class charged the highest premium rate normally contains occupations involving substantial manual labor.

Unfortunately, the risk variable that is the hardest to measure (and is probably the most significant) is the motivation to work. A highly motivated insured with some health problems may be a better risk than a healthy insured who is willing to turn a questionable condition into a claim. The problems associated with malingering make it more difficult for the underwriter to judge the disability risk than to judge the life risk. It is crucial that no applicant be allowed to purchase a higher benefit amount than is justified by his or her income.

1.4 Overview

We will explore some of the mathematics of disability income insurance in the U.S. and Canada. In particular, Section 2 describes the underlying model and the basic notation used throughout the paper. Section 3 describes the concept of *claims costs* and contrasts it with *benefit cash flows*. The *statutory active life reserve* and the *disabled life reserve* are discussed in Section 4. Section 5 describes two approaches to calculating profits: the *claim cost profit model* and the *statutory profit model*. Section 6 provides an overview of the techniques used in pricing disability income insurance: the loss ratio technique, the percentage of premium profit method, and the asset share and the return on investment techniques. Section 7 shows how to calculate cash flows. Section

8 deals with waiver of premiums. Section 9 investigates the impact of relaxing some of the assumptions (in Section 2) on the model. Section 10 reviews a sample asset share calculation presented by Bluhm and Koppel (1988, Chapter 4). The appendix contains a numerical example.

For an overview of the mathematics of disability income insurance in Europe, see Gregorius (1993), Hertzman (1993), Mackay (1993), and Segerer (1993).

2 The Model

Consider a closed block of individual disability income insurance policies (to be described in Section 2.1). This block is assumed to consist of homogeneous business cells in which all policyholders have common parameters (characteristics) such as issue age, occupation class, elimination period, waiting period for waiver of premiums, and gender. Each cell is defined by the value of θ, a vector of parameters. To be precise, θ = (issue age, gender, occupation class, elimination period, waiting period for waiver of premiums). For example, a cell may consist of policies sold to female surgeons with issue age 45, a 30 day elimination period, and a 180 day waiting period for waiver of premiums. This yields θ = (45, female, surgeon, 30, 180).

Throughout the rest of this paper, the symbol θ will be used to denote a particular business cell. We will develop functions and expressions that are dependant on θ. In most of these cases, θ appears as a subscript.

2.1 The Policy

Consider an individual disability income insurance policy with the following features:

1. The policy is sold only to those individuals who are active (not disabled) at the time of issue.

2. An insured receives benefits if and only if the insured is disabled for a continuous period of time in excess of the elimination period. The definition of disability is not important here, suffice to say that the definition chosen will affect the probabilities of disability and recovery.

3. Premiums are not necessarily level and are paid annually.

4. There is a waiver of premium if the insured remains disabled for a period of time in excess of the waiting period for waiver of premiums. Premium payments resume upon recovery.
5. Benefits are paid up to age 65.
6. The unit of benefit is $1/12 per month.

2.2 Mortality, Morbidity, Recovery, and Lapses

There are three sources of decrement at work on the *in-force population*: (i) voluntary lapsation; (ii) death while on claim; and (iii) death while not on claim. The mortality rates of insureds on claim are different from those of insureds not on claim. In addition, the mortality rates of insureds who have been on claim and who since have recovered may be different from those who have never been on claim. Finally, the voluntary lapse rates of insureds who have been on claim are almost certainly lower than the voluntary lapse rates of insureds who have never been on claim. Those who have benefited from the contract are more likely to hold onto it.

In practice, the disability income insurance actuary uses a single set of tables, called *lapse tables*, to decrement the in-force population. These tables express the three sources of decrement mentioned above as a single aggregate source of decrement. Separate lapse tables for insureds who have been on claim and insureds who have never been on claim are not used.

The in-force population can be divided into two subpopulations: the active population and the claim population.

1. The *active population* is the population exposed to the risk of disablement. There are three sources of decrement and one source of increment on the active population:

   (a) Voluntary lapsation;

   (b) Death while not on claim;

   (c) Migration from the active population to the claim population; and

   (d) Recovery from claim (the one source of increment).

   We will assume initially that we can identify the active population with the in-force population. That is, the population exposed to risk will be calculated by taking the in-force population at time of issue and decrementing only for voluntary lapsation, death on claim, and death not on claim.

2. The *claim population* consists of those persons who are receiving disability benefits. Thus, disabled persons who are *not* receiving a benefit are not considered as being on claim. There are two sources of decrement and one source of increment on the claim population:

   (a) Death while on claim;

   (b) Recovery from claim; and

   (c) Going on claim (the one source of increment).

When measuring the size of the active population during the year, migration to the claim population and return from claim to the active population will be ignored. We make this assumption for two reasons. First, the claim population is very small relative to the active population, so little error results from identifying the active population with the entire population in-force. Second, tracking the continuous two way migration between active and claim population is a tedious process, and we want to keep the model as simple as possible. In Section 9.2, we will describe a way to track this two way migration at discrete intervals.

A set of morbidity tables with two decrements (death on claim and recovery from claim) is used to project increments and decrement to the claim population. Claim incurral is the sole source of increment. As with lapse rates, the disability income insurance actuary does not use a multiple decrement table. Instead he or she uses claim termination rates that combine termination due to death and recovery.

The rate of claim incurral for insureds who have never filed a claim is undoubtedly different from the incurral rate for those who have filed claims in the past. In practice, the disability income actuary uses aggregate rates of claim incurral drawn from a single set of morbidity tables that do not distinguish the two groups of insureds. In addition, because policyholders are assumed to pay premiums annually, voluntary lapsations only occur at policy anniversaries. While premiums are being waived for a disabled insured, the policy cannot lapse.

Given these three decrements (mortality, morbidity, and lapses) and the increment (recovery), a separate combined mortality-disability-lapse table is used for each occupation class and gender. The mortality-recovery table for disabled lives will be a select table with age at time of disability and duration of disability as the parameters in the table.

For ease of computation, the following assumptions are made:

1. The incidence of claim is distributed uniformly throughout the year. Thus, given that there are $E_\theta(n)$ units of claim at the start of the nth policy year and that $r_\theta(n)$ is the incidence rate of claim,

then the expected number of claims in the interval $(t, t+dt)$, with $0 < t < 1$, is $E_\theta(n) r_\theta(n) dt$.

2. Deaths and lapses in the in-force population occur at the end of the policy year.

3. The population on claim is so small relative to the active population that we initially will assume that the units of insurance exposed to risk are unaffected by claim incurrals and recoveries.

The effects of relaxing these assumptions are investigated in Section 9.

2.3 Notation

Some of the more basic symbols used in the model will be defined and assumptions for the model presented.

$\theta =$ A vector of parameters that characterize each business cell;

$i =$ The valuation rate of interest;

$v =$ The annual discount factor, i.e., $v = 1/(1+i)$;

$x =$ The issue age, $x = 15, 16, 17, \ldots$;

$z =$ The attained age $= x + n - 1$;

$n =$ The policy year of disablement, $n = 1, 2, 3, \ldots$;

$m =$ The current policy year $m = n, n+1, \ldots$;

$e =$ The length of the elimination period (measured in years). It is the minimum length of time the insured must be disabled in order to qualify for benefits;

$w =$ The length of the waiting period (measured in years) to qualify for waiver of premium, with $w > e$;

$b =$ The length of the benefit period (measured in years). Usually the benefit period extends to age 65. Note, $b > w > e$;

$r_\theta(n) =$ The incidence rate of claim in the nth policy year for a policy with parameter θ, i.e., the probability that an active insured (currently age $x+n-1$) from the business cell with parameter θ becomes disabled and remains disabled for at least the length of the elimination period;

$E_\theta(n) =$ The number of units[7] of insurance exposed to risk in the nth policy year for a policy with parameter θ;

$P_\theta(n) =$ The annual premium per unit of insurance in-force during the nth policy year for a policy with parameter θ;

$s_\theta(y,n) =$ The probability that a policyholder (with parameter θ) who is disabled in the nth policy year and remained disabled throughout the elimination period will stay on claim (i.e., receive benefits) for at least y consecutive years into the future. Also, $s_\theta(y,n) = 0$ when y is greater than the length of the benefit period;

$BC_\theta(n,m) =$ The benefit cash outflow in the mth policy year to claimants (with parameter θ) who are disabled in the nth policy year; and

$W_\theta(n,m) =$ The amount of premium waived at the beginning of the mth policy year on claimants (with parameter θ) who are disabled in the nth policy year.

3 Claim Cost

Before introducing the concept of a claim cost, let us review the way $PDB_\theta(n)$, the projected death benefit cash outflow in the nth policy year on a single life insurance policy with parameter θ, is calculated. Once mortality and lapse assumptions are chosen for pricing purposes, the life actuary easily can determine $PDB_\theta(n)$ as follows:

$$PDB_\theta(n) = q_\theta(n)DB_\theta(n)$$

where $q_\theta(n)$ is the mortality rate in the nth policy year, and $DB_\theta(n)$ is the number of dollars of death benefit exposed to risk during the nth policy year for a policy with characteristics θ.

For a disability income product, however, the calculation is more complex because the benefit is an annuity, not a single lump sum payment. The pricing actuary simplifies the problem by calculating a *claim cost.* For policies issued at age x, the claim cost associated with the nth policy year, $CC_\theta(n)$, is given by:

$$CC_\theta(n) = v^e r_\theta(n)E_\theta(n)a_\theta(z)$$

[7]A unit of insurance is defined as $1 of annual benefit paid monthly until recovery or the end of the benefit period.

where $a_\theta(x)$ is the actuarial present value of an annuity paying \$1 per year (paid monthly) after the elimination period ends to a life who is disabled at age z in the nth policy year.[8] In other words, a claim cost associates with a policy year the present value of all future monies that will be paid to a policyholder whose date of disablement is in that year.

In life insurance, future death benefits in a closed group of cohorts are a function of two random processes: lapses and mortality. In disability income insurance, the extent of future claim payments in a closed group of cohorts depends on lapses, the incidence of claims, and the severity of claims. The fact that the claim severity is random is one of the reasons why the claim experience on disability income insurance is inherently more volatile than the claim experience on life insurance. It is also one of the reasons why disability income insurance has more in common with property/casualty insurance than it has with life insurance.

Lapses, the incidence of claims, and the severity of claims vary by age, gender, occupation class, elimination period, policy age, and the contract's definition of disability. They also are affected by many other factors, some of which the pricing actuary cannot quantify easily (such as the state of the economy, for example). The claim incidence rate increases with age. In addition, the claim incidence rate decreases as the elimination period increases; it is higher for women than for men, at least when women are in their child-bearing years—it is possible this relationship reverses at advanced ages. The greater the physical stress of an occupation, the higher is the claim incidence rate.

The severity of a claim (i.e., the annuity factor) also varies with the benefit period and the interest rate. The seriousness of a claim generally increases with age. On the other hand, the benefit period shrinks as the insured ages. A policyholder with a benefit period to age 65, for example, can stay on claim for ten years if disabled at age 55. He or she can stay on claim only for five years if disabled at age 60. This can lead to a curious pattern: as a block of policies ages, claim costs first will increase. As insureds reach their late fifties or early sixties, claim costs can decrease as the shrinking benefit period causes claim severity to become smaller.

Unlike claim incidence, claim severity rises as the elimination period increases. This is because a long elimination period screens those claims that would have closed relatively early. For example, in comparing an insured with a 30 day elimination period to an insured with a 90 day elimination period, the latter must be disabled more seriously

---

[8] By convention, $CC_\theta(n)$ is valued as of the date of disablement.

than the former to get on claim. But if they get on claim, the latter is expected to remain on claim longer. The author frequently has observed that claim severity in the less skilled occupation classes is less than that in the highly skilled occupation classes. In other words, blue and gray collar workers return to work faster than white collar workers. This is the case even when differences in elimination period and age distribution are accounted for.[9]

Let $H_\theta(z)$ denote the claim cost per unit at attained age z, then:

$$H_\theta(z) = v^e r_\theta(n) a_\theta([z + 1/2] + e, b) \tag{1}$$

where $a_\theta([z+1/2]+e, b)$ = the actuarial present value of an annuity of \$1 per year (paid monthly) for at most b years starting at age $z+1/2+e$ to a life (with parameter θ) disabled at age $z + 1/2$. Age at disability is taken to be $z + 1/2$ because it is assumed that, on average, disability begins in the middle of the policy year. The annuity starts at $z+1/2+e$ because payment commences after the elimination period is completed. The annuity is contingent on the insured remaining disabled. The claim cost now can be rewritten as

$$CC_\theta(n) = E_\theta(n) H_\theta(z). \tag{2}$$

4 Reserves

First let us introduce two important items that appear on the balance sheet of the disability income insurer: the active life reserve and the claim reserve. Roughly speaking, the active life reserve is that part of the liability for future claims (yet to occur) that must be prefunded, and the claim reserve is the liability for claims that already have been incurred. In more precise language, the active life reserve is the expected present value of future claim costs minus the expected present value of future net premiums. It is analogous to the policy reserves held on life insurance contracts. The claim reserve is another item that makes disability income insurance similar to property/casualty insurance. It is the expected present value of all future payments, both contingent and noncontingent, that will be made on claims that have been incurred. For a more thorough discussion of the active life reserve and the claim reserve, see Bartleson (1968) and Shapland (1988, Chapter 5).

[9]One explanation may be that the higher claim incidence rate among blue collar workers causes the lower severity. High claim incidence means there are many claims for conditions that are not serious, so they close quickly. Another is that insureds in the lower skilled occupations have less generous contracts, giving them greater incentive to return to work.

The formula for the *statutory active life reserve* at the beginning of the nth policy year, $V_\theta^{(aa)}(x,n)$, on a single unit of insurance issued at age x and continuable until attainment of age 65 is:

$$V_\theta^{(aa)}(x,n) = \sum_{k=0}^{65-z} v^k {}_kp_z^{(aa)} [v^{1/2} {}_{1/2}p_{z+k}^{(aa)} H_\theta(z+k) - P_\theta^*] \qquad (3)$$

where H_θ is defined in equation (1), and ${}_kp_z^{(aa)}$ is the probability that an active life age in cell θ (with issue age x) survives k years.[10] To avoid cumbersome notation, the θ is not shown; P_θ^* is the statutory net level premium payable to age 65 for a unit of insurance issued an active life age in cell θ.

Again, a claim is assumed to be incurred, on average, in the middle of the policy year; hence, the presence of $1/2$ in the exponent of v and the subscript of the probability of survival p. We assume that the reserve is established on a level premium contract. The terminal active life reserve is calculated by multiplying the above reserve per unit by the number of units in-force. As equation (3) is to be used to calculate statutory reserves, values of H, v, p, and P will be specified by state regulation and may bear little resemblance to the actuary's best guess assumptions about interest rates and future morbidity.

The claim reserve is more complex. At any given time, the population of insureds on claim will be the sum of several closed cohorts. Each cohort is defined by the amount of time it has been on claim. Let $R_\theta(n,m)$ be the statutory claim reserve at the beginning of the mth policy year (or the end of the $(m-1)$st policy year) on those claimants disabled during the nth policy year. The total claim reserve at the beginning of the mth policy year, $R_\theta(m)$, is the sum of the claim reserve on $m-1$ cohorts,

$$R_\theta(m) = \sum_{n=1}^{m-1} R_\theta(n,m).$$

Now $R_\theta(n,m)$ is a function of two quantities: (i) the number of units of insurance disabled in the nth year still on claim at the beginning of the mth year; and (ii) the claim reserve established on a single unit disabled in the nth year and still on claim at the beginning of the mth year. Later we will develop the formula for the number of units disabled in the nth year still on claim at the beginning of the mth year. The claim

[10] Once a person is insured, a reserve is maintained for her or him. Even if she or he is on claim, an active life reserve is maintained for her or him because she or he could go on claim again after recovery.

reserve factor for a single unit of insurance *disabled* in the nth year and *still on claim* at the beginning of the mth year is $V_\theta^{(ii)}(n,m)$, i.e.,

$$V_\theta^{(ii)}(n,m) = \sum_{k=1}^{12(b-(m-n-e-1/2))} \frac{1}{12} v^{k/12} {}_{k/12}p^{(ii)}_{[z+1/2]+m-n-1/2} \quad (4)$$

which is the present value of a \$1/12 per month annuity. The notation ${}_{k/2}p^{(ii)}_{[z+1/2]+t}$ denotes the probability an insured who became disabled at exact age $z+1/2$ and has remained disabled for t years will remain on claim for at least another k months. The term $12(b-(m-n-1/2-e))$ is the number of months remaining in the benefit period. Again, we assume that disablement occurs on average in the middle of the policy year. The values of $p^{(ii)}$ and v are specified by statutory regulation and may not coincide with the pricing actuary's best guess estimate of what will happen.

Let $D_\theta(n,m)$ be the number of units of insurance disabled in the nth year and still on claim at the start of the mth policy year; then

$$R_\theta(n,m) = V_\theta^{(ii)}(n,m)D_\theta(n,m). \quad (5)$$

5 Profits

We present here two general approaches to measuring profit (for a homogeneous group of policies with parameter θ) in disability income insurance. Each provides a formula for annual recognition of profit. It should be evident by the end of this discussion that the two approaches do not recognize the same profit year by year. Throughout this section, we will drop the subscript from the symbols to reduce clutter. We must remember, however, that the totals in this section apply to the business cell with parameter θ.

5.1 The Claim Cost Profit Model

Under a claim cost pricing model, the total profit from cell θ in the nth policy year, $PROF_n^{(cc)}$, is given by the following formula:

$$PROF_n^{(cc)} = P_n + I_n - EXP_n - COM_n - CC_n - (ALR_n^{(e)} - ALR_n^{(b)}) \quad (6)$$

where P_n is the total premium earned; I_n is the total investment income; EXP_n is the total expenses; COM_n is the total commissions from cell θ; CC_n is the total claim cost incurred during the nth policy year; $ALR_n^{(e)}$ is the total active life reserve at the end of the year; and $ALR_n^{(b)}$ is the total active life reserve at the beginning of the nth policy year. Premiums, expenses, and commissions are calculated in much the same way as they are for life insurance. The *ALR* terms are calculated by multiplying equation (3) by the number of units of insurance in-force, i.e.,

$$\begin{aligned} ALR_n^{(b)} &= V_\theta^{(a)}(n)E_\theta(n) \\ ALR_n^{(e)} &= V_\theta^{(a)}(n+1)E_\theta(n+1). \end{aligned}$$

We uncover a serious flaw in the claim cost pricing model when we try to calculate investment income, I_n. The reader may be tempted to calculate it as:

$$I_n = i \times (P_n + ALR_n^{(b)} - EXP_n - COM_n - v^{1/2}CC_n) \tag{7}$$

where i is the assumed rate of interest. Again, we have assumed that claim costs are incurred in the middle of the policy year. It is a mistake to subtract the entire value $v^{1/2}CC_n$, however, because CC_n is a lump sum representing a series of cash flows that may be spread over many future years. Only the portion that is disbursed in the current year should be subtracted; the insurer is free to invest the remainder until payment is due. Because the claim cost pricing model does not divide that lump sum into money disbursed now and money disbursed later, it cannot correctly allocate investment income by policy year. The statutory profit model corrects this flaw.

5.2 The Statutory Profit Model

Statutory book profit, $PROF_n^{(s)}$, is given by the equation:

$$\begin{aligned} PROF_n^{(s)} &= P_n + I_n - EXP_n - COM_n - BEN_n \\ &\quad - (CR_n^{(e)} - CR_n^{(b)}) - (ALR_n^{(e)} - ALR_n^{(b)}) \end{aligned} \tag{8}$$

where BEN_n represents the actual benefits paid in cell θ, $CR_n^{(b)}$ and $CR_n^{(e)}$ are the total claim reserve at the beginning and the end of the nth policy year. The only apparent difference between equations (6) and (8) is that claim cost has been replaced by benefits paid plus the change in claim reserve. Are equation (6) and equation (8) equal? They typically will not be. Claim costs are projected using *pricing* mortality-morbidity tables and interest rates. Statutory claim reserves are measured using *statutory* mortality-morbidity tables and interest rates. It is highly unlikely that quantities derived from different assumptions will be equal.

Under the statutory profit model, investment income is:

$$I_n = i \times (P_n + ALR_n^{(b)} + CR_n^{(b)} - EXP_n - COM_n - v^{1/2}BEN_n). \quad (9)$$

Notice that in equation (9) investment income is counted on all monies held in reserve and that the active life reserve earns interest, but the claim reserve is ignored. In equation (9), benefits paid in a given year are subtracted from revenue before applying an interest rate, but money for future benefits earns interest until the benefits are paid. If the pricing actuary needs to project the year by year pattern by which statutory profit emerges, claim costs cannot be used. The pricing actuary instead must calculate benefit cash flows.

6 Overview of Pricing Techniques

This section discusses some of the methods used to determine premium rates for disability income insurance. The list is by no means exhaustive.

6.1 The Loss Ratio Technique

Before describing this technique, we must define what is meant by the *loss ratio.* The term loss ratio usually is understood to be the fraction of the policyholders' premiums that is returned in benefits. That sounds simple enough, yet there is great confusion about what a loss ratio is and significant disagreement about how it is calculated. A loss ratio can be retrospective or prospective. It can be applied over the entire life of a block of business or to a single experience year. It can be calculated with GAAP reserves, statutory reserves, natural reserves, or no reserves at all. It can be calculated using a realistic interest rate, a statutory interest rate, or no interest rate at all. Claim settlement expenses can be added to the numerator of the loss ratio (property/casualty insurance) or play no part in the calculation of the loss ratio (health insurance). (See Pharr, 1979, with discussion.)

For our purposes, the term *loss ratio* is defined to be the ratio of the expected present value of future benefits that will be paid over the business cell to the expected present value of future premiums that will be collected from the business cell. This is a prospective lifetime loss ratio. When calculating the relevant present values, the author recommends that we discount for voluntary lapsation as well as mortality and that we use a realistic rate of interest, not the statutory valuation rate. This view is not shared by all actuaries.

Once the question of how to calculate a loss ratio is settled, the loss ratio technique becomes the simplest approach of all. It can be broken down into three steps:

1. Calculate the present value of future claim costs on a single unit of insurance.

2. Calculate a level net premium rate by dividing the result of step 1 by an annuity factor.

3. Calculate a gross premium by dividing the net premium from step 2 by a target loss ratio. For example, assume the present value of future claim costs is \$50. Assume the present value of an annuity due of \$1 per year over the life of the insurance contract is \$5. Assume the target loss ratio is 50 percent. Then $50/(5 \times 0.5) = \$20$ is that gross premium rate for which the expected present value of future benefits divided by the expected present value of future premiums is 50 percent. Note that in step 1 you could calculate the present value of future benefit cash flows in place of future claim costs. Claim costs, however, are suited perfectly for this technique.

There are advantages and disadvantages to the loss ratio method. The most important advantage is its simplicity. The formulae are easy to understand, and no assumptions for future expenses are needed. One doesn't even need to know precisely what the commission scales are. This simplicity is also its most important disadvantage; the difference between gross and net premiums may not be sufficient to cover expenses and commissions. Nevertheless, if a company's expenses or commissions are unacceptably high, the pricing actuary may be compelled to use the loss ratio technique.

A minimum loss ratio is required by law or regulation in most states. Taking the minimum loss ratio as a pricing target, the actuary can solve for the maximum premium that legally can be charged. Once the actuary knows how much premium is left after paying benefits, he or she can solve for target expense and commission levels down to which actual levels should be managed. Another advantage is that if the target loss ratio is set high enough, the method automatically produces a rate manual whose anticipated loss ratio exceeds the required minimum. This has not been a significant advantage of late. Anticipating a certain loss ratio is one thing, experiencing it is another. Despite all the confusion regarding correct loss ratio calculation, one unambiguous lesson has emerged from the last half of the 1980s: the least of the industry's worries is that the loss ratio will be too low.

6.2 Percentage of Premium Profit Method

This method is similar to the equation method in life insurance except that the present value of future death benefits is replaced by the present value of future claim costs. The pricing actuary takes as his or her target a certain percentage of the present value of premiums that will go to profit. The actuary projects the present value of claim costs, expenses, etc., and solves for that premium rate at which a sufficient percentage of premium will be left to meet or exceed the target set. Claim costs are suited perfectly to this technique. Again, the present value of future claim costs could be replaced by the present value of future benefit cash flows to obtain the same result.

Because the percentage of premium profit method makes explicit provision for commissions, expenses, and taxes, it is superior to the loss ratio technique. A major shortcoming of the percentage of premium profit method is that it does not quantify the risk/reward relationship. How great a percentage of premium profit is sufficient to compensate the insurer for taking the disability risk? How does a risky disability income portfolio with a 10 percent of premium profit compare with a risk-free Treasury bill paying a 3 percent return?

As noted earlier, disability income insurance has volatile claim experience. In part this is due to the extra random process in morbidity, the claim severity, and the fact that the insured exercises some control over morbidity, possibly electing to be on claim rather than being put on claim by forces beyond his or her control. Disability income insurance is risky business for a variety of reasons, and the faint of heart are driven from the marketplace. A pricing method should produce a rate manual that is not merely profitable, but is more profitable than one for a less risky line of business.

6.3 The Asset Share Technique

As with life insurance, the profit target for disability income insurance can be set as a certain asset share by a certain policy year. Alternatively, the target could be set as a certain level of surplus by a certain year, where *surplus* could be defined as the difference between the asset share and the total statutory reserve. Claim costs are not appropriate for the calculation of asset shares. One cannot take the asset share equation for life insurance, substitute claim cost incurred for death benefits paid, and consider the result to be the disability income insurance asset share equation.

An asset share is the accumulation, per unit of insurance in-force, of all cash inflows to date minus all cash outflows to date. Every variable in the equation for asset share represents a cash flow. A claim cost is not a cash flow, however; it is a lump sum assigned to a single policy year and is equal to the present value of a series of cash flows that may spread over several future years. In the first policy year, the claim cost is higher than the benefit cash flow because the claim cost includes payments that will be made in future years. A claim cost pricing model understates the first year asset share. There may be subsequent years in which the claim cost is less than the benefit cash flow because some claimants disabled in prior policy years will be collecting benefits in the current year; each year's claim cost measures payments only to claimants whose disabilities commenced in that year.

Bluhm and Koppel (1988) present a sample asset share calculation. An attempt is made to calculate benefit cash flows using claim costs and changes in claim reserve. After developing our own model to project benefit cash flows, we will discuss some of the problems inherent in their method and show how it can be rehabilitated.[11]

6.4 The Return on Investment Technique

With this method, the pricing actuary projects future book profits and then solves for the return on investment (ROI). The ROI is the discount rate at which the present value of renewal year profits equals the loss in the first policy year. The profit target is a threshold ROI. In the author's opinion, this technique is superior to the other methods presented above because it quantifies the insurer's reward for bearing the significant risk of competing in the disability income marketplace. The rational investor only increases risk if he or she has a reasonable expectation of a higher return (and insurance company shareholders are, we presume, rational investors).

If the risk of selling disability income is higher than that of selling term life, then the insurer is entitled to expect a higher ROI from the former product than from the latter. The ROI pricing technique not only allows the insurer to compare disability income to other products, but allows the insurer to tailor individual rates to each particular risks that it bears. For example, the insurer bears more risk when it sells a contract with a lifetime benefit period than when it sells a contract with a benefit period of only one year. The rates should be set so that the

[11]These comments are not intended to detract from the quality of the articles contained in this excellent text. The author highly recommends the O'Grady textbook.

insurer's expected ROI on lifetime contracts is higher than that on one year benefit period contracts.

It is here that we come to the most serious defect of a pricing model based on claim costs instead of cash flows. While claim costs can be used to project the present value of profit over the entire life of a block of business, they cannot be used to project the year by year pattern by which statutory book profit emerges. Consequently, a claim cost pricing model cannot measure ROI correctly. This is because claim costs are calculated with pricing assumptions, while statutory claim reserves are calculated with assumptions specified by regulatory authorities.

If pricing assumptions are more liberal than statutory assumptions (not a given in today's morbidity environment), then claim costs will be lower than benefits paid plus change in claim reserve in the early policy years. At this time a claim cost pricing model will overstate statutory book profit. In later years the inequality will reverse as money released from the claim reserve makes the real book profit higher than that predicted by a claim cost pricing model. Therefore, if pricing assumptions are more liberal than statutory valuation assumptions, a claim cost pricing model will recognize profit earlier than would emerge under statutory accounting and will overstate the ROI that will be realized.

In summary, if an insurer places any importance on estimating the asset shares or the ROI of a new disability contract it contemplates introducing, the actuary must translate pricing assumptions into cash flows, not claim costs. The remainder of this paper will be devoted to a model for doing this.

7 Cash Flows on the Base Policy

We now develop a model to calculate $BC_\theta(n,m)$, the value of the benefit cash outflow in the mth policy year to claimants with dates of disablement in the nth policy year. As an example, consider a business cell consisting of disability income policies that will expire five years after issue. Let

$$\begin{pmatrix} BC_\theta(1,1) & BC_\theta(1,2) & BC_\theta(1,3) & BC_\theta(1,4) & BC_\theta(1,5) \\ & BC_\theta(2,2) & BC_\theta(2,3) & BC_\theta(2,4) & BC_\theta(2,5) \\ & & BC_\theta(3,3) & BC_\theta(3,4) & BC_\theta(3,5) \\ & & & BC_\theta(4,4) & BC_\theta(4,5) \\ & & & & BC_\theta(5,5) \end{pmatrix}$$

be the matrix of benefit cash flows.

The total benefit cash flow in the first year is $BC_\theta(1,1)$. In the second policy year, money is paid to some claimants with disablement dates in the first policy year, $BC_\theta(1,2)$, and money is paid to claimants with disablement dates in the second policy year, $BC_\theta(2,2)$. The total benefit cash flow in the second policy year is $BC_\theta(1,2) + BC_\theta(2,2)$, the sum of the entries in the second column of the array. Likewise, the cash flow in the third policy year is the sum of entries in the third column, and so forth for all other policy years. It is interesting to note what happens if we move across a row of the array rather than down a column. If the series of cash flows $BC_\theta(1,1)$, $BC_\theta(1,2)$, $BC_\theta(1,3)$, $BC_\theta(1,4)$, and $BC_\theta(1,5)$ are discounted for interest back to policy issue, the result is the claim cost in the first policy year. The total cash outflow in the mth policy year, $TCO_\theta(m)$, is the sum of entries in the nth column, i.e.,

$$TCO_\theta(m) = \sum_{k=1}^{m} BC_\theta(k,m).$$

In addition, the total claim cost in the nth policy year, $TCC_\theta(n)$, is the result of discounting the entries in the nth row back to the nth policy year, i.e.,

$$TCC_\theta(n) = \sum_{m=n}^{\infty} v^{m-n+1/2} BC_\theta(n,m).$$

This illustrates the flexibility of a cash flow model. From cash flows you can calculate claim costs. From claim costs you cannot calculate cash flows.

Next we turn our attention to the calculation of $BC_\theta(n,m)$ for a unit of benefit. This task is divided into the three cases shown below. **For simplicity, the benefit of \$1 per year is assumed to be paid continuously.**

Case 1, m = n : $BC_\theta(n,m)$ is the benefit cash flow in the nth policy year to claimants with dates of disablement in that same year. $E_\theta(n)$ units of insurance are exposed to risk at the start of the nth policy year. The number of those units that will go on claim during the time period $(t, t+dt)$ (with $0 < t < 1$) by completing the elimination period is $r_\theta(n)E_\theta(n)dt$. At time $t+e$ these claimants will begin to accrue benefits. The number of units disabled during the time interval $[t, t+dt]$ that are still on claim at time $t+e+y$ is $r_\theta(n)E_\theta(n)s_\theta(y,n)dt$. Each unit of insurance that is on claim at time $t+e+y$ is paid dy during the interval $[t+e+y, t+e+y+dy]$. The final equation is:

$$BC_\theta(n,n) = r_\theta(n)E_\theta(n) \int_{t=0}^{1-e} \int_{y=0}^{1-e-t} s_\theta(y,n)\,dy\,dt. \qquad (10)$$

Note the upper limits of integration. If an insured becomes disabled after time $t = 1 - e$, he or she will not complete the elimination period before the end of the nth policy year. If an insured begins to accrue benefits at $t + e$, then when $y > 1 - (t + e)$, he or she will have reached the end of the policy year.

In practice, a mathematical expression for $s_\theta(y, n)$ may not be available. Hence, we must use tabulated values from weekly or monthly claim termination rates. Therefore, equation (10) is handled best by numerical integration. An example is provided in the appendix.

Case 2, m = n + 1 : The calculation in this case is more complex than the previous case because there are two classes of claimants disabled in the nth policy year. One class consists of those who completed the elimination period before the end of the nth policy year. This class began to accrue benefits before the $(n + 1)$st policy year began. The time of disablement for all members in this class is $t < 1 - e$. The other class consists of those with time of disablement $t > 1 - e$. Members of the latter class will not satisfy the elimination period and hence will not begin to accrue benefits until after the $(n + 1)$st policy year has begun.

For claimants who complete the elimination period before the end of the nth policy year, the benefit cash outflow in the $(n + 1)$st policy year is:

$$BC_\theta^{(1)}(n, n+1) = r_\theta(n) E_\theta(n) \int_{t=0}^{1-e} \int_{y=1-e-t}^{2-e-t} s_\theta(y, n)\, dy\, dt. \qquad (11)$$

The integrand has remained the same as in Case 1, only the limits of integration have changed. As already pointed out for Case 1, a claim beginning at time $t + e$ will have lasted for $y = 1 - (t + e)$ years by the end of the nth policy year. Note that time 0 is the start of the nth policy year, 1 represents the end of that year, 2 represents the end of the $(n+1)$st policy year, and $2-(t+e)$ is the time on claim for a claimant disabled at time t and persisting at least to the end of the $(n+1)$st year.

For those claimants disabled so late in the nth policy year they do not complete the elimination period until after the beginning of the $(n + 1)$st year, the cash flow in the $(n + 1)$st year is

$$BC_\theta^{(2)}(n, n+1) = r_\theta(n) E_\theta(n) \int_{t=1-e}^{1} \int_{y=0}^{2-e-t} s_\theta(y, n)\, dy\, dt. \qquad (12)$$

It follows that

$$BC_\theta(n, n+1) = BC_\theta^{(1)}(n, n+1) + BC_\theta^{(2)}(n, n+1). \qquad (13)$$

Case 3, m > n + 1 : This is the simplest case of the three. Given, as we have assumed, that the elimination period is not more than one year, by the end of the $(n+1)$st policy year all claimants with disablement dates in the nth policy year will have satisfied the elimination period. Thus

$$BC_\theta(n,m) = r_\theta(n)E_\theta(n)\int_{t=0}^{1}\int_{y=m-n-e-t}^{m+1-n-e-t} s_\theta(y,n)\,dy\,dt. \tag{14}$$

8 Waiver of Premium Cash Flows

Waiver of premiums can be modeled as a cash flow to the insured where the benefit is his or her premium; that is, we assume claimants pay their premiums and then immediately receive reimbursements from their insurer. Because some insurers do not pay commissions on waived premiums, the financial impact of the waiver benefit may be less than that of 100 percent reimbursement. We assume the benefit is equivalent to 100 percent reimbursement and that the waiting period for waiver is less than one year.

Calculating $W_\theta(n,m)$ is easier than calculating $BC_\theta(n,m)$. While claim payments can occur at any time during the policy year, in our model premiums can be waived only on policy anniversaries, making $W_\theta(n,n) = 0$. Thus, we only need to evaluate single integrals rather than double integrals.

Case 1, $\mathbf{m = n+1}$: $W_\theta(n,n+1)$ is the amount of premium waived at the beginning of the $(n+1)$st policy year on insureds with disablement dates in the nth policy year. Insureds disabled in the nth policy year divide into two classes, those disabled at time $t < 1-w$ and those disabled at time $t > 1-w$. If an insured's disablement occurs at time $t > 1-w$, she or he will not satisfy the waiver waiting period by the time the $(n+1)$st premium is due, but she or he will be reimbursed once the waiting period is completed. If an insured's time of disablement is $t < 1-w$ and he or she still is disabled when the $(n+1)$st premium is due, it will be waived.

First we will handle the case in which $t > 1-w$. By our assumption of uniform distribution of claim incidence, the probability that an insured exposed during the interval $1-w < t < 1$ will go on claim is $w \times r_\theta(n)$. But $r_\theta(n)$ merely gives the incidence rate of a disability lasting at least e units of time. In order to go on waiver, the insured must stay disabled an additional $w-e$ units of time. (Recall our assumption that $e \le w$.) Thus, the full probability that an insured will become disabled during the interval $[1-w, 1)$ and remain disabled long enough to satisfy the waiver of premium waiting period is $w \times r_\theta(n)s_\theta(w-e,n)$.

Those insureds who were disabled in the interval $(t, t + dt)$ with $t < 1 - w$ will go on claim at time $t + e$ and merely need to stay on claim for $1 - (t + e)$ units of time (until the start of the $(n + 1)$st policy year) to have their $(n + 1)$st premium waived. The probability this happens is $r_\theta(n)s_\theta(1 - e - t, n)$. Thus, the total premium waived on both of these types of insureds is:

$$\begin{aligned} W_\theta(n, n+1) &= P_\theta r_\theta(n) E_\theta(n)[w\, s_\theta(w - e, n) \\ &\quad + \int_{t=0}^{1-w} s_\theta(1 - e - t, n)dt]. \end{aligned} \tag{15}$$

Case 2, m > n + 1 : As w is not greater than one year, all claimants disabled in the nth policy year who still are disabled at the beginning of the mth policy year will be on waiver. These insureds will have remained on claim for $m - (n + t + e)$ years. The premium waived is:

$$W_\theta(n, m) = P_\theta r_\theta(n) E_\theta(n) \int_{t=0}^{1} s_\theta(m - n - e - t, n)dt. \tag{16}$$

In Section 2.3, we asserted that $s_\theta(t, n) = 0$ when t is greater than the length of the benefit period. The pricing actuary must bear in mind that the benefit period for waiver of premium need not be equal to the policy's base benefit period. If periods are not equal, then the survivorship function used to project $W_\theta(n, m)$ will be different from that used to project $BC_\theta(n, m)$.

Even a contract with a benefit period as short as one year typically will permit the insured to remain on waiver until he or she attains the age at which the policy expires. If a claimant still is disabled after he or she reaches the end of the benefit period, the waiver of premium provision will keep the policy in-force. The policy is still of value because he or she later may recover, return to work, resume payment of premiums, and then go on claim again. The benefit period is a limit on the amount of time the insured can collect for a single claim, not a limit on the total time the insured may collect during the life of the policy.[12]

9 Modifications to the Model

With slight modifications, the model can be adapted to situations that do not fit all of the assumptions listed in Section 2.

[12]During a claim audit, the author came across some insureds who twice had collected benefits successfully for the maximum length of time.

9.1 Nonannual Premium Payments

If premiums are paid j times per year, the units of insurance exposed to risk no longer will be constant between policy anniversaries. In this case, we need to make $E_\theta(n)$ a function of the variable t as well as the variable n. For the time being, let us ignore the effects of claim incidence and recovery. In this case, let $q_\theta^{(w)}(n)$ be the policy withdrawal rate in the nth policy year and $E_\theta^{(j)}(n,t)$ be the expected number of units in-force t $(0 \le t \le 1)$ years after the start of the nth policy year given that premiums are paid j times per year. It easily is seen that

$$E_\theta^{(j)}(n,t) = E_\theta^{(j)}(n,0) \times \left(1 - \frac{[j \times t]}{j} q_\theta^{(w)}(n)\right), \tag{17}$$

assuming lapses occur at the time of premium payment. The $[y]$ notation refers to the greatest integer less than or equal to y.

9.2 Claim Incidence and Recovery

Insured lives constantly are migrating between the population exposed to risk for being on claim and the actual population on claim. In Section 2 and in equation (17) we chose to ignore this continual decrement and increment under the assumption that in any given policy year the population on claim is small relative to the active population. If this assumption is relaxed, however, the expected exposure is given by the following:

$$E_\theta^{(j)}(n, \frac{k}{j} + t) = (1 - t\, r_\theta(n))E_\theta^{(j)}(n, \frac{k}{j}) \quad \text{for } 0 \le t < 1/j \tag{18}$$

$$E_\theta^{(j)}(n, \frac{k+1}{j}) = (1 - \frac{1}{j}(r_\theta(n) + q_\theta^{(w)}(n)))E_\theta^{(j)}(n, \frac{k}{j}) \tag{19}$$

for $k = 1, 2, \ldots, (j-1)$. Again, we have assumed the uniform distribution of claims hypothesis.

Tracking the inflow of insurance units as claimants recover and return to work is more complex. In order to track this inflow continuously, we need an aggregate rate of claim recovery. Remember that the claim population consists of distinct cohorts of individuals who were disabled at different attained ages and at different policy durations. The rate at which a cohort recovers is a strong function of how long that cohort has been on claim. In general, the longer an insured has been on claim, the less likely he or she is to recover in the near future.

The aggregate inflow of recovering insureds is a mixture of lives from cohorts that are recovering at different rates.

To simplify matters, we will assume that premiums are paid once per year, i.e., $j = 1$. In addition, we assume that the net inflows and outflows over one year's time from policyholder mortality, claim incidence, and claim recovery are so small that we can wait until the end of the policy year to count them. So, instead of tracking the two way migration continuously, we only need to do so on policy anniversaries.

Define $D_\theta(n, m)$ to be the number of units of insurance disabled in the nth policy year and still disabled by the beginning of the mth policy year.

Case 1, $\mathbf{m > n + 1}$: An insured disabled between time $n + t$ and $n + t + dt$ must have been on claim at least $m - n - e - t$ units of time in order to be disabled at the beginning of the mth policy year. Therefore:

$$D_\theta(n, m) = r_\theta(n)E_\theta(n) \int_{t=0}^{1} s_\theta(m - n - e - t, n)dt. \tag{20}$$

Case 2, $\mathbf{m = n + 1}$: This case is different because some disabled insureds may not have completed the elimination period by the time the mth policy year has begun. In this case the value of $D_\theta(n, m)$ is:

$$D_\theta(n, m) = r_\theta(n)E_\theta(n)[e + \int_{t=0}^{1-e} s_\theta(1 - e - t, n)dt]. \tag{21}$$

The first term in equation (21) accounts for those insureds who become disabled during the interval $(1 - e, 1)$ in the nth policy year. The net change in exposure due to incurrals/recoveries is the sum of the change in the number of units on claim and the number of claims terminated in the preceding year by death.

9.3 Residual Disability

Up to this point we have assumed that a unit of insurance on claim is paid at the rate of \$1 per year. If the insured is on residual[13] disability, this may not be the case. Many disability income contracts sold today provide a residual disability benefit. Under such contracts, an insured can collect some fraction of his or her full benefit if a disability causes the individual to lose a portion of income but does not completely remove him or her from the work force. If an insured suffers a 60 percent loss of income because a disability renders him or her able to work only

[13]In Goldman (1990) the term *residual* is used in a different sense. Group and individual terminology are not always equivalent.

part time or unable to do all of the duties of the occupation, he or she can collect 60 percent of the policy's full benefit. The benefit is a function of income lost, not amount of time unable to work. It is this fact that distinguishes the residual disability benefit from a partial disability benefit. A partial benefit is based on time lost rather than income lost. An insured who can put in a full day's work but loses income because he or she cannot perform certain key duties could qualify for a residual benefit.

To deal with residual disability benefits, we define a residual benefit function, $\rho_\theta(y, n)$ to be the fraction of the total benefits paid y years from now to persons who were on residual disability in policy year n. For example, suppose 75 percent of all claimants begin their claims on total disability and the remaining 25 percent begin on 50 percent of the base benefit, then

$$\rho_\theta(0, n) = 0.75 \times 1 + 0.25 \times 0.5.$$

Once $\rho_\theta(y, n)$ is known, we compute $BC_\theta(n, m)$ by multiplying $s_\theta(y, n)$ and $\rho_\theta(y, n)$ in the integrands. For example, in the case of residual disability, equation (14) becomes:

$$BC_\theta(n, m) = r_\theta(n)E_\theta(n) \int_{t=0}^{1} \int_{y=m-n-e-t}^{m+1-n-e-t} \rho_\theta(y, n) s_\theta(y, n)\, dy\, dt.$$

10 Analysis of Alternate Method

As mentioned earlier, Bluhm and Koppel (1988, Chapter 4, pp. 83-88) present a sample asset share calculation that purports to calculate benefit cash flows using claim costs and statutory claim reserve changes. We will point out deficiencies in this method and indicate how to correct the method. In addition, we will compare the Bluhm and Koppel method and the method presented in this paper.

Recall the notation used in Section 5.2. The equation used by Bluhm and Koppel can be restated as

$$BEN_n = CC_n - (CR_n^{(e)} - CR_n^{(b)}). \tag{22}$$

There are two problems with this equation. First, if the CR terms are the statutory reserve, then the equation is not based on assumptions that purport to be realistic. Statutory assumptions are supposed to be more conservative than realistic assumptions in order to ensure reserves contain a safety margin. For equation (22) to be correct, the CR

terms must denote the natural claim reserve, a reserve based on realistic assumptions.[14]

The second problem is that the change in claim reserve needs to be adjusted for interest. Equation (22) now will be corrected. Let $R_\theta(n, m)$ be the natural claim reserve at the beginning of the mth policy year on insureds disabled in the nth policy year, let $CC_\theta(m)$ be the claim cost incurred in the mth policy year, and let i be a realistic rate of interest. We will calculate $BC_\theta(n, m)$, the benefit cash flow in the mth year to insureds disabled in the nth year.

Case 1, n < m : Consider $R_\theta(n, m)$ to be the current balance of a fund established to pay benefits to those insureds disabled in the nth policy year. Withdrawals are made to pay benefits. Interest is added, but no other deposits are made because the original balance of the fund exactly matches the present value of the benefits to be paid to the cohort of claimants. Then:

$$R_\theta(n, m+1) = (1+i)R_\theta(n, m) - (1+i)^{1/2}BC_\theta(n, m).$$

We accumulate the reserve at the start of the year for a full year of interest and then subtract the money withdrawn to pay benefits (accounting also for the half year of interest lost when the withdrawal is made). The result is the fund balance at the end of the year. Rearranging this equation yields:

$$BC_\theta(n, m) = -(1+i)^{1/2}[\nu R_\theta(n, m+1) - R_\theta(n, m)]. \tag{23}$$

Case 2, n = m : At the start of the mth policy year no one could have been disabled in the mth policy year. Thus, the fund balance on this empty cohort is zero. A deposit must be made to the fund when the cohort is established. That deposit is the claim cost. This gives

$$R_\theta(m, m+1) = (1+i)^{1/2}[CC_\theta(m) - BC_\theta(m, m)]$$

which yields

$$BC_\theta(m, m) = -[\nu^{1/2}R_\theta(m, m+1) - CC_\theta(m)]. \tag{24}$$

[14]This criticism does not strictly apply to the sample calculation presented by Bluhm and Koppel because in their example the claim reserves are calculated using pricing assumptions. For claims less than two years old, the valuation actuary is allowed to measure claim reserves using experience assumptions in place of the statutory valuation table. The Bluhm and Koppel example is for a disability income policy with a benefit period of two years. For benefit periods longer than two years, if the pricing actuary wishes to use this method, he or she will be obliged to calculate two sets of reserves: one realistic, the other statutory.

Summing the $BC_\theta(n,m)$ terms from $n = 1$ to $n = m$ yields the correct expression for BEN_m (as opposed to the expression in equation (22)), i.e.,

$$BEN_m = CC_\theta(m) - (1+i)^{1/2}(\nu CR_n^{(e)} - CR_n^{(b)}). \tag{25}$$

Equation (25) may appear to be a lot easier to evaluate than those equations with double integrals in Section 7. This is not necessarily the case. To use the method in Section 7, the pricing actuary must calculate a two dimensional array $BC_\theta(n,m)]$. To project statutory book profit the pricing actuary also must project statutory claim reserves, among other things. To use equation (25), the pricing actuary must calculate a two dimensional array $R_\theta(n,m)]$ of natural claim reserves as well as an array of claim costs. If statutory claim reserves are different from natural reserves, the pricing actuary must project statutory claim reserves separately. The number of quantities to be calculated using equation (25) is larger, not smaller.

References

Atkinson, D.B. "Introduction to Pricing and Asset Shares." Study Note 210-25-90. Schaumburg, Ill.: Society of Actuaries, 1990.

Bartleson, E.L. *Health Insurance Provided Through Individual Policies.* Chicago, Ill.: Society of Actuaries, 1968.

Bluhm, W.F. and Koppel, S. "Individual Health Insurance Premiums." In *Individual Health Insurance,* (edited by F.T. O'Grady). Schaumburg, Ill.: Society of Actuaries, 1988: 57-91.

Conning & Company. *Non-Cancellable Disability Income Insurance, Decision Time.* Strategic Study Series, 1993.

Goldman, R. "Pricing and Underwriting of Group Disability Income Coverages." *Transactions of the Society of Actuaries* XLII (1990): 171-215.

Gregorius, F.K. "Disability Income Insurance in the Netherlands." *Insurance: Mathematics and Economics* (1993): 100-116.

Halpern, E. "An Approach to the Pricing and Valuation of a Social Insurance Supplement to Disability Income Policies." *Transactions of the Society of Actuaries* XXI (1979): 533-542.

Hertzman, E. "Summary of Talk on 'Work of the CMI PHI Sub-Committee'." *Insurance: Mathematics and Economics* (1993): 117-122.

Kidwell, W.D. "Individual Disability Income Benefits." In *Individual Health Insurance,* (edited by F.T. O'Grady). Schaumburg, Ill.: Society of Actuaries, 1988: 35-55.

London, D. *Survival Models and Their Estimation,* 2nd ed. Winsted, Conn.: ACTEX Publications, 1988.

Mackay, G. "Permanent Health Insurance: Overviews and Market Conditions in the U.K." *Insurance: Mathematics and Economics* (1993): 123-130.

O'Grady, F.T. (ed.) *Individual Health Insurance.* Schaumburg, Ill.: Society of Actuaries, 1988.

Pharr, J.B. "The Individual Accident and Health Loss Ratio Dilemma." *Transactions of the Society of Actuaries* XXXI (1979): 373-406.

Rejda, G.E. *Social Insurance & Economic Security,* 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1984.

Segerer, G. "The Actuarial Treatment of Disability Risk in Germany, Austria, and Switzerland." *Insurance: Mathematics and Economics* (1993): 131-140.

Shapland, R. "Reserves and Liabilities for Individual Health Insurance." In *Individual Health Insurance,* (edited by F.T. O'Grady). Schaumburg, Ill.: Society of Actuaries, 1988: 93-115.

Soule, C.E. "The Disability Income Industry Prepares for the Next Millennium." *Journal of the American Society of CLU & ChFC* (September 1993): 40-48.

Appendix

The following example will demonstrate how the formulae can be evaluated to estimate cash flows. We project the amount paid in benefits in the first policy year and the amount of premium waived at the beginning of the second policy year. $E_\theta(1) = 1000$ units of insurance each with \$100 per month of benefits (\$100,000 per month in-force), $e = 1/12$ (30 day elimination period), $w = 1/4$ (90 day wait to qualify for waiver of premium), $r_\theta(1) = 0.03$, and tabulated values of $s_\theta(y, 1)$ are given below:

Table A1

Data on s_θ

y	$s_\theta(y,1)$	y	$s_\theta(y,1)$
0/12	1.00	6/12	0.38
1/12	0.80	7/12	0.36
2/12	0.66	8/12	0.34
3/12	0.54	9/12	0.33
4/12	0.44	10/12	0.32
5/12	0.40	11/12	0.31
6/12	0.38	12/12	0.30

We can use equation (10) and a repeated trapezoidal rule (with monthly intervals) to calculate $BC_\theta(1, 1)$.

$$BC_\theta(1,1) = \frac{1,200}{144} r_\theta(1) E_\theta(1) \sum_{k=0}^{11-12e} \sum_{j=0}^{11-12e-k} \frac{1}{2}[s_\theta(\frac{j}{12}, 1) + s_\theta(\frac{j+1}{12}, 1)].$$

From the data given, $BC_\theta(1, 1) = \$9,255$.

Next we will estimate $W_\theta(1, 2)$, the amount of premium waived at the start of the second policy year. Assume the annual premium rate, P_θ, is \$10 per unit. From equation (15),

$$\begin{aligned} W_\theta(1,2) &= P_\theta r_\theta(1) E_\theta(1)[w s_\theta(w - e, 1) \\ &\quad + \frac{1}{12} \sum_{j=0}^{11-12w} \frac{1}{2}[s_\theta(1 - e - \frac{j}{12}, 1) + s_\theta(1 - e - \frac{j-1}{12}, 1)] \\ &= \$132.00. \end{aligned}$$