Time-Consistent Health Insurance

John H. Cochrane

Time-Consistent Health Insurance

John H. Cochrane

University of Chicago

Currently available health insurance contracts often fail to insure long-term illnesses: sick people can suffer large increases in premiums or denial of coverage. I describe insurance contracts that solve this problem. Their key feature is a severance payment. A person who is diagnosed with a long-term illness and whose premiums are increased receives a lump sum equal to the increased present value of premiums. This lump sum allows him or her to pay the higher premiums required by any insurer. People are not tied to a particular insurer or a group, and the improvement is free: insurance companies can operate at zero economic profits, and consumers can pay exactly the same premium they do with standard contracts.

I. Introduction

Currently available health insurance contracts do not fully insure many long-term illnesses, such as AIDS, cancer, senile dementia, heart disease, or organ failure. Many people who get such diseases face ruinous increases in premiums. Others lose their health insurance by losing their job or their spouse or by exceeding a lifetime cap on benefits. Some are bankrupted by health expenses; some are unable to get further medical care. Many other kinds of insurance do not cover long-term risks in the same way, for example medical malpractice and product liability insurance.

The absence of effective long-term health insurance is perhaps *the* most important issue driving health care regulation proposals, including the Clinton plan and congressional proposals. To the voting, insured, middle class, the plans offer "health insurance that *can't be taken away.*" This absence is also a central motivation for plans offered by academics (see Enthoven and Kronick 1989, pp. 34–35; Himmelstein et al. 1989; Pauly et al. 1991, pp. 14–15; and esp. Diamond 1992, pp. 1238–39).

Neither the authors of health regulation plans nor their critics have focused on the standard question for a proposed regulation: What, precisely, is the market failure? Or is the absence of effective long-term health insurance due instead to regulation or poor court enforcement and ex post reinterpretation of long-term contracts? Plan authors typically just assume that markets cannot provide long-term health insurance, or make anecdotal reference to textbook asymmetric information stories. Critics have focused on the distortionary side effects of the plans and their financing provisions (see, e.g., the *Journal of Economic Perspectives* "Health Care Reform" symposium; Newhouse 1994; Cutler 1994; Aaron 1994; Pauly 1994; Zeckhauser 1994; Diamond 1994; Poterba 1994). But these and other critics have not answered the planners' central challenge: How can long-term insurance be provided *without* intrusive and distorting regulation?

Here, I try to answer this question. I show why current health insurance contracts cannot provide long-term insurance in a competitive environment. I describe *time-consistent* contracts that can provide long-term insurance, and I discuss how time-consistent contracts might be implemented in practice. I anticipate some objections, and I offer an explanation for the fact that they have not already been implemented. The bottom line of the analysis is that markets can provide long-term health insurance, and *de*regulation is the likely policy route to achieve it.

II. Overview

A long-term standard insurance contract should provide insurance for long-term illness. In its simplest form, the consumer agrees to pay a constant premium and the insurer promises to pay health expenses, cross-subsidizing the expenses of those who turn out to be sick from the premiums of those who turn out to be healthy. (For the purposes of this paper, it does not matter whether the "consumer" is an individual or an employer-based or other group; whether the consumer pays for insurance directly or whether the employer does so on his or her behalf; or whether the "insurer" is a health insurance company or an insurer-provider such as a health maintenance organization.)

But suppose that the consumer gets a long-term illness. He is now a long-term liability of the insurer, so the insurer has a strong incentive to get rid of him. Current contracts are not in fact long-term contracts because the insurer can respond to this incentive. It can increase an individual's premiums or deny a renewal of the contract. Devices such as lifetime caps on health expenditures and pre-existing conditions clauses further limit coverage of long-term illness.

It is tempting to simply shore up insurers' obligations in long-term contracts: outlaw premium increases and pre-existing conditions clauses, mandate renewability, and so forth. This change will not solve the problem, however, because *consumers* cannot be held to long-term contracts. Suppose that a consumer turns out to be healthier than average. This consumer now owes the insurer a long-term stream of net payments: the prearranged premium is higher than his expected health costs. The insurer needs to bind him to the contract to pay the expenses of the sick, but the courts will not and arguably cannot force healthy consumers to stay with the original insurer forever, or pay damages for leaving. A competing insurer can woo the healthy consumer away at a lower premium, so the original insurer is left with the lemons. The original insurer is then forced to limit coverage of the sick or it will go bankrupt.

Furthermore, long-term contracts require *lifetime* ties in order to insure long-term illness. Consumers cannot change insurers, even for reasons unrelated to health, such as a move, marriage or divorce, job change, retirement, or changing preferences over quality and convenience of care. If sick, a consumer depends on the lifetime commitment of *one* insurer, since no other will take him. If well, he must be bound to his insurer, to cross-subsidize the sick.

For these and other reasons, the health insurance literature recognizes that long-term contracts are poor vehicles for insuring long-term health risks (see Diamond 1992, pp. 1238–39). Contract theory also recognizes the defects of long-term contracts. Fortunately, it finds that in many situations (including some with moral hazard), a long-term contract can be replaced by a sequence of carefully crafted short-term contracts (see Malcomson and Spinnewyn 1988; Fudenberg, Holmstrom, and Milgrom 1990; Rey and Salanie 1990). The sequence of short-term contracts must be *time-consistent,* or *renegotiation-proof.* They must satisfy a *participation constraint:* each party must be willing to sign the next contract, no matter what happens. (Kocherlakota [1994] shows how participation constraints can result in suboptimal contracts.) A sequence of short-term contracts with this feature constitutes a *self-enforcing,* long-term contract. Typically, Pareto-optimal, time-consistent contracts require a series of state-contingent *severance payments*. The methods in this paper are taken from finance,

where replacing long-term contingent claims with dynamic trading in short-term securities is a fundamental technique.

For example, suppose that the long-term debts between insurer and consumer are periodically settled or marked to market. If the consumer has gotten sick, the insurer pays him the increased present value of his expected lifetime health costs and is now free to charge an actuarially fair premium. If the consumer has gotten healthier, he pays the insurer the decreased present value of lifetime health costs and is free to leave or demand a lower premium. Now, both sides are happy to sign a new contract since the premiums are actuarially fair. Whether the consumer stays or changes to a new insurer, long-term illness is insured, since the severance payment exactly compensates for changes in premiums.

Since long-term debts are periodically settled, the contract is time-consistent and can provide long-term health insurance. It implements the Pareto-optimal or contingent-claim outcome. Furthermore, consumers do not depend on the long-term commitment of a single insurer, they are not stuck in jobs, and they do not face termination of insurance or disastrous rises in out-of-pocket expenses if they lose their jobs, get divorced, move, or change insurers for any reason. This freedom to change insurers should enhance competition and product variety. Finally, the improvement is free: the consumer's total payments are exactly the same as in an enforced long-term contract. In the place of a cross-subsidy to the sick, the healthy consumer now pays an actuarially fair premium against the chance that *he* gets a long-term illness, and hence a severance payment.

This simple implementation is not practical since consumers cannot be forced to pay insurers ex post if they do not get sick. This difficulty can be avoided if each consumer has a special account that can be used only to pay health insurance premiums and pay or receive severance payments. Every period, the consumer pays a constant amount into the account, and the account pays a premium to an insurer for one-period insurance. Competition requires that sick people pay higher premiums and healthy people pay lower premiums. If a person is diagnosed with a disease that raises his premiums, the insurer pays into the account a lump sum equal to the increase in the present value of premiums. If he gets healthier so that his premiums decline, the account pays the insurer a lump sum equal to the decline in the present value of premiums. The arithmetic, presented below, shows that there is always enough in the account to make any required severance payment.

The account may be used only for health insurance payments because, as long as the sum paid *by* the insurer when a consumer got

sick is located in the account, it is easy to require that the consumer pay the lump sum *back* to the insurer if he gets healthier. If the lump sum were paid directly to the consumer, it might be hard to get it back. The consumer might spend the money and declare bankruptcy. Finally, one hopes that courts will enforce an insurer's right to receive payments from an account that is explicitly set up for that purpose, while they may not enforce severance payments taken directly from consumers.

The time-consistent contract provides "premium insurance"—insurance against rises in premiums—as well as "health insurance"—insurance against the uncertain component of one-period health expenditures. Premium insurance does not have to be provided by the health insurer; financial services companies could offer insurance against the event that a person's health insurer raises his premiums. Therefore, time-consistent insurance can be offered by simply adding such premium insurance contracts to existing health insurance.

There are many additional ways to implement time-consistent contracts. If insurers are successfully forbidden from raising premiums or limiting coverage for the sick, severance payments could happen only when a consumer decides to change insurers. If the special accounts I described above are unworkable, the contract could state that the current insurer will pay a new insurer to take the consumer if he is sick, or have the right to receive a payment from a new insurer if the consumer is healthier than average. Contracting costs and the vagaries of court enforcement will determine which of these or other implementations of the time-consistent contract are chosen. The essence is just some enforceable mechanism for settling long-term debts.

By contrast, most policy proposals herd consumers into large pools, outlaw health-based premiums and pre-existing conditions clauses, and attempt to outlaw selection based on health. These proposals require a heavily regulated system that enforces a uniform product and eliminates competition, because economic incentives to select are not eliminated. Insurers must be effectively prevented from subtly and cleverly trying to improve their pool of customers or from discreetly providing lower levels of care for sick and expensive consumers. Worse, they must be prevented from competing for the healthy. If they just focus their marketing and advertising to healthier groups, the pooling solution will break down. Similarly, healthy consumers must be stopped from trying to join better groups.

The time-consistent contract described above most closely resembles proposals to allow medical savings accounts. But the accounts in

current proposals do not provide *insurance;* each person's lifetime resources are still reduced one-for-one by his lifetime health expenses. Time-consistent contracts add insurance to medical savings accounts. Money is added to the accounts of people who get sick and is drawn from the accounts of those who stay well or get better.

III. Optimal Health Insurance Contracts

Figure 1 presents the timing of events. In the beginning of each period, wealth W_t is evaluated, premium payments p_t can be made, and the consumer earns income e. Then information about the consumer's health is revealed, including current health costs x_t as well as information about future health costs. State-contingent payments y_t can be made, health costs x_t are paid, and finally the consumer consumes c_t. The consumer earns interest $1 + r$ between periods. The term E_t denotes the expectation conditional on time t information before x_t is revealed, and E_{t+1} refers to time t information after x_t and any other news are revealed. Therefore, an expression such as $E_t x_t$ means the expected value of x_t in the first half of period t, before x_t is revealed.

Consumers maximize a standard intertemporal utility function:

$$\max E_0 \sum_{t=0}^{\infty} \beta^t u(c_t). \tag{1}$$

For simplicity, the discount factor β equals $1/(1 + r)$. Different values or varying interest rates complicate the algebra without changing the basic point. I treat uncertain lifetimes below. Since budget constraints are linear, we can separately treat health insurance and insurance against other shocks. For this reason, I simplify the model to a constant labor income e.

I assume that insurers are risk neutral and competitive, and can borrow or lend at the interest rate r. Risk neutrality follows when individual illness is a perfectly diversifiable risk. Section V argues that the results are not substantially altered with imperfect credit markets or nondiversifiable risks. For simplicity, I also focus on contracts with

Time t Time $t + 1$

| Evaluate W_t | Earn e pay p_t | Learn x_t | Pay $x_t, y_t,$ consume c_t | earn $1 + r$ |

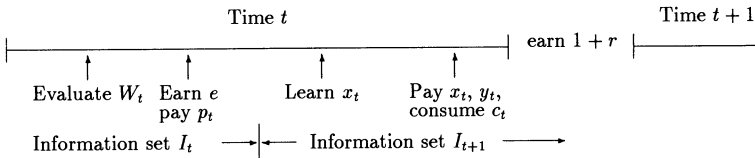Information set I_t ⟶ ⟵ Information set I_{t+1} ⟶

FIG. 1.—Timing of events

zero economic profits. Larger profits or administrative expenses are straightforward extensions.

Pareto-optimal allocations maximize the consumer's objective for each value of the insurer's objective. In this model, Pareto-optimal allocations give constant consumption streams: given the present value of net payments to consumers, the insurer cares nothing about rearranging payments to provide a constant consumption stream. Given a constant consumption stream and concave utility, there is no way to increase the consumer's utility without raising the present value of the consumption stream.

We can most easily derive the Pareto-optimal allocation by finding the allocation that results from complete contingent-claims markets. At time 0, the consumer sells claims to his income stream and buys contingent claims to cover health expenses and consumption. Since insurers are risk neutral (aggregate marginal utility is independent of an individual's health in the underlying general equilibrium), the time 0 value of contingent claims equals their discounted expected present value. Thus the consumer's time 0 budget constraint in a contingent-claim market is

$$E_0 \sum_{t=0}^{\infty} \beta^t c_t = W_0 + E_0 \sum_{t=0}^{\infty} \beta^t (e - x_t).$$

The first-order conditions to the consumer's optimization problem—maximize utility (1) subject to this constraint—specify a constant consumption level c at every date and in every state. Solving the budget constraint with constant consumption, we find that

$$c = r\beta W_0 + r\beta E_0 \sum_{t=0}^{\infty} \beta^t (e - x_t).$$

Time 0 contingent-claims contracts are (among other impracticalities) not *time-consistent;* they do not satisfy a *participation constraint.* As soon as health status is revealed, healthy consumers and the insurers of sick consumers will withdraw. Since both parties are free to abandon the contract at the end of any period, a time-consistent contract must be equivalent to a series of one-period contracts. Therefore, we search for a market structure of one-period contracts that implements the Pareto-optimal or contingent-claims allocation.

We need only two contracts or securities, one-period insurance and riskless period-to-period saving (bank accounts). For insurance, the consumer pays a premium p_t in the first part of the period (see fig. 1). The insurer then pays this period's health expenses x_t plus a potentially state-contingent severance payment y_t, whose value is deter-

mined below. This severance payment is the key to the paper and the innovation that allows one-period contracts to insure lifetime health expenses.

Since insurers are risk neutral and competitive, the insurance premium must equal the expected value of payments:

$$p_t = E_t(x_t + y_t). \tag{2}$$

The insurer must expect to make zero profits from each consumer, period by period. Sick consumers must pay higher premiums or insurers will try to get rid of them; healthy consumers must pay lower premiums or other insurers can woo them away. Competitive insurers cannot cross-subsidize in the absence of two-sided commitment to long-term contracts.

Now we determine the severance payment y_t. After all payments are made, consumption equals the time t present value of resources:[1]

$$c_t = r\beta\left(W_t - p_t + y_t - E_{t+1}\sum_{j=1}^{\infty}\beta^j x_{t+j}\right) + e. \tag{3}$$

If consumption is to be constant, we must have $c_t = E_t c_t$. If we take expectations of equation (3), the unexpected severance payment must equal the innovation in the present value of health expenses:

$$y_t - E_t(y_t) = (E_{t+1} - E_t)\sum_{j=1}^{\infty}\beta^j x_{t+j}. \tag{4}$$

The market for one-period loans means that the consumer's intertemporal budget constraint is

$$W_{t+1} = (1 + r)(W_t + e + y_t - p_t - c_t). \tag{5}$$

One can verify that consumption is constant over time by combining this equation with the consumption decision rule, equation (3), and the value of severance payments y_t, equation (4).

The one-period, zero-profit condition, equation (2), and the full-insurance condition, equation (4), do not uniquely determine the optimal contract. If we add one dollar to the premium, p_t, and one dollar to the payment, y_t, neither condition is affected. Therefore, all Pareto-optimal (fully insuring), time-consistent, zero-profit contracts have the form

$$p_t = E_t(x_t) + b_t \tag{6}$$

---

[1] Precisely, eq. (3) gives consumption if contingent-claim markets are opened in the second half of period t. Since we are implementing a contingent-claim outcome, this expression must also give consumption in our restricted market structure.

and

$$y_t = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \beta^j x_{t+j} + b_t, \qquad (7)$$

where b_t is an arbitrary amount in the time t information set. The quantity b_t can be thought of as a bond. The consumer pays a premium equal to one-period expected health costs, $E_t(x_t)$, plus the bond; the bond is then returned along with the severance payment.

The choice $b_t = 0$ gives the simplest contract. Each period's premium equals expected health care costs in that period:

$$p_t = E_t(x_t). \qquad (8)$$

The severance payment y_t is simply the revision in the present value of health expenses, by equation (7):

$$y_t = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \beta^j x_{t+j} = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \beta^j p_{t+j}. \qquad (9)$$

Armed with this severance payment, a sick consumer can pay higher premiums with no change to his consumption stream. The second equality, derived from equation (8), shows that the severance payment is also the innovation in the present value of premiums. Therefore, we do not need to measure health expenses. The insurer's announced schedule of premiums provides all the information needed for the contract.

This contract requires that a consumer who gets unexpectedly *healthier* must make an ex post severance payment *to* the insurer, equal to the unexpected *decline* in his health care expenses. Such payments may be hard to collect. Even one-period contracts may not be enforceable against consumers; we may wish to impose a participation constraint that the consumer can abandon the contract at any time, even in the middle of a period.

The health account described in the Introduction solves this problem. The consumer pays an amount q_t into the account in the first part of each period. The account pays health insurance premiums, $p_t = E_t(x_t)$, and pays or receives severance payments, y_t. Therefore, health account balances A_t evolve as

$$A_{t+1} = (1 + r)(A_t + q_t - p_t + y_t). \qquad (10)$$

The consumer keeps the rest of his wealth in a savings account with balances K_t, so $W_t = K_t + A_t$. This component of wealth evolves as

$$K_{t+1} = (1 + r)(K_t + e - q_t - c_t).$$

Since income is constant, we might as well specify a constant out-of-pocket payment $q_t = q$. The account may be used only for health care or insurance payments, so the present value of out-of-pocket payments q_t must equal the present value of health expenses. Therefore, a constant payment q implies that q equals the time 0 flow present value of health expenses:

$$q = r\beta E_0 \sum_{j=0}^{\infty} \beta^j x_j. \tag{11}$$

In practice, income and health expenses typically rise through time, so it might make sense to specify a rising schedule of out-of-pocket expenses.

This system is still Pareto-optimal and time-consistent. Nothing has changed; we have just split the consumer's wealth into two accounts. Now, we check that the health account balance is always nonnegative. Then the contract never requires ex post out-of-pocket expenses; if the account is bonded to the contract, the consumer can leave at any time, including in the middle of a period.

Combining (11), (10), and (9), we find the health account balance:[2]

$$A_t = (1 + r)^t A_0 + E_t \sum_{j=0}^{\infty} \beta^j x_{t+j} - E_0 \sum_{j=0}^{\infty} \beta^j x_j. \tag{12}$$

Therefore, if the consumer enters the account healthy—if the present value of health expenses at the beginning of the contract is as low as it can be—then the amount A_t in the account will always be nonnegative, even if the consumer posts no initial bond, $A_0 = 0$.

We have made no restriction on the time path of health expenses or expected health expenses. Therefore, the fact that health costs

---

[2] Equation (12) obviously holds at time 0. Supposing it holds at time t, I show that it holds at time $t + 1$. From eqq. (8), (9), and (10),

$$\beta A_{t+1} = A_t + q - x_t + (E_{t+1} - E_t) \sum_{j=1}^{\infty} \beta^j x_{t+j};$$

from eqq. (11) and (12),

$$\beta A_{t+1} = (1 + r)^t A_0 + E_t \sum_{j=0}^{\infty} \beta^j x_{t+j} - E_0 \sum_{j=0}^{\infty} \beta^j x_j$$

$$+ r\beta E_0 \sum_{j=0}^{\infty} \beta^j x_j + \beta E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+1+j} - E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}$$

$$= (1 + r)^t A_0 - \beta E_0 \sum_{j=0}^{\infty} \beta^j x_j + \beta E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+1+j}.$$

Canceling β and rearranging, we get eq. (12).

typically rise with age does not matter. If an old consumer reverts to perfect health, there is a large change in the present value of health expenses, but by then the account has built up a large balance.

A larger apparent difficulty results from the possibility of death. In our setup, death is a state of perfect health, since health expenditures will be zero forever after. As death approaches, the present value of health expenses declines. But at death, equation (12) seems to specify a large negative account balance.

However, consumption and income should also be zero after death, not constant as specified by the model so far. The Appendix presents a generalized version of the model that includes the possibility of death. The consumer chooses consumption that is constant if he is alive and zero if he is dead. A time-consistent, Pareto-optimal sequence of one-period contracts can again implement the optimum. The severance payment now includes the market value of an annuity as well as the market value of lifetime health expenses.

With the possibility of death, the health account balance of an alive consumer who starts with $A_0 = 0$ generalizes from equation (12) to

$$
\frac{A_t}{E_t \sum\limits_{j=0}^{\infty} \beta^j a_{t+j}} = \frac{E_t \sum\limits_{j=0}^{\infty} \beta^j x_{t+j}}{E_t \sum\limits_{j=0}^{\infty} \beta^j a_{t+j}} - \frac{E_0 \sum\limits_{j=0}^{\infty} \beta^j x_j}{E_0 \sum\limits_{j=0}^{\infty} \beta^j a_j}, \tag{13}
$$

where $a_t = 1$ if the consumer is alive at time t and $a_t = 0$ if he is dead at time t. The expression $E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}$ is the value of a one-dollar annuity and captures the changing probability of death.

This formula verifies that declining present values of health expenses due to higher probabilities of death do *not* trigger out-of-pocket payments or negative balances in the health account. For example, if health expenses are a constant x when alive, but their present value can change with changing probabilities of death, then equation (13) simplifies to

$$
\frac{A_t}{E_t \sum\limits_{j=0}^{\infty} \beta^j a_{t+j}} = \frac{x E_t \sum\limits_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum\limits_{j=0}^{\infty} \beta^j a_{t+j}} - \frac{x E_0 \sum\limits_{j=0}^{\infty} \beta^j a_j}{E_0 \sum\limits_{j=0}^{\infty} \beta^j a_j} = 0.
$$

The account has a constant balance for *any* probability of death.

The contract with an account amounts to a choice of a bond b_t in the setup of equations (6)–(7). If we drop the distinction between the

consumer and the account, the consumer pays a premium equal to current wealth in the account plus q. This amount is equal to current expected health expenses plus a large bond, b_t:

$$p_t = q + A_t = E_t(x_t) + b_t. \tag{14}$$

This bond is so large that the severance payment, y_t, always goes from insurer to consumer. If there is no change in health, the entire amount b_t is returned, to be posted as a bond again the next period. In practice, there is no point in having an actual payment move back and forth every period. *Rights* to the account are exchanged between insurer and consumer each period instead.

A time-consistent contract does not need to have severance payments every period. For example, the contract could specify a constant payment q per period, and the insurer pays health cost x_t. Both sides have the *right* to have the contract marked to market and an account created at any time, but will typically do so only rarely.

IV. Costs and Comparison with Other Contracts

One might think that time-consistent contracts are more expensive than other contracts. This turns out not to be true. Time-consistent contracts require no increase in payments relative to enforced standard contracts or the guaranteed renewable contracts described by Pauly, Kunreuther, and Hirth (1992).

Standard Contracts

The standard contract has no severance payment. There is a constant premium, $\bar{p}$. For firms to make zero profits, the present value of premiums must equal the present value of health care expenditures. Hence,

$$\bar{p} = r\beta E_0 \sum_{j=0}^{\infty} \beta^j x_j.$$

A standard long-term contract, if enforced on both parties, *is* Pareto-optimal (in the absence of product variety and competition considerations). Since the payments are constant, the consumption stream is constant. However, unless illness is entirely transitory—if $E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}$ can differ from $E_0 \sum_{j=0}^{\infty} \beta^j x_j$—the standard contract does not give time t zero expected profits and so is not time-consistent.

The health account contract described above was set up so that the payment each period is $\bar{p}$. Thus it obviously has the same cost as the

standard contract. More generally, *every* time-consistent, Pareto-optimal contract has the same cost, ex post, as the standard contract. It must. The essence of insurance is that ex post wealth does not depend on losses.

This statement can be verified as follows. Starting with (6)–(7), we have

$$\beta^j(p_j - y_j) = \beta^j E_j(x_j) - (E_{j+1} - E_j) \sum_{k=1}^{\infty} \beta^{j+k} x_{j+k}$$

$$= E_j \sum_{k=j}^{\infty} \beta^k x_k - E_{j+1} \sum_{k=j+1}^{\infty} \beta^k x_k.$$

Summing over j, we obtain

$$\sum_{j=0}^{\infty} \beta^j(p_j - y_j) = E_0 \sum_{j=0}^{\infty} \beta^j x_j = \frac{\bar{p}}{r\beta}.$$

The left-most expression is the ex post present value of payments in a time-consistent contract. It equals the time 0 expected present value of health expenses and the present value of standard contract premiums.

Of course, a fully insuring contract is more expensive than a zero-profit standard contract on which the insurer can default as soon as one suffers a long-term illness! Nonetheless, overall health care expenditures may or may not increase. Ex post uninsured consumers currently find some alternative sources of financing—savings, charity, or the government—rather than forgo all health care. Insured care is often thought to be cheaper than care for the uninsured, so overall expenses could decline.

Guaranteed Renewable Contracts

Pauly, Kunreuther, and Hirth (1992) advocate guaranteed renewable contracts to provide long-term insurance. In these contracts, the consumer always has the right to continued insurance at a prearranged premium. Guaranteed renewable contracts do not feature the severance payment, so all consumers of the same age must pay the same premium. To keep the healthy from defecting to a competing insurer, the premium charged to all must be the same as that charged to a healthy person. Therefore, consumers must prepay the expected value of the rise in premiums that will occur if they become sick.

Sick consumers must depend on and enforce the long-term commitment of their current insurers in a guaranteed renewable contract,

whereas they are free to change in time-consistent contracts. And time-consistent contracts can be arranged to specify exactly the same payments as guaranteed renewable contracts. Again, the improvement is free.

Pooling

The majority of health insurance is currently provided through group plans or pools. Initially, group plans seem strange to an economist: what function do insurers serve except to form pools and diversify risk? But a system in which only pools, formed on a characteristic independent of health status, can be insured helps standard contracts to provide long-term insurance. Pools bind consumers to long-term contracts: if only a pool can be insured, ex post healthy consumers cannot defect.[3]

However, pooling provides imperfect long-term insurance in a number of ways. Healthy individuals must be prevented from obtaining individual insurance at a lower, actuarially fair, rate. The tax deduction for employer-provided insurance may help to keep healthy individuals in employer-run pools from doing so. (At last, a reason for this much-disparaged deduction!) Labor contracts in which the employer contributes to a group plan, but will not contribute to a privately chosen plan or pay higher wages to consumers who choose such plans, have the same effect. These provisions have obvious distortionary consequences.

Furthermore, pool formation and movement into and out of a pool must be based on events that are independent of health status. They are not, and this is why employer-based groups are now losing long-term insurance. The ability to get or keep a job is obviously correlated with long-term illness. Since it is illegal to vary wages with health status, firms have an incentive to select healthier workers, and firms with healthy workforces can woo healthy workers away from competitors with less healthy workforces.

Most important, the stronger and larger the pool, the less product

---

[3] It is often claimed that the prevalence of group insurance reflects instead economies of scale: higher administrative costs for servicing individuals or greater bargaining power of pools (see, e.g., Diamond 1992, p. 1234). But if this is the reason for pools, why don't individuals or third companies form pools to buy health insurance? Or why don't competitive insurance companies form the pools, i.e., sell individual insurance at the pool rate in the first place? The alternative story in the text gives an answer: pools must be formed on characteristics unrelated to health status. If one could form pools for the purpose of getting or providing health insurance, selection based on health status would start, and the system would no longer provide long-term insurance.

variety, competition, and resulting market discipline. It must be so: the *point* of the pool is to keep insurers from competing for healthy individuals or groups. Imagine the effect on the car market if everyone in a city or place of employment had to purchase exactly the same make and model of car!

Most current health regulation proposals seek to provide better long-term insurance by strengthening the pool mechanism, enlarging pools, and limiting freedom to change pools or insurers. The health alliances, community rating, employer mandate, limited options, and other features of the Clinton plan are devices to strengthen the pool mechanism. Strengthening and enlarging pools is also the essence of academics' plans, such as Enthoven and Kronick (1989) and Diamond (1992).

But the difficulties of providing long-term insurance and the resultant impetus to continue expanding the pools and limiting competition will continue. Most plans propose location-based pools, but location is not uncorrelated with health status. Retired people move to Florida and Arizona, drug users live in inner cities, and people can move or change legal residence in order to pay lower premiums and receive better care. And many long-term illnesses may still not be covered in large regulated pools, since levels of treatment will depend on administrators' ideas of cost effectiveness or lobbying by patients with specific diseases, rather than individuals' preferences for health versus other expenditures. Furthermore, as Weisbrod (1991) argues emphatically, further regulation of health care may dramatically reduce the rate of technical improvement in medicine.

In summary, the only pool that can provide complete long-term insurance is nationalized health insurance with mandated individual participation, because this is the only system that can really bind healthy consumers to a long-term contract and eliminate selection. This solution has zero competition, flexibility, and product variety. As pools are made smaller, competition and product variety increase, but the amount of long-term insurance decreases. Time-consistent contracts provide long-term insurance and allow competition at the individual level.

V. Extensions and a Few Objections

Measuring Health Expenses

Time-consistent contracts do not require a difficult and possibly contentious computation of expected health expenses. Changes in premiums can determine payments to the health account, since competition forces premiums to reveal expected health expenses. And if premi-

ums do not change, the consumer does not need a severance payment to insure long-term illness, no matter what happens to expenses.

Calculating the change in the present value of future premiums is a little harder, but not impossible, since premium schedules are public information.[4] And even a rough accounting, say a severance payment equal to 20 times the change in premiums or an annuity that pays the higher premiums each period, can provide a workable approximation and a big improvement over no insurance. Similarly, home insurance is viable and useful, even though the exact value of a home and its contents is difficult to measure.

Imperfect Credit Markets

Time-consistent contracts can be implemented without credit markets. Diversified, competitive insurers earn zero profits each period, ex post as well as ex ante. Individual consumers save, but do not borrow, in premium accounts. The net amount in these accounts is constant, so a company that provides premium accounts does not need access to capital markets either.

Start-up Problems

What about consumers who are *not* healthy when they first purchase insurance? They can sign on to the contract described above, paying the same amount per period $\bar{p}$ as everyone else, if they deposit in their account the amount that would be there if they had started the contract healthy. By equation (12), this is the current present value of their health expenses, less those of a healthy person, $E_\tau \sum_{j=0}^{\infty} \beta^j x_{\tau+j} - E_0 \sum_{j=0}^{\infty} \beta^j x_j$, where τ is the date on which the contract is signed.

It is likely that consumers can pay a much smaller deposit. Part of the account is a bond against the event that the consumer becomes healthier. If there is no chance that the consumer will revert to perfect health, there is no need for the corresponding portion of the bond. For example, if the consumer cannot get any healthier, he can pay each period the flow present value of his (higher) lifetime expenses, $\bar{p}_\tau = r\beta E_\tau \sum_{j=0}^{\infty} \beta^j x_{\tau+j}$, and deposit nothing in the account.

Alternatively, the contract can specify that the health account is simply not debited if the consumer becomes healthier than he is at the start of the contract. With this specification, the contract no longer

---

[4] See Feenberg and Skinner (1992) for a good example of dynamic health cost calculations.

provides perfect insurance. If the consumer becomes healthier than at the beginning of the contract, he will be able to consume more since his premiums decline and there is no offsetting severance payment. On the other hand, premium insurance will no longer be free. Consumers will have to be charged a higher premium to cover the missing severance payments. The contract is now an option and has positive value.

It is likely that this increase in premiums is small. In practice, most people with long-term illnesses are very unlikely to revert to perfect health, that is, become eligible for the low health insurance premiums of people with no history of disease. Therefore, the probability that they do so, times their lowered present value of health expenses, is quite small and would result in a small increase in insurance premiums.

The government may wish to provide initial accounts or subsidize higher payments per period for those born with genetic defects or poor family histories or those who are already sick when a time-consistent system starts. As in all insurance, there is a good argument that government policy should insure events that happened before contracts could be signed. Since the government is currently partially liable for the chronically ill, through Medicare, Medicaid, and other programs, the net cost may not be large. And subsidies based on long-term health status are less distorting than regulatory proposals to shore up long-term insurance.

Quality Variation

Individuals desire a large variation in quality and other attributes of health care, as with all goods. Poor people generally choose lower-quality care for lower premiums. Others vary in their trade-offs among cost, convenience, promptness of appointments, treatment by many specialists versus a single familiar general practitioner, desired level of treatment for specific conditions (professional athletes are willing to pay for much more extensive treatments of injuries than economists), willingness to suffer a restricted choice of physicians, and so forth.

Variations in quality and other attributes can be accommodated fairly easily in time-consistent contracts. The severance payment equals the change in present value of the *current* insurer's premiums. The consumer can change to a new insurer of the same quality at no out-of-pocket cost. However, if the consumer decides to change to a higher quality of insurer, he must pay the higher premiums out of pocket. Unobservable changes in preferences are not insurable.

Nondiversifiable Shocks and Technical Change

There are potentially important nondiversifiable shocks to health care expenses—events that do not average to zero over the insurer's customer base. Epidemics and natural disasters come quickly to mind, but regulatory surprises and unexpected improvements in technology are perhaps the most important nondiversifiable shocks in practice. Imperfectly diversifiable shocks do not rule out insurance. Insurance companies take risks, and large risks can be insured: hurricanes, oil spills, satellite launches, and so forth. Nondiversifiable shocks do raise some subtleties, and I examine three in turn.

Bankruptcy

Insurers may declare bankruptcy after a large negative shock. This problem can be addressed, as it is now, by requiring insurers to maintain loss reserves and capital requirements. Insurance contracts specifically exclude events so large that there is no hope of solvency, such as war. Year-to-year innovations in aggregate health expenses are not larger than other currently insured nondiversifiable risks, or risks in other industries. There is nothing peculiar about long-term illness or health care in this limited liability problem.

Risk Premiums

With nondiversifiable shocks, insurers may require somewhat more than actuarially fair premiums. Faced with such premiums, sufficiently risk-neutral consumers may choose not to insure or to partially insure. But if risk premiums are small and consumers not too risk neutral, they may choose to buy essentially full insurance anyway.

Since insurers are public companies, risk premiums are determined in financial markets. Risks that are not diversifiable in the insurance sense may well be diversifiable in the finance sense: uncorrelated with the factors such as market return that drive expected returns in capital markets. If so, insurers will still act risk-neutrally. If not, the covariance of a risk with asset market factors determines its premium. Year-to-year variation in aggregate health expenses is small compared to other risks in the economy, and not highly correlated with those risks, so nondiversifiable health risk premiums are likely to be small, or at least not much larger than risk premiums in other, thriving, insurance markets.

A quantitative example follows. Cutler (1992) analyzes nondiversifiable technological risks in nursing home care. He estimates that profit rates in nursing home insurance are on the order of five per-

centage points larger than other, more diversifiable, insurance. Five percent is about as high a figure as one can hope to defend on the grounds of nondiversifiable risk in capital markets.[5] Now, in the simplest model,[6] the consumer buys insurance until

$$\frac{\text{wealth if sick}}{\text{wealth if well}} \simeq \left(\frac{\text{probability of loss}}{\text{premium per dollar coverage}}\right)^{1/\text{risk aversion coefficient}}$$

With risk aversion of one (log utility) and a 5 percent risk premium, consumers will insure until sick wealth is only 5 percent below well wealth. If risk aversion is 10 or more, as suggested by the asset pricing literature, consumers will insure to the point at which sick wealth is only 0.5 percent less than well wealth: they will buy essentially full insurance. But nursing homes cost roughly $100 a day, or $36,500 a year. Without insurance, sick wealth is likely to be 50 percent or less of well wealth. Thus nondiversifiable risks do not account for the unpopularity of nursing home insurance. (Good alternatives are that the contracts are not time-consistent or that consumers plan to transfer or spend down assets and rely on Medicaid.) Similar numbers are likely to apply for other long-term insurance.

On reflection, it is not surprising that risk premiums and risk aversion have small effects. Insurance policies already charge as much as 50 percent loadings for reasons unrelated to nondiversifiable risk premiums, and consumers buy them. An extra 5 percent loading is unlikely to have huge effects.

---

[5] For example, the capital asset pricing model states that expected returns obey

$$E(R^i) - R^f = \beta_{im}[E(R^m) - R^f],$$

where R^i is the return on a given security or investment project, R^f is the risk-free rate, R^m is the market return, and β_{im} is the regression coefficient of R^i on R^m. The market risk premium $E(R^m) - R^f$ is about 7 percent, so a profit rate of 5 percent over the risk-free rate requires a regression coefficient of nursing home technology risk on the market return of about $5/7 = 0.7$. Cutler estimates the standard deviation of the present value of an insurance company's liability at 4–14 percent, and the standard deviation of the market return is about 17 percent. To generate a 0.7 regression coefficient, then, we need the highest standard deviation estimate and the unlikely assumption that nursing home technology risk is almost perfectly correlated with the market return ($\beta_{im} = \rho_{im}\sigma_i/\sigma_m$). If a 5 percent risk premium is taken over a normal rate of return that is higher than the Treasury-bill rate, even more extreme assumptions are required.

[6] There are two states, sick s and well w. State s occurs with probability π. There are two dates, and consumption equals wealth at the second date. A premium p pays $1.00 in the s state. The consumer's first-order condition is

$$\frac{u'(c_s)}{u'(c_w)} = \frac{p/(1-p)}{\pi/(1-\pi)}.$$

With constant relative risk aversion preferences, $u'(c) = c^{-\gamma}$, and for small p and π we can ignore the $(1-p)/(1-\pi)$ term, so this expression implies $c_s/c_w \simeq (p/\pi)^{1/\gamma}$.

Technical Change

Expected improvements in technology, like any other expected event, do not trouble insurance contracts. *Unexpected* technical change is the only potential problem.

Technology does not automatically raise costs. Almost by definition, improvements in technology imply declines in the *price* of treatment. When a cure for a previously untreatable disease is discovered, the price declines from infinity to some possibly large value. This decline in price, together with an elastic demand for medical care, can result in increased expenditures. Newhouse (1992) argues that this story explains the bulk of health care expenditure growth. When demand elasticities are low, improved technology lowers costs, such as when a drug treatment is discovered to replace surgery.

Technological innovations are basically events that trigger a changed desire for quality. Given that a heart transplant is possible, consumers are willing to pay for a policy that gives more generous payments to patients with heart disease.

Under current contracts, premiums are changed to accommodate increased use of new technology, and the consumer bears most of the risk. Time-consistent contracts can at least incorporate new technology in the same way. As with other changes in quality, the contract can state that individual customers who are reclassified in the current premium structure receive severance payments. After insurance payments have been made, the insurer can announce an across-the-board premium change to accommodate different treatments or expenditures, which does not trigger a severance payment.

In fact, time-consistent contracts may smooth the adaptation to new technology. In a long-term contract, consumers are wary that the insurer that raises premiums is trying to avoid long-term debts to individuals or groups that have become sicker than average, rather than adapt to new technology. With a time-consistent contract, consumers who do not like the new premiums and level of care, or who simply think that the insurer is trying to gouge them, can take their severance payment to another insurer that charges the original premiums and provides the original level of care.

However, we cannot insure against technological risks by keying severance payments to any changes in premiums, including those that adapt to new technology. Consumers would claim that a new technology should be used for every ache, that expenses and premiums should skyrocket, and that a huge severance payment should be made so that they can pay for the much higher level of health expenses. Insurers would claim the opposite.

The problem with changes in technology is not time-consistency or

private information. The problem is contracting costs. If contracting costs were zero, the contract could specify in advance the level of expenses for every health state, in every possible state of technology. Then consumers would choose not to pay for contracts that provide too lavish treatment when new technology is introduced.

Contracting costs are obviously not zero, and contracts contingent on undreamed-of inventions obviously impractical. The question is, To what extent can feasible contracts or institutions approximate this contingent-claim result? Many other contracts are successfully written and enforced even though every contingency cannot be spelled out. In this case, we just need some mechanism for deciding how much health expenses should adapt to unexpected changes in technology. The standard solution is appeal to a disinterested third party. Contracts could index the adoption of new technology to standards promulgated by the American Medical Association, other private organizations, independent bodies set up for the purpose, or government rules such as Medicaid reimbursement rules. Then severance payments keyed to any premium increases would provide at least some insurance against unexpected technical change.

Adverse Selection, Moral Hazard, and Participation

Private information is a standard objection to any insurance contract. Private information about health at the beginning of the contract can cause adverse selection. Private information about actions taken during the contract or about what state has actually occurred can cause moral hazard.

Selection is not "adverse selection." The stories of sick consumers who lose their insurance represent selection based on *public* information, and so are not evidence for a private-information failure of the long-term health insurance market.

Adverse selection has not been documented to cause specific failures of health insurance contracts, and the absence of long-term insurance in particular. Do individuals, armed with private information about their aches and pains, really know much more than a doctor, armed with a medical history and simple tests? The answer is not obvious. If anything, the health economics literature stresses the opposite conclusion. Pauly (1986, pp. 650–51) notes several aspects of current health insurance that are at odds with adverse selection models. Among others, the fact that insurers do not now condition on easily observable indicators of health status argues against a market right up against an information constraint, and the fact that most people *are* insured argues against a lemons model in which only the sickest get insurance. Cutler (1992, pp. 35–37) argues similarly that

private information does not account for the failure of the nursing home insurance market.

Private information about lifestyle choices that affect long-term health is also a doubtful explanation for the lack of long-term insurance. Many lifestyle choices alleged to influence long-term health risks are in fact observable: eat too much, don't exercise, and you get fat; it is easy to tell who smokes, uses intravenous drugs, and so forth. And the influence of lifestyle choices on the incidence of disease is (alas) not that great. Many long-term illnesses, such as Alzheimer's disease, are not related to any known controllable and potentially unobservable actions by the individual, so moral hazard does not explain why they should not be insured.

Private information about ex post health cannot account for the lack of long-term insurance. There is no question that someone with, say, cancer or AIDS actually has the disease and needs *some* treatment. While overuse of medical services resulting from the fact that doctors and patients know more than the insurer about the patient's health is an important issue, it should doom insurance according to the observability of an illness, not according to the persistence of an illness. Mental illness is poorly insured, precisely because its severity is hard to measure. Many long-term illnesses have clear diagnoses and narrow ranges of treatment, and yet are still not insured.

In summary, textbook private-information stories are not easy explanations for the absence of long-term health insurance, especially given that short-term health insurance does exist. At a minimum, the analysis of time-consistent contracts in a public-information setting is a good starting point for understanding how and what private information is really at the bottom of inadequate health insurance. At a maximum, private information, though the subject of an enormous and fascinating literature, may just not have much to do with this particular economic problem.

More recently, contract theory has found an explanation for imperfect insurance under perfectly symmetric information, in a participation constraint. In Kocherlakota's (1994) example, two people share a constant income that will be divided randomly and publicly between them. They can write contracts, but either side can always revert to autarky. Therefore, the optimal sharing rule must partially reward the lucky agent, to keep his utility above what he could get by eating the lucky draw and withdrawing from the contract. The result is imperfect insurance despite complete information. I have argued above that it is exactly the consumer's inability to commit not to defect to a competing insurer if he turns out to be healthy—a state observed by both sides—that makes the standard contract unravel. Therefore, the branch of contract theory that studies imperfect commitment or

participation constraints under symmetric information seems a much more useful parable for the failures of long-term health insurance. In these models, the Pareto optimum is again achievable if one side can be forced to honor state-contingent severance payments. This paper presumes that insurers can be held to one-period contracts and shows how such payments can be arranged.

Why Don't We Observe Them Already?

If better contracts are available, why have they not already been instituted? Imperfect credit markets, measurement of health status, variation in quality, nondiversifiable shocks, technical change, risk aversion, and private information do not pose insurmountable obstacles to the implementation of time-consistent contracts, so none of these factors explains why we do not already see such contracts. In addition, time-consistent contracts do not seem outlandishly costly to write.

Contracts may simply have adapted imperfectly to changed medical and economic circumstances. A generation ago, health expenses were mostly temporary. Health care was largely devoted to treating injuries and some infectious diseases. Other illnesses, such as cancer, usually led to rapid and inexpensive death or to chronic but untreatable and hence inexpensive conditions. Only recently has technology changed to allow long-term, expensive treatment of persistent illnesses.

Furthermore, health insurance was less competitive a generation ago. Large insurers could cross-subsidize sick customers from the premiums of healthy customers. As competition increased, new entrants vigorously searched for healthy customers, leaving the older insurers with only the sick or old. Insurers were forced to respond by charging higher premiums for the sick. In the same way, telephone companies have been forced to stop cross-subsidizing local from long-distance telephone service following the breakup of American Telephone and Telegraph.

Contracts did adapt. Group plans now predominate, whereas most people bought individual health insurance a generation ago. Since pooling is a partial solution to the time-inconsistency of long-term contracts, we can read this transition as an evolutionary adaptation to the increasing need for and difficulty of providing time-consistent long-term insurance.

But contracts are now stuck at a local maximum. More pooling will hurt competition and product variety, whereas less pooling will imply less insurance of long-term illness. The *optimal* contract cannot be found by local variation about existing contracts. It requires the simultaneous institution of severance payments and publicly acknowledged risk-rating.

In addition, health insurance is already a highly regulated market, and current regulations or the fear of future regulation may discourage the radical experimentation required to arrive at time-consistent contracts. An insurer that proposed the extreme experience rating of time-consistent contracts would certainly wake up insurance regulators! Epstein (1994) argues that regulatory and legal impediments account for many pathologies of insurance markets. In particular, he argues that courts often reinterpret insurance contracts ex post, judge the merits of each clause separately rather than how the clauses fit together to form a reasonable contract, and will not enforce severance payments or bond forfeiture against consumers. He finds that the fear of court reinterpretation has eliminated innovation in product warranties of a much smaller scale than the change to time-consistent health insurance contracts.

In summary, it seems at least possible that time-consistent contracts can be implemented, even though they have not yet been implemented. A well-documented story for the absence of long-term and time-consistent insurance is an important topic, requiring a detailed study of the regulatory and competitive history of the health insurance market that is obviously beyond the scope of this paper. The correct story for the absence of long-term health insurance will also have important policy implications. If the story is an unintended pathology of regulation, it will argue for a careful *de*regulation of insurance markets.

VI. Concluding Remarks

I have described Pareto-optimal, time-consistent health insurance contracts. These contracts *fully* insure consumers against all health risks, even long-term, expensive risks that are not insured under current contracts. The contracts feature severance payments equal to the present value of premium changes or, equivalently, insurance against changes in premiums. Consumers are not tied to insurers, so the disasters that befall sick consumers who now lose their health insurance do not occur, and this freedom promotes competition and product variety in health care and insurance. Surprisingly, contracts with this feature require no more payments than standard health insurance contracts, but merely a rearrangement of the rights to which those payments give rise.

The Clinton plan, most congressional proposals, and most regulation plans advanced by academics take exactly the opposite approach. At a most basic level, the plans force insurers and consumers into long-term contracts rather than specify a time-consistent structure. The plans herd people into large pools, whereas individuals can purchase time-consistent contracts. The plans try to stop insurers from

using health information to set premiums, whereas time-consistent contracts allow extreme rating. The plans force healthy consumers into the system to pay for the sick, whereas everyone pays actuarially fair premiums in a time-consistent contract. To avoid competition for healthy consumers, the plans must mandate much uniformity of health care and severely limit competition, whereas time-consistent contracts allow any amount of competition and product variety. Most important, the plans all feature a large regulatory structure. They must, to keep people from responding from obvious economic incentives: to keep insurers and providers from trying to improve their pool or reduce levels of care, and to keep individuals from trying to get better care at lower prices. Time-consistent contracts are a decentralized market solution, and it is likely that deregulation will be required to implement them.

Appendix

Contracts with Death

I generalize the contracts described in the text to allow a nonzero probability of death. Let $a_t = 1$ if the consumer is alive at time t and $a_t = 0$ if he is dead. The value of a_t is revealed in the second half of period t. Utility is zero if he is dead, so the consumer maximizes

$$\max E_0 \sum_{j=0}^{\infty} \beta^j u(c_j) a_j.$$

He has a constant income when alive, but income also drops to zero when he dies. Thus time t income is ea_t. Health expenses are also zero when he is dead.

Insurers are still risk neutral, and I maintain the simplifying assumption $\beta = 1/(1 + r)$. Thus time 0 contingent-claim values are equal to expected present values. The consumer's budget constraint in a time 0 contingent-claim market is

$$E_0 \sum_{j=0}^{\infty} \beta^j c_j = W_0 + E_0 \sum_{j=0}^{\infty} \beta^j (ea_j - x_j).$$

In addition, consumption at any date and in any state must be positive. The first-order conditions direct the consumer to consume a constant amount if alive, and zero when dead: $c_t = ca_t$. Plugging this value in the budget constraint and solving, we obtain

$$c = \frac{W_0 - E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j} + e. \tag{A1}$$

We support this optimum with a sequence of spot markets. Introduce a state-contingent payment y_t in the second half of each period. The value of this payment, which the consumer will have to pay in the first half of the period, is $E_t y_t$. To simplify notation, the consumer pays x_t, and y_t includes payments for current health expenses as well as severance payments. After the first such payment, y_0, the consumer would choose c_0 in a spot contingent-claim market as

$$c_0 = \frac{W_0 + y_0 - E_0 y_0 - E_1 \sum_{j=0}^{\infty} \beta^j x_j}{E_1 \sum_{j=0}^{\infty} \beta^j a_j} + e \qquad \text{(A2)}$$

if he lives, and zero if dead.

We determine y_0 so that c_0, so determined, equals the constant value c. Equating (A2) with (A1) and rearranging, we obtain

$$y_0 - E_0 y_0 = \left(\frac{E_1 \sum_{j=0}^{\infty} \beta^j a_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j} - 1 \right) W_0 + E_1 \sum_{j=0}^{\infty} \beta^j x_j - \frac{E_1 \sum_{j=0}^{\infty} \beta^j a_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j} E_0 \sum_{j=0}^{\infty} \beta^j x_j.$$

The same logic holds for any time period t, so

$$y_t - E_t y_t = \left(\frac{E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} - 1 \right) W_t$$

$$+ E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+j} - \frac{E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}. \qquad \text{(A3)}$$

When $a_t = 1$ for all time, this expression reduces to equation (9). If the consumer dies, $a_t = 0$, and $x_t = 0$, then he forfeits all wealth, $y_t - E_t y_t = -W_t$. Since the consumer chooses zero consumption if dead (I have not built in a bequest motive, though this is a simple extension), the contract naturally specifies no wealth after death.

The payment y_t can be viewed as a combination of two securities: one pays $E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+j}$ with price $E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}$. It is the combination of one-period health insurance (payoff x_t with premium $p_t = E_t x_t$) and premium insurance or severance payment (payoff $E_{t+1} \sum_{j=1}^{\infty} \beta^j x_{t+j}$ with price $E_t \sum_{j=1}^{\infty} \beta^j x_{t+j}$). It is a lifetime health insurance contract that is continually marked to market. The second security has payoff $E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}$ and price $E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}$. It is an annuity that is also continually marked to market. The expression for

$y_t - E_t y_t$ is composed of information set t quantities, which may be interpreted as portfolio weights, times these payoffs.

The consumer also has a risk-free saving or borrowing opportunity, so that

$$W_{t+1} = (1 + r)(W_t + y_t - E_t y_t + ea_t - x_t - c_t). \tag{A4}$$

It takes a few lines of algebra to verify that, with this evolution equation, equation (A3), and the decision rule for consumption, the time t version of equation (A1), consumption will be constant (e.g., from period 0 to the first half of period 1) as long as the consumer lives, and zero otherwise.

To generalize the health account contract, I specify a constant out-of-pocket payment q, equal to the time 0 flow present value of health expenses when the consumer is alive:

$$q_t = a_t q = a_t \frac{E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j}.$$

Payments and balances are separated into a regular savings account used for consumption smoothing and a health account. Their balances evolve as

$$\begin{aligned}
K_{t+1} &= (1 + r)[K_t + y_t^K - E_t y_t^K + (e - q)a_t - c_t], \\
A_{t+1} &= (1 + r)(A_t + y_t^A - E_t y_t^A + qa_t - x_t),
\end{aligned} \tag{A5}$$

where payments y_t are split up as

$$y_t^K - E_t y_t^K = \left(\frac{E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} - 1 \right) K_t,$$

$$y_t^A - E_t y_t^A = \left(\frac{E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} - 1 \right) A_t + E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+j} \tag{A6}$$

$$- \frac{E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}.$$

One can verify that $y_t^A + y_t^K = y_t$, that $K_t + A_t = W_t$, and that the W_t evolution equation (A4) is satisfied. The savings account K now also has a state-contingent payment, which simply reflects marking to market the value of annuities that a permanent-income consumer uses when lifetime is uncertain.

We can use equation (A5) to track the balance in the health account. Starting from $A_0 = 0$, we obtain

$$\frac{A_t}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} = \frac{E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} - \frac{E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j}$$

if the consumer is alive, and $A_t = 0$ if the consumer is dead at time t. This is equation (13) in the text. To derive this equation, note that it holds at time 0: $A_0 = 0$. Then suppose that it holds at time t and that the consumer is alive at the end of time t, and show that it holds at time $t + 1$. From equation (A5) and (A6),

$$\beta A_{t+1} = A_t + y_t^A - E_t y_t^A + q - x_t,$$

$$\beta A_{t+1} = E_{t+1} \sum_{j=0}^{\infty} \beta^j a_j \left(\frac{E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} - \frac{E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j} \right)$$

$$+ E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+j} - \frac{E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+j}}{E_t \sum_{j=0}^{\infty} \beta^j a_{t+j}} E_t \sum_{j=0}^{\infty} \beta^j x_j + \frac{E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j} - x_t,$$

$$\beta A_{t+1} = \left(1 - E_{t+1} \sum_{j=0}^{\infty} \beta^j a_j \right) \frac{E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j} + \beta E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+1+j},$$

$$A_{t+1} = E_{t+1} \sum_{j=0}^{\infty} \beta^j x_{t+1+j} - E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+1+j} \frac{E_0 \sum_{j=0}^{\infty} \beta^j x_j}{E_0 \sum_{j=0}^{\infty} \beta^j a_j}.$$

Dividing by $E_{t+1} \sum_{j=0}^{\infty} \beta^j a_{t+1+j}$, we obtain the result.

References

Aaron, Henry J. "Issues Every Plan to Reform Health Care Financing Must Confront." *J. Econ. Perspectives* 8 (Summer 1994): 31–44.

Cutler, David M. "Why Doesn't the Market Fully Insure Long-Term Care?" Working Paper no. 4301. Cambridge, Mass.: NBER, March 1992.

————. "A Guide to Health Care Reform." *J. Econ. Perspectives* 8 (Summer 1994): 13–30.

Diamond, Peter A. "Organizing the Health Insurance Market." *Econometrica* 60 (November 1992): 1233–54.

————. "Two Improvements on the Clinton Framework." *J. Econ. Perspectives* 8 (Summer 1994): 61–66.

Enthoven, Alain C., and Kronick, Richard. "A Consumer-Choice Health Plan for the 1990s: Universal Health Insurance in a System Designed to Promote Quality and Economy." *New England J. Medicine* 320 (January 5, 1989): 29–37; (January 12, 1989): 94–101.

Epstein, Richard. "Mortal Peril: The Reform of Health Care in America." Manuscript. Chicago: Univ. Chicago, Law School, 1994.

Feenberg, Daniel, and Skinner, Jonathan. "The Risk and Duration of Catastrophic Health Care Expenditures." Working Paper no. 4147. Cambridge, Mass.: NBER, August 1992.

Fudenberg, Drew; Holmstrom, Bengt; and Milgrom, Paul. "Short-Term Contracts and Long-Term Agency Relationships." *J. Econ. Theory* 51 (June 1990): 1–31.

Himmelstein, David U., et al. "A National Health Program for the United States: A Physicians' Proposal." *New England J. Medicine* 320 (January 12, 1989): 102–8.

Kocherlakota, Narayana R. "Efficient Bilateral Risk Sharing without Commitment." Manuscript. Iowa City: Univ. Iowa, 1994.

Malcomson, James M., and Spinnewyn, Frans. "The Multiperiod Principal-Agent Problem." *Rev. Econ. Studies* 55 (July 1988): 391–407.

Newhouse, Joseph P. "Medical Care Costs: How Much Welfare Loss?" *J. Econ. Perspectives* 6 (Summer 1992): 3–21.

————. "Symposium on Health Care Reform." *J. Econ. Perspectives* 8 (Summer 1994): 3–12.

Pauly, Mark V. "Taxation, Health Insurance, and Market Failure in the Medical Economy." *J. Econ. Literature* 24 (June 1986): 629–75.

————. "Universal Health Insurance in the Clinton Plan: Coverage as a Tax-Financed Public Good." *J. Econ. Perspectives* 8 (Summer 1994): 45–54.

Pauly, Mark V.; Danzon, Patricia; Feldstein, Paul; and Hoff, John. "A Plan for 'Responsible National Health Insurance.' " *Health Affairs* 10 (Spring 1991): 5–25.

Pauly, Mark V.; Kunreuther, Howard; and Hirth, Richard. "Guaranteed Renewability." Manuscript. Philadelphia: Univ. Pennsylvania, 1992.

Poterba, James M. "A Skeptic's View of Global Budget Caps." *J. Econ. Perspectives* 8 (Summer 1994): 67–74.

Rey, Patrick, and Salanie, Bernard. "Long-Term, Short-Term and Renegotiation: On the Value of Commitment in Contracting." *Econometrica* 58 (May 1990): 597–619.

Weisbrod, Burton A. "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment." *J. Econ. Literature* 29 (June 1991): 523–52.

Zeckhauser, Richard. "Public Finance Principles and National Health Care Reform." *J. Econ. Perspectives* 8 (Summer 1994): 55–60.