Abstract
In this paper, I analyze four alternative policies that address "premium risk": the risk that health insurance premiums will increase if an individual acquires a chronic condition. They are: premium risk insurance, community rating, anti-discrimination laws for large firms, and government subsidies. I show that community rating and anti-discrimination laws provide incomplete protection, and have the potential to create welfare-reducing distortions in incentives, while subsidies can be designed that mimic the efficiency of premium risk insurance. While the economic concepts behind these conclusions are well-documented and well-understood, I believe my own model adds value in the form of simplicity and ease of comparison. There is also an unambiguous policy implication: if we are concerned with premium risk, the complex legislation that has been introduced so far should be abandoned in favor of a new subsidy program.

* Special thanks are owed to Professors Kenneth Kuttner and Barbara Craig for their helpful and insightful comments, as well as to Professors Cleeton, Craig, Fernandez, Gaudin, Kasper, and Kuttner for all that I have learned in four great years as an Oberlin Economics major. I thank Terri Pleska, all of my other professors at Oberlin, and my fellow Economics majors. Finally, my most profound gratitude is reserved for my parents, my sister, and for Stephanie.
1. Introduction

In the last several decades, improvements in medical technology and public health have resulted in a new phenomenon: widespread, expensive chronic conditions. Whereas previously most healthcare treated acute conditions—after becoming ill, you were either cured, you died, or you suffered untreatable symptoms—chronic conditions that can require treatment for decades have become a much more important factor in determining expenditures on healthcare. According to the CDC\(^1\), more than 90 million Americans (nearly 30% of the population) have some form of chronic illness, and they account for more than 75% of the nation’s total spending on medical care. 70% of U.S. deaths and 33% of under-65 life-years-lost are attributed to chronic conditions. Some of the most common include cardiovascular disease, cancer, diabetes, Alzheimer’s, hypertension, asthma, and arthritis. Other diseases, like AIDS, are becoming less like acute and more like chronic conditions all the time.

With traditional health insurance contracts, which are renewed yearly, chronic conditions pose a particular risk: “premium risk”\(^2\) is the risk that an individual’s health insurance premiums will increase if she acquires a chronic condition. For insurers to find their business worthwhile, they must charge a premium for insurance that is at least as great as their expected costs of paying for care. Since a person with a chronic condition generally has higher expected costs than the overall population, insurers will justifiably want to charge her higher premiums.\(^3\) If they were restricted to only charge her the same premium as healthy consumers, they would expect to lose money on her contract and

---


\(^2\) As far as I can tell, the phrase “premium risk” is attributable to Kifmann, Mathias (2002).

\(^3\) This market imperfection was first pointed out by Arrow (1963), pages 963-964. See also Cutler and Zeckhauser (1999).
would prefer not to insure her at all. Even worse, if they tried to charge healthy consumers higher than actuarially fair premiums, these consumers could be lured away by other insurers offering coverage at lower rates. Competition forces insurers to charge risk-rated, actuarially fair premiums if they wish to survive. Given the skewed distribution of healthcare expenditures, premium risk is potentially huge. In 1996, 1% of Americans accounted for 27% of total healthcare spending, 5% accounted for 55% of spending, while the bottom 50% of the population accounted for only 3% of spending. Almost 54% of the top 1% of spenders were under 65.4

"Premium risk," interpreted in various ways, has drawn considerable attention from policymakers, especially during the Clinton administration. It was specifically addressed in Title 1 of the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which prohibited group insurance plans from charging a higher premium or denying coverage to a member of the group on the basis of health status, and limited exclusions for pre-existing conditions.5 Many state legislatures have imposed similar restrictions on the individual insurance market, including guaranteed issue requirements: coverage cannot be denied on the basis of health status; and rate bands: restrictions on the variability of premiums.6 Of course, guaranteed issue and rate bands must both be enforced for either to be effective, otherwise insurers could charge a low premium but deny coverage, or offer coverage only at an exorbitant premium. Similarly, regulations like these have begot the need for many other costly- and difficult-to-enforce rules. States have tried to monitor the way insurers market their products to prevent them from

---

6 For a comprehensive description of state regulations, see Kofman and Pollitz (2006).
purposefully targeting healthy clients and avoiding unhealthy ones. “Mandated benefits” endeavor to prevent insurers from differentiating policies by limiting coverage; for example, they may require plans to pay for diabetes supplies (46 states), Pap smears (27 states), mammograms (50 states), or nurse practitioner services. Variation across states increases the cost of compliance.

Although commentators on healthcare policy often emphasize three goals of insurance regulation: “affordability,” “availability,” and “adequacy,” from an economics perspective we can understand these three goals to really be just one: affordability. There is no reason why insurers would not be willing to offer any reasonable amount of coverage—for the right price. Only when prices are restricted will insurers have an incentive to exclusively offer incomplete plans. Premium risk may be insurable, but only *ex ante*: before a chronic condition is revealed. After a chronic condition is known to exist, the idea of affordability becomes one of subsidy: anyone who thinks insurance premiums should not increase because a consumer acquired a chronic condition—assuming that consumer had not purchased premium risk insurance when well—is saying that the consumer should be allowed to purchase insurance below cost. HIPAA and the state regulations try to accomplish this subsidization through price controls. A much more efficient means would be through an explicit subsidy.

Both state and federal legislatures have implemented subsidy programs in various forms, although none of these is particularly effective. Subsidy programs—which tend to be small and under-funded, especially on the state level—target uninsured, low-income individuals with chronic conditions. The goal of eliminating premium risk is often mixed up with efforts to redistribute income and increase the number of Americans with
insurance. As of 1999, “State High-Risk Pools” had been set up in 28 states and provided insurance to 105,000 people, a tiny fraction of the already small individual insurance market. Pools are operated by state governments and provide insurance to qualified “uninsurable” consumers at subsidized prices. Despite the subsidization, most high-risk pools charge premiums that are considerably higher than average rates in the individual market, offer limited benefits, and have pre-existing condition exclusion periods. States have struggled to fund these pools, relying on general revenues, designated funds (such as taxes on tobacco products), and, most commonly, assessments on private insurers. This last source can have the perverse effect of exacerbating high prices in the individual market and distorting incentives.8

Several states sponsor reinsurance programs for the small-group and individual insurance markets. Reinsurance subsidizes insurers’ highest-cost clients, although the amount of coverage varies widely depending on the state. As of 2004, insurers enrolled in Massachusetts’ plan paid for an individual’s claims up to $5,000 in a year, 10% coinsurance on the next $50,000, and nothing above $55,000. Insurers in Connecticut were covered for all claims above a $5,000 per person deductible. In New York, the state reinsurer covered 90% of annual claims between $5,000 and $75,000 per person. In some states, insurers are charged premiums that are large enough for the reinsurer to break even, while in others the reinsurance pool is subsidized with general revenue. Even in an unsubsidized pool, reinsurance acts as a tax transfer, shifting costs from unhealthy pools with high claims to healthy pools that pay excess premiums. Because enrolled

7 See Achman and Chollet (2001).
8 For example, the Employee Retirement Income Security Act of 1974 (ERISA) prohibits taxes on self-insured employer plans, which places insured plans at a competitive disadvantage.
insurers have less of an incentive to raise premiums if a client acquires a chronic condition, premium risk is transferred from the individual to the state. Reinsurance programs have several shortcomings, including variability across states, entrance requirements like low income and lack of insurance, and vulnerability to adverse selection when insurers choose which pools to enroll.

Health Savings Accounts\(^{10}\) (HSAs) were established in 2003 and allow consumers to purchase private high-deductible health insurance with untaxed income, up to a maximum contribution of $2,600 for individuals or $5,150 for families. Qualified plans must have deductibles of at least $1,000 for an individual, $2,000 for a family, and out-of-pocket maximums of $5,000 individual/$10,000 family. Before HSAs were introduced, there were several similar tax-advantaged accounts in place: Flexible Spending Accounts, Medical Savings Accounts, and Health Reimbursement Arrangements. Although the exact tax treatment, eligibility, allowable uses, and ownership of these accounts vary, they all share the same basic principle: extend the tax deduction given to employer-sponsored insurance to encourage uninsured consumers to purchase individual or small-group insurance on their own. Although HSAs may help chronically ill consumers pay for health insurance, they do almost nothing about the problem of premium risk. They provide no incentive for an insurer to not charge risk-rated premiums. Although HSAs can be carried over across years, and earn interest tax-free, the accounts are really too small to help a healthy consumer save up in case she eventually acquires a chronic condition. Even if the accounts were larger, saving is an inefficient way to deal with premium risk: consumers who remain healthy will have saved too much, while chronically ill consumers probably will not have been able to save

\(^{10}\) See Kofman (2004).
enough. Another concern is that HSAs may increase risk segmentation as healthy consumers have an increased incentive to switch away from employer-based pools in favor of cheaper high-deductible plans.

The United States is unique among developed nations in the way it pays for healthcare. In countries like Canada, the United Kingdom, Finland, and Australia, all citizens are obliged to have some form of health insurance. Since these “universal healthcare systems” are generally immune to adverse selection, community rating is enforced and premium risk is not a problem. A similar situation prevailed in the first two-thirds of the 20th century in the U.S., when a relative monopoly by Blue Cross allowed it to charge a single premium, regardless of health status. Competition, despite its many benefits, has made voluntary community rating unsustainable.

As of 2004, over 15% of the U.S. population, or nearly 46 million Americans, had no health insurance and had to pay for their care themselves (or receive charity care). About 27% of Americans had insurance coverage through public programs—most notably Medicare, which covers people over 65, those with disabilities, and anyone with end-stage renal failure (a chronic condition), and Medicaid, which is for low-income children, parents of dependent children, and the elderly (nursing home care). Government employees, the active-duty military, and veterans also receive government-financed insurance or government-provided care (through the Veteran’s Administration), and made up about 4% of the population in 2004. By far the most common source of insurance was employer-sponsored plans, which provided coverage to nearly 60% of the population. The main reason for the popularity of job-based insurance is that compensation in the form of insurance benefits, unlike wages, is tax-free to the employee.
Only about 6% of the population purchased their primary insurance directly in the individual market. Besides unfavorable tax treatment, the individual insurance market is disadvantaged by high administrative costs (caused by higher selling costs of policies) and a potentially severe adverse-selection problem.\(^\text{11}\)

Because of the high proportion of Americans in either group or public insurance, where premiums are generally not risk-rated, it may seem that very few people are subject to premium risk. Note that, in the absence of regulation, any consumer outside of Medicare would be subject to premium risk: premiums in the individual and group market could be risk-rated, high-cost individuals could be excluded from plans, and Medicaid enrollees might be forced to transfer to other plans as their life situation changed. For the uninsured, the possibility of a chronic condition would add to the uncertainty of their medical expenses. Even with HIPAA, group plans are allowed to exclude treatments or diseases from coverage as long as the exclusion is applied uniformly to all employees. State regulation varies considerably depending on the state, and regulations of the individual market can often be circumvented. More importantly, even if state regulations and HIPAA were more complete, consistent, and worked together perfectly to eliminate premium risk, a much more efficient solution could be reached with a well-designed government subsidy.

The kind of subsidy advocated here is quite different from those discussed above. For a subsidy program to deal optimally with premium risk, it should be organized by the federal government, so that all Americans can benefit and so that risk can be spread out

\(^{11}\) Note that these percentages—from the U.S. Census Bureau’s 2005 Current Population Survey—do not sum to 100% because of double-counting when individuals had more than one kind of insurance during the year, which may lead to bias, especially in the number of uninsured Americans. See U.S. Department of Health and Human Services, (2005) “Overview of the Uninsured in the United States.” http://aspe.hhs.gov/health/reports/05/uninsured-cps/.
across the largest possible pool. Beneficiaries should not be qualified by level of income or lack of insurance: there are surely more efficient ways to redistribute income, and higher-income taxpayers will be much more likely to support the plan if they stand to benefit from it directly. The program is almost assured to be redistributive anyway, since lower-income households tend to have more chronic conditions, and, depending on the financing mechanism, tend to pay less in taxes. The subsidy should be understood to be a kind of insurance, the same as premium risk insurance. Taxpayers are willing to pay for the government PRI because they know that, if they become chronically ill, they will be protected: the premium they pay for insurance will not increase net of the subsidy. Although the only explicit goal of this program is to eliminate premium risk, it is also likely to increase the number of Americans with insurance by decreasing adverse selection and allowing premiums to approach their actuarially fair level. The exact subsidy payout mechanism should be the one least susceptible to fraud or misuse, probably a fixed disease-specific payment directly to the insurer.

In section 2, I review the literature on private premium risk insurance and the reasons it has not been widely adopted. In section 3, I introduce a model to explain and compare four alternative policies for dealing with premium risk: premium risk insurance, community rating, anti-discrimination laws in group policies (HIPAA), and a government subsidy. In section 4, I comment on the effects of changing some of the model’s assumptions, combining different policy options, and related issues. Section 5 summarizes my conclusions.
2. Premium Risk Insurance

In an ideal world, premium risk would not exist. Insurers would offer a single insurance contract that would be agreed to by all consumers before birth, and would cover all of their lifetime healthcare expenses. Insurance companies would be committed to charging a predetermined sequence of premiums regardless of the health status of the consumer, and consumers would be obliged to pay those premiums. Since the full future path of premiums would be decided before any unknown health characteristics or chronic conditions were revealed, there would be no premium risk.

There are many reasons a lifelong insurance policy is not possible. Insurance companies can go bankrupt, consumers can move away from their insurers, babies cannot sign contracts, and so on. The biggest obstacles, however, are time inconsistency and adverse selection. If an insurer agrees to a premium that turns out to be below its costs, for example if it is unlucky enough to end up with a pool of clients that have lots of high-cost chronic conditions, it will have a strong incentive to renege on its commitment by canceling the policies or going out of business. Similarly, if a consumer agrees to premiums that are above her expected costs, for example because she ended up much healthier than average, she will want to stop paying her expensive premiums and switch to another insurer. If only unhealthy consumers remain in the original insurance pool, premiums will have to increase or the insurer will become insolvent. In order for lifelong insurance to work, healthy consumers must remain in insurance pools to subsidize the premiums of unhealthy consumers, even after their health states are revealed to them. Judicial attempts to enforce unfavorable contracts are likely to be ineffective and costly.
For premium risk insurance to really be feasible, contracts must be designed that consumers willingly uphold.

Pauly Kunreuther, Hirth (hereafter PKH) (1995) and Cochrane (1995) have both derived “optimal” contracts that make premium risk insurance (PRI) possible. These papers share the same objective: to divide the lifetime insurance contract into a series of one-period contracts that will be purchased by both healthy and unhealthy consumers in all periods, and that allow insurers to break even. The idea is to make all consumers “pre-pay” higher future premiums before they are aware of who will actually end up with a chronic condition. If these contracts existed, they would represent an efficient, market-based way to eliminate premium risk. There have been many explanations offered as to why they are rarely, if ever, seen in the U.S.

PKH (1995) envision a “Guaranteed Renewable” (GR) insurance contract. Insured consumers are guaranteed the right to renew their insurance at the end of the contract term, and into the future, at a predetermined rate of premium growth. The premiums are structured such that low-cost consumers will choose to purchase insurance in all periods, the same premium is charged regardless of health status, and insurers break even. In PKH’s model, total costs in the last period are paid in the first period, total costs in the second-to-last period are paid in the second period, and so on. In the early periods, premiums exceed annual costs. Healthy individuals are willing to buy insurance because it is actuarially fair, considering the risk that they will become unhealthy in the future.

---

12 Insurance contracts claiming to be “guaranteed renewable” are actually quite common in the U.S., and GR protections are even mandated by most states and by HIPAA. However, a GR designation is usually deceptive in this case. The HIPAA rule does not restrict premium increases over time, or even require that premiums be the same for all individuals in a rating class. Stricter state regulations can often be bypassed by raising premiums for an entire GR pool and then recategorizing healthy consumers into new pools. For GR to be meaningful, it must specify a future rate of premium growth at the time of purchase. GR provisions that do not restrict premiums are more like “guaranteed issue” requirements, which are useless in isolation.
Premiums decline over time, so that by the last period, they have fallen to the one-period actuarially fair level for a healthy consumer. By contract, those consumers that become high-cost are guaranteed insurance at the same premium as the healthy consumers. If a low-cost consumer drops out of the insurance pool at any time, it will have no effect on the premium or the insurer’s zero level of profits. If a high-cost consumer drops out, the insurer’s profits will increase. Although the model predicts premiums that decrease over time, other factors, like age-related increases in risk over time, may make the observed premium sequence flat or increasing. Guaranteed renewability is implicit in large-group insurance policies under HIPAA.

Rather than committing consumers to a particular insurer, Cochrane’s (1995) contract requires a severance payment at the end of each period equal to the change in the present discounted value of a consumer’s expected future costs. If a consumer becomes healthier, she makes a payment to her insurer, and if she becomes less healthy, she receives a payment. One-period insurance can then be offered in each period at the risk-rated actuarially fair premium, and consumers can use their severance payments to pay for higher premiums if they acquire a chronic condition. The PRI need not be provided by the same company as the health insurance, or even by a health insurer. The cost of this “time-consistent” PRI is the same as in PKH: the actuarially fair premium, given the future severance payments a consumer has some probability of receiving or paying. Cochrane expects that these severance payments will be kept in a special account created for that purpose, so that the consumer need only pay a fixed amount into the account every period, and severance payments are received or paid automatically.
Problems have been demonstrated with both kinds of PRI contract, especially by the authors themselves. Neither contract can do anything for consumers who start out with a chronic condition before any PRI contracts are available. Similarly, PRI will not help consumers with a congenital predisposition to illness. PRI premiums must be adjusted for initial differences in expected costs or these contracts will be no more viable than standard community-rated insurance. There may be limited demand for PRI if consumers systematically underestimate their own probability of illness because of cognitive bias. Although an individual with a chronic condition may be highly aware of chronic care issues, healthy young people are probably oblivious to the risks they face. Moral hazard may be an issue if consumers with PRI are less likely to make lifestyle changes, like losing weight or quitting smoking, that could decrease their risk of a chronic condition in the future. PRI could be subject to adverse selection if consumers have private information about their own health state before they purchase PRI. Perhaps most importantly, the existence of alternative policies like anti-discrimination laws and rate bands limit (or at least are perceived to limit) the premium risk consumers face, and make them less likely to purchase PRI.

There are other problems specific to either PKH's (1995) or Cochrane's (1995) contract. Guaranteed renewability suffers from lock-in. Insurers are committed to provide a predetermined level of coverage at a single price to both high- and low-cost consumers. Those consumers that turn out to be high-cost are essentially bound to the insurer because they have already pre-paid premiums. If they switched to another insurer, they would be charged a risk-rated premium. This reduces competition and consumer choice. It is problematic if a consumer wishes to change insurers for any
reason, such as relocation or a change of employer. Insurance companies have an incentive to exploit lock-in by raising premiums for an entire pool (if the policy allows for some flexibility in pricing), or by skimping on quality (e.g. by reinterpreting the coverages or limiting the network of care providers). This behavior may be limited by a careful drafting of the insurance contract, although any such commitment will be time-inconsistent, and perhaps unenforceable.

The supply of GR coverage will be limited by intertemporal risks that cannot be diversified. By committing to a future level of coverage, the insurer exposes itself to the risk of medical inflation or expensive new technology. Unlike the interpersonal risk of illness (except in the case of epidemics), this kind of risk cannot be pooled. Medical inflation will affect all insured individuals in the same way. Insurers may have different estimates of probable future costs. Firms with low estimates—especially if they are highly constrained with respect to the premiums they can charge or the coverage they must offer—may risk bankruptcy. If an insurer goes out of business for any reason, high-cost clients will have pre-paid premiums for no reason. There is an incentive for insurers to collect excess PRI premiums and then immediately drop out of the insurance market. The need to fix coverages ahead of time makes it difficult to adapt to new medical technology.

Frick (1998) provides another reason that guaranteed-renewable contracts may not be possible: consumer capital market constraints. The PKH model assumes perfect capital markets, no consumer discounting, and zero interest rates. Frick’s model eliminates consumer’s ability to borrow to pay high initial premiums, and assumes zero interest rates and positive consumer discounting, observing that full GR protections will

---

never be purchased in this case (the discount rate acts as a tax on pre-paid premiums). For consumers with relatively low discount factors—high discount rates, i.e. “impatient” consumers—guaranteed-renewability may not be purchased at all.\textsuperscript{14}

Cochrane’s PRI does not result in insurer lock-in. Cross-subsidization through severance payments equalizes consumers’ net expected future expenses at the end of every period. Special PRI accounts are attached to consumers, so they are free to change insurers at the end of any period. There is less concern that an insurer will go bankrupt because of the reduced time frame. Insurers can always readjust premiums to one-period actuarially fair levels, so medical inflation can be incorporated over time.\textit{Ex post} bilateral severance payments mean that PRI premiums can remain level over time, eliminating the need for excessive or impossible borrowing to pre-pay premiums.

On the other hand, Pauly, Nickel, Kunreuther (1998) show that the existence of Medicare, which puts a time limit on future cost risk resulting from a chronic condition, can cause level-premium PRI to unravel. Cochrane’s contract (as well as the subsidy system I propose) requires that changes in expected healthcare expenses be measurable, which is not necessary with guaranteed renewability. Of course, changes in expected expenses should be observable through changes in premiums. If premiums do not change, no PRI benefits need be paid. Trouble may arise in estimating changes in future expected expenses (to determine the appropriate severance payment), although Cochrane believes a reasonable approximation could be found. Perhaps a more serious problem is that the information contained in premiums may be obscured if a consumer switches her

\textsuperscript{14} Herring and Pauly (2006b) use an empirical simulation to determine an optimal path of GR premiums that increases with age and approximates observed premium schedules in the individual market. This finding implies that Frick’s critique may not be relevant, and even that GR contracts exist and are effective in the individual market.
insurance plan to one with a different level or quality of coverage. PRI should only reimburse changes in premiums for the same contract, not changes in preferences toward higher or lower quality insurance. Similarly, changes in premiums resulting from medical inflation must be separated from those caused by changes in health status.

It is possible that PRI contracts have simply not had sufficient time to evolve, given the relatively recent appearance of widespread, expensive chronic conditions. Cochrane’s PRI in particular would require radical change from currently offered health insurance contracts, including a complicated new system of severance payments, the creation of special PRI accounts, and institutions to assure fairness and prevent fraud. Insurance companies may be loath to introduce the new contracts because of fear of future regulation, for example rules aimed at preventing them from fully risk-rating premiums, and because of the risk that courts will not enforce PRI contracts, especially politically-sensitive provisions like consumers making severance payments to insurers when they become healthier.

3. Model

I consider a two-period model. At the beginning of period 1, all consumers are assumed to be identical, and none has a chronic condition. I assume there is only one chronic condition and all consumers have a probability \( p \) of acquiring the condition during period 1. I assume \( 1 > p > 0 \). “Healthy” consumers (i.e. those without the chronic condition) have healthcare costs \( \varepsilon \), where \( \varepsilon \) is a continuous random variable that takes on a value in the range \( 0 \leq \varepsilon < \infty \) according to the probability density function
The expectation of healthcare costs for a healthy consumer is $m_0 = \int_0^\infty f(\varepsilon) d\varepsilon$. I assume the insurance market is competitive and there are no administrative costs: premiums are actuarially fair, and for a healthy consumer the premium is $m_0$. On average, a portion $p$ of consumers will be "unhealthy" (i.e. they have the chronic condition) in period 2. Unhealthy consumers have healthcare costs $\mu$, a continuous random variable that occurs in the range $0 \leq \mu < \infty$ according to the probability density function $g(\mu)$. The expectation of $\mu$ is $m_1 = \int_0^\infty \mu g(\mu) d\mu$ where $m_1 > m_0$. The presence of a chronic condition can affect the uncertainty of healthcare costs as well as the expected value of those costs. Since what we are really interested in is premium risk, and insurers are assumed to be risk-neutral, the difference between expected costs of "healthy" and "unhealthy" consumers is the only relevant effect.\(^\text{15}\)

I. Series of Events:

1. Purchase insurance for period 1
2. Purchase premium risk insurance?
3. Reveal chronic condition
4. Purchase insurance for period 2?
5. Switch employers?
6. Collect subsidy?
7. Period 1
8. Period 2

\(^{15}\) Note that this is also the reason I distinguish between "high-cost" and "low-cost" consumers, rather than using the more common terminology "high-risk" and "low-risk."
II. Sample cost-shock distributions

Healthy

prob.

Unhealthy

prob.

I assume utility is a function of wealth left over after paying healthcare costs and insurance premiums, $U(w - \varepsilon / \mu)$ without insurance or $U(w - m_{0,1})$ with insurance (the slashes denote an either/or choice), and that consumers are Bernoulli expected-utility maximizers. Consumers begin each period with $w$ in endowed wealth, and wealth is not transferable between periods. I assume consumers are risk-averse, resulting from (positive) diminishing marginal utility of wealth: $U' > 0$ and $U'' < 0$. Absent moral hazard, consumers will always prefer a fixed loss (the actuarially fair insurance premium) to an uncertain loss of the same expected magnitude. Consumers are assumed to discount period 2 outcomes by a factor $\delta$. I assume that insurers and consumers have the exact same perfect information about a consumer's health state with respect to the chronic condition. I consider five policy alternatives: A. do nothing (insurance is offered in each period at the risk-rated actuarially fair premium); B. there is premium risk insurance à la Cochrane (1995) or PKH (1995); C. there is government-enforced community rating in both periods; D. there are large-firm employers with implicit guaranteed renewability.

---

16 See "Note 1: Savings" in the next section.
(e.g. because of HIPAA); E. there are transfer payments to the chronically ill. In section 4, I comment on the effects of adding savings, moral hazard, *ex ante* heterogeneity, additional periods, administrative costs, or endogenous changes in income to the model. I also discuss combining policies (C) and (D), which would be most similar to actual policy, the possibility of PRI despite the existence of policies (C), (D), or (E), and potential concerns about the subsidy system.

**A. Do Nothing**

This is the simplest scenario. Premiums in each period are equal to a consumer’s expected loss, conditional on her health state. Consumers purchase insurance in all periods because they are risk-averse and insurance is actuarially fair. The first period premium is just $m_o$ for all consumers. The second period premium is $m_o$ for the $1 - p$ consumers who don’t acquire the chronic condition, and is $m_1$ for the $p$ consumers who do. Before health states are revealed, expected utility for the representative consumer is:

$$U[w-m_o] + (1-p)\delta U[w-m_o] + p\delta U[w-m_1].$$

**B. Premium Risk Insurance**

As in PKH (1995), premium risk insurance requires that the premium schedule match the cost schedule in reverse chronological order, so the last period is financed first. This allows insurers to charge premiums that are no higher than what consumers will consider actuarially fair (with respect to the entire time frame) in all periods. Average

---

17 For a graphical analysis of this scenario, and the utility loss from premium risk, see Figure 1 in the Appendix.
healthcare costs in period 2 are \( p m_1 + (1 - p) m_0 \). Insurers earn interest on unspent premiums at a rate \( r \), so the excess costs of unhealthy consumers in period 2 require excess premiums in period 1 of \( \frac{p(m_1 - m_0)}{1 + r} \). In Cochrane’s (1995) conception, this excess premium will be used to finance a severance payment of \( m_1 - m_0 \) to the “unhealthy” portion \( p \) of consumers at the end of period 1, which they can then use to purchase insurance at the actuarially fair premium \( m_1 \). In PKH’s (1995) conception, the excess premium allows the insurer to charge a premium \( m_o \) to all consumers in period 2, which is actuarially favorable to unhealthy consumers, and actuarially fair to healthy consumers. Either mechanism produces \textit{ex ante} expected utility of:

\[
U\left[w - m_o - \frac{p(m_1 - m_0)}{1 + r}\right] + \delta U[w - m_o].
\]  

(2)

An uncertain loss in period 2 has been converted into a certain loss in period 1. Compared to the scenario in (A), consumers are made better off by purchasing PRI if:

\[
U[w - m_o] - U\left[w - m_o - \frac{p(m_1 - m_0)}{1 + r}\right] < p \delta \left(U[w - m_o] - U[w - m_1]\right).
\]  

(3)

If inequality (3) does not hold, no one will buy premium risk insurance, so consumers cannot be made worse off by its availability. In this case consumers would choose to “self-insure,” and expected utility would be the same as in (A). Inequality (3) shows that for consumers to demand PRI, the utility foregone in period 1 to pay the extra premium (the left-hand side) must not outweigh the increase in discounted expected utility in period 2 (the right-hand side). Rearranging terms, we have:

\[
p \delta \left(U[w - m_o] - U[w - m_1]\right) - U[w - m_o] + U\left[w - m_o - \frac{p(m_1 - m_0)}{1 + r}\right] > 0.
\]  

(4)
Taking the derivative of the left-hand side of (4) with respect to \( r \) and \( \delta \), we can verify that premium risk insurance becomes more attractive with increasing \( r \) and increasing \( \delta \):

\[
\frac{\partial}{\partial r} = U'[w - m_0 - \frac{p(m_1 - m_0)}{1 + r}] \left( \frac{p(m_1 - m_0)}{(1 + r)^2} \right) > 0 \tag{5}
\]

\[
\frac{\partial}{\partial \delta} = p(U[w - m_i] - U[w - m_0]) > 0. \tag{6}
\]

Higher interest rates allow insurance companies to charge lower excess premiums, and less discounting (\( \delta \) closer to 1) means consumers are more concerned about the risks they face in period 2. Both of these factors make PRI more desirable.

PKH (1995) assume zero interest rates and no discounting, \( r = 0 \) and \( \delta = 1 \).

Cochrane (1995) makes the more general assumption \( \delta = \frac{1}{1 + r} \). Using Cochrane's value for \( \delta \), making the simplifying substitutions \( w - m_0 = w_1 \) and \( m_i - m_0 = x \), and rearranging terms, inequality (3) becomes:

\[
U[w_i - p\delta x](1 - p\delta)U[w_1] + p\delta U[w_i - x]. \tag{7}
\]

Inequality (7) is, of course, always true for a risk-averse consumer: it is just the standard preference for a fixed loss over an uncertain loss of the same expected magnitude.

However, note that if \( \delta < \frac{1}{1 + r} \), consumers will not necessarily choose to purchase PRI.

As per Frick (1998), if consumers have sufficiently high discounting (relative to the interest rate)\(^{18}\) they will not find it worthwhile to sacrifice current consumption to

\(^{18}\) Note that Frick (1998) actually assumes a zero interest rate but positive discounting, which seems unrealistic.
eliminate a future risk. Given \( r \), the particular range of \( \delta \) over which (3) will hold will depend on the degree of risk aversion (the shape of the utility curve).\(^{19}\)

**C. Community Rating\(^{20}\)**

In this scenario, the government requires that an insurer charge the same premium to all consumers in an age- or geographical-based group regardless of health status. While strict community rating is rarely seen in the U.S., partial community rating is a common feature of state insurance regulations.\(^{21}\) Where community rating is enforced, insurance companies can often find ways around it, for example by offering plans with less complete coverage that attract healthier consumers or through targeted advertising.

In the extreme case that risk groups can be completely segregated into different plans, the results will be the same as in (A): a “separating equilibrium.” I model the opposite extreme: insurers can charge only one premium and must accept anyone who demands insurance at that price.

Under community rating, the actuarially fair premium in period 1 remains \( m_o \), and all consumers purchase insurance. If both healthy and unhealthy consumers remain in the pool in period 2, insurers will charge the break-even premium \( m_o + p(m_i - m_o) \).

This premium is actuarially favorable to the unhealthy consumers, so they will definitely purchase insurance. It is actuarially unfavorable to healthy consumers, who have expected costs \( m_o \). Since I am assuming that no other insurer can lure the healthy consumers away with a lower premium—it would not be able to keep out the unhealthy

---

\(^{19}\) For a graphical analysis, see Figure 2 in the Appendix.

\(^{20}\) Rothschild and Stiglitz (1976) first observed the adverse selection results of parts (C) and (D).

consumers—the healthy consumers face a tradeoff. They can either bear the full risk of their uncertain healthcare costs, $\varepsilon$, or they can pay the excess premium $p(m_1 - m_0)$ for insurance. If the former happens, insurers are left with Akerlof’s (1970) “lemon’s market”, and premiums must increase to $m_1$. Expected utility is then given by:

$$U[w - m_0] + (1 - p)\int_0^{\infty} U[w - \varepsilon]f(\varepsilon)d\varepsilon + p\delta U[w - m_1].$$

(8)

If unhealthy consumers remain in the insurance pool, expected utility is:

$$U[w - m_0] + \delta U[w - m_0 - p(m_1 - m_0)].$$

(9)

Comparing (9) with (1), we see that consumers are made unambiguously better off under community rating compared to the “do nothing” outcome in (A) if unhealthy consumers remain in the insurance pool. The comparison is analogous to (7):

$$U[w - m_0 - p(m_1 - m_0)] > (1 - p)U[w - m_0] + pU[w - m_1].$$

(10)

Community rating may even be better than the optimal contracts of (B) if $\delta < \frac{1}{1 + r}$, since the cross-subsidization occurs in period 2 rather than period 1.

Comparing (8) to (1), we can readily see that consumers are made unambiguously worse off under community rating if the healthy consumers drop out of the pool.\(^{22}\) The unhealthy consumers end up in exactly the same predicament as in (A)—paying $m_1$ for insurance—while the healthy consumers have to bear their own cost risk. The assumption of risk aversion ensures that:

\(^{22}\) Whether community rating actually results in adverse selection is a somewhat disputed empirical question. Herring and Pauly (2006a) find some evidence of selection: higher premiums and a larger portion of uninsured consumers in community-rated individual markets as compared to unregulated markets, although the differences are relatively small in both cases.
Although (9) is strictly preferable to (8) *ex ante*, healthy consumers decide whether to remain in the insurance pool *after* their health state is revealed to them. They will continue to purchase insurance only if:

\[
U[w-m_0] \geq \int_0^\infty U[w-\varepsilon]f(\varepsilon)d\varepsilon. \tag{11}
\]

The definite reduction in utility from paying the excess premium \(p(m_1-m_0)\) must not exceed the average loss in utility from bearing the risk of \(\varepsilon\).\(^{23}\) The value of the right-hand side of (12) will depend on the distribution of \(\varepsilon\) (the probability density function \(f\)) and the degree of risk aversion. The greater the difference \(m_1-m_0\), and the more unhealthy consumers there are (larger \(p\)), the less likely it is that a healthy consumer will be willing to subsidize the unhealthy consumers’ premiums.

**D. Large-Firm Employers\(^{24}\)**

The HIPAA legislation prohibits employers from denying coverage or charging a different premium to an employee on the basis of health state. For a large firm, this restriction amounts to implicit guaranteed renewability for current employees, and a kind of safety net for unhealthy individuals who may seek a job at the firm for the sake of low-cost insurance. HIPAA restrictions are less effective in a small firm where the premiums of all employees may increase significantly, or the firm may drop coverage altogether, if there is even one employee with a chronic condition. I assume that the only way a

\(^{23}\) See Figure 3 in the Appendix.

\(^{24}\) Pauly, Kunreuther, Hirth (1995) describe this large-firm/individual-market switching procedure in some detail.
healthy employee in a large firm can avoid subsidizing the premiums of unhealthy employees is by leaving the firm: if an employee could just drop out of the insurance pool the situation would be the same as in (C). I also assume that firms have no way to discriminate against unhealthy job applicants.

I introduce some *ex ante* heterogeneity in consumers. All consumers are either self-employed or employed with a large firm.\(^{25}\) I assume \(\pi\) of them work for the large firm and the other \(1 - \pi\) are self-employed, where \(1 > \pi > 0\). If self-employed, a consumer can only purchase insurance at the risk-rated actuarially fair premium; premium risk insurance is assumed to be infeasible. If employed with the large firm, insurance must be purchased at average cost for the group. The initial allocation of consumers to employers is assumed to be optimal,\(^{26}\) and every consumer faces a cost \(c_i\) of switching, which may be different for different consumers. The subscript "\(i\)" indexes the set of consumers.

Utility in period 1 is the same for both groups: \(U[w - m_o]\). If there is no switching between employers, premiums in period 2 are \(m_o + p(m_i - m_o)\) for large-firm employees, \(m_o\) for the healthy self-employed, and \(m_i\) for the unhealthy self-employed.

We can simplify the notation by substituting \(w - m_o = w_i\) and \(m_i - m_o = x\).\(^7\) *Ex ante*,

\(^{25}\) I assume there is only one large firm. If there were many large firms with heterogeneous employee pools, inefficient switching/selection may occur between large firms as well.\(^26\) This assumption implies that, ignoring health insurance considerations, no one could be made better off by switching employers, which is the same as asserting \(c_i > 0\) for all \(i\). There may be instances when a consumer would prefer to switch employers for reasons other than lower insurance premiums. In this case, the link between the employer and health insurance could result in job lock for people with chronic conditions, especially in the absence of "portability" regulation like HIPAA Title I. See Stroupe, Kinney, Kniesner (2001).

\(^{27}\) This simplification is useful in this section because of the other complications involved. It could be made in any other section, although I believe it would detract from the main points.
employees of the large firm are obviously better off than the self-employed since they have implicit premium risk insurance:

\[ U[w_1 - px] > pU[w_1 - x] + (1 - p)U[w_1] \tag{13} \]

Total utility is:

\[ U[w_1] + \pi \delta U[w_1 - px] + (1 -\pi)\delta(pU[w_1 - x] + (1 - p)U[w_1]), \tag{14} \]

which is better than the outcome in (A) but not as good as (B) since only \( \pi \) consumers have the GR protection.

Individuals can still switch employers after their health states are revealed but before period 2 premiums are paid. Unhealthy self-employed consumers have an incentive to join a large firm (I will call this “type 1 switching”), and healthy large-firm employees have an incentive to become self-employed (type 2 switching). Initially, type 1 switching will be done by consumers for whom:

\[ U[w_1 - px] - c_i > U[w_1 - x], \tag{15} \]

and type 2 switching by consumers for whom:

\[ U[w_1] - c_i > U[w_1 - px]. \tag{16} \]

The employees with the lowest costs of switching are, of course, the most likely to switch. If any consumers switch though, selection will begin to play a role. The self-employed population will become relatively healthy, while the large firm will have a disproportionate number of high-cost workers. The more the healthy and unhealthy consumers become segregated, the higher the premium in the large firm and the more unattractive this firm is for all consumers. Let \( S_1 \) equal the proportion of unhealthy consumers who have switched from being self-employed to a large firm—initially those that satisfy (15). Let \( S_2 \) equal the analogous proportion of healthy consumers who have
switched from a large firm to being self-employed. The break-even (excess) premium in
the large firm is then:

\[ \theta x = \frac{\pi p + p(1 - \pi)S_1}{\pi(1 - S_2(1 - p)) + p(1 - \pi)S_1}. \]  \hspace{1cm} (17)\textsuperscript{28}

Note that if \( S_1 = S_2 = 0 \) we return to the large-firm premium without switching,
\[ \theta x = px. \]
We can confirm that \( \theta \) is increasing in \( S_1 \) and \( S_2 \) for \( 0 < S_1 < 1 \) and
\[ 0 < S_2 < 1: \]

\[ \frac{\partial \theta}{\partial S_1} = \frac{(1 - \pi)p}{\pi(1 - S_2(1 - p)) + (1 - \pi)S_1p} - \frac{p(1 - \pi)(\pi p + (1 - \pi)S_1p)}{(\pi(1 - S_2(1 - p)) + (1 - \pi)S_1p)^2} > 0 \]  \hspace{1cm} (18)\textsuperscript{29}
\[ \frac{\partial \theta}{\partial S_2} = \frac{\pi(1 - p)(\pi p + (1 - \pi)S_1p)}{(\pi(1 - S_2(1 - p)) + (1 - \pi)S_1p)^2} > 0. \]  \hspace{1cm} (19)

The more general forms of (15) and (16) are:
\[ U[w_i - \theta x] - c_i > U[w_i - x] \]  \hspace{1cm} (20)
\[ U[w_i] - c_i > U[w_i - \theta x]. \]  \hspace{1cm} (21)

III. Movement of workers

<table>
<thead>
<tr>
<th>Large Firm</th>
<th>Self-Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial workers: ( \pi )</td>
<td>Initial workers: ( 1 - \pi )</td>
</tr>
<tr>
<td>Healthy: ( \pi (1 - p) )</td>
<td>Healthy: ( (1 - \pi)(1 - p) )</td>
</tr>
<tr>
<td>Unhealthy: ( \pi p )</td>
<td>Unhealthy: ( (1 - \pi)p )</td>
</tr>
</tbody>
</table>

\textsuperscript{28} This is just the number of unhealthy workers who end up in the large firm divided by the total number of
workers in that firm (which is reduced by \( S_2 \) and increased by \( S_1 \)), times the excess costs of unhealthy over
healthy consumers.

\textsuperscript{29} This simplifies to:
\[ 1 - \frac{\pi p + (1 - \pi)S_1p}{\pi(1 - S_2(1 - p)) + (1 - \pi)S_1p} > 0; \]
\[ 1 - S_2(1 - p) > p; \]
\[ 1 > S_2. \]
As consumers switch employers, the premium in the large firm increases, altering
the tradeoff other consumers face when considering a switch. As $S_2 \to 1, \theta \to 1$ and (21) is
more likely to hold; the more healthy workers leave the large firm, the greater the
incentive for other healthy workers to do the same, even those with higher costs of
switching. On the other hand, (20) is less likely to be true; only unhealthy workers with
very low switching costs will seek to join the large firm since the premium is nearly the
same as the risk-rated premium they pay in the individual market. Total utility is:

$$U[w_i] + \delta L U\left[w_i - \frac{N_L}{L} x\right] + \delta (N_s U[w_i - x] + H_s U[w_i]) - C$$

(22)

where $C$ is the average switching cost incurred, $L$ is the total number of consumers who
end up in the large firm, $N_L$ is the number of unhealthy large-firm employees, $N_s$ is the
number of unhealthy self-employed consumers, and $H_s$ is the number of healthy self-
employed consumers.

Taking derivatives of (22), ignoring $C$, we observe that both kinds of switching
involve a tradeoff. There is a benefit to the consumer who switches but there is a
disutility to all other employees of the large firm, whose premiums increase:

$$\frac{\partial}{\partial S_1} = \delta p (1 - \pi) \left( U\left[w - \frac{N_L}{L} x\right] - U[w - x] + U'\left[w - \frac{N_L}{L} x\right] \left(\frac{N_L}{L} - 1\right) x\right) > 0$$

(23)$^{30}$

$$\frac{\partial}{\partial S_2} = \delta \pi (1 - p) \left( U[w] - U\left[w - \frac{N_L}{L} x\right] - \frac{N_s}{L} x U'\left[w - \frac{N_L}{L} x\right]\right) < 0$$

(24)

Type 1 switching, where unhealthy self-employed consumers join the large firm,
increases social utility on the margin: it moves us closer to a “pooling equilibrium.”

---

$^{30}$ For a proof of inequality (23), see the Appendix. An analogous proof applies to (24).
Type 2 switching is utility-reducing on the margin: it moves us closer to a “separating equilibrium.” Utility is maximized when \( S_1 = 1 \) and \( S_2 = 0 \), in which case total utility is:

\[
U[w_1] + \delta \left( p (1 - \pi) + \pi \right) \left[ w_1 - \frac{p}{\pi (1 - p) + p} x \right] + (1 - p) (1 - \pi) U[w_1].
\]  

(25)

This is not as good as the community-rating outcome in (C) when healthy consumers did not drop out of the pool—or the result in (B) assuming \( \delta \geq \frac{1}{1 + r} \)—even without including average switching costs. Pooling is incomplete: the outcome is between the pure “separating” and “pooling” equilibriums. All of the original large-firm employees, and all unhealthy consumers, end up working for the large firm. They pay a premium

\[
\frac{p}{\pi (1 - p) + p} x,
\]

which is larger than the premium in (9)\(^{31}\) because healthy consumers who started out self-employed stay that way, and do not contribute to subsidizing the premiums of unhealthy consumers. Note that \( S_1 \) and \( S_2 \) are not exogenously determined. As more switching occurs, type 1 switching becomes less likely and type 2 switching becomes more likely, which only increases the probability that the end result will be considerably worse than the ideal large-firm scenario in (25).

---

\(^{31}\) \( \frac{p}{\pi (1 - p) + p} x > px; \pi (1 - p) + p < 1; \pi < 1 \)

---

\( E. \) Government Subsidy

If the market-based premium risk insurance contracts of (B) are not feasible, the best policy alternative is a government-enforced transfer of wealth from healthy to
unhealthy consumers. If well designed, such a subsidy could produce the same outcome as with the market-based optimal contracts, while avoiding the potential selection and labor market distortion problems of (C) and (D). Assume there is a total net transfer of $t$ that occurs in period 2. Healthy individuals each pay $\frac{t}{1-p}$ and unhealthy consumers each receive $\frac{t}{p}$. Both groups purchase full insurance in all periods at the risk-rated premium. Expected utility is:

$$U[w-m_o] + (1-p)\delta U\left[w-m_0 - \frac{t}{1-p}\right] + p \delta U\left[w-m_1 + \frac{t}{p}\right]. \tag{26}$$

Maximizing (26) with respect to $t$, we can verify that the optimal transfer equalizes wealth across states of the world; there should be a payment of $(1-p)(m_1 - m_o)$ to unhealthy consumers funded by a tax of $p(m_1 - m_o)$ on healthy consumers:

$$\frac{\partial}{\partial t} = -U'[w-m_0 - \frac{t}{1-p}] + U'[w-m_1 + \frac{t}{p}] = 0 \tag{27}$$

$$t = p(1-p)(m_1 - m_o). \tag{28}$$

Exactly as in (B), the premium risk has been eliminated: the loss is certain in all periods. In period 1 it is $m_o$ and in period 2 it is $m_o + p(m_1 - m_o)$. Since the government has the ability to finance the subsidy in any period, it is unconstrained by interest rate or discounting considerations. Consumers are made unambiguously better off compared to the situation in (A).

\[32\] van de Ven, et al. (2000) reach the same conclusion. Although their approach and perspective are quite different from my own, their empirical simulations generally confirm the findings here.
4. Model Notes

Note 1: Savings

Depending on the interest rate, consumers can be made better off in all scenarios if they are allowed to save and borrow. In (B), PRI’s entire excess premium has to be financed in period 1 to ensure that healthy consumers contribute. In (A), unhealthy consumers can only pay their higher premiums in period 2. Welfare would be improved if individuals could smooth their consumption over time. Assume consumers can carry net savings of \( \gamma \) between the two periods at the same interest rate insurance companies earn, \( r \). If consumers have a total expense \( P \) that occurs in period 1, as with premium risk insurance, utility is:

\[
U\left[ w - P - \frac{\gamma}{1+r} \right] + \delta U\left[ w + \gamma \right].
\] (29)

Optimal savings requires that:

\[
\frac{\delta}{\delta \gamma} = -\frac{1}{1+r} U'\left[ w - P - \frac{\gamma}{1+r} \right] + \delta U'\left[ w + \gamma \right] = 0
\] (30)

\[
U'\left[ w - P - \frac{\gamma}{1+r} \right] \over U''\left[ w + \gamma \right] = \delta(1+r).
\] (31)

The ratio of marginal utility in the first period to that in the second period must be proportional to the interest rate and inversely proportional to the discount rate. If \( \delta = \frac{1}{1+r} \), marginal utilities should be equated in the two periods, which means the level of consumption should be the same: consumers should borrow \( \frac{P(1+r)}{2+r} \) from the second period to pay for half of their excess premium in the first period. In this case, it is quite
irrelevant when a loss occurs; consumers will always distribute it evenly between periods. Combining this result with the one in (E)—that wealth evenly distributed between states of the world is optimal—we can readily see that the premium risk insurance of (B) and the subsidy of (E) might both represent optimal policies. Since capital markets may be constrained—consumers may not be able to borrow at \( r \), it is possible that \( \delta \neq \frac{1}{1+r} \)—PRI will not always be optimal depending on our assumptions.

Since the government can collect taxes in any period, and borrow at the market-lowest interest rate, the subsidy system will always be optimal.\(^{33}\)

If we allow savings in scenario (A), utility is:

\[
U\left[w - m_o - \frac{\gamma}{1+r}\right] + (1-p)\delta U[w - m_o + \gamma] + p \delta U[w - m_1 + \gamma].
\]  

(32)

Assuming \( \delta = \frac{1}{1+r} \), savings are optimal when:

\[
U\left[w - m_o - \frac{\gamma}{1+r}\right] = (1-p)U\left[w - m_o + \gamma\right] + p U\left[w - m_1 + \gamma\right],
\]

(33)

which implies some positive amount of savings such that \( \gamma < \frac{m_1 - m_o}{1+\frac{1}{1+r}} \). Though ex ante expected utility is improved, healthy consumers will inevitably have saved too much and unhealthy consumers too little.\(^{34}\) In my model, subsidizing savings through a tax

---

\(^{33}\) Insurers may be able to earn an interest rate that is higher than what the government could earn, for example because they are allowed to invest in equities. I assume the government could either borrow and subsidize the (mandatory) purchase of private PRI in period 1, or else the government is unconstrained in its investing.

\(^{34}\) If consumers had known what their health states would be, they would have equated consumption in the two periods. For a healthy consumer, this implies no saving. For an unhealthy consumer, a little algebra verifies that
deduction (e.g. with health savings accounts) would increase the amount of saving. Financing the deduction with a fixed levy on all consumers, however, would reduce efficiency: all consumers would save more, while incomes would be reduced dollar-for-dollar with the tax benefit. Distorted incentives would result in too much saving by all parties: savings were already optimal without the tax incentive. If the deduction were financed by a tax only on healthy consumers, the result would be the same as in (E): forced premium risk insurance. Healthy consumers would subsidize unhealthy consumers’ premiums.

Note 2: Moral Hazard

There are two kinds of moral hazard, in the sense of Pauly (1968), that could be relevant to the problem of premium risk. On the one hand, consumers with insurance may purchase care that is valued below marginal cost, since they do not bear the full expense. If consumers are more likely to engage in this “consumption moral hazard” when they have a chronic condition, the difference between “unhealthy” and “healthy” premiums, \( m_1 - m_0 \), may be excessive. The informational content of premium differentials is distorted. There is also a risk of “behavior moral hazard”: consumers may be less likely to undertake costly behavioral changes that could reduce their risk of illness, such as changes in diet or exercise, knowing their premiums will not increase if they acquire a chronic condition. Reducing either form of moral hazard would require

\[
y = \frac{m_1 - m_0}{1 + \frac{1}{1 + r}}
\]

this implies saving \( y \), which is an inefficiently high amount \textit{ex ante}.

De Meza (1983) points out that increased consumption of healthcare by insured consumers can result from income effects, rather than moral hazard, and is not necessarily inefficient.
PRI with coinsurance or only a partial public subsidy. It is also worth noting that the government will have an incentive under the subsidy system to invest in programs that reduce long-term healthcare expenditures; in contrast, Cochrane’s premium risk insurance providers would have no such incentive: they are not necessarily responsible for paying future benefits to current clients.

Note 3: Ex Ante Heterogeneity

In my model, the only ex ante heterogeneity appears in (D), where consumers can have different costs of switching employers. Including this heterogeneity and no other is an arbitrary decision; surely consumers differ on all of the dimensions in the model. They may have different attitudes toward risk (differently shaped utility functions), different wealth and income, different discount factors, and so on. In (B), consumer heterogeneity may result in only partial uptake of PRI if some consumers have particularly high discount rates. The potential for adverse selection in (C) is essentially the same as in (D): had I allowed for consumers with different degrees of risk aversion in (C), we would have had the “gradual” selection outcome of (D), rather than an all-or-none choice. Only some consumers would have self-insured, and community-rated insurance would become less and less attractive the more healthy consumers left the pool. The subsidies in (E) would still be preferred by all consumers as long as we retained the assumption of risk aversion.

A more damaging kind of heterogeneity for PRI would be if consumers differed in their ex ante probability of illness, \( p \), or in the cost shocks they faced in different states of the world, \( \varepsilon \) and \( \mu \). If consumers and insurers had the same information about
these variables, future premium risk would still be insurable. The actuarially fair
premium charged for PRI would differ depending on the underlying risk. Only the
uncertain part of costs could be insured. For example, if some people already had a
chronic condition at the beginning of period 1, PRI would do nothing for them. The
problem is more severe if information is asymmetric. Consumers would be subject to
Rothschild and Stiglitz's (1976) adverse selection, perhaps resulting in a separating
equilibrium or, in the extreme case, Akerlof's (1970) lemons market, where only high-
cost consumers purchase PRI. The subsidy program could still be implemented in either
case. However, with this kind of *ex ante* heterogeneity, subsidies take on a slightly
different character, and may not be preferred by all consumers. If subsidies were used to
equalize premiums, consumers who knew themselves to be healthier than average when
the program began—ignoring possible externalities—would have no incentive to
subsidize the premiums of consumers who started out unhealthy. Previously, all of the
premium risk was in the future, and all consumers agreed to the subsidy because of the
chance that they themselves would benefit. In this case, healthy consumers know they
will benefit less than average, and the benefit may not be worth the cost. The subsidy
program is essentially community-rated PRI, except that consumers can be forced to
participate.

*Note 4: Additional Periods*

Adding more periods to the model may mitigate the selection problems in (C) and
(D). In (B), multiple periods would require PKH's sequence of premiums that decline
over time as more information is revealed. Premiums remain "actuarially fair" with
respect to future risk. With positive costs of switching, between plans in (C) and employers in (D), additional periods could encourage consumers to remain in their current insurance pool even as they are revealed to be healthy. For example, a large-firm employee who did not acquire a chronic condition before period 2 would be more likely to switch to the individual market if that were the final period than if there were still a chance that she could acquire a chronic condition in periods 3 and beyond. She could switch to the individual market, and then switch back if she becomes chronically ill, but that would require incurring switching costs twice.

**Note 5: Administrative Costs**

Including administrative costs would affect several of the scenarios. In (B), if administrative costs are a fixed amount, for example selling costs per policy, they make it less likely that consumers will purchase PRI. If administrative costs vary with the amount of coverage, for example capital costs that are a percentage of claims, consumers will generally prefer PRI with a deductible. A variable loading charge acts as a tax on the insurance coverage, and consumers are better off bearing some premium risk themselves. The optimal subsidy of (E) will be similarly reduced if administrative costs are proportional to the size of the transfer.

Administrative costs are generally much higher in the small-group and individual markets—normally accounting for 30-40% of the total premium—than they are for large-group pools. This is mostly due to higher costs of selling policies and of determining consumers' relative risks. In (C), community rating could lower the expenses of

---

36 See Arrow (1963).
37 See Pauly and Nichols (2002).
“medical underwriting,” making individual insurance more attractive, and reducing
incentives for healthy consumers to self-insure (relative to the scenario outlined in the
model). In (D), higher administrative costs in the individual market make it less likely
healthy consumers will leave large firms, and more likely unhealthy consumers will join
them. This brings us closer to the “(partial) pooling equilibrium,” $S_1 = 1$ and $S_2 = 0$,
which is the best we can do with this policy.

Note 6: Income as a Function of Health

Earning potential is probably reduced by the presence of a chronic condition:
unhealthy consumers may be less productive, they may miss work because of illness or to
visit a doctor, and they may face employer discrimination. In my model, it is easy to see
that a reduction in an unhealthy consumer’s period-2 income would be the same as an
increase in $m_1$. Since the optimal tax transfer of (E) equalizes wealth across states of the
world, any decrease in income resulting from a chronic condition should be fully
compensated. Such compensation may not be desirable if lost earning potential is
unobservable, or if the transfer mechanism would distort incentives, for example by
discouraging individuals from working. Because of the existence of Medicaid, the
government program that insures low-income Americans, there is a possibility that
having income vary with health status would result in a selection of unhealthy consumers
into Medicaid, mitigating premium risk in private insurance markets.
Note 7: Community Rating and Large-Firms

Current policy toward premium risk, and the kind of policy most commonly advocated by (non-economist) commentators on the topic, is some combination of rate bands in the individual market and anti-discrimination laws in the group market. Combining (C) and (D) will not discourage healthy consumers in the individual market from self-insuring. It may mitigate the adverse selection problem of (D). In particular, if both groups start period 2 with the same proportion of unhealthy consumers, there will be no incentive for anyone to switch employers in the first place. The single premium in the large firm will be the same as the community-rated premium in the individual market. However, this could be an unstable equilibrium. Healthy consumers can still choose to become self-employed and self-insured. If they stop purchasing insurance altogether the societal outcome will be even worse than the worst possible switching outcome in (D), $S_2 = 1$, which is the complete “separating equilibrium.” In this case, unhealthy consumers are left paying risk-rated actuarially fair premiums in both the individual and group market, while healthy consumers bear their full period-2 cost risk $c$, in addition to incurring switching costs. Alternatively, differences in the initial risk profile of each group could lead to higher premiums in one or the other and potentially costly switching by both healthy and unhealthy consumers into the lower-cost pool until premiums equalized.

Since a large-scale move to self-insurance appears unlikely, a combination of policies (C) and (D) can seem like an attractive option. There are many other reasons why a government subsidy would be a superior policy. Rating restrictions are usually implemented and enforced on the state level. Rate bands differ greatly across states, and
some states impose no restrictions at all. Without PRI, people in low-restriction states will bear at least some of their premium risk, and there is a possibility of inefficient selection across states (healthier consumers move to low-restriction states, unhealthy consumers do the opposite). Even if national rate bands were adopted, there are significant problems of enforcement.

The purpose of rate bands is to force insurers to sell policies to high-cost clients at an expected loss, which creates a strong incentive for insurers to target their marketing efforts toward healthier consumers, and do everything they can to avoid taking on unhealthy clients. There may be relevant social-equality considerations since low-income and several minority groups tend to have more chronic conditions. There will also be an incentive for insurers to risk-rate by varying the level and type of coverage offered to different clients. Preventing this would require costly mandated-benefits legislation. Almost all observers agree that some rate variability is required for insurance markets to survive; in particular, some geographic areas have particularly expensive or inexpensive medical care, and elderly consumers generally require more care regardless of health status. Since both geographic region and age can be correlated with the presence of chronic conditions, it is difficult, if not impossible, for government agents to distinguish between allowable and inappropriate premium differentials. Finally, this line of regulation reduces competition and product differentiation, which may encourage all insurers to skimp on quality, while inhibiting innovation. Imagine what any other product market would look like if the government legislated the product’s price and all of its features. These problems are entirely avoided with the proposed subsidy program. A similar argument applies to proposals for a system of “universal healthcare.”
**Note 8: Premium Risk Insurance and Other Policies**

In scenarios (C), (D), and (E) above I implicitly assume that there is no private PRI available. In (C), PRI would be illegal since it would require insurers to charge differential premiums in period 2 depending on whether a consumer purchased PRI in period 1. In (D), large-firm employees could be forced to pre-pay premiums in period 1. Consumers could still game the system by being self-employed in period 1 and joining large firms in period 2 if they ended up unhealthy, receiving PRI benefits without having paid for them. In (E), consumers will have no incentive to purchase PRI since they already receive it from the government.

**Note 9: Government Inefficiency**

The government subsidy system has the same potential drawbacks as most public programs. There is no opt-out mechanism. Since administrative costs prevent the government PRI from being offered at its actuarially fair value, consumers who are nearly risk-neutral may prefer to bear their own premium risk. Consumers who know themselves to be particularly healthy ex ante may not be willing to subsidize premiums of less fortunate compatriots. Ideally, the subsidy system should be financed with a fixed per-capita tax levy. This may be politically unpalatable, and alternative financing mechanisms like increased income taxes may be distortionary. A public program could be more administratively inefficient than market-based PRI would be, and it may permanently prevent such PRI from arising. The public system could be subject to fraud, and it may be difficult or costly to determine who should receive what level of benefits (although the same problem exists with market-based PRI). Insurers would need to be

---

paid lump-sum, health-status-contingent benefits to prevent distorted incentives at the margin, similar to Medicare’s “diagnosis-related groups” (DRG) system. The government would have to be careful to curb insurers’ incentives to exaggerate diagnoses. Finally, a “slippery slope” may be a concern if the government extends its new authority as a premium-risk insurer into less appropriate interventions.

5. Conclusions

The current approach of public policy toward premium risk is misguided. It emphasizes premium restrictions in the individual insurance market and anti-discrimination laws in the group market. At best, these regulations provide only partial protection against premium risk. At worst, distorted incentives result in selection and welfare losses. Private premium risk insurance, despite its potential benefits, has so far failed to appear. This may result from moral hazard, adverse selection, medical inflation risk, lock-in problems, or the uncertainty of future regulation. There may be a lack of demand for PRI because of excessive consumer discounting or a cognitive bias that makes healthy consumers underestimate their chances of acquiring a chronic condition. The existence of alternative regulations may further limit the appeal of private PRI. Even if it existed, PRI could do nothing for consumers who already have chronic conditions, or who are congenitally predisposed to them. Subsidy programs implemented thus far have been incomplete. They have suffered from under-funding, adverse selection, and eligibility rules that limit their effectiveness and obscure their purpose. The optimal subsidy program should be viewed as government-provided premium risk insurance: a private market for PRI has failed to develop and the government is in an excellent
position to fill the gap. Premium risk should continue to be an issue of highest priority in reforming the U.S. health insurance market. The most appropriate intervention would be to correct the specific market failure, while avoiding new distortions.
References


Appendix

What follows is a proof of the inequality in (23). An analogous proof applies to (24).

Using a second-order Taylor series expansion, we find that:

\[ U[w_i - x] = U\left[w_i - \frac{N_L}{L} x\right] - \left(x - \frac{N_L}{L} x\right) U'\left[w_i - \frac{N_L}{L} x\right] \left(\frac{x - \frac{N_L}{L} x}{2}\right) + \frac{(x - \frac{N_L}{L} x)^2}{2}\ . \]

This can be rewritten:

\[ U\left[w_i - \frac{N_L}{L} x\right] - U[w_i - x] + \frac{(x - \frac{N_L}{L} x)^2}{2} U'\left[w_i - \frac{N_L}{L} x\right] = \left(x - \frac{N_L}{L} x\right) U'\left[w_i - \frac{N_L}{L} x\right] \]

Since \( U'' < 0 \), we can make the substitution \( \omega = -\frac{(x - \frac{N_L}{L} x)^2}{2} U'\left[w_i - \frac{N_L}{L} x\right] > 0 \). Therefore,

\[ U\left[w_i - \frac{N_L}{L} x\right] - U[w_i - x] = \left(x - \frac{N_L}{L} x\right) U'\left[w_i - \frac{N_L}{L} x\right] + \omega \]

Substituting this into (23) results in:

\[ \left(x - \frac{N_L}{L} x\right) U'\left[w_i - \frac{N_L}{L} x\right] + \omega + U'\left[w_i - \frac{N_L}{L} x\right]\left(\frac{N_L}{L} - 1\right)x > 0 \]

Simplifying, we have: \( \omega > 0 \),

which is true.
For simplicity, the following graphs are shown for the case where there are only two possible revelations of one-period health states: perfect health and illness. The curve is a concave utility-of-wealth function. Both healthy and unhealthy consumers can be in perfect health, in which case they have wealth \( w \). The lines marked “healthy” and “unhealthy” connect the utility function at \( w \) with its level at wealth in the “illness” states of the world. Illness events are more costly for unhealthy consumers, and they occur with higher probability. The unmarked dots are a distance from \( w \) along their respective lines equal to the probability of illness for each type of consumer. Both types of consumer purchase insurance after their health state is revealed. The expected loss for healthy consumers is \( m_0 \), and they end up at point A. For unhealthy consumers, the expected loss is \( m_1 \), and they end up at point B. Ex ante expected utility is denoted by C. The “premium risk” has an actuarially fair value of \( p(m_1-m_0) \), where \( p \) is the probability a consumer will be “unhealthy.”
Expected utility in the "Do Nothing" scenario, where premium risk insurance (PRI) is not available, is denoted by A. With PRI, consumers pay a guaranteed premium of $m_0 + p(m_1 - m_0)$ regardless of health status, which results in expected utility at B. Because all premium risk has been eliminated, consumers are better off with PRI, ignoring discounting considerations.
If community rating is enforced, healthy consumers are left with a choice: go uninsured and bear their full one-period cost risk (point A, utility level $U_1$), or subsidize the premiums of unhealthy consumers (point B, utility level $U_2$). If $U_1 > U_2$, there will be a full separating equilibrium, with healthy consumers at A and unhealthy consumers at C. Unhealthy consumers are no better off than they would have been without community rating, and healthy consumers are worse off.