



University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

Spring 5-16-2011

The Impact of Medical Spending Growth on Guaranteed Renewable Health insurance

Robert D. Lieberthal
liebertr@wharton.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Economics Commons](#)

Recommended Citation

Lieberthal, Robert D., "The Impact of Medical Spending Growth on Guaranteed Renewable Health insurance" (2011). *Publicly Accessible Penn Dissertations*. 325.
<http://repository.upenn.edu/edissertations/325>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/325>
For more information, please contact libraryrepository@pobox.upenn.edu.

The Impact of Medical Spending Growth on Guaranteed Renewable Health Insurance

Abstract

I examine the problem of writing guaranteed renewable health insurance in the presence of medical spending growth. Prior research suggests that the growth and difficulty in forecasting future medical costs is an impediment to multiperiod health insurance, where contract reserves are used to pay a portion of the benefits in later years of the contract. Medical spending growth is an input to calculating the magnitude of premiums and reserves, so setting up reserves to pay future claims involves forecasting spending growth. Hedging assets can ameliorate the investment problem by providing assets that automatically adjust to unexpected shocks in spending growth.

I expand an existing model of guaranteed renewability in an economy with risk to show the specific ways that medical spending growth enters the premium and reserve functions. I treat stochastic trend as a factor the insurance company can predict with error. I utilize aggregate and individual level insurance spending data and financial returns data to analyze whether medical trend can be hedged with existing assets. I separate trend into predictable and error components and analyze the correlation between the error component and return on assets. I find that medical spending growth is predictable with error over short and medium time horizons. I find that there is no significant correlation between asset returns and forecast errors across several broad asset classes.

The combination of partially predictable spending growth and the absence of a hedging asset imply that insurers should be using reserves to manage the macroeconomic risk of spending growth. The load for reserving for trend is an up-front cost in addition to the up-front expense of guaranteed renewability. Insurers should use a diversified investment strategy for reserves rather than one targeted at trying to match spending growth. I conclude by noting the positive and negative effects of the newly passed health reform law (PPACA) on guaranteed renewable health insurance and other health insurance arrangements that require contract reserves and policies that shift health care spending onto public plans.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Managerial Science and Applied Economics

First Advisor

Mark V Pauly

Keywords

Guaranteed renewability, health insurance, health finance, medical inflation

Subject Categories
Economics

THE IMPACT OF MEDICAL SPENDING GROWTH ON GUARANTEED
RENEWABLE HEALTH INSURANCE

Robert D. Lieberthal

A DISSERTATION

In

Health Care Management & Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

In

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2011

Supervisor of Dissertation

Mark V. Pauly, Professor of Health Care Management, Business and Public Policy, Insurance and Risk Management, and Economics

Graduate Group Chairperson

Eric Bradlow, Professor of Marketing, Statistics, and Education

Dissertation Committee

Scott E. Harrington, Professor of Health Care Management, Insurance and Risk Management

Gregory Nini, Assistant Professor of Insurance and Risk Management

Jessica A. Wachter, Associate Professor of Finance

Dedication

I dedicate my dissertation to my grandparents, Molly and Hirsh Lieberthal and Charles and Esta Rosenthal. All are of blessed memory. Everything that they were allows me to be who I am.

Acknowledgments

My dissertation committee provided me with a tremendous amount of guidance and feedback. I am grateful to Mark Pauly for all the help he has given me during the dissertation process as well as my doctoral experience leading up to the dissertation process. I am grateful to Scott Harrington for the feedback he provided on my dissertation and on my dissertation proposal. I am grateful to Greg Nini for his input on my dissertation and his availability to talk informally about the job market and my plans for the future. I am grateful to Jessica Wachter for her input on my dissertation and for helping me to make finance a part of my research.

I am grateful for financial support I received in the form of a dissertation grant from the Agency for Healthcare Research and Quality (grant number R36-HS018835-01). I am grateful to Brenda Harding at AHRQ, who helped me in all stages of the application process.

I am grateful for financial support I received from Wharton through the Fifth Year Dean's Fellowship.

I am grateful to the National Bureau for Economic Research for providing MarketScan

claims data. I am grateful to Jean Roth who was instrumental in getting me up to speed with the data and the NBER system.

I am grateful for support I received from the Wharton Health Care Management department and the Leonard Davis Institute for Health Economics.

I am grateful to Joanne Levy for all of her help and support during my time at Wharton. I am grateful to her help during my application to the program, her advice on classes and prelim exams, and all of her guidance during the dissertation process. I appreciate the way that her door is still open for me.

I am grateful to Drew Griffen, and Dominic Draye, Tanguy Brachet, and many others who read or commented on various versions of my thesis.

I am grateful to my family for all their support. I am grateful to my parents who helped get me to today. I am grateful to my brother and sister, who always believe in me. I am grateful to my wife, who heard more about my dissertation than anyone else. I am grateful to my son, who reminds me that there's more to life than work.

ABSTRACT

THE IMPACT OF MEDICAL SPENDING GROWTH ON GUARANTEED RENEWABLE HEALTH INSURANCE

Robert D. Lieberthal

Mark V. Pauly

I examine the problem of writing guaranteed renewable health insurance in the presence of medical spending growth. Prior research suggests that the growth and difficulty in forecasting future medical costs is an impediment to multiperiod health insurance, where contract reserves are used to pay a portion of the benefits in later years of the contract. Medical spending growth is an input to calculating the magnitude of premiums and reserves, so setting up reserves to pay future claims involves forecasting spending growth. Hedging assets can ameliorate the investment problem by providing assets that automatically adjust to unexpected shocks in spending growth.

I expand an existing model of guaranteed renewability in an economy with risk to show the specific ways that medical spending growth enters the premium and reserve functions. I treat stochastic trend as a factor the insurance company can predict with error. I utilize aggregate and individual level insurance spending data and financial returns data to analyze whether medical trend can be hedged with existing assets. I separate trend into predictable and error components and analyze the correlation between the error component and return on assets. I find that medical spending growth is predictable with error over short and

medium time horizons. I find that there is no significant correlation between asset returns and forecast errors across several broad asset classes.

The combination of partially predictable spending growth and the absence of a hedging asset imply that insurers should be using reserves to manage the macroeconomic risk of spending growth. The load for reserving for trend is an up-front cost in addition to the up-front expense of guaranteed renewability. Insurers should use a diversified investment strategy for reserves rather than one targeted at trying to match spending growth. I conclude by noting the positive and negative effects of the newly passed health reform law (PPACA) on guaranteed renewable health insurance and other health insurance arrangements that require contract reserves and policies that shift health care spending onto public plans.

Contents

1	Introduction	1
1	Defining long term health insurance	1
2	Problems specific to GR	6
2.1	Prefunding required	6
2.2	Prefunding may be expensive	8
2.3	Prefunding may be risky	9
2.4	Problems in related insurance products	11
2.5	Comparison between GR and annual renewal insurance	12
3	Importance of GR for insuring long term health	14
3.1	Serially correlated medical spending	14
3.2	Spending is growing	15
3.3	Best possible prefunding solution	15
3.4	Time limits of protection	17
3.5	General application to long term health insurance	17
3.6	Policy implications	18

4	New contributions	21
4.1	Estimate the growth and variance of medical spending	21
4.2	Effect of spending trends on guaranteed renewable plans	22
4.3	Determining the optimal funding policy	24
4.4	Defining the residual risk	24
5	Approach	25
2	Literature review	27
1	Theoretical background	28
1.1	Guaranteed renewability as a one-sided commitment	28
1.2	Financial management of guaranteed renewability	33
1.3	Motivation for hedging	35
1.4	Theory of optimal reserves	37
1.5	The role of double taxation	39
1.6	Ideal investment for guaranteed renewability	40
1.7	Securities market incompleteness	42
2	Empirical literature	44
2.1	Guaranteed renewability is an established benefit	44
2.2	Actuaries and state regulators use a class average standard	45
2.3	Possibility of successful guarantees	46
2.4	Medical spending level and growth	46
2.5	Hedging liabilities	49

2.6	Securities to hedge medical spending	51
2.7	Other plans with guarantees have failed	53
3	Guaranteed renewable insurance in an economy with risk	55
1	Introduction	55
2	Model economy	57
2.1	Discrete risky economy	57
2.2	Aggregate risky economy	60
3	Guaranteed renewable solution	63
3.1	Fair premium	64
3.2	Contract reserve	65
3.3	Final period contract	66
4	Alternatives to guaranteed renewability	67
4.1	Single premium insurance	67
4.2	Spot insurance	68
4.3	Enforceable long term contracts	70
5	Limits of guarantees	71
5.1	Temporal limits of guaranteed renewability	71
5.2	Investment policy limitations	72
6	Extensions of the guaranteed renewable model	73
6.1	Discounting	74
6.2	Death	75

6.3	Finite time horizons	76
7	Lapsation	77
7.1	High type lapsation	77
7.2	Low type lapsation	78
7.3	Random lapsation	78
8	Partial insurance	79
9	Final remarks	80
10	Tables and figures	82
4	Investment policy for guaranteed renewability	88
1	Introduction	88
2	Numerical example	89
2.1	Numerical parameters	90
2.2	Duration mismatch under discounting	91
2.3	Matching duration with an investment policy	93
3	Mathematical duration of this contract	94
3.1	Duration of loss payments	94
3.2	Duration of premium payments	95
3.3	Asset liability management for a static two period model	95
3.4	Trend in a two and a three period model	97
3.5	Trend in a three period model	99
4	Trend in general guaranteed renewable insurance	101

4.1	Notation for trend	101
4.2	The effect of deterministic trend	102
5	Single period investing	104
5.1	One period example	104
5.2	Investment problem setup	105
5.3	State contingent claims	106
5.4	Stock and bond investment	107
5.5	Risk free bond	108
6	Multiperiod investment policy	108
6.1	General investment problem	109
6.2	Two period model	111
6.3	Investment policy with changing interest rates	112
7	Conclusion	115
5	Analysis of aggregate insurance premiums	116
1	Motivation	117
1.1	Predictability of spending growth	117
1.2	Finding hedges	118
1.3	Determining the unhedged portion of risk	120
2	Aggregate data sources	121
2.1	National Health Expenditure Survey	121
2.2	Aggregate medical care statistics	123

2.3	Securities returns	125
3	Results	126
3.1	Regression analysis	126
3.2	Adaptive expectations analysis	128
4	Implications	131
4.1	Insurance plan specific data	131
4.2	Optimal reserves	132
4.3	The effect of health reform	133
5	Tables and figures	134
5.1	Summary statistics	135
5.2	Price data	139
5.3	Employment data	142
5.4	Securities data	144
5.5	Medical spending regression results	147
5.6	Adaptive expectation results	151
5.7	Figures	155
6	Analysis of individual level insured spending	158
1	Introduction	158
2	MarketScan data	159
2.1	Population selected	160
2.2	Spending data used	161

3	Spending analysis	162
3.1	Level of spending	162
3.2	Trend of spending	165
4	Analyzing trend	167
5	Hedging trend	170
5.1	Returns on assets	170
5.2	Excess returns on assets	172
6	Remaining differences in volatility	175
6.1	Importance of volatility	175
6.2	Differences by sex	176
6.3	Differences by age	177
6.4	Implications for long term health insurance	178
7	Tables and figures	180
7	Effect of spending growth on guaranteed renewable reserves	198
1	Determinants of the size of reserves	198
1.1	Rationale for reserving against spending growth	198
1.2	Consequences of no hedging assets	199
1.3	Difficulties for regulators in observing reserves	202
1.4	Difficulties for insurers in observing parameters	205
2	Reserves and the problem of insurer follow through	206
2.1	Role of reserves in one-sided commitment	206

2.2	Reserves and the incentives to renege	208
2.3	Reserves and ruin	209
3	Impact of spending growth on reserving	210
3.1	Size of spending growth	210
3.2	Variability of spending growth	211
3.3	Deterministic numerical example	212
3.4	Stochastic numerical example	213
3.5	Effect of less predictability	215
4	Solutions derived from problems of reserving	216
4.1	Limits to reserves	216
4.2	Possible policies	218
5	Conclusion	219
5.1	Reserves and overall uncertainty	219
5.2	Pricing without correlated assets as a benchmark	220
5.3	Other impediments and costs to reserving	221
5.4	Risk neutral discounting	223
6	Tables	226
8	Conclusion	229
1	Introduction	229
2	Policies for improving guaranteed renewability	230
2.1	Making additional asset classes	230

2.2	Partial guarantees may be possible	232
2.3	More transparent accounting could help	233
3	Future research	234
3.1	Applications to other insurance programs	234
3.2	Patient Protection and Affordable Care Act	235
4	Final remarks	238
4.1	Spending growth path may change in the future	238
4.2	Variance could be a problem if leverage is high	239
4.3	Implications for other types of insurance	240
A Theory appendix		241
1	Proofs	241
2	Maximization problems	244
2.1	State contingent claims–zero losses	245
2.2	Risky stock–single period solution	246
3	Infinite period premium	247

List of Tables

3.1	Measure of types	83
3.2	Measure of types over time	84
5.1	Growth in nominal medical spending, nominal GDP, and population	135
5.2	Per capita nominal insurer expenditures, rate of change (%)	136
5.3	Unit root test of per capita nominal insurer expenditures, 1971-2008	137
5.4	Unit root test of per capita nominal insurer expenditures, 1982-2008	138
5.5	Consumer Price Index, log monthly change (%)	139
5.6	Producer Price Index, sectors, log monthly change (%)	140
5.7	Producer Price Index, sectors, log monthly change since 1992 (%)	141
5.8	Health industry employment, monthly log change (%)	142
5.9	Health industry employment, monthly log change (%)	143
5.10	Bond monthly nominal returns (% , continuous log basis)	144
5.11	Stock monthly nominal returns (% , continuous log basis)	145
5.12	Correlations of monthly inflation and nominal returns, 1982–2008	146
5.13	Per capita nominal insurer expenditures regressed on lagged expenditures	147

5.14	Per capita nominal insurer expenditures regressed on lagged expenditures	148
5.15	Per capita nominal insurer expenditures regressed on inflation and employ- ment growth	149
5.16	Per capita nominal insurer expenditures regressed on multiple variables	150
5.17	Adaptive expectations average rates for per capita nominal insurer expen- ditures, 1982–2008	152
5.18	Adaptive expectations forecast rates for per capita nominal insurer expen- ditures, 2009	153
5.19	Adaptive expectations errors for per capita nominal insurer expenditures, 1982–2008	154
6.1	Covered population by age group, 2000-2006	181
6.2	Covered population by age group and sex, 2000-2006	181
6.3	Total nominal spending per member per day, by year	182
6.4	Plan nominal spending per member per day, by year	182
6.5	Total nominal spending per member per day, by month	182
6.6	Plan nominal spending per member per day, by month	183
6.7	Nominal spending per member per day, weekend vs. weekday	183
6.8	Total nominal spending by sex	183
6.9	Plan nominal spending by sex	184
6.10	Unit root test of daily nominal spending by population	184
6.11	Daily log change in total nominal spending by demographic group	185

6.12	Average population by group	185
6.13	Nominal spending regression	186
6.14	Nominal spending regression	186
6.15	Regression results by demographic group for total nominal spending	187
6.16	Regression results by demographic group for plan nominal spending	188
6.17	Daily nominal return less daily nominal trend, weekend included	189
6.18	Excess nominal return regressed on residuals	190
7.1	Three period guaranteed renewable reserve–no trend	227
7.2	Three period guaranteed renewable reserve with trend	228

List of Figures

3.1	Evolution of Membership	85
3.2	Two period model with insurance	86
3.3	Set of acceptable premiums under full, zero-load insurance	87
3.4	Acceptable premiums for insuring each type	87
5.1	Growth rate of nominal GDP and nominal medical spending per capita . . .	156
5.2	Nominal insured medical spending per capita, year on year rate of change .	157
6.1	Count of members by age	191
6.2	Count of members by age and sex	191
6.3	Per capita drug experiences by date	192
6.4	Per capita outpatient experiences by date	192
6.5	Total spending per capita	192
6.6	Plan spending per capita	193
6.7	Histogram of total spending per capita	193
6.8	Histogram of plan spending per capita	193
6.9	Drug spending per capita	194

6.10 Inpatient spending per capita 194

6.11 Outpatient spending per capita 194

6.12 Outpatient, out-of-pocket deductible payments per member 195

6.13 Autocorrelation for daily trend 195

6.14 Partial autocorrelations for daily trend 195

6.15 Scatter plot of residuals 196

6.16 Histogram of residuals 196

6.17 Autocorrelation for daily trend 196

6.18 Partial autocorrelations for daily trend 197

6.19 Outpatient spending per capita by males 197

6.20 Outpatient spending per capita by females 197

Chapter 1

Introduction

1 Defining long term health insurance

The serial correlation in medical spending is an important reason to have long term health insurance. Eichner, McClellan and Wise (1998) find that “the median period of unusually high expenses is only about 4 years,” which suggests that the overall autocorrelation of individual spending is not substantial. However, splitting medical expenses between “transitory” and “persistent”, De Nardi, French and Jones (2009) find that the persistent component explains only one-third of variance in medical expenses, but that the autocorrelation of the persistent component is 0.922 “. . . so that innovations to the persistent component of medical expenses have long-lived effects. Most of a household’s lifetime medical expense risk comes from the persistent component.” The degree of persistence is consistent with the notion that chronic conditions that arise randomly and persist make individuals difficult to insure. If consumers could prefund the possibility of entering a persistent high cost state

through long term insurance, they would reduce exposure to a major risk.

Long term health insurance policies finance medical costs over a long time horizon. Unlike annually renewed health insurance, health insurance contracts with multiyear obligations, such as guaranteed renewable health insurance and long term care insurance, cover the ongoing high costs of illness. There is a high degree of serial correlation in medical spending, which makes such long term protection attractive for individuals. The upside of a long term contract is relative certainty for the buyer and the insurer over the cost of health insurance in future periods, as well as savings on the search and underwriting costs of new insurance policies. The downside for insureds is high upfront costs needed to set up the commitment underlying the contract.

The recent changes to health insurance law enacted as part of the Patient Protection and Affordable Care Act will both ameliorate and exacerbate the problems of long term health insurance. The mandate to purchase health insurance could make guaranteed renewability more attractive for insurers if it increases persistence in the nongroup insurance market. It could also decrease the attractiveness of guaranteed renewability by restricting the range of allowable medical loss ratios for nongroup health insurance (U.S. Congress 2010). It could also increase or decrease the attractiveness of guaranteed renewable health insurance depending on whether the law succeeds in bending the cost curve (Orzag 2009).

Insurance companies that want to write long term health contracts face problems dealing with macroeconomic risks unrelated to the health of the pool of insured individuals. Medical spending growth has been large and variable for at least the post-World War II pe-

riod. The large annual increases in medical spending mean that the flat premium structure used in term life insurance insurance plans is unrealistic. The looser premium guarantee of guaranteed renewability is one alternative with lower up front premiums. Guaranteed renewability could still be out of reach because of the large early year premiums required to finance future benefits, or individuals could be uncertain about their willingness to remain with the same insurer in the future, in which case future premium guarantees would be less valuable.

Managing the forecast errors in medical spending may not be easy. Long term insurance relies on prefunding, so individuals pay more than their expected costs in early years, and less than their expected costs in later years. The mismatch between premiums and claims means that long term health insurance policies are highly leveraged, so small errors can cause large disruptions. In theory, insurers could perfectly match premiums and claims through immunization. Guaranteed renewable insurer involves the insurer selling a long tailed risk for up-front payments, so the insurer could invest reserves in high duration assets like zero coupon bonds to match the payment of claims with the payoff from investments. In practice, immunization involves an accurate forecast in spending growth and assets that covary with medical spending. Whatever the cause of spending growth, it is important to know the long term trend in growth and the variance of the growth rate around the long term trend. Insurers would also want to know the “resistance point” where (if) the trend rate bends back toward the growth in the overall economy (Getzen 2007), or when a new trend in growth will start. It may be better to recognize that there is some limit to the length

of guarantee that private insurers can manage.

I analyze how the existence of spending growth impacts long term health insurance. Growth rates can be managed through ideal zero coupon bonds, while full hedging of spending variance requires a securities market that is complete with respect to medical spending growth risk, or a third party willing to take on spending variance for a price. I show that insurers can use bond immunization techniques to manage the flows from guaranteed renewable insurance, and that it helps insurers to fulfill their commitment to consumers and deter shirking of their commitment.

I address the impact of spending growth and uncertainty about spending growth on long term insurance in the ideal pricing and investment policies. I show how the existence of spending growth impacts guaranteed renewable health insurance. Growth rates can be managed through ideal zero coupon bonds, while hedging of spending variance requires a securities market that is complete with respect to the risk of medical spending growth.

I show evidence that aggregate spending is partially predictable using measures of prices and quantity, but not asset returns. In the aggregate, spending can be imperfectly forecast. I show how the variance differs by subpopulations through the use of daily, individual level medical spending from the MarketScan database. I then match the improved spending forecasts from microdata to higher frequency asset returns, to determine whether more active hedging strategies work.

I use medical care spending and price data calculate the rate of change and variance of the rate of change in medical care prices and total spending. I use securities data to

calculate the rate of change and variance of assets available for investment. I also calculate the correlation of the expected rate of change of medical costs and assets. I assess the extent to which the medical spending growth increases the required reserves under the plans, as well as the potential uncertainty of future premiums and insurer surplus.

Insurers offering long term insurance require tools and public policy changes to write guaranteed renewable insurance. Insurers face problems because of asymmetric taxation of gains, and the investment policies for best managing spending growth could generate large gains. If gains are taxed as income or subject to double taxation, then insurers impose even more loading on already expensive benefits. Insurers could offer not to risk rate individuals in most states of the world, but could impose partial risk rating in some states of the world. Partial insurance of this sort would require regulatory approval, and could aggravate the signaling problem. Another form of partial insurance would be pure indemnity insurance with predefined indemnity, where insurers use defined contribution techniques to manage spending growth. Similar provisions would include benefit limits that do not allow for new technologies. Pure indemnity and strict benefit limits are probably not possible under current insurance regulation.

Insurers also need to minimize the trading and investment maintenance costs associated with investment strategies, which may require new asset classes. It is also possible that there are some additional assets that could prove useful for hedging long term medical costs. Some of these assets would be issued by the government, such as Treasury bonds indexed medical price inflation. Others would include private contracts, such as health

insurance futures. I analyze the extent to which these solutions would be helpful by including these hypothetical assets in my forecasts of medical spending growth. I also discuss the general advantages of assets that can hedge medical spending growth as investments that individuals might find useful. The other advantage of additional hedging assets would be to allow individual investors to use personal investments as a partial substitute for guaranteed renewable and other long term health insurance plans.

2 Problems specific to GR

2.1 Prefunding required

Individual health insurance coverage is guaranteed renewable, which leads to initially higher premiums that are less volatile over time. The level of premiums also declines over time relative to annual renewal insurance. Guaranteed renewable health insurance combines protection against current and future costs of spells of illness into one long term contract. Consumers prefund future benefits by paying initial insurance premiums that are higher than the prevailing spot premiums. In return, insurance companies commit to rating based only on class experience, and to avoid individual reunderwriting. Other forms of health insurance, such as long term care insurance, retiree medical plans, and Medicare, feature similar protections or prepaid funding. The cost advantage of group insurance over individual insurance has led to low persistence in individual plans, so it is hard to tell if guaranteed renewability is binding on most insurers. There is some evidence that insurers

are setting up the required contract reserves for guaranteed renewability. For example, insurers in North Carolina have been forced to refund reserves meant to stabilize premiums that will not be needed after the PPACA comes into full effect (Young 2010). There have been also been allegations that high premium growth for new insureds—duration rating—represents a violation of guaranteed renewability (Bluhm 1993).

Guaranteed renewable insurance is only supportable with an initial level of prefunding. Guaranteed renewable health insurance combines protection against premium increases stemming from illness with premium increases stemming from the future cost of medical claims. The mechanism for binding the hands of the currently low risk insureds is the reserve fund, funded by greater than spot market premiums. The reserve allows those who become high risk to pay actuarially favorable rates relative to the spot market without driving the low risks out of the guaranteed renewable pool. To make make the contract zero profit in expectation, the reserve is calculated based on the shadow price of purchasing annual renewal insurance for high types for the remaining term of premium protection specified in the plan.

The value of the shadow benefits rises with the general level of medical spending. Plans could limit their exposure to rising medical spending, for example through a pure indemnity arrangement, but in practice they do not because this exposes the insureds to more risk. It is particularly hard to limit benefits in guaranteed renewable plans, since retroactively removing previously allowable benefits could violate the guarantee.

Health insurers utilize deductibles, copayments, and other individual demand side be-

havioral techniques to define the level of insurance coverage, but not the growth in benefits. Therefore, the value of the benefit in guaranteed renewable insurance is linked to the rise in medical spending. Insurers that fail to prefund the portion of high risk claims generated by the increase in spending growth are giving low types to drop out without funding adequate reserves, violating the premium guarantee. Insurers that charge adequately for prefunding future spending growth could be charging rates that are unaffordable, or seem exorbitant when compared to spot rates.

2.2 Prefunding may be expensive

The magnitude of the rise in medical spending means that prefunding guaranteed renewable insurance is relatively more costly than spot premiums in the current period. Buyers may perceive guaranteed renewable insurance as expensive unless insurers can promise lower future premiums due to the return on guaranteed renewable reserves.

The return on reserves is a function of the expected time until claim payments, which is in turn a function of the explicit length of the premium guarantee as well as the overall health of the insured population. A less healthy population, or one with a relatively short term of premium guarantee, has a shorter tail of claims, so reserves should be invested in lower risk, lower return assets. Previous research has not considered the duration of guaranteed renewable insurance claims, which should be a crucial element of how reserves are invested, and how the investment strategy then determines the difference between the expected return and the expected growth of medical spending.

2.3 Prefunding may be risky

The variability of the rise in medical spending means that prefunding alone is not enough—insurers need to absorb shortfalls or surpluses caused by unanticipated spending growth and adjust future premiums on a prospective basis. The variability also means that guaranteed renewable premiums will fluctuate, possibly severely. In theory, spot rates should fluctuate to an even greater extent, although the relationship between the variance of spot and guaranteed renewable health insurance rates has not been widely explored. Assuming the risk pool is large enough to apply the law of large numbers to claims experience is a tenuous assumption for individual health insurance. Applying the law of large numbers to medical spending, or even medical prices, is difficult because of the small sample of applicable time series data. The per capita insurance spending data I use is annual, and is only available starting in 1971. For insurers, data before a certain date may not be useful because of a trend break. For annual renewal insurance, the amount of data available should not be a problem, but for a guaranteed renewable policy that can run for ten years, there may not be adequate data to fit forecasting models.

Asset returns and spending growth are both subject to shocks. Even relatively safe bond returns have a standard deviation at least double their mean return. The skewness of asset returns tend to be negative, because surprises are often below average returns. The standard deviation of medical spending increases are also large relative to the mean. Medical spending growth is relatively unskewed, although the skewness of medical cost price change are positive, meaning that surprises tend to be above the mean. The differences

in the first four moments of asset returns and medical cost growth point to the unknown correlation between assets and spending. If spending growth and assets, or a blend of assets, are highly correlated, then investments of reserves can be tailored to minimize the problems of shortfall. The case of general inflation, where selling stocks short is a partial hedge, is one example where a similar risk can be partially hedged by shorting stocks (Bodie 1976). If, on the other hand, assets returns are not correlated with medical spending growth, then the best insurers can do is to choose the level of risk that they are willing to take in their investments, set premiums accordingly, and absorb adverse (and advantageous) experience from incorrect predictions.

The question is then how much of the unexpected growth in spending can be shifted to annual premium increases. Premiums are prospective, so insurers would have to absorb retrospective losses. In theory, whatever increases affect spot insurance prospectively should be borne by individuals with guaranteed renewable insurance, because the guaranteed renewable premium is based on what current low type insureds would expect to pay for insurance in the future. The problem in practice is how to price based on the future expectation of low type risks. If enough low types leave the pool, naive class averaging could impute too much of the costs of high types into the “guaranteed renewable” premium. Medical spending growth will tend to inflate the problem, especially if it is unhedged. Hedging does not solve the problem, but it does give insurers a larger reserve to use to charge more level premiums over time.

2.4 Problems in related insurance products

The rising rate of medical spending, its variance, and the possibility that it cannot be hedged, relate to problems with other insurance products. Retiree medical benefits have been harmed by unanticipated spending growth, as well as failure to provision for future costs with adequate prefunding. The long term care insurance market has been hampered by a perception that insurance companies may not be able to fulfill a promise for benefits many years in the future (Cutler 1993). More generally, the investment decisions of those planning for retirement and in retirement are hampered by uncertainty over the future cost of medical care (Stewart 2008), as well as the return on capital (Ball and Mankiw 2007).

Long term care insurance and retiree medical insurance are specifically addressed in the PPACA. The change to long term care insurance is to add a new, government sponsored long term care insurance arrangement through the CLASS Act (U.S. Congress 2010). It is currently unclear what the premiums and benefits of the CLASS Act will be, or whether it will supplant private long term care insurance. The change to retiree medical insurance is to allow retirees who are not yet 65 to purchase health insurance through the health insurance exchanges. In the interim, the government will provide reinsurance on a “bridge” basis for those employers currently providing retiree medical insurance coverage (The White House 2010). In both cases, it is conceivable that the government will take over medical spending risk from the private sector.

The experience with inflation hedging is relevant for the problem of buying private protection for the long term rise in medical spending. As with general inflation, it is unclear

whether any entity other than the government can credibly offer long term contracts that pay off in proportion to the rate of medical inflation. The investment policy to back such a hedge may be complex and involve short selling, and therefore best implemented by institutional investors on behalf of individuals. The specific problem of hedging against changes in general inflation was solved in the U.S. (and many other countries) by the Treasury, which issues TIPS bonds whose price is linked to inflation (Bureau of the Public Debt 2008b). The value of these bonds can be seen in their popularity at auction, although recently, the Treasury has reduced the issuance of these securities (Irving 2009). No such asset, or known combination of assets, exists for medical inflation or spending growth. Alternative assets, such as health insurance futures and health reinsurance, have not shown promise for hedging. Institutions and individuals other than insurance companies also have a difficult time managing risks related to rising, variable medical spending.

2.5 Comparison between GR and annual renewal insurance

The guaranteed renewable model of insurance also shares features with annual renewal health insurance. Both types of insurance are vulnerable to a firm renegeing on its commitments, either to increase short term profits or due to financial distress. As a result, individuals and regulators will have to expend some costly effort to scrutinize either type of insurance contract. Insurers will have to hold short term reserves to manage the flow of funds for either type of product. In the end, reserves, financial capital, intangible capital, and reputation are the main ways of dealing with the market imperfections that lead to firms

not upholding their commitment.

The main difference between guaranteed renewable and annual renewal insurance is the amount of information insurers can use in setting rates each year. In annual renewal insurance, insurers can individually risk rate their contracts subject to the general restrictions on the use of underwriting data ¹. In guaranteed renewable insurance, insurers are restricted in that they cannot charge different amounts to individuals in the same pool.

In addition, the amount that they charge all members should not reflect the raw expected experience of the pool. In the pure GR model, insurers should charge a price based on the future expenses of the original members of the plan who remain healthy, even if those members leave for a different insurance plan. In the actuarial model of class average, the community rate is based on the average expected experience, adjusted for the credibility of the group. While there is a debate over the use of factors such as duration for rating contracts, what differentiates the guaranteed renewable contract is the insurer's commitment to community rating within the insurance pool, and the commitment not to recoup retrospective losses through future premiums.

It is as a result of the commitment that the insurers collect the additional frontloaded premium implied by the model. The pure GR model implies an exact formula for frontloading; insurers may not use those exact formulas. What separates the guaranteed renewable contract from the annual renewal contract is that it is both prudent and actuarially required that the insurer collect frontloaded premiums. Therefore, the insurer in the guaranteed

¹I discuss new forms of restrictions on underwriting under the Patient Protection and Affordable Care Act in chapter 8

renewable contract is collecting premiums against a future risk, subject to a partially predictable amount of cost growth, and must therefore impute the volatility of cost growth into the calculation of current premiums.

3 Importance of GR for insuring long term health

3.1 Serially correlated medical spending

The serial correlation in medical spending is an important reason to have long term health insurance. The more permanent health shocks are, the more the cost of illness is due to higher future costs relative to current costs. Herring and Pauly (2006) identify the effect of the presence of illness on expected claims based on the novelty of diagnosis, based on the observation of Eichner et al. (1998) that “the median period of unusually high expenses is only about 4 years.” Splitting medical expenses between “transitory” and “persistent”, De Nardi et al. (2009) find that the persistent component explains only one-third of variance in medical expenses, but that the autocorrelation of the persistent component is 0.922 “... so that innovations to the persistent component of medical expenses have long-lived effects. Most of a households lifetime medical expense risk comes from the persistent component.” The degree of persistence is consistent with the notion that preexisting conditions make individuals difficult to insure. If people and insurers could prefund the possibility of entering a persistent high cost state, they could reduce exposure to a major risk.

3.2 Spending is growing

The main reason to investigate the ability to hedge long term medical claims is that medical care spending is growing. The rate of growth continues to outpace GDP growth, and it is unclear how far the trend will continue. Forecasts that promise a leveling out of medical trend rely on a “resistance point”, where medical spending growth stops outpacing general spending growth (Getzen 2007). Whatever the cause of spending growth, it is important to know the long term trend in growth, the variance of the growth rate around the long term trend, and the “stopping point” where (if) the trend rate bends back towards the growth in the overall economy, or when a new trend in growth will start. Modeling these three aspects of spending growth are crucial for planning for the future claims arising from health insurance.

3.3 Best possible prefunding solution

My research addresses the best possible solution to the difficulty in prefunding guaranteed renewable insurance. The best mix of assets to fund the expected level of inflated future claims, given realistic assessments of the trend rate of spending increases and reasonable asset returns, minimizes the amount of front loading required for insured, and maximizes the amount of protection that insurers can offer. The best mix of assets also tries to match the volatility of asset returns and spending growth in guaranteed renewable claims, minimizing the amount of unpredictable premium changes for insureds and minimizing the unexpected losses and surpluses for insurers. It is likely that to some extent, the goals of

maximizing surplus and minimizing the variance of surplus conflict, in which case it is important to understand the trade-offs between risk and reward that insurers will make.

It is also important to determine the limits of prefunding. I examine the residual risk of insurers and insureds to compare guaranteed renewable insurance with alternative insurance arrangements, such as annually renewable insurance. The limits of protection also allow regulators to improve their supervision of firm performance and insurer solvency, and may argue for solutions such as greater capital cushions for health insurers. The results also inform wider health reform by showing how reliable guaranteed renewable individual insurance is as way to expand insurance. I also shed light on the usefulness of new assets, such as Treasury bonds indexed to the medical portion of the Consumer Price Index, in prefunding guaranteed renewability and, in general, helping to hedge the risk from the long term growth in medical costs.

The limit on prefunding is the optimal distribution of risk. Some risk should be borne by insurers. Some risk will remain with individuals. Individuals will retain idiosyncratic risks, and will retain other risks where the cost of shifting the risk to an insurer exceeds the benefit of offloading the risk. Insurers will also retain some risk while potentially offloading other risks, either to private or public reinsurers or to capital markets. All entities, including the government, will have to recognize some limit to the risk they can bear.

3.4 Time limits of protection

Guaranteed renewable health insurance provides protection for multiple years. The term of the contract can vary. “ Model Regulations adopted by the NAIC define a guaranteed renewable policy as one that has the right to continue in force by the timely payment of premiums until at least the later of: Age 50, or Five years from its date of issue” (Robbins and Winslow 2010). The potential length of protection raises the question of whether any company can reliably offer such protection, and why more limited forms of guaranteed renewability are not available. Limited guaranteed renewability, in the sense of a time limit on premium guarantees, may be an alternative if long term spending trends are truly unpredictable. The length of protection is also linked to the problem that individuals with guaranteed renewable insurance who become high risk are effectively locked into their insurance company.

3.5 General application to long term health insurance

The results of my research also have more general application to other forms of long term health insurance. A number of private long term health plans currently exist—guaranteed renewable insurance for those under 65 with individual insurance, and a combination of retiree medical care, and long term care insurance for working or retired individuals at a range of ages. Public plans, such as Medicare, can also be considered long term insurance plans, although individuals do not prefund their own coverage on the basis of their personal risk level. The remainder of costs are funded through out of pocket spending. The solution

for funding guaranteed renewable health insurance can also be applied to health insurance for those 65 and over managed by private firms, such as long term care insurance.

The ability and desirability of individuals prefunding future medical costs also addresses the issue of low levels of annuitization. Sinclair and Smetters (2004) explain low observed levels of annuitization by demonstrating that an annuity is a risky asset for individuals exposed to a health shock, which makes individuals need a lump sum for health costs and reduces the expected value of their annuity. Guaranteed renewable insurance deals with this problem, especially in the model of (Cochrane 1995), where a health shock would lead to the insurer increasing the value of the individuals' health insurance account by the cost of the shock. Individuals who pay the large front loading for guaranteed renewable health insurance could still face problems if they are liquidity constrained. If a negative health shock were associated with large costs not covered by health insurance, such as loss of income, then the individual might prefer to have a lump sum in the state of the world where he gets sick. As a result, mandating guaranteed renewability could make some types of individuals worse off even though it could make their health insurance more affordable.

3.6 Policy implications

An important policy audience from my work is the state regulators of individual health insurance. I address regulator concerns about whether insurance companies can fulfill the promise of guaranteed renewability. I also provide some guidance as to the prudent man-

agement of investments by insurance companies. I also address regulator concerns about excessive upfront and renewal premiums, buildup of excessive reserves, and low medical loss ratios. I offer guidance, though in a broader context, to other similarly regulated insurance products, such as long term care insurance.

Creating additional assets is another implication of the possibility of incomplete markets. One suggestion is for the Treasury to sell bonds linked to particular components of inflation, such as the Medical Care portion of the CPI-U (Jennings 2006). The GAO has targeted inflation protected securities as a way to finance the increasing fiscal deficit at a low cost (Irving 2009), so branching out from general inflation to components of inflation may be a practical policy. I find that medical inflation as defined by the CPI-U is not correlated with general inflation in a way that is both useful for hedging long term spending growth and not already provided by existing products.

Another possibility is bonds linked to nominal GDP, by the public sector, possibly in partnership with the private sector (Shiller 1993, Griffith-Jones and Sharma 2006). Such bonds would be useful to the extent that nominal GDP growth is correlated with spending growth, especially if the unpredictable components of the two time series are correlated. I also assess the possibility that health insurance futures, mortality swaps, and other products plausibly correlated to medical spending variance offer similar forms of novel protection (Cox and Schwebach 1992, Brady 2008).

The idea with medical inflation or nominal GDP bonds would be the same as for long duration government bonds. 30 year government bonds allow individuals or firms with a

certain, long dated liability, to invest in an asset with the same risk profile as the liability. For example, long dated government bonds are exposed to interest rate risk—if interest rates go up, the value of the bond goes down. A firm might want to hedge a 30 year liability with a similar interest rate risk—if interest rates go up, the value of the liability goes down, and vice versa. Presumably, medical inflation or nominal GDP bonds would have the same risks as the underlying index, and would be priced by the market just like ordinary government bonds. Then the government would bear the risk of higher medical inflation or nominal GDP growth, but would reap the benefits of paying a lower interest rate.

Insurers offering long term insurance require tools and public policy changes to offer premium protection. Insurers face problems because of asymmetric taxation of gains, and flight-to-quality from financial distress. Insurers could offer not to risk rate individuals in most states of the world, but could impose partial risk rating in some states of the world. Partial guaranteed renewability of this sort would require regulatory approval. Another form of partial guaranteed renewability would be pure indemnity insurance, where insurers use defined contribution techniques to manage spending growth. Similar provisions would include benefit limits that do not allow for new technologies. Insurers could also benefit from credible sources of economy-wide medical spending growth when creating rating classes of consumers. The problem with any index is the insurance pool specific nature of the risk of medical claims.

4 New contributions

I aim to improve the understanding of medical spending growth and how it affects the feasibility of long term health insurance products. I determine the expected rate of growth in medical spending, with new estimates from aggregate and insurance plan data. I describe the extent of the connection between trend in medical spending and trends in asset returns, interest rates, and other macroeconomic variables. I describe the variance of spending growth, and the extent to which it is correlated with changes in the rate of return on assets and interest rates. I show how the results on spending growth affects the premiums and reserves of guaranteed renewable health insurance policies.

I show the degree to which medical spending growth exacerbates the mismatch of premiums and claims already present in guaranteed renewable insurance and raises the prospect for insolvency, or failure to follow through with the guarantee, due to unforeseen shocks. I find the best investment policies to mitigate the problems created by medical spending growth. I show the residual risk that is borne by insurance companies and insureds from incomplete hedging. I show how the residual risk can be managed through the use of contract reserves. I show how the results for guaranteed renewable health insurance can be extended to other long term health insurance products.

4.1 Estimate the growth and variance of medical spending

I first establish the reasonable expectation of the rate of insured medical spending growth. I produce new forecasts of the trend in medical spending. I produce three distinct forms

of forecast. As a first approximation for the magnitude of growth, I calculate the summary statistics of consumer and producer experience with medical inflation, medical employment, and total spending growth. I generate simple forecasts for the long term growth rate with adaptive expectations forecasts. I create more ind-depth forecasts that integrate medical prices and spending with an array of macroeconomic variables, such as GDP and general inflation, as well as returns on the stock and bond markets. I show how different models produce different forecasts of growth rate, as well as how sensitive the forecast is to the historical time horizon.

I establish the reasonable expectation about the variance of medical spending growth. I review the error rate in prior forecasts and my new forecasts of the trend in medical spending. I show the variance in the summary statistics, adaptive expectations, and multivariate models of spending growth. I show how different models produce different forecasts of variance, as well as how sensitive the forecast is to the time horizon. I also make simple assessments of market completeness by showing how well existing assets can hedge monthly and annual variance in the growth of spending through covariance between asset returns and spending variance.

4.2 Effect of spending trends on guaranteed renewable plans

Using a range of forecasts for spending trend, I show how the magnitude of medical spending growth determines the components of guaranteed renewable insurance. I show how the trend rate interacts with the distribution of losses and the term of guaranteed renewable

protection to determine the long term inflows and outflows of guaranteed renewable plans. I show how different levels of spending growth change the duration and convexity of claims and premiums, and how they affect the level of prefunding required to support guaranteed renewable premiums.

I show how the timing of insurance claims and the variance of medical spending drive the variance of guaranteed renewable premiums. I determine the variance in claims, premiums, and final reserves given the variation in the growth of cost. I show that while the variance from population loss parameters can be driven to zero through the law of large numbers, there is not enough data to resolve the variance arising from increases in medical spending in the same fashion. I show how increasing the time of premium guarantee can actually decrease the variance of the timing of claims in the absence of stochastic growth, but that with a varying growth rate longer term claims more difficult to forecast.

The factor determining the long term effect of spending growth on reserves is the time series characteristic of medical spending growth. If each year's growth rate is drawn from a random distribution, and there is no serial correlation in growth rates, then variations should average out over time. Above average trend rates that occur in the early years of a guaranteed renewable insurance policy will be more damaging than those that occur in later years, but overall the variations should not affect the ability of insurers to pay claims. If the growth rates are serially correlated, then a single year of above trend growth will tend to translate into multiple years of above trend growth. In this case, the rate of medical spending growth could generate losses, or gains, in several years of the policy until the

prospective premiums reflect the higher expected spending growth rates. The importance of the funding policy is to match asset returns and spending growth to the greatest extent possible, so that the time series properties of medical spending growth have less of an impact on the solvency of the policy.

4.3 Determining the optimal funding policy

I demonstrate the ideal funding policy that maximizes the expected rate of return on assets conditional on the long run rate of spending growth and variance. I show the mix of assets that best matches spending growth, and the expected difference between the long run return on assets and the long run growth in spending. I also show how the loss distribution of the insured population, the term of premium protection, and the growth rate in spending change the best investment policy, and thereby affect the expected rate of return on reserves.

I contribute to the literature on market incompleteness by determining how well hedging works. I evaluate the hedging available through raw covariance between spending and assets. I estimate whether the theoretical investment policies implied by my results can be realistically implemented. I show that the investment policy may require costly monitoring, frequent rebalancing, or short selling in order to function properly.

4.4 Defining the residual risk

I characterize the residual risk that insurers face under guaranteed renewable insurance, and where the risk comes from. For insurance companies, I show the residual risk of shortfall

for companies utilizing the ideal investment policy. I show how the growth rate, loss distribution, and term of protection interact to affect the probability and size of shortfall. I also show the effect on the reserve of positive and negative forecast errors. I determine whether the size of the shortfall is small enough to be borne by investors, or large enough to give rise to a realistic chance of reserve inadequacy. To the extent that depleting the reserves is possible, I show how it can be mitigated or avoided through solutions such as capital requirements, partial rather than full guaranteed renewability, time limits on the term or premium protection, or restrictions on the populations offered premium guarantees.

5 Approach

I use time series data for medical spending and asset prices to determine the excess return on reserves for guaranteed renewable insurance contracts. I use medical spending data, supplemented with price and quantity data, to demonstrate the substantial growth rate and unpredictability of medical spending. I use securities data to calculate the rate of change and variance of assets available for investment. I also calculate the correlation of the expected rate of change of medical spending and available assets to show the lack of good hedging assets for guaranteed renewable contracts.

I discuss the implications of the undiversifiable portion of the risk of medical spending growth. It is also possible that there are some additional assets that could prove useful for hedging long term medical spending. Some of these assets would be issued by the government, such as Treasury bonds indexed to medical price inflation. Others would in-

clude private contracts, such as health insurance futures, reinsurance contracts, and hedge funds. I analyze the extent to which these solutions would be feasible by including these hypothetical assets in my forecasts of medical spending growth. I also discuss the general advantages of assets that can hedge medical spending growth as investments that investors might find useful.

I show how reserving for guaranteed renewable insurance is affected by the time series properties of medical spending growth when compared to annual renewal insurance. Insurers must use contract reserves to manage medical spending growth, and the size of the reserves differs by how far in the future the guarantee runs. I discuss the place of regulators in improving the solvency of guaranteed renewable insurance, which involves focusing on capital.

Chapter 2

Literature review

In this chapter, I review the prior literature on guaranteed renewability and hedging medical spending growth. I first discuss the theoretical models of guaranteed renewability in the economic and actuarial literatures. The main difference is the use of one-sided commitment in economics versus the use of class average rates in actuarial science. I show how guaranteed renewability is part of a larger class of health insurance products that involve long term (more than one year) insurance.

Next, I review prior empirical findings on the feasibility of guaranteed renewability and the challenge of making it work. I review the history of guaranteed renewable health insurance contracts, and the possibility of firm ruin. While related products have failed to pay claims, the main concern about guaranteed renewability has been the possibility that insurance companies do not fully keep their promises.

Finally, I review the literature on medical spending growth as a macroeconomic risk unrelated to the health of the insured population. I examine past attempts to insure against

the risk of medical spending growth, showing that they have mostly fallen short. Looking at the history of hedging macroeconomic risks shows the approaches used, both in the use of existing assets and the creation of tailored assets, as well as their limitations. Drawing on this literature, I show the past negative implications of a lack of hedging assets for guaranteed renewable and other forms of long term health insurance contracts.

Finally, I review the ways that, in the past, long term health insurance has avoided dealing with macroeconomic risk. Strategies include the use of exogenous lapsation, as well as encouraging lapsation by the “right” insureds. Other ideas include the use of reinsurance or social insurance to supplement or replace private insurance. I show that the nature of the risk, especially how systematic it is, determines how feasible these alternative solutions are. The inability for individual firms to deal with the macroeconomic nature of the risk has motivated a great deal of the literature of medical spending growth as a health policy, and public finance, problem.

1 Theoretical background

1.1 Guaranteed renewability as a one-sided commitment

In one-sided commitment models, the firms cannot break preexisting contracts, but consumers always have the choice of exit. A definition of the one-sided commitment comes from Phelan (1995), where the motivation is the solution to a repeated moral hazard game: “This typical firm also takes as given that any agent-customer it contracts with will be able

to quit out of the insurance contract at the beginning of any future period and receive a new contract from another firm delivering this same market utility, w^* .” Another motivation for one-sided commitment is a model where consumers are discovering their type, which can be static, as in Harris and Hölmstrom (1982), or dynamic, as in guaranteed renewable insurance models (Pauly, Kunreuther and Hirth 1995). In such a model, information is symmetric but incomplete, so consumers are simply shifting the risk that they are, or become, the expensive type onto an insurer who pools over enough individuals to manage the population risk.

The major advantage of the consumer’s ability to exit is it allows long term implicit contracts when explicit contracts are unenforceable (Bull 1983). The consumer’s ability to exit gives them a way to discipline firms that act avariciously: the option of walking away is the solution the problem where courts can only enforce one period contracts (Phelan 1995). However, the possibility of exit means that in some states of the world, it is rational for most consumers to exit and enter into new contracts, while a few will not want or be able to exit. For example, when interest rates drop, many people will prepay their fixed rate mortgages and take up new loans, while those whose credit scores have dropped since taking the mortgage will be unable to take advantage of lower rates. Such behavior can lead to perverse outcomes, since the impossibility of exit for the remaining consumers takes away one avenue consumers have to keep firms honest.

The scope of action of the two parties in a one-sided commitment mechanism dictates the preventative activities they should undertake. The customer makes an additional up-

front payment to secure the commitment, and then need only monitor the contract terms at the initial offering and at renewal, and outside options, exiting into a new contract or self-insurance when either option is preferable (Krueger and Uhlig 2006).

The firm that makes the commitment is supposed to uphold its end of the contract. The spirit of a one-sided commitment dictates that the firm undertake reasonable actions to ensure that it will be able to follow through on the commitment in the future, but preventative actions are not required by most one-sided commitment models. For example, if bankruptcy or firm ruin would lead to forced renegotiation of the contract, the firm should take actions to avoid these forms of financial distress if the cost is less than the expected benefit of avoiding ruin. However, whether the firm is obligated to take actions beyond fulfilling the duties of the contract is debatable, although the costs of firm ruin include not only damage to directly effected consumers but negative externalities for the economy (Bernanke 1981).

The place of insurance regulation is to ensure that insurers have adequate surplus and reserves to maximizing the probability of honoring the terms of the contract. Munch and Smallwood (1980) found that in property-liability insurance, capital requirements for domestic insurers reduced the number of insolvencies. The entire effect came from the effect of the regulations on smaller firms, which had more difficulty entering the markets with higher capital requirements. The greatest cost to consumers came from the absence of specialized forms of insurance offered by smaller insurers. In the context of guaranteed renewable health insurance, companies are offering a more homogeneous product, so the

availability of specialized products might not be a concern. The feature of the Munch and Smallwood (1980) model where insurers endogenously choose their probability of insolvency is applicable: companies that choose higher level of reserves are choosing a lower probability of insolvency due to factors outside their control.

It is possible that circumstances will make upholding the contract impossible for the firm. The build-up of contract reserves on long term insurance contracts almost always means that, on a cash basis, the contract will generate “losses” in later years. Insurers that try to find ways to hold on to the reserves in a manner that violates the commitment could expose themselves to legal retribution, as well as a loss of reputation. Both legal consequences and reputation have limits, especially if low risk consumers choose to lapse out of the contract, while high risk types cannot find a better contract at a competitor i.e. adverse retention (Altman, Cutler and Zeckhauser 1998). While profit seeking against “locked-in” borrowers may violate the spirit, if not the letter, of the contract, the temptation for profit may be hard for the insurer to resist. The results of Munch and Smallwood (1980) suggest that insurers that anticipate the possibility of exiting from a market may choose a low level of reserves and a holding company firm structure that make the costs of exit lower. Ex ante, parties who accept the commitment may look for the committing firm to undertake actions that make later pernicious behavior unprofitable.

Both economic models of guaranteed renewability in health insurance are based on one-sided commitment by consumers. The original model of guaranteed renewability as part of an insurer pool of nongroup insurance contracts is Pauly et al. (1995). Their model entails

healthy people paying initial premiums above their expected claims level. In exchange for prefunding future costs, premiums do not change for people who enter a more expensive “high risk” state. The premiums are computed so that they always represent an actuarially fair contract for the remaining life of low risk consumers.

A mathematically equivalent model was proposed by Cochrane (1995) and recently extended (Cochrane 2009), with a health status insurance fund that rises when people become less healthy and sinks when they become more healthy. In his model, the rises and falls correspond to the shadow price of purchasing spot insurance for the remaining lifetime of the consumer. The models are mathematically the same, but differ in whether the contracts are considered separately or as part of a risk pool. The model of Pauly et al. (1995) assumes that individuals will stay with their insured for their lifetimes, or for an extended period until they are eligible for another form of insurance such as Medicare. The model of Cochrane (2009) assumes that individuals will often want to use the health status fund to switch insurers.

Other models of insurance premium guarantees in areas such as term life insurance also use a one-sided commitment model. In all models, guaranteed renewable insurance is a second best insurance contract (Hendel and Lizzeri 2003). It is inferior in that the first best solution, fully level payments, produce a lower interperiod variance; even agents unmotivated by the risk sharing aspects of the contract may be motivated by the consumption smoothing arrangement (Krueger and Uhlig 2006). The same is true in general of one-sided commitments, which are inferior to long term two-sided commitments in terms of

spreading consumption, but superior in that they are easier to enforce on consumers.

The guaranteed renewable solution relies on competition, and symmetric information. As in Rothschild and Stiglitz (1976), “Free entry and perfect competition will ensure that policies bought in competitive equilibrium make zero expected profits. . .”, although unlike their model, I am mainly concerned with a group of consumers and a multiperiod model. In addition, cooperate bargaining assumes away the incentive that insurers would have to “game the system” by inducing those who become high types to drop their coverage and lose their portion of the surplus. Similarly, it relies on symmetric information, which removes the possibility of adverse selection (Akerlof 1970), and the existence of a single, defined loss, which alleviates the problem of moral hazard.

1.2 Financial management of guaranteed renewability

Guaranteed renewable plans are leveraged in the sense that the liability firms take on is much larger than the initial cash premiums paid by individuals. The main effect of the growth in medical spending and the tendency for individuals to get sicker over time is to move the bulk of insurance payments to the future for any set of buyers. Later cohorts will enter as new insureds with an expected stream of claims that matches the updated data on spending growth for a new, and therefore healthy, pool of lives. Front loading for the commitment mechanism has the effect of transferring the bulk of premium payments from the future to the present. Discounting has the effect of transferring both inflows and outflows to the present.

Managing guaranteed renewable insurance means matching the timing premium payments and investment returns with the timing of claims payments. The measurement of the timing of claims and premiums in guaranteed renewable insurance shows how reserves should be invested to provide a flow of payments that matches the expected timing of claims. When "...the value of future benefits at any time exceeds the value of any appropriate future valuations net premiums at that time" the contract reserve fund brings them into balance (Bluhm 2007, p. 175).

Insurance companies could in theory offer conversions or charge inadequate upfront premiums in a way that leads to a violation of the premium guarantee (Bluhm 1993). Violating the guarantee could happen through neglect rather than strategic action by insurers. Actuaries agree that the effect of increasing health care spending due to increased prices or increased quantity should be passed through to insureds, so it could be difficult to distinguish between appropriate and inappropriate rate increases. It might also be better for insurers to set a ceiling on the level of unexpected spending growth that they will absorb, such as the 5% inflation protection offered with long term care insurance (Lutzky, Corea and Alecxih 2000). Guaranteed renewability lacks the transparency of noncancellable term life insurance contracts that lock in a fixed premium for a given period of time.

Another cause of dislocation to guaranteed renewability is variance of the claims and returns streams. Not only do both claims and returns have variance, but they may be correlated. The insurer would prefer that they be positively correlated, or at least uncorrelated, so that above trend growth in spending is matched by above trend growth in assets. The

possibility of positively correlated assets and liabilities relies on the completeness of the securities market. A dynamically complete securities market allows people to insure themselves against bad events with assets that pay off when those bad events occur (Huang and Litzenberger 1988, p. 179). Heaton and Lucas (1996) consider the possibility of certain idiosyncratic risks, such as labor income, that cannot be hedged in any way. Ultimately, whether a given risk can be hedged is an empirical question.

Insurance companies play a role as an intermediary for consumers who do not have the expertise or scale to engage in active trading. The use of whole life insurance as an investment vehicle is an example of the way that insurers can act as investment managers and insurers (Campbell 1980). One trend has been decreasing consumer interest in the general investment management function of insurers, but the role of insurers in managing and hedging risks related to their core insurance functions remain. Insurers have access not only to hedging markets through primary and derivatives markets, but access to wholesale reinsurers who do not market their products to individual consumers. While reinsurance can significantly reduce risk, it might only be provided to primary insurers at a significant load over the cost of the risk (Cummins, Dionne, Gagne and Nourira 2008).

1.3 Motivation for hedging

Return maximization is not the main goal of the investment policy for funds supporting the contract reserves. Reserve formulas are based on using the same discount rate for the present value of future benefits and premiums (Bluhm 2007, p. 169), so investments should

immunize the risk that changes in interest rates affect the present value of inflows and outflows differently. Duration, which measures "...a weighted average term-to-maturity of a security's cash flows" (Fabozzi, Pitts and Dattatreya 1997, p. 85), can be used as an approximation for the timing of payments. Convexity, which measures "...the rate of change of duration as yields change" (Fabozzi et al. 1997, p. 94) can be used as an approximation of the sensitivity of the timing to changes in interest rates. Matching the duration and convexity of inflows and outflows in guaranteed renewable plans could reduce dislocation caused by differences in the timing of claims and investment returns, but at the cost of foregoing profitable investments.

Insurance company hedging can also be motivated by the cost of the insurer's financial distress. Insurance company bankruptcy leads to large costs for consumers and providers as claims go unpaid, or there is a long delay in the payment of legitimate claims; even healthy companies may use payment delays as a financial management tool (Connor, Wholley, Feldman and Riley 2004). Firm insolvency will destroy firm specific assets such as intangible capital (Munch and Smallwood 1981).

In the extreme, long term health benefits for retirees have been reduced or eliminated under the bankruptcies of large corporations (Larson 2009). Reserves and other assets go towards paying current claims, but are unavailable to fund the explicit or implicit guarantees of future benefits. The possibility of forcible renegotiation also harms those companies who are near bankruptcy, or those perceived to be near bankruptcy, because consumers know that the bankruptcy option is more valuable to such firms (Doherty and Garven 1986). High

probability of bankruptcy can lead to a “flight to quality” that damages the constructed one-sided commitment underlying long term implicit contracts (Cummins and Lewis 2003). The possibility of nonpayment can also lead to providers refusing to provide benefits to insured consumers, in which case high risk consumers are locked in to an insurer that cannot provide promised benefits (Patel and Pauly 2002).

The effect of taxes on insurance companies is another reason for hedging. Corporate tax schedules are often graduated, and also penalize gains more heavily than they forgive losses. As a result, insurance companies will try to smooth their earnings in order to minimize their tax bill, and investment management is the way to engage in smoothing (Garven 1992). Guaranteed renewable health contracts have several financial aspects that make nonlinear taxation likely to affect premiums. Front loaded premiums will lead to the appearance of large profits in the early years of the contract. The accumulated reserves in the middle years of the contract generate large investment gains that will tend to make profits look bigger than they are. The payment of claims in excess of premiums in later years gives the appearance of loss, but insurance companies have difficulty fully recouping their tax losses.

1.4 Theory of optimal reserves

In the model of solvency regulation of Munch and Smallwood (1981), insurers that increase their reserves are seen as safer, and thus can attract more policyholders, while those that decrease their reserves can take advantage of limited liability in the case of firm ruin.

When regulators set a minimum level of reserves, it is a binding constraint on those firms that would otherwise choose a low level of reserves (in the stylized model of Munch and Smallwood (1981), reserves tend towards either zero or infinite for limited liability firms).

Regulators are particularly interested in the reserving strategy for longer term lines of insurance, i.e. those where claims arrive well after the payment of premiums. Actuaries are required by regulators to set up contract reserves for "... contracts with respect to which, due to the gross premium pricing structure at issue, the value of the future benefits at any time exceeds the value of any appropriate future valuation net premiums at that time."(Bluhm 2007, p. D-8) Reserve adequacy refers to premiums themselves, but also the investment returns they can generate, and is judged relative to the size of expected claims. Expected claims are affected by the characteristics of the insured population, as well as the tendency for claims to grow over time due to general rising medical costs (Pauly 2003). The magnitude of claims in the future is compounded by the rate of growth in medical spending, so future claims are more expensive (on a nominal basis). The reserve pool is calculated based on discounting to supplement future premiums that are actuarially favorable, so required reserves are lower the higher the discount rate (Bluhm 2007, Ch. 6).

The optimal reserve for a guaranteed renewable health insurer does differ other lines of insurance in that the maximum possible benefit is infinite. While some health insurance plans utilized lifetime maximums to set a ceiling on the loss, the maximums were essentially infinite, and are now banned under the new Patient Protection and Affordable Care Act (i.e. the PPACA; see chapter 8). The absence of a maximum is not an impediment

to utilizing the theoretical work on optimal reserves—in the case of Munch and Smallwood (1981), their results lead to the extreme outcomes of either zero or infinite reserves. The regulator then has a role in setting a floor on reserves which is a binding constraint in the world where the insurer would otherwise choose zero reserves.

The tax and bankruptcy motivation for hedging are also important factors in the reserve setting function of insurance regulators. Insurers may hedge because the firm and its managers find bankruptcy costly. However, some of the costs insurers impose through bankruptcy are borne by others, so regulators use reserve requirements to ensure that the existence of limited liability does not induce overly risky behavior by firms. Regulators can also utilize state guarantee funds for the purpose of spreading the cost of insurer failure over many firms, but the structure of the guarantee funds can induce risky behavior as well (Cummins 1988). While regulators can induce insurers to set their reserves closer to the optimal level, more intense regulation of rates can cause higher loss ratios (Harrington 1984), which would mitigate the effect of higher reserves¹. As with rate regulation, reserve regulation adherence is costly and difficult to observe (Harrington and Nelson 1986), so even if regulators have a target level of reserves for insurers regulators will not be able to verify that insurers hold the specified level of reserves.

1.5 The role of double taxation

Insurers may hold a suboptimally low level of reserves due to the double taxation of investment gains. Insurers may hold their reserves in the equity of other companies, evaluat-

¹The PPACA is explicitly designed to lower premiums and raise the loss ratio.

ing the stream of future profits based on the taxes paid by those companies (Modigliani and Miller 1958). Firms that make equity investments face the problem of double taxation, when the investment pays taxes as a corporation and then the investing firm pays taxes on the equity gains they realize through dividends or capital gains (Litzenberger and Van Horne 1978, e.g.). As a result, the insurer's valuation of the possible equity investment for the reserve fund may be lower than that of other investments. The adverse effect of double taxation on reserves is a problem in catastrophe insurance, and is one explanation for the attraction of alternative investments including catastrophe bonds and insurance futures (Harrington, Mann and Niehaus 1995, Harrington and Niehaus 2003). The investment policy for reserves should take into account both the hedging and double taxation of gains problems for managing invested reserves.

1.6 Ideal investment for guaranteed renewability

The ideal investment for managing long term medical claims would be pure contingent claims on the path of medical spending. A contingent claims would offer specific payouts in specific states of the world (Huang and Litzenberger 1988, p.119). For example, the payoff could be equal to a unit of medical care when the per capita spending for a homogeneous group of insured individuals is a given value. Then, such several such claims can be combined to match the liabilities faced by an insurer in different states of the world. Securities that pay off in money rather than in units of medical care would be equivalent (Arrow 1964), as well as a better way of conceptualizing of the investments of the

insurer.

Generally, insurers face uncertain liabilities due to a range of possible future payouts. The uncertainty over payouts can be a question of both timing or magnitude (Cummins 1974). The asset policy that hedges the liability at any time would be to invest the present value of future liabilities in a mix of claims of different maturities in proportion to the expected timing of claims payments (Bodie 1991). When the policy is written, the investment policy would involve borrowing or seeking equity investment, since the initial premiums would be less than the total liability of the insurance policy (by definition, this is a restatement of the fact that the reserve is positive and larger than actuarially fair front loaded guaranteed renewable premiums). Then, the insurance company would realize the investment return as claims came due, with the remaining amount of investment still equal to the future liabilities. The stream of premiums would be transformed into a stream of investment returns that would perfectly match the inflated claims of the insurance company. The strategy may or may not be feasible, depending on whether the underlying claims are available as investments and the transactions costs of hedging.

The simplest form of hedging is immunization. Say that the nominal claim is certain and the only risk is the effect of uncertain changes in interest rates on the real value of the liability.

Suppose the sponsor has an obligation to pay \$100 N years from now. It can immunize this liability by investing \$ 100^{-rn} in zero-coupon bonds maturing in N years (where r is the continuously compounded risk-free rate of interest).

To immunize an obligation of \$100 due in 20 years at a risk-free interest rate of 8 per cent per year would cost \$20.19 today ($= 100e^{-0.08 \times 20}$). The pension plan will be fully funded. No additional funds will be required from the corporation (corporate shareholders) to cover any underfunding, nor will there be a surplus. (Bodie 1991, Leibowitz 1986)

The immunization strategy assumes that the liability is well known, which is the case for pensions, life insurance, and other pure indemnity contracts with large risk pools. It also requires the ideal asset: zero coupon bonds timed to coincide with the timing of obligations. There is no way that the managers of the pension fund can improve on this strategy unless they have the ability to earn above average returns in the stock market.

1.7 Securities market incompleteness

A complete securities market would provide the securities needed to hedge any risk. A complete securities market for state contingent claims would allow the purchaser to receive a prespecified quantity of a commodity in a given state of the world (Debreu 1959, pp.98-102). These securities could also be long running, complex securities, where delivery is in money rather than a commodity (Hirshleifer 1964). These markets are then said to be complete in the sense that any Pareto optimal allocation of claims can be achieved with these securities (and any allocation achieved in equilibrium is Pareto optimal) (Rubinstein 1975).

The possibility of securities market incompleteness is motivated in part by the positive

aspects of a complete securities markets. One concern about financial regulation is that it could prevent the creation and sale of innovative securities (Merton 1995), which is fundamentally a concern about regulation causing incompleteness in the securities markets. There are many examples of assets whose trading is restricted or prohibited, such as foreign currencies, which exist to combat market failures at a price of increased transactions costs to enter these markets. The social cost of these restrictions and prohibitions is, in part, the barrier to efficient allocation of risk throughout the economy.

The possibility of securities market incompleteness is also motivated by empirical puzzles in the asset pricing literature. The original puzzle is the equity risk premium, which is the observation that the difference between observed equity and risk free returns is too large to be explained in a original complete securities market model (Mehra and Prescott 1985). The focuses for the source of incompleteness include idiosyncratic risk factors and transactions costs, although these factors have not brought a satisfactory explanation for the equity premium puzzle. For example, Telmer (1993) finds idiosyncratic risk factors cannot explain asset pricing puzzles, and Aiyagari and Gertler (1991) find that transaction costs cannot explain the puzzle and match the observed level of liquid assets.

One candidate for the idiosyncratic risk factor is labor market risk. Labor market risk may be difficult to hedge because of moral hazard, which ties back into the literature on contracting under uncertainty through the literature on contracts based on observed outcomes alone (Hölmstrom 1979). Heaton and Lucas (1996) considers the effect of idiosyncratic risks finding that transactions or security market limits can explain the equity

premium puzzle. The incompleteness of markets leads to basis risk, when the securities being used and the liability being hedged are not perfect matches (Cummins, Lalonde and Phillips 2004, e.g.).

In the insurance market, reinsurance can be thought of as an important asset class. Reinsurance is simply a bet against the insurer's own risk profile. Reinsurance may also be incomplete or partial, in that the purchaser must keep some of the risk in order to reduce moral hazard (Jean-Baptiste and Santomero 2000). Reinsurance will therefore be likely to provide some of the hedging needed, but not all. Also, these contracts are not traded on an open market, leading to a risk of default by the reinsurer. Possible reinsurer default motivates securitization of reinsurance contracts (Modu 2007), including catastrophe bonds as a specialized form of securitizing reinsurance (Harrington and Niehaus 2003).

2 Empirical literature

2.1 Guaranteed renewability is an established benefit

Accident and disability companies have been offering guaranteed renewability since 1915 (Bartleson et al. 1963, p. 33). As medical spending has increased, the guarantees have shifted from fully level premiums to promises that premium increases will not be related to changes in individual health status. Guarantees against re-underwriting are valuable enough to potentially make individual insurance better protection than small group insurance for many people, especially the sick (Pauly and Lieberthal 2008).

The actuarial theory of guaranteed renewability is tied to plan experience. The key protection in actuarial and state mandated guaranteed renewability is that “. . . rates will change only on a class rating basis, without consideration of the experience of health status of the individual”; HIPAA’s less stringent definition allows insurers two ways to end the option to renew coverage: offer conversion for all individual policies or cancel all policies in the state (Bluhm 2007, p. 68). As a result the HIPAA definition of guaranteed renewability and the state insurance regulation definition may differ.

2.2 Actuaries and state regulators use a class average standard

The actuarial class average rating definition of guaranteed renewability may not adequately protect consumers in practice. One concern is that class averaging allows for duration rating, where insurance premiums rise steeply in the first five policy years. The rise in premiums by duration may cause the low persistence of individual insurance (Wachenheim 2006). Currently, the correct contract reserve for guaranteed renewability is an open question. Bluhm (1993) contends that insurance company failure to adequately prefund the contract through reserves is a violation of the premium guarantee.

Actuarial science does not allow for insurers to inflict the totality of claims experience on insureds. The major constraint on the use of experience is credibility: the inflated claims experience of plans, especially those with small risk pools, should be blended with broader claims experience to bring the variance of premiums closer to the variance of the population (Mahler and Dean 2001). Insurance companies then manage the risk that they

cannot pass on to insureds through financial management.

2.3 Possibility of successful guarantees

Most individual insurance policies are guaranteed renewable to various degrees through HIPAA and state insurance law, which demonstrates that guaranteed renewable insurance is affordable under class average rating (Patel and Pauly 2002). One recently introduced insurance policy, UnitedHealth Continuity, allows those who are currently in good health to add an explicit guaranteed renewable option to their existing health insurance contract, or buy the option alone, although this product is still untested (Abelson 2008).

Prior economic studies of guaranteed renewable insurance and health status insurance have shown that the protection could be affordable under adequate levels of prefunding. Herring and Pauly (2006) show how, under the current distribution of medical claims in the population, insurers can provide individual health insurance that is fully guaranteed renewable to age 65. Upfront costs are affordable, and premiums increase only due to overall increases in medical spending, as well as aging effects. Their model is predicated on deterministic medical spending growth of GDP+1%, and a 3% real rate of return on reserves.

2.4 Medical spending level and growth

Medical care is a large fraction of total consumption, and includes provision of both goods and services. Providers include services from physicians in hospital and out patient set-

tings, as well as other health care providers, such as pharmacists. The cost of goods includes pharmaceuticals, diagnostic equipment, and medical providers. The cost of delivering insurance (administration) is another cost of medical care for those who have insurance. Payers include public plans, such as Medicare and Medicaid, private insurance plans, and out-of-pocket expenses by consumers.

One aggregate collection of medical spending and growth forecast comes from the Medicare. Medicare collects aggregate figures for medical spending, and break the spending down by payer and provider. They also project spending levels for 10 years, based both on current law assumptions and reasonable alternative assumptions. The most recent estimate pegs medical spending at 17.3% of GDP in 2009, rising to 19.6% of GDP by 2019 (Sisko, Truffer, Keehan, Poisal, Clemens and Madison 2010). The major sources of uncertainty in the study include the effect of the Patient Protection and Affordable Care Act and the overall path of economic growth. Medicare (Caldis 2009) and the CBO (Elmendorf 2010, p. 35) produce a long term forecast based on a computable general equilibrium model that does account for feedback effects. However, this forecast is heavily dependent on long term demographic forecasts.

Forecasting the rate of growth in medical spending has met with limited success. Prior work suggests that expected inputs to a prediction of medical spending are the overall inflation rate (Levit, Smith, Cowan, Lazenby and Martin 2002), technological improvements (Cutler, McClellan and Newhouse 1998), and the level and growth of private and public insurance for medical goods and services (Peden and Freeland 1998, Finkelstein

2007). More generally, if health care is a superior good, it will comprise a greater portion of spending as incomes increase (Hall and Jones 2007). However, these structural factors have not increased the accuracy of predictions, partially because they rely on forecasting contemporaneous macroeconomic time series, such as GDP, which are themselves hard to forecast long term (Kitchen and Monaco 2003).

On the supply side, health care may be a sector with relatively low increases in productivity. Low productivity growth, combined with inelastic demand for medical care, causes large shifts of labor into health care and/or increasing wages in health care when compared with high productivity growth sectors (Baumol 1967). That means that health care is increasing as a share of the economy, but does not necessarily help with forecasting shocks to health care spending. It actually implies that forecasting is more difficult, since shocks to medical spending could come from innovations in health care or from productivity shocks in other sectors of the economy.

More limited work on the predictability of medical inflation also shows the difficulty of forecasting. Ewing, Piette and Payne (2003) forecast the CPI-U, Medical Care price series and its components to generate net medical discount rates (NMDR). However, they do not analyze the change in NMDRs or the factors affecting them. Prices alone may not be enough if the variance in spending mainly comes from quantity, but the decomposition of the variance in total spending is not well known. Bundorf, Royalty and Baker (2009) use the MarketScan commercial claims data and conclude that “most spending growth was driven by outpatient services and pharmaceuticals...”, and that, within those types of encounters,

the growth was driven primarily by quantity, not price.

What is important to insurers is the effect of general inflation and spending trends on the growth in spending on insured services. The National Health Expenditure Accounts includes the per capita cost of private insurance and Medicare. The costs are available on a total basis, which includes changes in all benefits, and a common benefits basis. The all benefits basis is the most relevant, because quantity changes are an important part of medical spending growth (Bundorf et al. 2009). Bundorf et al. (2009) also note that studies of spending growth have generally focused on Medicare, and especially physician services within Medicare. Studies such as Sisko et al. (2010) have focused on the rate of growth in spending in the overall economy. The medical inflation series in the CPI includes insurance spending as a component, but is also focused on economy-wide medical inflation (Bureau of Labor Statistics 2010).

2.5 Hedging liabilities

The classical example of hedging is in the elimination of the risk that raw materials will unexpectedly change price.

“The typical illustration to show the advantage of organized speculation to business at large is the use of the hedging contract. By this simple device the industrial producer is enabled to eliminate the chance of loss or gain due to changes in the value of materials used in his operations during the interval between the time he purchases them as raw materials and the time he dis-

poses of them as finished product shifting this risk to the professional speculator.” (Knight 1921, p.256).

The idea of the simple form of unloading of risk is to allow the producer to specialize in his area of expertise, while buffeting him from shocks that are outside his control. Insurers should be able to offer similar similar products, but moral hazard could make these policies infeasible (Shiller 1993, p. 2).

Transactions costs can also interfere with market making in products designed to allow for hedging. Entering into a hedging position involves exchanging business risk for financial risk. As a result, any firm offering hedging products must use a combination of capital and market mechanisms to reassure traders that the financial risk is (substantially) less than their business risk. In order to avoid expensive levels of capital, futures market makers can use a combination of margin accounts and daily contract settlement to eliminate counterparty risk (Kolb and Overdahl 2006, pp. 12-15). The Treasury Inflation Protected Securities (TIPS) program is an example where zero sum hedging did not induce enough interest from traders, but positive sum hedging through government issued bonds does (Dowd 1994, Irving 2009). However, it is unclear whether the value in TIPS is that they are bonds rather than futures contracts or that they are U.S. government backed instruments.

2.6 Securities to hedge medical spending

Investment strategies to capitalize on the results of spending prediction models are not well established. Medical prices are certainly hard to hedge. Jennings, Fraser and Payne (2007) find that health care mutual funds covering all health care, as well as subsectors in the industry, are a poor hedge for the CPI-U, Medical Care series. No TIPS bonds exists for the CPI-U, Medical Care series alone. Health insurance futures could allow hedging of insurance price risk and, as a by-product, produce market estimates of future medical inflation (Cox and Schwebach 1992). Other health care securities, such as the bonds that finance hospital construction, may be too thinly traded to include in a realistic investment policy.

In general, equities may be a poor or partial hedge for inflation based liabilities. Shorting common stocks, and buying commodities are two strategies that act as partial hedges for general inflation (Bodie 1976, Bodie and Rosansky 1980, Cooper and Kaplanis 1994). The Treasury's TIPS program now allows for perfect CPI-U hedging, with the continuing basis risk that the government's definition of inflation may not match that of the consumer. Inflation indexed bonds alleviate the need for complex investment strategies, such as short sales, as well as providing a market-based forecast of future inflation (Emmons 2000).

Investing in assets with returns correlated with medical spending growth may be difficult because claims on much of the medical care industry are unavailable to most investors. Most physicians operate as part of office based practices, not in hospitals (Cooper 2008). Not-for-profit hospitals account for 70% of all hospital beds (Cutler 2000, p. 291). Phar-

maceutical companies' profits may be driven more by patents and the timing of innovation than pricing or volume (Danzon 1999). Cutler (1993) also suggests that for these reasons, the market does not insure long term health risks.

If the drivers of medical spending are related to relative productivity between health care and other sectors of the economy, then there may be a wider range of hedging assets. It may be that manufacturing, which Baumol (1967) mentions, would be a good hedge (possibly as a short). If the demand side theory of Hall and Jones (2007) explains some of the shocks in spending, then bets on the economy would be good hedges.

Unfortunately, the stock market is generally a poor reflection of economic activity. This is especially true in the short term (Pearce and Roley 1985). Over a longer time horizon of 3-5 years, there is some evidence that, for example, wages are correlated with stock returns (Lucas and Zeldes 2006), so that long term hedging may be possible. It is difficult to find correlations between wages and stock returns in order to fund pension plans, such as those provided to public workers (Novy-Marx and Rauh 2008). Long term health insurance may be a more difficult product to manage than pensions, since the promise is a basket of goods rather than a fixed amount per month until death.

The absence of a hedge in the public securities market does make the risk uninsurable. Private firms could bundle related assets together to create the hedging asset. The risk of medical spending growth could be uncorrelated with the returns to market securities, which would mean that they have a beta of zero (Huang and Litzenberger 1988). Hedging assets are unnecessary if the plan can build up a reserve that, with investment returns, will pay

off the claims of the contract with a high likelihood. It is not known at what point medical spending risk becomes uninsurable in the sense that it is too costly to build up an adequate reserve.

2.7 Other plans with guarantees have failed

Plan failures in other health plans with premium guarantees are indicators that planning for medical spending is important. Retiree medical plans have caused bankruptcies of firms in the U.S. steel and automobile industries, because plans were not well funded (Larson 2009). The same is true of state and local governments, which are estimated to be \$1.5 trillion underfunded in total retiree medical plan obligations (Zion and Varshney 2007). Health insurance arrangements have also failed to pay claims in the past, as was the case with the Multiple Employer Welfare Arrangements (MEWAs) in the 1980s (McDonald 1992). The Actuarial Standards Board has noted the unique concerns about the "...level funded structure of LTC insurance and the long potential lags between receipt of premiums and their disbursement as benefits ...” (Actuarial Standards Board 1999).

The failure of other long term health plans shows the problems inherent in managing long term health liabilities. The failures result from the fact that long term care and retiree medical plans are borrowing short, and lending long. The result is that unless insurers can match the short duration payments with long duration claims, they expose themselves to large losses and the possibility of ruin (Lamm-Tennant 1989). In the case of postretirement benefits, employers may not even have a fund dedicated to future benefits: “An employer’s

practice of providing postretirement benefits may take a variety of forms and the obligation may or may not be funded.” (Financial Accounting Standards Board 1990). Whether funded or unfunded, these companies are not only shouldering future claims, but also the risk arising from the growth in medical spending. In the next section, I show how long term medical plans must be time limited, and how they are exposed to increasing medical spending risk.

Chapter 3

Guaranteed renewable insurance in an economy with risk

1 Introduction

I restate the model economy of Pauly et al. (1995). In this economy, agents face the possibility of a single loss in each period (intraproduct risk). The probability of loss is path dependent, meaning that there is an interperiod risk of becoming more loss prone. After making an assumption that long term contracts cannot be enforced, Pauly et al. (1995) propose guaranteed renewable insurance as the optimal contract ¹. To simplify the risk insured, the model is predicated on the law of large numbers in that the size of the population insured must be large enough so that the expectation and other moments of loss do not

¹Cochrane (1995) uses a mathematically equivalent setup and derives a contract with identical premiums and payouts.

contain an individual component.

Guaranteed renewability is the optimal solution under two key assumptions. First, insureds and insurance companies are playing a cooperative game. Second, explicitly multiperiod contracts are not enforceable. Given these assumptions, guaranteed renewability is the variance minimizing single period contract. It is superior to other one-sided commitment mechanisms, such as full prepayment of premiums, as well as period-by-period arrangements. Importantly, the solution is invariant to the proportion of individuals of the two types at any given time in the model. The fact that multiperiod contracts are ruled out means that the solution is second best.

The guaranteed renewable mechanism has two important limitations. The first is that it relies on a finite time horizon. This may be an explicit fixed horizon or a contract with a stochastic, but probabilistically finite, time horizon. However, in the latter case, the time horizon must come from a source other than lapsation. For example, death and discounting could be sources of limitation to the series of future payments.

I show how the model can be extended to allow for a finite time horizon or decrements that limit the scope of payments. I focus on discounting, and show how it is mathematically equivalent to medical inflation but different than death or a finite time horizon. Given these extensions to the Pauly et al. (1995) model, I motivate the importance of investment policies to manage these extensions, a subject I explore in detail in chapter 4. I develop formulas for the duration and convexity of premiums and claims, and show that the guaranteed renewable contract is prone to payment mismatches. I extend the the model for medical spending

growth and discounting, which have the same effect in opposite directions. That is, medical spending growth increases the need for prefunding, while discounting reduces the need for prefunding. As a result, the difference between medical trend and the return on reserves is a crucial determinant of the size of premiums.

2 Model economy

This section relies heavily on Pauly et al. (1995). The main assumption is that agents are either “low” or “high” risk and that insurance companies offer insurance through zero profit insurance contracts.

2.1 Discrete risky economy

In any period, agents face a single loss of amount L with probability p_i (i.e. a Bernoulli trial). The probability of loss is a function of the path of prior claims. Agents are considered to be a “low loss” type, and have a probability of loss p_L , while their claims history includes no incurred losses. A single loss makes the agent a “high loss” type, and in ensuing periods the probability of a loss is $p_H > p_L$. Knowing an agent’s type and whether he experienced a loss in the prior period is fully informative of his type in the current period.

In any given period, the two types of agents face the following lotteries:

$$\begin{array}{ll}
 (L, p_L; 0, 1 - p_L) & \text{Low type} \\
 (L, p_H; 0, 1 - p_H) & \text{High type}
 \end{array} \tag{3.1}$$

These lotteries can be generalized as random variables. The random variable Z is a function of the parameters for the probability and loss amounts. In addition, it is a function of the current type of the individual (“low” or “high”) as well as the single period over which the random variable is evaluated. $Z_L(0)$ is for a low type individual for the single period starting at time 0 (0 to 1), $Z_H(2)$ is for a high type individual for the single period starting at time 2, (2 to 3) and so on.

The transition matrix for agents is given by:

$$\begin{array}{cc}
 & \begin{array}{cc} L & H \end{array} \\
 \begin{array}{c} L \\ H \end{array} & \begin{pmatrix} 1 - p_L & p_L \\ 0 & 1 \end{pmatrix}
 \end{array} \tag{3.2}$$

I illustrate this evolution of types for a given individual in figure 3.1. In section 2.2, I show the implications of this transition structure for the composition of the population.

In addition to the single period losses faced by low types, they face the additional risk of becoming high types. Low type agents’ risk is split between the intraperiod and interperiod risks. High types retain their character for the entirety of the model, so that this is an absorbing state. This certainty means they only have intraperiod risk.

A single period risk exchange contract is the technology available for agents to reduce their idiosyncratic risk. The other actor in this economy, the insurance company, is the counterparty to the contract. I assume that the contracts are negotiated between insureds and insurers under a cooperative bargaining framework (Nash 1950), and that knowledge of the distribution of claims and agents’ type is fully known by all parties, but multiperiod

contracts are unenforceable ².

I consider only full insurance in this chapter. The main theoretical reason is that, under actuarially fair contracts, insureds will always choose full insurance (Schlesinger 1981). I show the two period model as a decision tree in figure 3.2. The figure shows sequential moves by agents and nature. At time 0, all individuals are low types. Then, each individual chooses insurance or uninsurance at full coverage. I denote the choice of insurance by Ins_i where $i \in \{L, H\}$ denotes the individual's type ³. I denote the choice of uninsurance by $Unins_i$ where $i \in \{L, H\}$ denotes the individual's type. Then, nature moves, generating a loss L with probability p_L and no loss with probability $1 - p_L$. Insured agents have no loss in either state of nature, while uninsured agents pay the loss amount L in the loss state. Then, agents make the same choice of insurance or uninsurance, this time given the possibility that they are now high types. I use dashed lines to indicate the portion of the figure for high types.

The premium function takes three parameter values that define the lottery agents face. The first parameter is L , the size of the loss. I constrain L only to be non negative, so $L \in (0, \infty)$. Second, there is a probability of loss p that comes from a binomial distribution (Bernoulli trial) for losses. There are two degenerate cases, $p = 0, p = 1$, but in general I will consider only nondeterministic probability distributions, $p \in (0, 1)$. The

²If multiperiod contracts were enforceable, companies would offer a flat premium structure (noncancelable contracts) that perfectly smoothed consumption, which is the first best contract.

³ Ins_i entails different premiums under different solutions. Since the premium will differ in the spot, guaranteed renewable, and prepayment cases, I leave the actual amount of the payment abstract.

third parameter defines the number of periods of insurance that are being provided. This is denoted N . For example, $N = 1$ is spot insurance.

I put all these parameters together into the spot premium function G_i when $N=1$:

$$G_i = p_i L \quad (3.3)$$

Here, I am using the type i instead of the usual subscript x for the age of the insured. This is because in the model the aging effect comes from the fact that, over time, individuals tend to become “high” types.

2.2 Aggregate risky economy

Given the risks faced by individuals, I show the aggregate level of loss in the economy. This section is predicated on the law of large numbers in that the size of the population insured must be large enough so that the expectation and other moments of loss do not contain an individual component.

I show the population proportion of the two types in table 3.1. Since there is no death or other decrement out of the population other than dropping coverage, the measure of the total population is always 1. In addition, the relative measure of types depends on p_L alone, due to the fact that the high type is an absorbing state.

Proposition 1. *Assuming no lapsation, the measure of high and low type individuals is driven solely by the probability distribution of low types.*

See Appendix A for the proof.

In addition, the ratio of low types for each high type tends to 0, possibly quite quickly depending on the value of p_L . For the three example values of p_L , 0.05, 0.10, 0.2, the number of periods it takes to get an equal measure of low and high types is not large.

The aggregate losses in any period are a function of the loss distribution and the risk profile of the population. I introduce the deferred single period insurance premium ${}_k|G_i$ that finances the expectation of the loss k periods from now for an individual *currently* of type i . ${}_k|G_i$ is the actuarially fair premium for this individual, based on the expectation over the population. If $k = 0$, this is the spot price, whereas if $k > 0$, this is a forward price. The first and second moments of this variable for spot one period insurance is given by:

$$\begin{aligned}
{}_0|G_L &= \mathbb{E}[Z_L(0)] = p_L L \\
\mathbb{E}[Z_L(0)^2] &= p_L L^2 \\
\text{Var}[Z_L(0)] &= p_L(1 - p_L)L^2 \\
{}_0|G_H &= \mathbb{E}[Z_H(0)] = p_H L \\
\mathbb{E}[Z_H(0)^2] &= p_H L^2 \\
\text{Var}[Z_H(0)] &= p_H(1 - p_H)L^2
\end{aligned} \tag{3.4}$$

These moments come directly from the binomial distribution formulas. Notice, however, that for high types, at any time N , ${}_0|G_H = {}_k|G_H$, and this is true for any moment of the distribution. This is a function of the fact that high types stay as high types, as well as the fact that there is no discounting.

For low types, this is not the case. Here, to find the moments of the distribution at later

periods, I have to use the measure of types obtained previously. For instance:

$$\begin{aligned}
{}_1|G_L &= \mathbb{E}[Z_L(1)] = p_L L(1 + (p_H - p_L)) \\
\mathbb{E}[Z_L(1)^2] &= p_L(1 + (p_H - p_L))L^2 \\
\text{Var}[Z_L(1)] &= p_L(1 + (p_H - p_L))(1 - p_L(1 + (p_H - p_L)))L^2
\end{aligned} \tag{3.5}$$

I calculate the expected loss k periods hence by using the measure of low and high types k periods in the future:

$$\begin{aligned}
{}_k|G_L &= p_L L(1 + \sum_{i=0}^{k-1} (1 - p_L)^i (p_H - p_L)) \\
\mathbb{E}[Z_L(k)^2] &= p_L(1 + \sum_{i=0}^{k-1} (1 - p_L)^{i-1} (p_H - p_L))L^2 \\
\text{Var}[Z_L(k)] &= p_L(1 + \sum_{i=0}^{k-1} (1 - p_L)^{i-1} (p_H - p_L)) \times \\
&\quad (1 - p_L(1 + \sum_{i=0}^{k-1} (1 - p_L)^{i-1} (p_H - p_L)))L^2
\end{aligned} \tag{3.6}$$

The existence of the high types could reduce or increase the variance of the losses in the population. The binomial variance, $p(1-p)$ is a quadratic that is maximized at $p = 0.5$. The crucial factor is how much higher the “high type” probability is than the low type. If $p_L < p_H \leq 0.5$, then high types have a higher variance in their losses than low types. Inversely, if $0.5 \leq p_L < p_H$, then high types have a lower variance. If $p_L \leq 0.5 \leq p_H$, then I can take the general rule that $|p_L - 0.5| < |p_H - 0.5| \Rightarrow \text{Var}[Z_L] > \text{Var}[Z_H]$ (the inverse is also true).

Proposition 2. $|p_L - 0.5| < |p_H - 0.5| \Leftrightarrow \text{Var}[Z_L(k)] > \text{Var}[Z_H(k)]$.

See Appendix A for the proof.

The crucial note from this solution is that, over time, the variance in spot and forward premiums comes from two sources. One is the overall variance embedded in the single period lottery. The second is the variance deriving from low types not knowing ahead of time what their future type will be. Low types could solve this problem by buying forward contracts for the deferred losses. However, buying all required forward contracts today implies a very unsmooth profile of premium payments over time (to say nothing about the affordability of such prepayments). I now discuss the solution to the balance between prepayment and premium smoothing, guaranteed renewability.

3 Guaranteed renewable solution

The following section restates the main results of Pauly et al. (1995), with extensions to the reserves and general actuarial notation. The argument for construction of the contract through backward induction, starting with the final period, but with the introduction of cooperative bargaining. Cooperative bargaining allows me to explain why insurers offer the guaranteed renewable contract without having to explicitly model the insurer's objectives. It is enough for me to show that the contract is second best for insureds, and that the first best contract is unenforceable.

3.1 Fair premium

Pauly et al. (1995) show that the guaranteed renewable premium at times 0 and $0 < k \leq N$ for a contract with an original duration of N periods is:

$$\begin{aligned} {}^{GR}G_{L:\overline{N}|} &= p_L L \left(1 + \sum_{j=0}^{N-1} (1 - p_L)^j (p_H - p_L) \right) \\ {}^{GR}G_{L:\overline{N-k}|} &= p_L L \left(1 + \sum_{j=0}^{N-k-1} (1 - p_L)^j (p_H - p_L) \right) \end{aligned} \quad (3.7)$$

What is significant is that the premium at time k for one period of GR insurance with a remaining term of $N - k$ periods is the same as the premium for *de novo* GR insurance with a term of $N - k$.

This contract is sequentially rational for all parties, and thereby becomes a long term contract structured out of single period contracts. The premium is known ahead of time, as opposed to spot insurance. This has the effect of reducing the variance in premium payments significantly.

The first two moments of the per period payments by insureds are:

$$\begin{aligned} \mathbb{E}[\text{Premium}] &= N^{-1} \sum_{j=1}^N {}^{GR}G_{L:\overline{j}|} \\ \mathbb{E}[\text{Premium}^2] &= N^{-1} \sum_{j=1}^N ({}^{GR}G_{L:\overline{j}|})^2 \\ \text{Var}[\text{Premium}] &= N^{-1} (\mathbb{E}[\text{Payment}^2] - N^{-1} \mathbb{E}[\text{Payment}]^2) \end{aligned} \quad (3.8)$$

The point here is that the premium is known ahead of time. This has the effect of reducing the variance in the payments significantly. The only remaining variance, as compared to the fully levelized option discussed below is the “aging” effect of the difference in low

and high type loss probabilities ($p_H - p_L$).

3.2 Contract reserve

The overall reserve is function of the reserve for high types and the measure of high types, since the low type reserve is always 0.

$${}_tV = \mu_{Ht}^t V_H \quad (3.9)$$

The reserves fund a cross-subsidy. The subsidy reduces the premium paid high type consumers on an *ex interim* basis ⁴. In any period, subsidies are a way of inducing an intertemporal transfer of wealth from those who stay in the low type to those who transition to the high type. Since these subsidies can also be thought of as a return of the reserves, proposition 3 is important: it shows that low types can never get subsidies, and thus the actuarially fair premium is a lower bound on what they will be expected to pay in any period.

Proposition 3. *The guaranteed renewable reserve for low types is 0 at all times.*

See Appendix A for the proof.

Proposition 4. *The guaranteed renewable reserve for high types has a lower bound of 0.*

See Appendix A for the proof.

The reserve ${}_tV_H = G_{H:\overline{N-t}} - {}_tG_{L:\overline{N-t}}$ is equivalent to a “cash settlement value” that Cochrane (1995) and Cochrane (2009) use to create health status insurance, which is guaranteed renewable insurance where reserves are owned individually through health status

⁴There can be no subsidies *ex ante* under the rule of actuarially fair contracts.

accounts. The account is the compensating payment that an insurer makes to a consumer to leave the insurance pool based on the shadow price of annual renewal (spot) insurance.

$${}_tV_H = (p_H - p_L)L \sum_0^{N-t-1} (1 - p_L)^t \quad (3.10)$$

The contract can only be implemented when the insurer anticipated the reserves required for current and future high types, or has adequate capital to make up the difference. The guaranteed renewable premium builds up adequate reserves in the prior periods to implement the final period contract I describe. It does so in the cheapest possible way (when cost is a function of the variance) under the assumption of single period contracts.

3.3 Final period contract

I first solve from the final period, showing that it is a special case where the only possible contract is the spot market price for low types.

Proposition 5. *The only possible contract in the core in the final period is G_L , the spot premium for low types.*

See Appendix A for the proof.

The insurer decides on this level of reserves in the penultimate period by observing that the only possible premium in the final period is the low type premium. Therefore, the insurer must charge enough in the penultimate period to build up these amount of losses. I consider two situations: one where the insurer is insuring only low types (i.e. a *de novo* two period guaranteed renewable contract) and one where there is a mix of high and low

types (i.e. where $N > 2$). In the either case, the only supportable equilibrium is the same premium, because otherwise low types will deviate. Again, this relies on the same recursive argument that there are enough reserves built up in the prior period to prefund the premiums of high types, or that insurers have some capacity to absorb losses due to adverse experience.

4 Alternatives to guaranteed renewability

I briefly review the alternative solutions to the multiperiod problem of providing insurance when types change. I calculate the interperiod variance of premium payments by insured individuals under different contractual arrangements. I show that the guaranteed renewable solution dominates all solutions that rely on period-by-period contracting, even though it is still inferior to a contract that fully levelizes premiums over all periods.

4.1 Single premium insurance

If the only constraint is that insurers cannot enforce long term contracts, other available long term contracts is fully prepaid insurance. The difference between the two types of contracts is that fully prepaid contracts generate higher initial reserves along with the higher premiums. Spot contracts, in contrast, generate no reserves, only realized gains or losses.

The premium schedule for full prepayment is:

$$\sum_{j=0}^{N-1} j |G_{L:\overline{N}}| * N = G_{L:\overline{N}}^1 \quad \text{at time 0}$$

$$0 \quad \text{in all other periods}$$

The main problem with this single premium arrangement is that the variance of premium payments by the insured is extremely high considered over multiple periods. The average payment in any given period and the variance is given by:

$$\mathbb{E}[Payments] = N^{-1} G_{L:\overline{N}}^1$$

$$\mathbb{E}[Payments^2] = N^{-1} (G_{L:\overline{N}}^1)^2$$

$$\text{Var}[Payments] = N^{-1} (1 - N^{-1}) (G_{L:\overline{N}}^1)^2$$

This is much higher than the intertemporal variance of the fully level payments solution, which is 0, or the GR solution. In this way, the level payments solution second order stochastically dominates the up front premium payment solution. Spot insurance also stochastically dominates full prepayment, based on the fact that the range of the spot payments is within the range of the single premium payments.

4.2 Spot insurance

The spot premium solution to this game is the period-by-period solution. I compute the set of one period premiums spot and forward premiums $Prem_i^k$ that a risk-neutral agent of type i will accept for insurance k periods hence. I will denote this set as \mathcal{P}_i .

$$\mathcal{P}_i = \{Prem_i^k : Prem_i^k \leq {}_k G_i\} \quad (3.11)$$

I also show the set of acceptable contracts in figure 3.3. The shaded area represents contracts that risk-neutral agents facing the lottery $(L, p; 0, 1 - p)$. This area includes the actuarially fair contracts on the line $(0, p_i L), (L, 0)$, which agents are strictly indifferent between accepting and rejecting. I assume that they accept these contracts.

I also consider the set of contracts that insurers will write. While writing insurance is not a choice in this model, I assume that insurers will not write actuarially favorable contracts. I therefore compute the set of one period premiums $Prem_i^k$ that the insurance companies will accept. I denote this set as \mathcal{P}_I .

$$\mathcal{P}_I = \{Prem_i^k : Prem_i^k \geq {}_k G_i\} \quad (3.12)$$

I also show the set of acceptable contracts in figure 3.4. Here, I show the acceptable contracts for both types of agents, L, H . The shaded area represents contracts that risk-neutral agents facing the lottery $(L, p_i; 0, 1 - p_i)$. This area includes the actuarially fair contracts on the line $(0, p_i L), (L, 0)$, which agents are strictly indifferent between accepting and rejecting. I assume that they accept these contracts.

In order to construct coalitions that are immune to deviation, I first investigate the overlaps of the sets of acceptable premiums.

$$\mathcal{P}_{SS}^* = \mathcal{P}_i \cap \mathcal{P}_I = \{{}_k G_i\} \quad (3.13)$$

If insurance companies are constrained to offer only one type of insurance contract ⁵, they will offer $Prem_H$, which high types will accept, and then low types are left out

⁵This concept may seem unlikely, but since I am constraining insurance companies to offer zero profit contracts, it is sensible to apply the solution concept to this simplest scenario.

of the coalition. If they can offer two types of contracts, they will offer two contracts, $Prem_H, Prem_L$ to high and low types (respectively) ⁶. In this case, the coalition contains all players. Since both contracts are zero profit, they are immune to individual deviation.

4.3 Enforceable long term contracts

The first best contract is a level premium in all periods. The premium in each period is:

$$N^{-1} \sum_{j=0}^{N-1} k |G_{i:\overline{N}}| = N^{-1} p_L L (N + \sum_{j=1}^{N-1} (N-j)(p_H - p_L)(1 - p_L)^{j-1}) \quad (3.14)$$

Since the premium is the same in every period, the interperiod variance in payments is 0. In the absence of a discount rate, this sums to an up front single premium, but with lower variance. Therefore, level premiums are the dominant contract. If insureds and insurers can sign unbreakable contracts, I can also show that this is immune to deviation, whether by starting with all low types or with a mixture of low and high types where each pays a levelized premium equal to their expected future claims.

Proposition 6. *Level premiums are the second order stochastically dominant contract. Fully prepaid contracts are second order stochastically dominated by all other forms of insurance.*

Proof. These are mean preserving spreads. The range of level premiums is nested in the range of GR premiums. The range of GR premiums is nested in the range of spot premiums. The range of spot premiums is nested in the range of single premium payments ⁷. □

⁶This is where I rely on full information.

⁷An alternative argument the the mean preserving spread is a recursive proof. If I choose any period other

This proof is by construction: I have to specify that individuals are unable to break the contract. I could make an alternative specification that individuals who break their contract must pay a penalty equal to their future premiums owed. Even better, I could define a penalty function based on the premium payable and the expected future losses. Denote the current period as k and the agent's current type as i .

$$Penalty_i^k = {}_t^L V_i - {}_t^{GR} V_i \quad (3.15)$$

where ${}_t^L V_i$ is the level premium reserve for type i individuals and ${}_t^{GR} V_i$ is the GR reserve.

Whether by forcing individuals to remain in the contract, or by enforcing a penalty, the level premium solution relies on outside enforcement of long-term contracts. However, a large part of the dynamic contracting literature derives from the fact that these contracts may be difficult to enforce, leading to implicit contracts (Bull 1983, e.g.) or futures contracts that are marked-to-market daily.

5 Limits of guarantees

5.1 Temporal limits of guaranteed renewability

There is a temporal limit to guaranteed renewable protection. In an infinite time model with no discounting, guaranteed renewability is no protection. The reason is that at some point, than the final one, and try to lower the contract premium by ε , then I have shown that I will have to raise subsequent premiums by ε . However, subsequent premiums are already at a maximum from the point of view of sequential rationality. Any higher and the low types will surely deviate. Therefore, this is the minimum variance contract with single period contracting.

all individuals will become high types in the absence of death (see section 6, and in the limit, high losses will swamp low losses. Therefore, the equilibrium must be a separating equilibrium, in which both types buy spot insurance that reflects their current probability of loss.

Proposition 7. $\lim_{k \rightarrow \infty} k|G_i = p_H L$

See Appendix A for the proof.

The proof extends to the situation where the high type is not absorbing, but there is some probability of becoming low risk again. In the limit, the guaranteed premium for low types is equal to the long run weighted average of premiums for the time spent in each state, conditional on starting as a high type. Therefore, the infinite time guaranteed premium is the same as the spot premium for high types, which is no protection at all.

5.2 Investment policy limitations

In the next section, I show how discounting can solve this temporal problem. This gives rise to an investment policy problem, in that the accumulation of a reserves is no longer trivial. Once I consider a dynamic model with time value of money, higher returns on risky assets, and other standard finance model characteristics, where the probability of ruin comes into play. This investment policy problem is a general one in insurance, where insurance companies have assets to back evolving liabilities, and would like their assets to evolve in such a way as to cover their liabilities. This question is intimately connected with the question of how affordable GR contracts are given the amount of front-loading

necessary in the real world (Herring and Pauly 2006). I investigate this question in greater detail in chapter 4, but I include it here as a limitation.

6 Extensions of the guaranteed renewable model

There are three ways of “solving” the time horizon problem: discounting, death, and a finite horizon. Discounting introduces the time value of money, reducing the present value of future losses. In order to use discounting alone, the relative magnitude of interest and loss rates by low types is the key factor. Death is a special type of discounting with the special property that it is more likely for high types than for low types. It not only introduces a shorter time span for losses, but it also makes high types cheaper than they would otherwise be. Finally, an explicit finite time horizon could be introduced, such as the recognition that at 65, individuals will be taken on to Medicare. The finite time horizon could also be probabilistic, such as random lapsation that eventually causes the risk pool to empty.

6.1 Discounting

With a constant discount rate $\beta < 1$, the premium in the k^{th} year of an N year policy is:

$$\begin{aligned}
 {}^{GR}G_{L:\overline{N-k}|} &= \beta[p_L L][1 + \beta(p_H - p_L) + \\
 &\beta^2(1 - p_L)(p_H - p_L) + \beta^3(1 - p_L)^2(p_H - p_L) + \\
 &\quad \dots + \beta^{N-k-1}(1 - p_L)^{N-k-2}(p_H - p_L)] \\
 \lim_{N \rightarrow \infty} {}^{GR}G_{L:\overline{N-k}|} &= \beta[p_L L][1 + \frac{\beta(p_H - p_L)}{1 - \beta(1 - p_L)}] \\
 \lim_{N \rightarrow \infty} {}^{GR}G_{L:\overline{N-k}|} &= \beta[p_L L][\frac{1 - \beta(1 - p_H)}{1 - \beta(1 - p_L)}]
 \end{aligned}$$

Note that in a certain sense, discounting is not a solution to the lack of protection in an infinite horizon model. All it does is to lower the real value of future medical payments. To the extent that this is true, the insured simply passes on these cost “savings” to the insured in the form of lower premiums under the zero profit contract assumption. This also implicitly assumes that there is a nominal interest rate but no inflation rate in the magnitude of loss L . More reasonable is the assumption that losses inflate by an amount equal to overall inflation, at the very least. Then, in a world with a long term real interest rate of 3% (Girola 2005), the infinite time GR contract could cost over 34 times the spot rate for low types ($\frac{1}{1-\beta}$). Even a very conservative assumption that $p_H = 0.2, p_L = 0.1$ leads to a premium 77% higher than the spot premium. Then, adding in the fact that medical inflation is even higher than general price inflation makes this even less likely to be a source of cost reduction.

The same is true of medical spending growth. Any change to the loss level, whether

deterministic or stochastic, is compounded because of the prepayment mechanism. If I drop the assumption of zero trend, or allow for trend to be a martingale, the contract continues in its guaranteed renewable form, but with the premium anticipating the future trend. If, over time, the loss for all types is subject to a linear transformation of the form:

$$L \mapsto aL \quad (3.16)$$

then the contract generalizes to:

$$G_{L:\overline{N}|} = p_L L \left(1 + \sum_{j=0}^{N-1} (a(1-p_L))^j (p_H - p_L) \right) \quad (3.17)$$

$${}_k|G_{L:\overline{N-k}|} = p_L a L \left(1 + \sum_{j=0}^{T-k-1} (a(1-p_L))^j (p_H - p_L) \right)$$

I use this simple transformation to model medical trend and discounting. The general result is that, since insureds are prepaying part of their future costs, even forecastable changes to medical costs must be factored into the premium function. This allows me to visualize the contract as an annuity or a collection of contingent claims.

6.2 Death

Death can be introduced probabilistically. If instead of a discount factor β there is a homogeneous probability of survival α , the mathematics are exactly the same as before. If the probability of death is correlated with the type, then it works differently than discounting. Say that high types die with probability $1 - \alpha > 0$, and death is a decrement that occurs at the beginning of the period, before the payment of premiums or experience of losses. In a three period model, the measure of individuals by period μ^t are given in table 3.2.

The expected losses by period are:

$$\mathbb{E}[Losses_1] = p_L L$$

$$\mathbb{E}[Losses_2] = p_L L((1 - p_L) + \alpha p_H) \mu$$

$$\mathbb{E}[Losses_3] = p_L L((1 - p_L)^2 + \alpha p_H(\alpha + (1 - p_L))) \mu^2$$

I have to include the measure of the population, which is no longer 1, in the premium function.

Death can also be introduced as a finite horizon. For instance, I could assume that no one lives beyond age 100. This is equivalent to another time limitation for the GR model, which is a finite time horizon.

6.3 Finite time horizons

A finite time horizon assumption is used in Pauly et al. (1995). There are several possible rationales for using a finite horizon model. Medicare will kick in at age 65. It is hard to enter into long term insurance contracts (or long term contracts in general). Since it does limit the premium to an amount less than the high type spot premium, and thus ensures that low type agents will have an incentive to engage in GR contracts. It is important to note that a finite model does not affect the decision rule by giving agents an incentive to act differently in the final period than in intermediate periods, since agents make the same insured/uninsured decision in all periods as long as they are motivated to purchase insurance in the first place.

7 Lapsation

Unlike discounting, death, or a finite time horizon, lapsation does not generate an upper bound on the guaranteed renewable premium. The reason is that even if all the high types lapse, the low types will eventually become high types. The effect of unanticipated high type lapsation is to generate gains for reserve pool. Low type lapsation has no effect on the reserve pool, since the reserve associated with each low type individual is zero. Lapsation also affects pricing if insurers are following a “class average” rule rather than a strict guaranteed renewability rule. In that case, lapsation drives the premium towards the high type premium unless all lapsing members are low type. Alternatively, the insurer could stick to the original guaranteed renewable rates based on what the premiums would have been without any lapsation.

7.1 High type lapsation

High type lapsation, or favorable retention, generates gains for the insurer. As an example, consider an N term policy after the end of the first year. There are a p_L measure of high types and a $1 - p_L$ measure of low types. Let us say that, due to some exogenous factor (such as emigration), the measure of high types is in fact αp_L with $\alpha \in (0, 1)$. The premium for the next period if the insurer were using this experience should be the same as in the death case. As a result, all remaining individuals face a lower premium than they would have without the lapsing individual.

However, it is not clear that the gains from the high type lapsation should be shared in

this way. The economic model dictates a path of premiums based on a “shadow” population of insureds, and the insurer could legitimately stick to this path of premiums, which will generate economic (positive) profits in expectation. The rationale for allowing the insurer to keep the excess reserve in this case is that we would want the insurer to follow this strategy in the opposite case, where low types lapse for exogenous reasons.

7.2 Low type lapsation

Low type lapsation is neither favorable nor adverse retention. Low types carry a zero reserve, so their decision to keep or drop coverage does not affect the profitability of the contract. However, under a class average rule for rates, they cause rates to climb and thereby generate gains for the insurer by leaving. The reason is that the reserve supports continuing the original guaranteed renewable pricing for high types without necessitating a rise in premiums. Again, under cooperative bargaining, the insurer’s incentive to engage in this behavior is assumed away, and class average rates are ruled out by guaranteed renewable rules. This is the downside of the class average rule as opposed to the economic guaranteed renewable rule.

7.3 Random lapsation

A third type of lapsation is random or “ ε ” lapsation. I am calling it ε lapsation to denote the possibility that individuals may drop coverage that it would be in their interest to keep (or that they are invariant to) due to some small error on their part or exogenous fluctuation

unrelated to health status. As with the cases of low and high type lapsation, random lapsation does not adversely affect the surplus position of the insurance pool, and enhances it to the extent that high types lapse. Again, the question of how the surplus is allocated is one that my model does not explicitly address because of the cooperative bargaining solution. Indeed, the problems of splitting the surplus are even thornier in a world where the pool is composed of a mix of high and low types, as would be the case in a mutual insurer with participating policies. In addition, class average rating is impossible as long as low types remain, since the class average rates are above guaranteed renewable rates that are keeping low types in the pool. In this way, the continuing existence of some low types in the pool is providing protection to current high types.

8 Partial insurance

Partial insurance is an extension that I do not consider. This is because I have utilized a full information game where deductibles are neither necessary nor desirable. Allowing for private information leading to adverse selection or moral hazard will cause heterogeneous income effects that destroy the pooling equilibrium. There is a negative income effect of inducing insurance companies to write partial insurance contracts, for instance with deductibles. However, it is worth noting that low types will want higher deductibles than high types (Schlesinger 1981, e.g.). Even if insurers could utilize perfect revelation mechanisms in order to sort insureds into the “correct” contracts, the transition from low type to high type would be associated with the shock of a higher premium associated with a lower de-

ductible (i.e. independent of the additional losses associated with high types), and so the protection from GR would in a sense be lessened.

9 Final remarks

I have shown in this chapter that guaranteed renewable insurance is a second best insurance contract. It is inferior in that the first best solution, fully level payments, produce a lower variance but are not enforceable by the insurance company on the insured. It is superior to all other solutions in that it is the variance minimizing single period arrangement.

I have also shown that the solution relies on a cooperative bargaining framework and complete information. The cooperative bargaining framework leads to insurers offering zero profit contracts, and to insureds accepting these contracts even if they are risk neutral. In addition, it removes the incentive that insurers would have to “game the system” by inducing high types to drop their coverage and lose their portion of the surplus. Similarly, it relies on complete information, which removes the possibility of adverse selection and moral hazard. This leads the insurer to set up a pool of homogeneous, low type insureds at the beginning of the contractual arrangement.

This pool must by necessity protect insureds over a limited time horizon. The fact that high type is an absorptive state means that the guaranteed renewable premium converges to the single period high type premium as time tends to infinity. Discounting, death, or an explicit term for the insurance all remediate this problem, but this comes from a lower overall value for the contract (in the discounting and death cases), and a reduced protection

for insureds (in the explicit term case). In addition, there is no allowance for discounting and medical inflation, which would aggravate the time horizon problem. This is a problem I deal with in the next chapter.

10 Tables and figures

				Proportion		
Type				p_L		
t	μ_L^t	μ_H^t	$\frac{\mu_L^t}{\mu_H^t}$	0.05	0.1	0.2
0	1	0	N/A	N/A	N/A	N/A
1	$1 - p_L$	p_L	$\frac{1-p_L}{p_L}$	19.000	9.000	4.000
2	$(1 - p_L)^2$	$p_L(1 + (1 - p_L))$	$\frac{(1-p_L)^2}{p_L(1+(1-p_L))}$	9.256	4.263	1.778
3				6.011	2.690	1.049
4				4.391	1.908	0.694
5				3.420	1.442	0.487
6				2.775	1.134	0.355
7				2.315	0.917	0.265
8				1.971	0.756	0.202
9				1.705	0.632	0.155
\vdots						
N	$(1 - p_L)^N$	$p_L(\sum_{i=0}^{N-1} (1 - p_L)^i)$	$\frac{(1-p_L)^N}{p_L(\sum_{i=0}^{N-1} (1-p_L)^i)}$	N/A	N/A	N/A

Table 3.1: Measure of types

Beginning of Period

t	μ_L^t	μ_H^t	μ^t
0	1	0	1
1	$1 - p_L$	αp_L	$1 - p_L(1 + \alpha)$
2	$(1 - p_L)^2$	$\alpha p_L(\alpha + (1 - p_L))$	$1 - p_L(1 - \alpha)[(1 + \alpha) + (1 - p_L)]$

Table 3.2: Measure of types over time

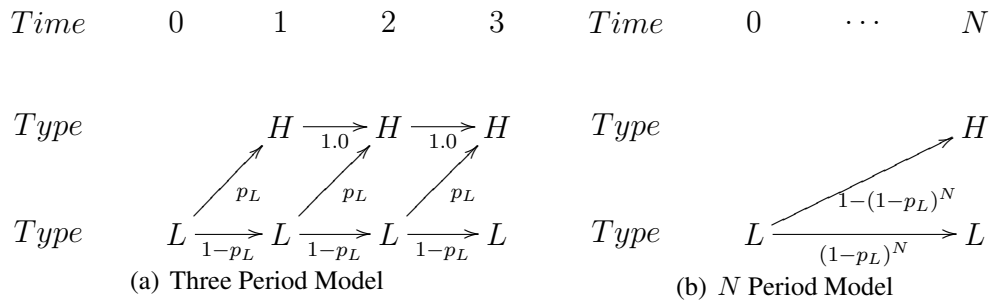


Figure 3.1: Evolution of Membership

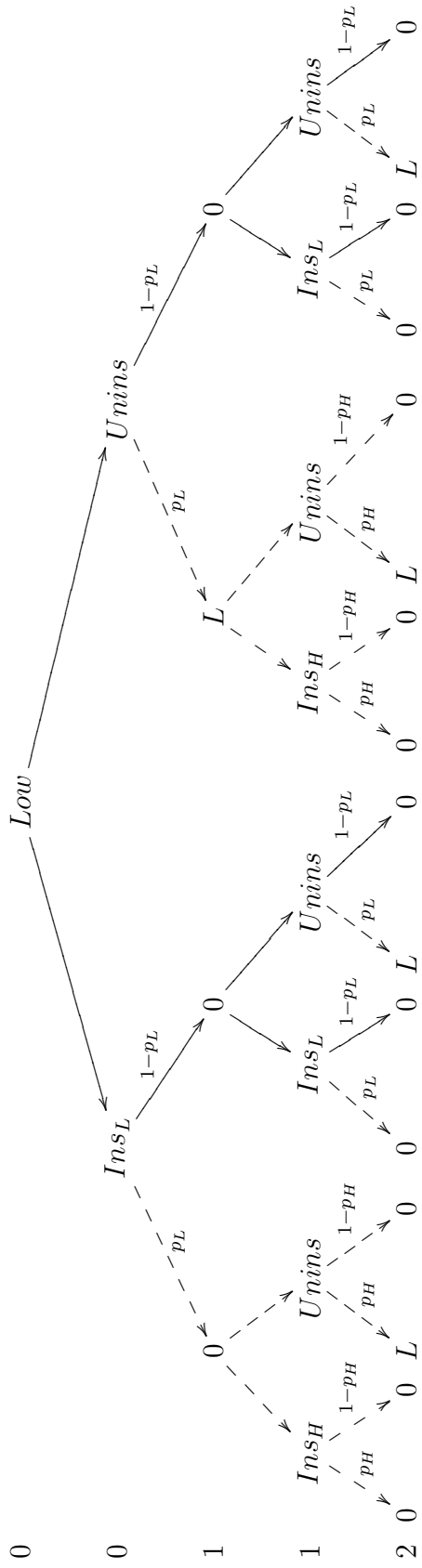


Figure 3.2: Two period model with insurance

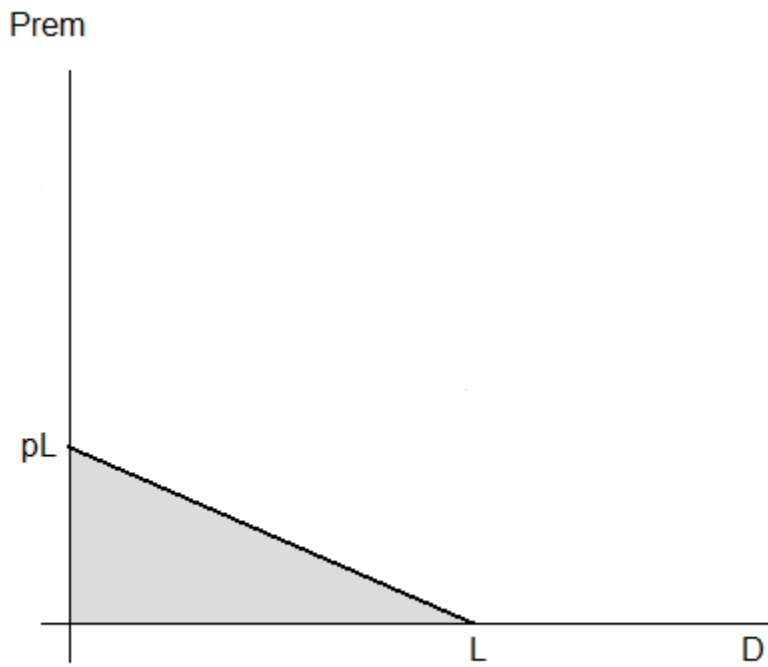
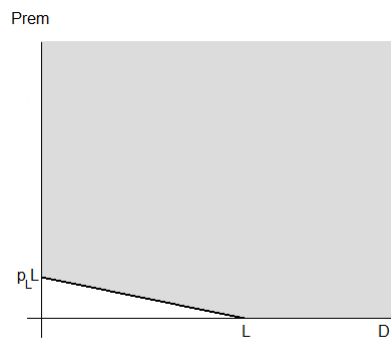
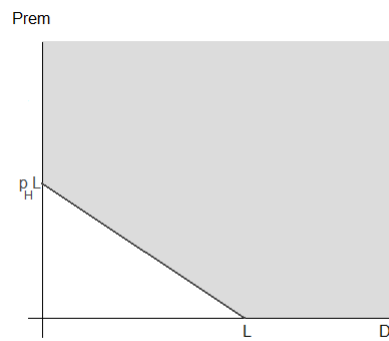


Figure 3.3: Set of acceptable premiums under full, zero-load insurance



(a) Acceptable premiums for insuring low types



(b) Acceptable premiums for insuring high types

Figure 3.4: Acceptable premiums for insuring each type

Chapter 4

Investment policy for guaranteed renewability

1 Introduction

In chapter 3, I show that guaranteed renewable insurance is the the optimal minimum variance insurance contract under the assumptions that the loss level does not change. The optimal contract involves front loading, serving to bind the insured to the insurer in the future. This front loading allows insureds whose probability of loss increases to pay premiums below the spot market amount, while making it sequentially rational for those with a low probability of loss to remain in the contract. This is despite the fact that insureds are free to lapse from this insurance at any time. I also show how the model can be extended to accommodate discounting, medical inflation, and other expansions of the model.

Discounting, in particular, shows that the main vulnerability of the contract is in the

mismatch of cashflows. The one-sided commitment mechanism is designed such that premiums are decreasing while benefit payments are increasing. This makes the duration of premium payments smaller than the duration of benefit payments. Actuarial standards, as well as finance, call for an investment policy to match these two cash flows.

The same is true of medical inflation. Any unexpected change to the loss level, whether deterministic or stochastic, causes losses (or gains) via multiple years of forecast claims. In this chapter, I extend the model without the assumption of zero medical inflation (trend). If I drop the assumption of zero trend, or even allow for trend to be a martingale, the contract continues in its guaranteed renewable form, but with the premium anticipating the future trend. However, there may not be a set of assets that exist to construct the correct investment policy due to market incompleteness. I explore this question empirically in chapters 5 and 6.

I demonstrate that the contract must be written with respect to the expected future size of losses. Pricing the insurance in any other way violates the assumption that the contracts do not cause losses for the insurer or allow low type insureds to profitably deviate in future periods. Whether the trend is deterministic or stochastic does not matter in the calculation of the zero premium.

2 Numerical example

A numerical example shows the mismatch between assets and liabilities when discounting is introduced. This mismatch shows up in standard finance primitives, including duration

and convexity. If the model admits dynamic elements, such as a discount rate subject to change, this mismatch can cause real gains or losses by the insurance company. Investment policy alone may not be enough to solve this problem.

2.1 Numerical parameters

I also fix values of the parameters in the model:

$$L = 1000$$

$$p_L = 0.04$$

$$p_H = 0.09$$

This gives the important probability “wedge”, $\Delta = p_H - p_L = 0.05$. In addition, I start with an interest rate $i = 2\%$ such that $\beta = v = 0.98039$. I also use an alternative rate $i' = 3\%$ such that $\beta' = v' = 0.97087$.

2.2 Duration mismatch under discounting

In the absence of discounting and inflation, The losses and premiums by period are:

$$\mathbb{E}[L_1] = 40$$

$$\mathbb{E}[L_2] = 42$$

$$\mathbb{E}[L_3] = 43.92$$

$$\sum \mathbb{E}[L] = 125.92$$

$$G_{L:\overline{3}|} = 43.92$$

$$G_{L:\overline{2}|} = 42$$

$$G_{L:\overline{1}|} = 40$$

$$\sum A = 125.92$$

Now, I derive the duration and convexity of the payments, given that losses are paid at the end of the period and premiums are paid at the beginning of the period:

$$D_L = 2.03$$

$$D_A = 0.97$$

$$C_L = 6.82$$

$$C_A = 2.57$$

The modified and Macaulay durations are the same because of the lack of inflation, discounting, or return on assets. These bonds do not have effective duration or convexity because the option for insureds to drop the insurance is worthless under “static” guaranteed

renewability. In a world where interest rates and medical trend do not change, the contract constructed with sequential rationality in mind is always preferred to other forms of insurance, as well as uninsurance, and the insured cannot improve by dropping the insurance. Nevertheless, insurer's side of the market is in some sense "attractive", because they are long a low duration contract and short a high duration contract.

Now, I compute the estimates with a 2% real interest rate. First, I recompute the present value of losses and premiums at time 0, assuming that premiums are payed at the beginning of the period and losses are paid at the end of the period:

$$\mathbb{E}[L_1] = 39.22$$

$$\mathbb{E}[L_2] = 40.37$$

$$\mathbb{E}[L_3] = 41.39$$

$$\sum L = 120.97$$

$$A_{L:\overline{3}|} = 42.95$$

$$A_{L:\overline{2}|} = 40.33$$

$$A_{L:\overline{1}|} = 37.69$$

$$\sum P = 120.97$$

While the nominal losses are the same, the present value of loss is lower, leading to a nominal premium stream of 42.95, 41.14 and 39.22. This contract sequence finances a positive reserve at interim periods and a 0 terminal reserve. In addition, the modified

durations are given by:

$$D_L = 1.98$$

$$D_P = 0.94$$

$$C_L = 6.49$$

$$C_A = 2.44$$

Discounting has the effect of compressing the difference between the two durations and the two convexities, though not by a lot.

2.3 Matching duration with an investment policy

There is an investment policy involving zero coupon bonds that will allow matching of assets and liabilities ¹. The 42.95 premium at the beginning of period 1 (time 0) is split so that 1.56 is invested in a two year zero coupon bond (security A) and the remainder (41.39) in three year zero coupon bonds (security B), allowing for the payment of third period losses (at time 3) and partial payment of second period losses (at time 2). The 41.14 premium at the beginning of period 2 (at time 1) is split between the payment of first period losses (at time 1) with the remainder (1.14) invested in a one year zero coupon bond (security C) that allows for the partial payment of second period losses at the end of the year (at time 2). The duration of total inflows is now 1.98, which matches the duration of outflows. The convexities are also matched with this asset policy. As long as the claims amounts and dates are given with relative certainty, then the risk arising from shifting interest rates can

¹There are an infinite number since the system is overidentified.

be managed.

3 Mathematical duration of this contract

Here I derive the mathematical duration for the claims and premiums under the contract.

The mismatch between the duration of claims and payouts under most insurance contracts is aggravated by the prepayment mechanism underlying the contract.

3.1 Duration of loss payments

The duration of loss payments for an N period GR contract without discounting is given by:

$$\frac{N + 1}{2} + \frac{\sum_{j=1}^{N-1} \frac{j}{2}(N - j)(p_H - p_L)(1 - p_L)^{j-1}}{N + \sum_{j=1}^{N-1} (N - j)(p_H - p_L)(1 - p_L)^{j-1}} \quad (4.1)$$

The duration with equal chance of loss in all periods is $\frac{N+1}{2}$. Since there is a greater chance of loss in later periods, the duration is increased by the second factor in equation 4.1.

The second factor in the duration of losses is also important because it is bounded. As $N \rightarrow \infty$, the fraction converges to a non-zero number. Therefore, as N grows large, the first part of the duration calculation dominates. This is in line with the result from chapter 3 that shows that the GR contract eventually converges to spot high type insurance. As the length of the contract increases, the timing of losses are eventually known with certainty. Any issues with the mismatched duration of the loss payments arise in intermediate periods.

The convexity of loss payments for an N period GR contract without discounting is

given by:

$$\frac{(N+1)(N+2)}{3} + \frac{\sum_{j=1}^{N-1} \frac{j}{3}(N-j)(N+j+3)(p_H - p_L)(1-p_L)^{j-1}}{N + \sum_{j=1}^{N-1} (N-j)(p_H - p_L)(1-p_L)^{j-1}} \quad (4.2)$$

3.2 Duration of premium payments

The duration of premium payments for an N period GR contract without discounting is

given by:

$$\frac{N-1}{2} - \frac{\sum_{j=1}^{N-1} \frac{j}{2}(N-j)(p_H - p_L)(1-p_L)^{j-1}}{N + \sum_{j=1}^{N-1} (N-j)(p_H - p_L)(1-p_L)^{j-1}} \quad (4.3)$$

The duration with equal premiums in all periods is $\frac{N-1}{2}$. Since there is a lower premium in later periods, the duration is decreased by the second factor in equation 4.1. The convexity of premium payments for an N period GR contract without discounting is given by:

$$\frac{(N-1)(N+1)}{3} - \frac{\sum_{j=1}^{N-1} \frac{j}{3}(N-j)(2N-j)(p_H - p_L)(1-p_L)^{j-1}}{N + \sum_{j=1}^{N-1} (N-j)(p_H - p_L)(1-p_L)^{j-1}} \quad (4.4)$$

The duration and convexity of the two flows differ, showing the need to manage the financing of the insurance product. Eventually, the duration and convexity of the streams of premiums and claims converge. The reason is that, for large N in the model without death, the guaranteed renewable insurance premium increases until it reaches the spot insurance of high types.

3.3 Asset liability management for a static two period model

In order to manage the liabilities, I construct an asset policy. The asset policy ensures that the duration (and higher order sensitivities) of the contract are equal. I focus first on the

discount rate case where β represents the risk free rate, under the assumption that an asset exists that generates a risk free annual return of β^{-1} . Then I solve the following system of equations:

$$\begin{aligned}
 Z_1^1 + Z_1^2 &= P_1 \\
 Z_2^0 + Z_2^1 &= P_2 \\
 Z_1^1\beta^{-1} + Z_2^0 &= L_1 \\
 Z_1^2\beta^{-2} + Z_2^1\beta^{-1} &= L_2
 \end{aligned} \tag{4.5}$$

I solve the weighting problem for splitting the first premium up among two zero coupon bonds (Z_0 denotes the situation where a portion of the premium payment is put toward payment of claims immediately). One pays off in one year and one pays off in two years. Since these are zeros, they have a duration equal to the time to payoff. They will sink in value with a rise in interest rates, just as the company is experiencing gains on the contract from the premiums being “too large” otherwise. A correct version of these weights matches the duration. There are three unknowns and four equations. This is because there is only one allocation of the second premium (it must all be invested in 1 period zeros), so this is a collinear equation. Solving the other three entails:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ \beta^{-1} & 0 & 1 & 0 \\ 0 & \beta^{-2} & 0 & \beta^{-1} \end{pmatrix} \begin{pmatrix} Z_1^1 \\ Z_1^2 \\ Z_2^0 \\ Z_2^1 \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \\ L_1 \\ L_2 \end{pmatrix}$$

This leads to the solution:

$$\begin{pmatrix} Z_1^1 \\ Z_1^2 \\ Z_2^0 \\ Z_2^1 \end{pmatrix} = \begin{pmatrix} \beta p_L L \\ \beta^2 p_L L (p_H - p_L) \\ 0 \\ \beta^2 p_L L \end{pmatrix}$$

There are, of course, an infinite number of solutions. However, I now recompute the new duration of inflows:

$$\begin{aligned} D_I &= \frac{\beta p_L L + 2\beta^2 p_L L (p_H - p_L) + 2\beta^2 p_L L}{\mathbb{L}} \\ &= D_L \end{aligned}$$

3.4 Trend in a two and a three period model

In a model with deterministic trend τ , the two period expected losses and contract premiums are given by the following equation:

$$\begin{aligned} \mathbb{E}[L_1] &= [p_L L (1 + \tau)] \\ \mathbb{E}[L_2] &= [p_L L (1 + \tau)^2] [1 + (p_H - p_L)] \\ P_1 &= [p_L L (1 + \tau)] [1 + (1 + \tau)(p_H - p_L)] \\ P_2 &= [p_L L (1 + \tau)^2] \end{aligned} \tag{4.6}$$

I can show that this contract is sustainable by backward induction.

At period 2, the contract represents the expected loss of the low types. Therefore, high types will certainly accept this contract. Low types are strictly indifferent between this

contract and a spot market contract, so I assume that they purchase this contract. Therefore, this contract is within the core of all insureds. This contract generates a loss to the insurer, in the second period, of:

$$\mathbb{E}[L_2] - P_2 = [p_L L(1 + \tau)^2][(p_H - p_L)]$$

In order to generate a zero profit contract on both sides, insureds must therefore pay expected losses in the first period plus the amount of the expected second period shortfall. In other words, $P_1 = \mathbb{E}[L_1] + \mathbb{E}[L_2] - P_2$. On a risk neutral basis, then, individuals at time 1, who are all low types, are strictly indifferent between taking the insurance and not taking the insurance, since the expected gain over the life of the contract is 0. However, individuals who are risk averse will prefer this contract to no insurance, and single period insurance, since it gives them a smoother pattern of consumption in expectation as I show in chapter 3. The portion of the second period premium that is front loaded into the contract is increased by the loading factor $1 + \tau$.

The front loaded nature of the contract stays the same under a deterministic trend assumption. The front loading still derives from the uncertainty over whether individuals will be in a high or a low loss distribution in the future. However, a trend, even if it is deterministic, applies to all levels of insurance. Therefore, it will affect all future contracts equally, and must be included in the front loading.

3.5 Trend in a three period model

By similar logic, the sequence of payments for the three period model is:

$$\mathbb{E}[L_1] = [p_L L(1 + \tau)] \quad (4.7)$$

$$\mathbb{E}[L_2] = [p_L L(1 + \tau)^2][1 + (p_H - p_L)] \quad (4.8)$$

$$\mathbb{E}[L_3] = [p_L L(1 + \tau)^3][1 + (p_H - p_L) + (1 - p_L)(p_H - p_L)] \quad (4.9)$$

$$P_1 = [p_L L(1 + \tau)][1 + (1 + \tau)(p_H - p_L) + (1 + \tau^2)(1 - p_L)(p_H - p_L)] \quad (4.10)$$

$$P_2 = [p_L L(1 + \tau)^2][1 + (1 + \tau)(p_H - p_L)] \quad (4.11)$$

$$P_3 = [p_L L(1 + \tau)^3] \quad (4.12)$$

In the third period, the insurance is actuarially fair for low types and actuarially favorable for high types. It generates a loss for the insurer of:

$$\mathbb{E}[L_3] - P_3 = [p_L L(1 + \tau)^3][(p_H - p_L) + (1 - p_L)(p_H - p_L)]$$

In the second period, the contract generates a loss for the insured (a gain for the insurer) in expectation. It is unfair for low types and favorable for high types when compared to the spot market price.

$$\mathbb{E}[L_2] - P_2 = -[p_L L(1 + \tau)^2][\tau(p_H - p_L)]$$

From the point of view of period 2, however, this contract is still actuarially favorable for high types on the spot market, because:

$$P_2 + P_3 < [p_H L(1 + \tau)^2] + [p_H L(1 + \tau)^3]$$

For low types from the point of view of period 2, this contract is in their core. The contract is actuarially unfair on a spot (one period) basis. Therefore, I show it is a fair contract on a two period basis:

$$\begin{aligned}
P_2 + P_3 &= [p_L L(1 + \tau)^2][1 + (1 + \tau)(p_H - p_L)] + [p_L L(1 + \tau)^3] \\
P_2 + P_3 &= [p_L L(1 + \tau)^2] + [p_L L(1 + \tau)^3][(p_H - p_L)] + [p_L L(1 + \tau)^2] \\
P_2 + P_3 &= [p_L L(1 + \tau)^2] + [p_L L(1 + \tau)^3][1 + (p_H - p_L)] \\
P_2 + P_3 &= [p_L L(1 + \tau)^2] + (1 + \tau)^3[p_L L + p_L L(p_H - p_L)] \\
P_2 + P_3 &= [p_L L(1 + \tau)^2] + (1 + \tau)^3[p_L L + p_L L p_H - p_L^2 L] \\
&= [p_L L(1 + \tau)^2] + (1 + \tau)^3[(1 - p_L)p_L L + p_L p_H L] \\
&= \{\mathbb{E}[L_2]|\text{Period 2 type is L}\} + \{\mathbb{E}[L_3]|\text{Period 2 type is L}\}
\end{aligned}$$

Therefore, from the point of view of period 2, the contract is actuarially fair for low types.

In the first period, a similar logic holds. The contract generates a single period loss for the insured of

$$\mathbb{E}[L_1] - P_1 = -[p_L L(1 + \tau)][(1 + \tau)(p_H - p_L) + (1 + \tau^2)(1 - p_L)(p_H - p_L)]$$

For high types, this contract is actuarially favorable.

Finally, I can show that the sum of all losses in expectation is 0:

$$\begin{aligned}
& (\mathbb{E}[L_1] + \mathbb{E}[L_2]) - (P_1 + P_2) = \\
& -[p_L L(1 + \tau)][(1 + \tau)(p_H - p_L) + (1 + \tau)^2(1 - p_L)(p_H - p_L)] \\
& \quad - [p_L L(1 + \tau)^2][\tau(p_H - p_L)] = \\
& -[p_L L(1 + \tau)^2(p_H - p_L)][1 + (1 + \tau)(1 - p_L) + \tau] = \\
& \quad -[p_L L(1 + \tau)^3(p_H - p_L)][1 + (1 - p_L)] = -(\mathbb{E}[L_3] - P_3)
\end{aligned}$$

This shows that the contract is actuarially fair in expectation over the life of the contract.

Low types at period 1 will also choose to take up this contract, as long as they are risk averse.

4 Trend in general guaranteed renewable insurance

4.1 Notation for trend

I first define a spending level and a spending growth rate. The spending level is an index of spending for a relevant population. I denote the spending level as R_t^M at time t . I denote the level of change, or spending rate, as τ_t^M . Equation 4.13 relates the level of spending and the change in spending:

$$\tau_t^M = \left(\frac{R_t^M}{R_{t-1}^M} \right) - 1 \tag{4.13}$$

The spending level is itself a composite of a number of factors: technology, preferences, and the age effect. As with inflation, technology is a determinant of the level of spending

through its effect on quality as well as the set of goods and services that is available. In addition, preferences determine the price level, since they influence what goods and services are in the average basket represented in the price level. Finally, medical prices have a special component known as the age effect. This is a special case of preferences, where the aging of the population induces a greater preference for medical care *ceteris paribus* (Lindh 2004, e.g.).

A special case I pay particular attention to is where the trend is not only deterministic, but constant throughout time. I denote trend as τ . This special case is useful because it allows me to expand the original guaranteed renewable model in a minor way. With this change, there are some facets of the model that change dramatically, while others do not change at all.

4.2 The effect of deterministic trend

When losses are subject to a trend rate τ , they are inflated by this amount annually. Assuming that there is no variance, it is possible to retain the guaranteed renewable features of the contract. In order to preserve sequentially rational contracts, insureds must prepay the expected future costs of high types at the future (inflated) loss level. This is true despite the fact that trend is perfectly forecastable (as shown in general in section 6.1).

The formula for the guaranteed renewable premium in an N period model at time k

with trend rate τ :

$$\begin{aligned}
 P_N(k) = & [p_L L(1 + \tau)] \times \\
 & [1 + (1 + \tau)(p_H - p_L) + (1 + \tau)^2(1 - p_L)(p_H - p_L) + \\
 & (1 + \tau)^3(1 - p_L)^2(p_H - p_L) + \dots + \\
 & (1 + \tau)^{N-k-1}(1 - p_L)^{N-k-2}(p_H - p_L)] \tag{4.14}
 \end{aligned}$$

The general principal is that consumers are paying for part of tomorrow's medical trend today. This is true because the guaranteed renewable insurance is a package of state contingent claims sold today, where some of those distant future claims cost more, on an inflation and discounting adjusted basis, than the near future claims. The state contingent claims the insurer is selling to the insured is conditioned on the state of the world for the insured (payoffs only occur when he incurs a loss), and for the state of the world for losses (there are different payoffs under the different possibilities for what the loss could be). Insurers may only want to hold part of the risk—the risk that the person incurs a loss—but not hold the risk related to the size of the loss. Selling the risk of the stochastic size of the loss is the hedging that insurers may want to engage in, as I discuss further in section 5.

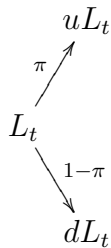
If trend is stochastic rather than deterministic, the premium is not necessarily changed. Under a rule of actuarially fair premiums, the variance in trend does not change the premiums at all. However, it does change the possible investment strategy of the firm. Without variance, the firm only needs to manage duration and convexity, which deals with asymmetric effects of changes in the discount rate (or the rate of return on assets more generally). With variance, the firm needs to consider whether asset returns and trend covary.

5 Single period investing

In this section, I solve the single period investment problem. This allows me to show an example for how to invest for a single period of insurance with a medical trend rate, as well as to develop a recursive investment policy for multiperiod problems.

5.1 One period example

Say that L evolves based on a binomial distribution. I now denote L with time based subscripts. The law of motion for trend rate τ is $L_{t+1} = \varepsilon L_t$. ε is an up or down motion, with $u > R > d, 0 < \pi < 1$. Therefore, τ evolves according to:



The two different states of nature imply different *ex post* realizations of medical inflation. On the up node, $\tau_t = u - 1$, and on the down node, $\tau_t = d - 1$. If the predicted overall change in medical costs is 0, the fair rate of return multiplier for this variable is $R = 1$.

If it were possible to invest in a contingent claim with a payout based on u, d , the no arbitrage condition would be

$$\pi = \frac{R - d}{u - d} \quad (4.15)$$

so $u\pi + d(1 - \pi) = 1$ (Huang and Litzenberger 1988, p.251). Therefore, $\mathbb{E}[L_1] = RL_0$, and this is just a mean preserving spread on L when $R = 1$. Then, the new present value

of losses and premiums, without a discount rate, is:

$$\mathbb{E}[L_1] = p_L LR$$

$$\mathbb{E}[L_1^2] = p_L L^2 (\pi u^2 + (1 - \pi) d^2)$$

$$\text{Var}[L_1] = p_L L^2 (\pi u^2 + (1 - \pi) d^2 - p_L R)$$

$$P_1 = p_L LR$$

$$PV(\mathbb{L}) = PV(\mathbb{P})$$

In general, the realization of trend is nonzero ². In a complete market, however, this should not be a problem. If individuals or firms can purchase contingent claims for all possible states of nature, or if there is securities market with a risky and a risk-free asset that match the evolution of medical inflation, then individuals or firms can purchase protection against the risk of stochastic trend. I show the result for the one period model where firms or individuals can use these securities to hedge their risk. I then proceed to that guaranteed renewable insurance is a bundle of these securities.

5.2 Investment problem setup

There is an initial wealth W_0 that comes from the premium paid and an initial liability $\mathbb{E}[L_1]$. Under fair insurance, $W_0 = P = \mathbb{E}[L_1]$. I solve the investment problem in the presence of three types of investments: a risk-free bond, state contingent claims, and a combination of a risk-free bond and risky stock.

²Except in the special cases where $u = 0$ or $d = 0$.

5.3 State contingent claims

Under state contingent claims, it is possible to totally immunize the firm against losses.

There is also a final realized payoff

$$L_1 = \begin{cases} uL_0 & P = \pi \\ dL_0 & P = 1 - \pi \end{cases} \quad (4.16)$$

Finally, the wealth is divided such that w_u is invested in the contingent claim that pays off in the up state and w_d is invested in the contingent claim that pays off in the down state.

$W_0 = w_u + w_d$. Denote $\delta = \frac{w_u}{W_0}$ as the fraction invested in the up asset. The final surplus is therefore $\frac{w_u}{\pi} - uL_0$ in the up state and $\frac{w_d}{1-\pi} - dL_0$ in the down state. The objective is to minimize the extent to which the surplus differs from 1, so I use a quadratic penalty function i.e.

$$\min_{\{w_u, w_d\}} \sum_{i=u,d} P_i \left(\frac{w_i}{P_i} - iL_0 \right)^2 \quad (4.17)$$

$$\text{s.t. } w_u + w_d = W_0 \quad (4.18)$$

I solve for w_u by substituting the constraint into the objective function. The solution is ³:

$$w_u = \pi W_0 + (1 - \pi)(R - d)L_0 \quad (4.19)$$

$$w_d = (1 - \pi)W_0 - (1 - \pi)(R - d)L_0$$

When $(1 - \pi)W_0 < (1 - \pi)(R - d)L_0$, this calls for short selling the down claim. I do not constrain investments to be positive i.e. I allow for short selling. This allows the surplus in both possible states of the world to be equal, which minimizes the sum of squares.

³For the full calculation, see Appendix A

Note that if the contract is fair, i.e. $W_0 = RL_0$, $w_u = (R - d(1 - \pi))L_0$, and this leads to 0 losses in all states (see Appendix A for a full calculation).

5.4 Stock and bond investment

In the case with a risk-free bond and a risky stock, it is also possible to achieve zero losses in all states. Here, I assume that the bond returns R and that the stock price evolves according to the same binomial distribution as the losses. Specifically,

$$S(1) = \begin{cases} uS(0) & P = \pi \\ dS(0) & P = 1 - \pi \end{cases} \quad (4.20)$$

The objective is again to minimize the extent to which the surplus differs from 1, so I use a quadratic penalty function. The initial wealth W_0 is divided between the amount invested in stock w_S and the amount invested in the bond w_B

$$\min_{\{w_S, w_B\}} \sum_{i=u,d} P_i (i w_S + R w_B - i L_0)^2 \quad (4.21)$$

$$\text{s.t. } w_S + w_B = W_0 \quad (4.22)$$

The optimal investment policy is to set the stock investment equal to the current loss level, i.e. $w_S = L_0$ (see the result in Appendix A). This is true even if the premium is unfair (in which case the “overpayments” are invested in safe bonds) or favorable (in which case the insurer borrows at the risk-free rate to invest in risky stock).

5.5 Risk free bond

Assume that the only asset in the economy is a risk free bond. The existence of only one, risk free, asset is a special case of market incompleteness. In this case, the entirety of the collected premium is invested in the bond. If the return to the bond is R , then the expected loss under the contract is

$$\pi W_0 R - u p_L L + (1 - \pi) W_0 R - d p_L L = R(W_0 - p_L L)$$

In other words, the spread between premiums and time 0 expected losses is increased by the factor R . Under fair insurance, this is still 0 at time 1, and be non-zero losses under both states of nature that obtain.

6 Multiperiod investment policy

In this section, I show how the single period solution to the investment problem can be applied to multiperiod guaranteed renewable contracts. I show how this allows insurance companies to dynamically adjust the premium in order to preserve the sequential rationality properties of the insurance arrangement. Under the correct investment scheme, with dynamically complete markets, insureds' premium will be based on the guaranteed renewable rates for the low types in all periods. If the premium for all insureds go up (or down), the premium must adjust accordingly to preserve the sequential rationality of the contract. This makes the investment policy path dependent.

6.1 General investment problem

The multiperiod problem is a more general form of the single period problem. At time t , after the payment of losses and collection of premiums, there is a surplus (or deficit) W_t that must be spread among various assets. This is for T period guaranteed renewability. Since there are only two outcomes, up and down, I need only two assets. I use only terminal assets to allow for market incompleteness later (in that non-terminal assets may not exist).

I repeat the single period results for trading state contingent claims. The two assets I use are the contingent claim for $T - t$ subsequent up and down moves. These assets cost π^{T-t} and $(1 - \pi)^{T-t}$ and have single period returns of $\frac{1}{\pi} - 1$ and $\frac{1}{1-\pi} - 1$ respectively. Finally, I denote the current expectation of future losses (liabilities) and premiums (assets) by $\mathbb{L}_t, \mathbb{A}_t$, respectively. The problem is now:

$$\min_{\{w_u, w_d\}} \sum_{i=u,d} P_i \left(\frac{w_i}{P_i} - i(\mathbb{L}_t - \mathbb{A}_t) \right)^2 \quad (4.23)$$

$$\text{s.t. } w_u + w_d = W_t \quad (4.24)$$

The solution is much the same as before:

$$w_u = \pi W_t + (1 - \pi)(R - d)(\mathbb{L}_t - \mathbb{A}_t) \quad (4.25)$$

$$w_d = (1 - \pi)W_t - (1 - \pi)(R - d)(\mathbb{L}_t - \mathbb{A}_t)$$

This is in contrast to the variance minimizing investment policy $w_u = \pi W_t, w_d = (1 - \pi)W_t$ which yields an expected return of 0 and a variance in the return of 0, since $W_{t+1} = W_t$. However, in the GR contract, future losses will always exceed future premiums since premiums are paid at the beginning of the period while losses are paid at the end of the

period. Further, the gap between losses and premiums will widen over time. Therefore, the relative portion of the investment in up securities will increase over time, increasing the variance of the investment strategy return as the terminal period approaches. Meanwhile, some losses are paid out of the reserves.

For a risky stock and a risk-free bond, the results are again similar to the one period case. I solve:

$$\min_{\{w_S, w_B\}} \sum_{i=u,d} P_i (i(w_S - \mathbb{L}_t) + (Rw_B))^2 \quad (4.26)$$

$$\text{s.t. } w_S + w_B = W_t \quad (4.27)$$

The solution is much the same as before:

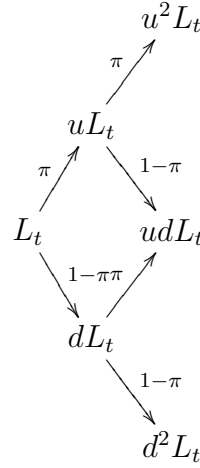
$$w_S = \mathbb{L}_t \quad (4.28)$$

$$w_B = W_t - w_S$$

As time goes on, the amount of future payments to be incurred decreases, and so the amount in the risky stock will decrease. As a proportion of total assets, however, it may stay the same, as in the case of actuarially fair premiums, where the insurance company is always 100% invested in “risky” stock.

6.2 Two period model

Now, I extend to model to two periods.



Then, the new present value of losses and premiums in the second period is:

$$\mathbb{E}[L_2] = p_L L(1 + (p_H - p_L))$$

$$\mathbb{E}[L_2^2] = p_L L^2(1 + (p_H - p_L))^2(\pi^2 u^4 + 2\pi(1 - \pi) + (1 - \pi)^2 d^4)$$

$$\text{Var}[L_2] = p_L L^2(1 + (p_H - p_L))^2(\pi^2 u^4 + 2\pi(1 - \pi) + (1 - \pi)^2 d^4 - p_L)$$

$$P_1 = p_L L(1 + (p_H - p_L))$$

$$P_2 = p_L L$$

$$PV(\mathbb{L}) = PV(\mathbb{P}) = p_L L(1 + (1 + (p_H - p_L)))$$

From the results for the single period model, I know how to invest the time 1 surplus. Whatever remains from time 0 premiums after paying off time 1 claims is invested in the manner described in the single period model. This is because the single period solution is valid for all surplus (or deficit) sizes, so I do not have to tailor the policy to the expectation of the surplus. In the case of GR premium, there is a positive surplus to invest.

If the market is complete, then the solution to the one period problem is the same as the

one for the two period problem. The amount invested at the two times will differ, because the first period premium is above the population average spot rate and the second period premium is below the population average spot rate.

Also note that the expectation of second period premiums will not be realized (unless $u = d^{-1}$) ex post. This is because charging $p_L L$ in the down node will give low types the incentive to deviate, given the spot market rate of $p_L u d L_t \pi + p_L d^2 L_t (1 - \pi) = d p_L L$. This is the maximum incentive compatible rate. This leads to the question of what the insurer charges in the up node. If this were a policy that combined “type change” insurance with trend insurance, they would pay $p_L L$. However, guaranteed renewability is not insurance against medical trend but simple against becoming the high loss type. Since the one period contracts premium is $p_L u L$, and the contract is valid for two periods, this is what all GR insureds pay. All insureds must still face the risk of the evolution in losses in the economy (or insure against this additional loss).

6.3 Investment policy with changing interest rates

Now I consider what the correct investment policy would be under changing interest rates. Say the interest rate changes from i to i' . If the insureds remained in the contract, then there would be no problem, except that the company would have to decide what to do with the extra percentage point of return on security C. However, the insureds will not remain in the contract with the static terms. Consider, in particular, the new two year guaranteed

renewable premiums for low type individuals:

$$\mathbb{E}[L_1] = 38.83$$

$$\mathbb{E}[L_2] = 39.59$$

$$\sum L = 78.42$$

$$G_{L:\overline{2}} = 40.72$$

$$G_{L:\overline{1}} = 37.70$$

$$\sum P = 78.42$$

The nominal premium stream for two year GR insurance for a low type individual is 40.72, 38.83. So the insurance company must offer this premium or risk losing all the low type individuals. In addition, if it offers this contract it is sure to make money. This is because the investment policy assumes higher cash flows than the company will ultimately receive. I demonstrate by following the prior investment policy and assuming interest rates do not change again.

Now let us investigate what happens when interest rates change. If interest rates were 3%, the three year premiums would have been 42.48, 40.72, and 38.83. The investments would have been 40.19 in a three year zero, 2.28 in a two year zero in the first period, and then 0.72 in a one year zero in the second period. If an insurer were pricing a two year GR contract at 3%, the contract would consist of premiums of 40.72, and 38.83, with 1.14 invested in a one year zero and 39.58 invested in a two year zero. This finances payments of 40, 42 at the end of periods 1 and 2. Therefore, low types have the incentive to deviate, but not high types, who face spot rates of 87.38 and 84.83 for single period insurance for

one and two periods hence, respectively.

The insurer will only be able to match these rates if the zeros are floating rate notes that reset each period. The investment strategy changes. The insurer now has only 0.72 to invest in a one year zero, and this investment will mature in one year. At the end of the second period, the insurer has $1.56 * 1.02 * 1.03 + 0.72 * 1.03 = 2.38$ in investment gains and an additional 38.83 in premiums. This leaves a shortfall of 0.79 that must be borrowed. At the end of the third period, the insurer has $41.39 * 1.02 * 1.03^2 = 44.79$ in investment gains, 43.92 in claim obligations, and $0.79 * 1.03 = 0.81$ in loan repayments. The net terminal reserve is 0.06.

This is not just a rounding error. When the interest rate changes to 5%, the terminal reserve rises to 0.18. However, a 1% interest rate leaves a terminal reserve of -0.05. The insurance company is short interest rates. The problem is that the numerical solution does not include the implicit option in the contract whereby the insureds have the option of dropping the contract. This option becomes worthwhile when interest rates fall, and hence the insurer must reprice the contract to prevent the option from coming into the money. However, this is the actuarially correct policy for reserving the contract.

The other option for the company is to allow the low types to lapse and charge the originally quoted rates to remaining high types. This will not cost the insurance company anything because the remaining low types have reserves of 0 associated with them (as shown in Proposition 3. However, the insurance company is still left with the problem of repricing. First, there is some gain to being able to invest 2% insurance premiums at

3% interest, so the gain has to be allocated to someone. Second, and more important, the insurance company is no longer offering the correct GR premium, but a higher premium due to changes outside the realm of the level and probability of loss. Class average rates also offer no guidance because they imply an even worse outcome: premiums approaching the new class average of $p_H L$. The question of the correct premium is not trivial even taking in to account the lack of strategic behavior by the insurer.

7 Conclusion

The investment policy problem is not trivial for guaranteed renewable plans. The frontloading of future costs into current premiums means that any variables that affect future loss amounts must be factored into the current premium. If they are not, the reserve is either too low, leading to strategic behavior by low types (dropping the policy and re-enrolling with a new GR plan), or the premium is too high, and then no low types will join the risk pool. If the variables affecting the size of future losses are stochastic, as is commonly the case with medical trend and interest rates, then the insurer will face investment gains and losses due to the macroeconomic variables (i.e. not due to the experience of the risk pool). A complete securities market would allow insurers to avoid taking the macroeconomic risks, but the completeness of securities markets for the risk of trend is an empirical question. In the next two chapters, I will address the question of whether assets currently available will allow insurers to hedge the risk of variable medical spending growth.

Chapter 5

Analysis of aggregate insurance premiums

I evaluate the usefulness of different securities for hedging aggregate medical spending growth. Mean growth rates in medical spending are high and variable. I use an adaptive expectations model to forecast medical spending growth, and calculate a series of unforecast growth (forecast errors). I show that returns to several asset classes that are natural candidates for hedging medical spending growth are uncorrelated with forecast errors. The implications are that I need to search for other, better hedges, or find better measures of medical spending growth.

1 Motivation

1.1 Predictability of spending growth

The predictability of spending growth is a key driver of its effect on guaranteed renewable health insurance. The guaranteed renewable premium is based on the expected future costs of insurance which does include trend (see section 4). When the trend is stochastic, the insurer still imputes the average expected trend into the policy. The insurer also factors the size of potential fluctuations into the contract reserve. The fair price for the insurance would include a risk premium for the additional risk the insurer is taking on.

The predictability affects the size of the contract reserve that the insurer has to hold against spending growth fluctuations. In a setting with stochastic trend, the insurer also has to be concerned about the credibility of the data used to predict future trends. The longer the time horizon for the guaranteed renewable contract, the more the insurer needs to know about the longer term stochastic properties of medical spending growth.

The specific time series properties of spending growth determine how to manage it within a guaranteed renewable insurance contract. If each year's growth is a draw from a distribution with noise, then gains from lower than expected years should mostly cover losses from higher than expected years. The point of reserves would be to cover shortfalls, especially if higher than expected spending occurred in the early years of the contract.

If spending growth is serially correlated, then it is more likely that several years of higher than expected growth could cause accumulated losses. In that case, the insurer would be interested in above average asset returns, and would be willing to sacrifice below

average asset returns in the situation where spending growth was below trend for several years.

The underlying data generating function for medical spending growth has almost certainly changed over time. The time since the last trend break is important for guaranteed renewability because, aside from any particular model of medical spending growth is how long any relationship will persist. The problem is determining the stationarity of the time series data, or whether there is a single consistent trend over time. Whether the data available can answer these questions is an important part of how long guaranteed renewable contracts could last.

1.2 Finding hedges

Prediction errors are not a problem if effective hedges exist. In the example of commodities, an individual that enters into a futures contract that guarantees a set price for the delivery of a commodity will not have an additional need to predict or reserve against changes in the commodity price. Similarly for medical spending, if there is an asset or group of assets that covary with the unpredicted portion of medical spending, then they would be a useful hedge. The amount of reserves a guaranteed renewable insurer would have to set aside would be decreased toward the amount needed to prefund predictable increases in medical spending. The higher the correlation between residual spending growth and asset returns, the better the hedge.

Separating trend from error is the first step in searching for a hedge for medical spend-

ing growth. This means fitting the observed series of insured spending to a model. I fit the spending growth time series to two models: a linear regression and an adaptive expectations model. The linear regression model is designed to assess the predictability of spending growth and the variables that might improve the fit of the regression model.

The adaptive expectations model is specifically designed to generate prediction errors and then find the correlation between errors and asset returns. The idea is that, for medical spending growth, there is a long term trend line and then errors around that trend line. The errors arise as a disturbance term that is random and uncorrelated across time. The size of errors around the trend line is not known, so deviations from the prior long term trend are factored into the long term trend based on an updating factor. The updating factor θ can range from 0 to 1, and is not known ahead of time. By estimating the equation and varying the parameter θ , I can test the correlation between assets and deviations from trend for a range of possible updating factors, giving me the correlations under different possible scenarios for how quickly the long term trend in spending growth is updated.

I search for hedges across a feasible set of assets. I am using broad equity and bond asset classes in order to find hedging assets that are generalizable. I also want to address the possibility that finding false positives for hedges becomes more likely as I search across more asset classes. Restricting myself to popular, broadly available asset classes, I have access to data with a long enough history to test against the spending data that I have. I also have access to specific return data on healthcare and healthcare subsectors, which allows me to test the proposition that healthcare assets are a good hedge for liabilities arising from

medical spending growth.

1.3 Determining the unhedged portion of risk

I am also investigating the extent of unhedged risk. Almost no hedge is perfect, so there is some gap between what prediction and hedging can do and the size of medical spending growth risk. There are several possible implications of unhedgeable risk in my results. One is that I did not include the right assets, either because I restrict the set of assets that I consider or because the assets do not exist. In either case, the size of the gap would determine how important it is to try new assets, examine policies that create new assets, or measuring the size of risk that cannot be managed with a hedging investment policy.

It is the role of reserves to manage unhedged risk. The size of the forecast errors informs the size of the contract reserves under guaranteed renewability. It might also inform the investment policy of the insurer. Medical spending growth is a shock that is common across all policy holders, so if the errors are large or persistent, then the insurer might want to utilize a more conservative investment policy.

There is also a point where the risk becomes uninsurable to an insurer, and possibly to any entity. For guaranteed renewability, this might include contracts where the class average guaranteed extends for a large number of years. For public insurance programs, the guarantee would be more in terms of how long a government could credibly commit to keep a health insurance program unchanged. The issue of insurability is tied to the availability of data, as well as the applicability of past data to predicting future spending

growth.

2 Aggregate data sources

2.1 National Health Expenditure Survey

The National Health Expenditure Survey tabulates data on total and per capita medical spending. The spending data is used not only by the Office of the Actuary at the Centers for Medicare & Medicaid Services (Sisko, Truffer, Smith, Keehan, Cylus, Poisal, Clemens and Lizonitz 2009), and private actuaries forecasting retiree medical plans (Getzen 2007). The data is available from 1960 to 2008. On a nominal basis, medical spending has grown rapidly since 1960, outpacing nominal GDP growth by almost 2.5% (see table 5.1).

CMS also surveys health plans to tabulate medical spending per insurance plan enrollee. The data includes per capital spending by private plans and Medicare, and further splits the data into all benefits and common benefits provided by both plans (for example, Medicare did not offer drug benefits until 2003 (Centers for Medicare & Medicaid Services 2011)). The data for par capita insurer expenditures is similar, growing at rates of 7-8% over the more recent 1982-2008 period (see table 5.2).

While private insurance growth has outpaced Medicare, both growth rates are substantial with a high degree of variance. The growth in spending is unskewed, and the kurtosis is close to normal. Also, there is not a single year of negative spending growth in the data. The advantage of the data over the general medical spending data is that it isolates

the spending on insured lives. It changes with general changes in the type and quantity of medical care delivered, which is part of the risk I want to measure. The disadvantage is that it is still aggregated across many types of private health insurance plans, which is above the ideal level of aggregation for my research question.

Both total spending and insurance premium growth are strongly serially correlated. The growth rate in total medical spending is correlated from year to year in a way that GDP growth is not (see figure 5.1). Premium growth per enrollee in private plans is also serially correlated. Using prior year trend alone explains 51% of the variation in the next year's spending, despite the fact that data runs only from 1970–2008, as shown in figure 5.2. The predictability in annual data suggests that it may be possible to forecast future medical spending growth, although it is also possible that the underlying properties of the time series of insurance growth rates has changed more than once.

The time series properties of medical spending growth suggest that there is not one consistent trend rate. For example, a unit root test for the period 1971–2008 shows that there is almost certainly a unit root in the total spending data, while there may or may not be a unit root in the change in spending time series (see tables 5.3 and 5.4). A unit root test for the period 1982–2008 shows that there is almost certainly a unit root in the total spending data, and while I can reject the possibility of a unit root in the Medicare change series, I still cannot reject the possibility of a unit root in the private spending change series. These tests must be interpreted cautiously, since there are few data points.

2.2 Aggregate medical care statistics

One source of data that could plausibly improve the prediction of spending growth is the Bureau of Labor Statistics' data on prices and employment. The data is more frequent and released more quickly than spending data. For aggregate data on inflation at the consumer level, I utilize the BLS' CPI-U data on consumer prices. The advantage of the CPI is that it is both data that I can use to evaluate the effect of inflation and the basis for the inflation protected securities sold by the U.S. Treasury (Bureau of the Public Debt 2008a).

The data on the overall price level is available on a monthly basis since 1913. BLS has produced a price index for medical care on a monthly basis since 1947. I also use producer level price information from the BLS' PPI index. The PPI for all commodities is available from 1913, the index for intermediate goods and the drugs and pharmaceuticals component are available from 1947, and industry data for physician offices from 1994 and general hospitals are available from 1993. All data is monthly.

While the price data features higher frequency than spending data, there are limitations of the data for measuring the change in price of medical goods and services. The total CPI-U and CPI-U, Medical Care series are available on monthly and annual bases (Bureau of Labor Statistics 2011). The data is also released only two to three weeks after the end of the month, meaning that it is very timely. The main problems of the CPI for medical care involve its design as a consumer price index. Since consumers do not directly pay for the majority of medical care, medical care is underrepresented in the CPI relative to its share of the economy. Medical care also contains innovations, which are hard to value as

part of overall inflation. It also contains health insurance premiums, and it is difficult to determine how premiums should be imputed into the prices that consumers face (Ford and Ginsburg 2001).

Average CPI-U medical inflation is approximately 50% higher than average general inflation during the postwar period, and 80% higher for the the more recent period (1982–2008). While the standard deviation of medical inflation is lower than that of general inflation, the skew is much higher, showing that shocks in in medical inflation are generally above trend (see table 5.5).

The PPI data shows similar trends to CPI and spending data. The change in the level of the PPI for all commodities is lower than the CPI, but the change in the cost of drugs and pharmaceuticals is more than twice the level of all prices since 1982 and more than 50% higher since 1993. The trend is mirrored in the hospital sector to a lesser extent, where input prices are higher than the overall change in prices, while change in inputs at doctors' offices are lower than the level of all prices. (see tables 5.6 and 5.7).

In employment, the growth of doctor's office and home health employment greatly outpaced the growth of general employment since 1982 (see table 5.8). Employer types include health care and social assistance employers (since 1990), and subindustries such as physician offices (since 1972), and home health care services (since 1985). More recently, health employment growth has outstripped general employment growth, with more moderate growth in hospitals and faster growth in doctor's office and home health employment (see table 5.9). Earnings, and payroll show a similar trend, with the payroll index for

education and health increasing much faster than general payrolls.

2.3 Securities returns

I use securities returns both as predictors and to provide hedging for spending growth. I show initial statistics on risk-free rates and bond returns in table 5.10. For the risk-free rate, I use one month Treasury bills. These are a standard in the literature because they are U.S. government securities of short duration, which eliminates default risk and inflation risk. The data comes from the Fama-French factor for risk-free rates (Fama and French 1993, Fama and French 2010). The return on bonds that I use as investments to hedge growth comes from ten year government bonds and Moody's index of AAA rated corporate bonds. Both are total return indices, and so contain interest and principal payments. The data come from the Global Financial Data Total Return database (Global Financial Data 2010).

I use the Fama-French factors to generate returns for the total market, the health care industry as a whole, and health care subsectors (Fama and French 2010). The Fama-French returns are value weighted and cover stocks on the NYSE, AMEX, and NASDAQ exchanges. The negative skews show the fact that shocks in the distributions of stock returns are often negative. It also appears that health sector returns "outperform" the market in the expected way—they have higher mean returns but also higher variance (see table 5.11).

Hedging growth with asset returns requires a separating medical trend from prediction errors. To the extent that growth predictably increases, insurers prefund future medical losses with actuarially fair premiums. For the spending data, I use regression rather than

simple correlations. However, simple correlations of nominal asset returns and total medical inflation are instructive. In table 5.12, I show the correlations for the high frequency time series and securities returns over the 1982–2008 period. The correlations show that medical inflation is most correlated with the risk-free rate (short term Treasury bills), general inflation, and health care services companies.

3 Results

3.1 Regression analysis

I assess the year-on-year predictability of per capita insurer spending growth by regressing current spending growth on lagged spending growth. My regression equation with only lagged spending growth as an explanatory variable is:

$$\text{Medical Spending}_t = \beta_0 + \beta_1 * \text{Medical Spending}_{t-1} \quad (5.1)$$

The medical spending growth that I am explaining with the regression is nominal medical spending. The reason is that in my model, insurance companies write contracts indexed to the total level of spending, not nominal spending. Insurance companies take trend as a given and pass it on to insureds to the extent allowed through prospective pricing.

Over the entire period my data covers, 1971–2008, one year lagged spending growth explains roughly half of the variation in current year’s spending growth. Over the more recent period 1982, prior growth is less predictive of current growth, with an adjusted R^2 of 39% (see tables 5.13 and 5.14). The lower predictability is not a function of the

variance of spending growth, which has remained constant relative to the average spending growth (which has fallen fairly constantly over the last 40 years). However, for the more recent period adding spending growth from two years ago improves the prediction of recent spending growth (adjusted R^2 of 43%) while slightly reducing the adjusted R^2 for the entire time horizon.

Dropping lagged spending growth leads to much less predictability. Using medical inflation and physician office employment leads to an adjusted R^2 of 47%, and only 32% over the recent time horizon (see table 5.15 for the recent results). In addition, while the medical inflation coefficient is significant, the physician office employment coefficient is not. While medical inflation is always significant for the full time horizon, and the regression without lagged spending for the recent time horizon, physician office spending is never significant. While removing the physician office regressor from the full time horizon model reduces the adjusted R^2 from 47% to 37%, it increases the adjusted R^2 in the recent time horizon from 32% to 35%.

I can improve the prediction of current spending by using lagged medical inflation and physician office employment. Using either variable in concert with two lags of medical spending produces an adjusted R^2 of 49% in the recent period. The coefficients on lagged employment growth and inflation are also have lower p-values. So the best model for predicting spending growth for the entire period is the model with two lags of spending growth and current medical inflation and physician office employment growth, and the best model in the more recent 1982–2008 period is the one with two lags of spending growth

and lagged physician office employment growth (see table 5.16).

Asset returns do not improve the prediction of current spending growth. Adding the risk free rate does not add to the predictive power of the regression. Adding the market return also does not improve the results, Adding health sector and subsector returns also do not improve the forecast, either because the number of data points is so small or because they are uncorrelated with spending growth. All of these results point to market assets being poor hedges for aggregate medical spending growth.

Adding one and two period lagged asset returns does not improve on the prediction of spending growth. The results holds across all of the asset classes I use for the recent period 1982-2008. There are two asset classes with significant results (p-values between 0.01 and 0.05) with lags: drugs and medical equipment. The results are insignificant over the longer 1973-2008 time horizon. Therefore, I believe that these significant results are artifacts of multiple hypothesis tests.

3.2 Adaptive expectations analysis

I use a simple adaptive expectations model to determine the long term expected rate of nominal medical spending growth. The long term rate of growth is equal to the prior long term rate of growth plus a linear “adjustment” for the prior difference between the expected

rate of growth and the experienced rate of growth. The updating equation is

$$\bar{D}(t) = \bar{D}(t - 1) + \theta[D(t - 1) - \bar{D}(t - 1)] \quad (5.2)$$

where

$\bar{D}(t)$ is the long term rate of growth

$D(t)$ is the experienced growth rate at time t

in the style of Bodie (1976). $d(t)$ are the unanticipated shocks. The updating factor θ for the adjustment ranges from 0 to 1. I calculate estimates of $\bar{D}(t)$, the expected rate of growth, over a range of values for θ . The choice of updating factor determines the time series of forecast errors in the model.

One test of the ability of assets to hedge medical spending growth is the effect of spending shocks on excess return. I use the following specification:

$$R_e(t) = \alpha_0 + \alpha_1 d(t) + \mu(t) \quad (5.3)$$

where (5.4)

$$R_e(t) = R(t) - \bar{D}(t)$$

in the spirit of Bodie (1976). $d(t)$ are the unanticipated shocks. If the coefficient on α_1 is significant, then the return index used to calculate $R(t)$ is a good hedge. The sign of the coefficient indicates whether the hedging position is long or short. I to use this method as a test of the ability and usefulness of stocks and other securities as a hedge on spending growth. I also use it as a way to test the relative effect of medical inflation and growth in the quantity of medical care on the long term rate of total spending growth.

The results for a grid of updating coefficients ranging from 0.10–1.00 is in tables 5.17 and 5.18, and I summarize the prediction errors in table 5.19. The expectation of long term medical spending growth has come down as the rate of growth in spending has decreased. For 2009, the predicted spending growth rates range from 3-7% depending on the updating parameter (θ). The reason is the strong moderation in spending increase rates from over 11% in 2002 to 4.5% in 2008. The smallest updating parameter ($\theta=0.1$) gives the smallest proportional variance. However, the mean error rates from the adaptive expectations model, even where θ is one, are negative on average with standard deviations half the size of the predicted spending growth rate for 2009. Both could be a function of the low number of observations of spending.

The expected values for January, 2009 medical inflation range from 0.1-0.2. The adaptive expectations model does not capture the above average inflation I observed in January of most years in the medical price series. The medical inflation series also has a mean forecast error of zero, with standard deviations approximately 50% of the mean expected inflation value. Based on forecast errors, the adaptive expectations model should capture medical inflation better than general inflation. Medical inflation from the CPI-U, and employment in doctors' offices from the CES (my measure of changes in quantity of care). Then I expressed the standard deviation of the forecast error as a percent of the average forecast rate of growth.

The results show that unanticipated shocks in medical spending do not feed through into contemporaneous nominal asset returns. I regressed excess returns to the overall market,

health care, and health care subindustries for the entire and recent periods. I used the full range of updating factors from 0.1 to 1.0, and almost none of the coefficients were significant. I searched over ten possible updating factors across eight asset classes (nine counting general inflation, which can be an asset class via TIPS bonds). I failed to find any significant results that are not also likely to be spurious correlations. I discuss the implications in the final section of this chapter.

I also used lagged asset returns which are generally insignificant with exceptions. Total and excess returns to the market lagged one year for small θ (0.1 or 0.2) are both correlated with errors, as are corporate bond returns both in total (all θ) and excess returns (θ 0.6 or less). Excess returns to short term government bonds ($\theta = 0.1, 0.9, 1.0$) and long government bonds ($\theta = 0.1, 0.2, 0.3$) are also correlated with the shocks I generated for the spending growth time series. All the correlations are also negative. This suggests that, as in Bodie (1976), above average medical spending growth may be bad for stock market returns. If true, the results would also indicate that stock returns are a leading indicator of the unpredictable portion of medical spending growth.

4 Implications

4.1 Insurance plan specific data

The results suggest that the search for any hedging assets might benefit from insurance plan specific data. A major limitation of the insured per capita spending data is that it

is available only at the aggregation of all insured costs. This might have the effect of dampening the variation to the extent that the correlations I am searching for do not appear. It would be better to have a consistent set of benefits across insured individuals, as well as be able to segment the population by demography to model the differences in spending across groups. Also, the data is only available annually, so I am using a relatively long time horizon for assets. A focus on a shorter time frame would deal with problems with the changing time series properties of medical spending and financial assets. I make use on data from a specific insured population in chapter 6.

4.2 Optimal reserves

My results also imply that reserve setting is more important than investment policy in the management of guaranteed renewable health insurance. I have already cited the problems of duration rating and lack of trust in health insurers. There may be some ways for insurers to signal that they are high quality insurers that hold high levels of reserves, although the signals that consumers get are often premiums and customer service. Also, if stock returns anticipate shocks to medical spending growth, it is possible that insurance companies could short the assets I used and then hold onto unanticipated positive returns as a bulwark against future spending growth. This is a fairly complex investment policy that would be difficult to monitor. As it is usually up to the state regulator to determine the funded position of insurance firms alongside the fairness of premiums. I address the issue of reserve setting more specifically in chapter 7.

4.3 The effect of health reform

The effect of the newly enacted health reform law will have two effects that will change the applicability of my results. First, the law may change the ability of insurers to create underwritten pools of lives for the purposes of guaranteed renewability. It will also affect the ability of insurers to price their premiums consistent with guaranteed renewable principles. I address the direct issues in chapter 8.

There is a second, indirect, potential effect of health reform. The law may change the time series properties of spending growth, returns on financial assets, and the relationship between these variables. As I have discussed, bending the cost curve was an explicit goal of health reform, so the effect of the PPACA on future spending growth is an intended consequence of the law. The effect of the law on healthcare asset returns is an anticipated consequence, but there could be an indirect effect on other financial asset returns. The effect would not be hard to imagine given that healthcare is such a large segment of the entire economy. Finally, the relationship between spending growth and asset returns is another unintended consequence of health reform. I have failed to find a correlation, but it is possible that the PPACA will cause financial asset returns to become more tied to the growth in medical spending. I consider the possibility a starting point for future research in this area.

5 Tables and figures

5.1 Summary statistics

Per Capita			
Statistic	Spending	GDP	Population
mean	8.19	5.87	1.03
sd	2.95	2.46	0.23
skewness	-0.04	0.40	0.59
kurtosis	2.03	2.40	3.68
min	1.91	1.56	0.46
max	13.72	11.34	1.60
p25	5.78	4.35	0.88
p75	10.50	7.66	1.14

Source: National Health Expenditure Survey, 1961–2008

Table 5.1: Growth in nominal medical spending, nominal GDP, and population

Statistics	All benefits		Common benefits	
	Medicare	Private	Medicare	Private
mean	7.08	7.91	6.08	7.33
sd	3.50	3.56	2.83	3.22
skewness	0.07	0.35	0.59	0.16
kurtosis	3.76	2.22	4.19	2.09
min	-1.50	1.87	0.08	2.18
max	15.21	15.33	14.10	13.69
p25	4.90	4.89	4.38	4.84
p75	8.96	10.39	7.37	9.62

Source: National Health Expenditure Survey, 1982–2008

Table 5.2: Per capita nominal insurer expenditures, rate of change (%)

Time series	Test statistic	p-value
Spending per enrollee		
Medicare		
All benefits	5.97	> 0.99
Common benefits	5.60	> 0.99
Private insurance		
All benefits	7.96	> 0.99
Common benefits	8.65	> 0.99
Change in spending		
Medicare		
All benefits	-3.08	0.03
Common benefits	-2.37	0.15
Private insurance		
All benefits	-2.34	0.16
Common benefits	-2.11	0.24

Table 5.3: Unit root test of per capita nominal insurer expenditures, 1971-2008

Time series	Test statistic	p-value
Spending per enrollee		
Medicare		
All benefits	3.79	> 0.99
Common benefits	2.69	> 0.99
Private insurance		
All benefits	4.34	> 0.99
Common benefits	5.04	> 0.99
Change in spending		
Medicare		
All benefits	-3.63	< 0.01
Common benefits	-3.54	< 0.01
Private insurance		
All benefits	-2.56	0.10
Common benefits	-2.34	0.15

Table 5.4: Unit root test of per capita nominal insurer expenditures, 1982-2008

5.2 Price data

Statistics	CPI-U	CPI-U, Medical
N	324	324
mean	0.25	0.44
sd	0.32	0.24
skewness	-1.26	0.69
kurtosis	10.69	2.79
min	-1.93	-0.06
max	1.25	1.08
p25	0.11	0.25
p75	0.42	0.59

Source: Bureau of Labor Statistics, 1982–2008

Table 5.5: Consumer Price Index, log monthly change (%)

Statistics	PPI, All	PPI, Drugs and Pharma
N	324	324
mean	0.17	0.40
sd	0.94	0.66
skewness	-1.37	4.95
kurtosis	11.75	56.77
min	-5.48	-1.23
max	2.94	8.04
p25	-0.16	0.05
p75	0.51	0.69

Source: Bureau of Labor Statistics, 1982–2008

Table 5.6: Producer Price Index, sectors, log monthly change (%)

Statistics	PPI, All	PPI, Drugs and Pharma	Hospital	MD Office
Start Date	1/31/1993	1/31/1993	1/31/1993	1/31/1994
End Date	12/31/2008	12/31/2008	12/31/2008	12/31/2008
N	192	192	192	180
mean	0.20	0.30	0.26	0.16
sd	1.13	0.68	0.39	0.42
skewness	-1.38	7.90	2.04	3.72
kurtosis	9.21	90.15	7.27	24.57
min	-5.48	-1.03	-0.18	-1.21
max	2.94	8.04	2.06	3.25
p25	-0.17	0.04	0.00	0.00
p75	0.69	0.41	0.35	0.19

Source: Bureau of Labor Statistics

Table 5.7: Producer Price Index, sectors, log monthly change since 1992 (%)

5.3 Employment data

Statistics	Non-Farm	Private	MD Office	Home Health
Start Date	1/12/1982	1/12/1982	1/12/1982	2/12/1985
End Date	12/12/2008	12/12/2008	12/12/2008	12/12/2008
N	324	324	324	287
mean	0.12	0.13	0.32	0.68
sd	0.18	0.21	0.20	0.71
skewness	0.10	0.19	0.58	0.06
kurtosis	7.24	7.88	5.23	3.90
min	-0.50	-0.59	-0.20	-1.64
max	1.23	1.46	1.29	2.88
p25	0.02	0.01	0.20	0.32
p75	0.23	0.26	0.43	1.09

Source: Bureau of Labor Statistics

Table 5.8: Health industry employment, monthly log change (%)

Statistics	Non-Farm	Private	Health	MD Office	Home Health	Hospital
N	227	227	227	227	227	227
mean	0.09	0.09	0.23	0.27	0.58	0.14
sd	0.15	0.18	0.10	0.15	0.71	0.11
skewness	-0.70	-0.76	0.45	0.04	0.04	-0.51
kurtosis	3.69	3.66	3.74	3.48	3.88	4.40
min	-0.50	-0.59	-0.08	-0.17	-1.64	-0.39
max	0.41	0.45	0.53	0.79	2.88	0.46
p25	-0.01	-0.01	0.17	0.18	0.24	0.06
p75	0.21	0.23	0.28	0.36	0.97	0.21

Source: Bureau of Labor Statistics, 1990–2008

Table 5.9: Health industry employment, monthly log change (%)

5.4 Securities data

Statistics	10 Year Government Bonds	AAA Corporate Bonds	Risk-Free Rate
mean	0.837	0.881	0.421
sd	2.421	1.909	0.207
skewness	0.219	0.524	0.425
kurtosis	3.752	5.536	3.218
min	-6.944	-4.733	0.020
max	8.640	8.549	1.130
p25	-0.650	-0.131	0.280
p75	2.346	1.868	0.540

Source: Global Financial Data, Fama-French factors 1982–2008

Table 5.10: Bond monthly nominal returns (% , continuous log basis)

Statistics	Market	Health	Health Services	Medical Equipment	Drugs
Start Date	1/31/1982	1/31/1982	1/31/1982	1/31/1982	
End Date	12/31/2008	12/31/2008	12/31/2008	12/31/2008	
N	324	324	324	324	324
mean	0.914	1.123	0.834	1.053	1.210
sd	4.481	4.833	6.922	5.291	5.014
skewness	-0.913	-0.193	-0.361	-0.501	-0.102
kurtosis	6.205	4.240	4.747	4.850	3.842
min	-22.540	-20.470	-31.500	-20.560	-19.100
max	12.850	16.540	20.490	16.310	16.370
p25	-1.695	-2.025	-3.510	-1.820	-1.890
p75	3.910	3.995	5.110	4.460	4.375

Source: Fama-French factors 1982–2008

Table 5.11: Stock monthly nominal returns (% , continuous log basis)

Variables	Stocks					Bonds			Inflation	
	Market	Health	Services	Equipment	Drugs	Government	10 Year	AAA		General
Market	1.000									
Health	0.805	1.000								
Services	0.032	0.032	1.000							
Equipment	0.719	0.772	0.024	1.000						
Drugs	0.781	0.993	0.030	0.714	1.000					
1 Month	-0.012	0.012	0.654	0.003	0.014	1.000				
10 Year	0.105	0.146	0.136	0.119	0.146	0.126	1.000			
AAA	0.188	0.171	0.150	0.148	0.168	0.102	0.595	1.000		
General	0.009	-0.034	0.204	-0.007	-0.037	0.259	-0.069	-0.108	1.000	
Medical	-0.036	-0.087	0.263	-0.074	-0.087	0.431	0.032	0.035	0.342	1.000

Table 5.12: Correlations of monthly inflation and nominal returns, 1982–2008

5.5 Medical spending regression results

Variable	Coefficient	SE	t-statistic	p-value	95% CI Min	95% CI Max
constant	2.50	1.3	1.92	0.06	-0.14	5.14
Spending Lag 1	0.72	0.12	6.07	< 0.01	0.48	0.96
N	38					
F-test	< 0.01					
Adj R^2	0.49					
Root MSE	3.26					

Source: National Health Expenditure Survey, 1971–2008

Table 5.13: Per capita nominal insurer expenditures regressed on lagged expenditures

Variable	Coefficient	SE	t-statistic	p-value	95% CI Min	95% CI Max
constant	2.58	1.29	2.00	0.06	-0.08	5.24
Spending Lag 1	0.64	0.14	4.50	< 0.01	0.35	0.93
N	27					
F-test	< 0.01					
Adj R^2	0.43					
Root MSE	2.7					

Source: National Health Expenditure Survey, 1982–2008

Table 5.14: Per capita nominal insurer expenditures regressed on lagged expenditures

Variable	Coefficient	SE	t-statistic	p-value	95% CI Min	95% CI Max
constant	1.75	1.89	0.93	0.36	-2.14	5.65
Med inflation	1.10	0.39	2.79	0.01	0.28	1.91
MD office	0.09	0.65	0.13	0.90	-1.26	1.43
N	27					
F-test	< 0.01					
Adj R^2	0.35					
Root MSE	2.86					

Source: National Health Expenditure Survey, BLS, 1982–2008

Table 5.15: Per capita nominal insurer expenditures regressed on inflation and employment growth

Variable	Coefficient	SE	t-statistic	p-value	95% CI min	95% CI max
constant	-0.09	1.57	-0.06	0.96	-3.35	3.17
Inflation	1.10	0.40	2.73	0.01	0.26	1.93
Premium growth						
Lag 1	0.62	0.16	3.84	< 0.01	0.29	0.95
Lag 2	-0.35	0.15	-2.26	0.03	-0.67	-0.03
MD office						
Lag 1	0.66	0.42	1.59	0.13	-0.20	1.53
N	27					
F-test	< 0.01					
Adj R^2	0.63					
Root MSE	2.16					

Source: National Health Expenditure Survey, BLS, 1982–2008

Table 5.16: Per capita nominal insurer expenditures regressed on multiple variables

5.6 Adaptive expectation results

	θ									
Statistics	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
mean	10.53	9.47	9.03	8.80	8.65	8.55	8.48	8.42	8.38	8.35
sd	2.19	2.57	2.72	2.85	3.01	3.16	3.32	3.46	3.60	3.73
skewness	0.22	0.42	0.42	0.36	0.33	0.32	0.33	0.34	0.34	0.33
kurtosis	1.59	1.86	1.97	1.89	1.80	1.76	1.79	1.88	2.02	2.19
min	7.52	6.32	5.49	4.98	4.64	4.47	4.38	3.50	2.66	1.87
max	14.19	14.35	14.37	14.34	14.35	14.44	14.62	14.86	15.16	15.49
p25	8.40	7.19	6.74	6.33	6.05	5.46	4.99	5.07	5.06	4.98
p75	12.22	11.61	11.71	11.21	11.28	11.63	11.29	10.96	10.66	10.57

Table 5.17: Adaptive expectations average rates for per capita nominal insurer expenditures, 1982–2008

	θ									
Statistics	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
	7.12	5.88	5.24	4.72	4.32	4.04	3.84	3.70	3.60	3.53

Table 5.18: Adaptive expectations forecast rates for per capita nominal insurer expenditures, 2009

	θ									
Statistics	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
mean	-2.62	-1.57	-1.13	-0.89	-0.74	-0.64	-0.57	-0.52	-0.48	-0.44
sd	3.21	3.29	3.27	3.22	3.16	3.10	3.05	3.01	2.99	2.97
skewness	-0.23	-0.13	0.02	0.14	0.22	0.28	0.33	0.36	0.38	0.41
kurtosis	3.65	3.91	4.10	4.41	4.75	5.01	5.10	4.99	4.76	4.48
min	-11.27	-10.51	-9.86	-9.34	-8.96	-8.66	-8.41	-8.19	-7.97	-7.73
max	3.70	5.48	6.76	7.60	8.10	8.33	8.33	8.13	7.76	7.21
p25	-3.94	-3.23	-2.64	-2.69	-2.45	-2.32	-2.27	-2.01	-2.31	-2.46
p75	-0.56	0.81	0.71	0.89	1.03	1.04	0.99	0.92	0.96	0.99

Table 5.19: Adaptive expectations errors for per capita nominal insurer expenditures, 1982–2008

5.7 Figures

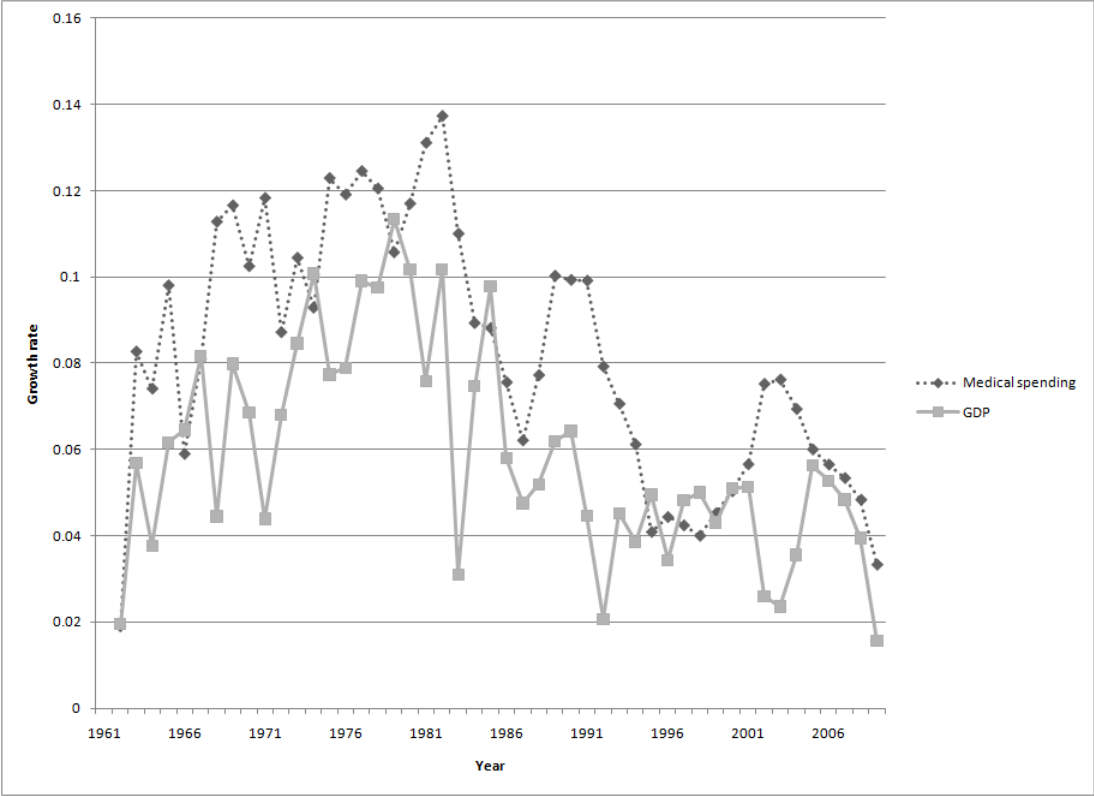


Figure 5.1: Growth rate of nominal GDP and nominal medical spending per capita

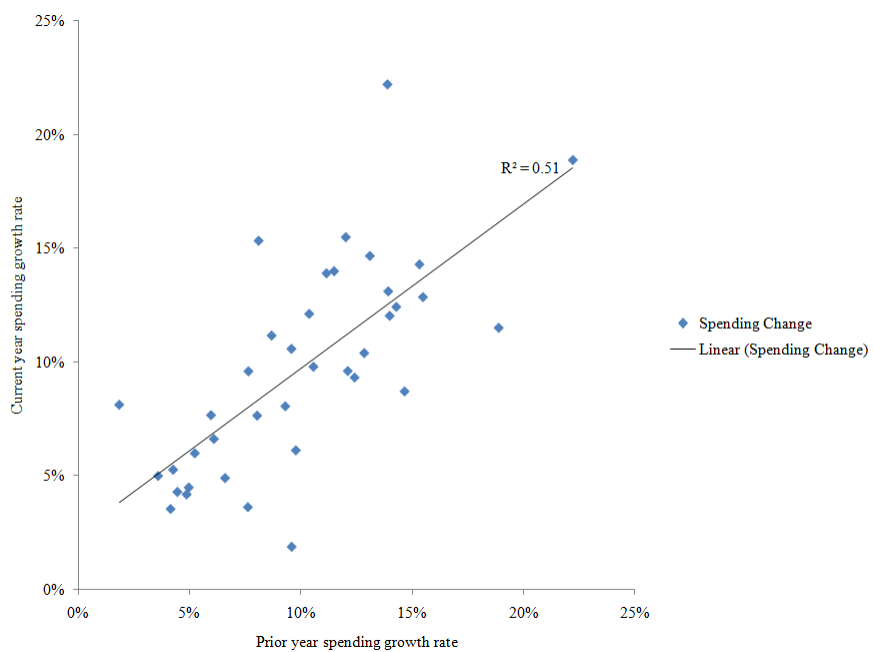


Figure 5.2: Nominal insured medical spending per capita, year on year rate of change

Chapter 6

Analysis of individual level insured spending

1 Introduction

In a previous chapter (chapter 5), I analyze the time series properties of annual average medical spending (specifically insurance premiums). The disadvantages of the level of aggregation are the inability to adjust for population level changes and the necessity of analyzing an annual time series. I am also unable to adjust for switching behavior in the aggregated annual data.

Demographic factors play a large role in the level of spending by health insurers. Aggregate data from the National Health Expenditure Survey shows that spending for females is higher than for males, and spending for those 65 and over is higher than for those aged 65 and over than for those under 65 (Cylus, Hartman, Washington, Andrews and Catlin 2011).

Spending also rises with age for adults, which translates into rising health insurance premiums for older individuals (Herring and Pauly 2006).

Now I focus on average growth rates for high frequency data of a population that I can segment by demographic factors. I can restrict the population to continuously insured individuals to negate the effects of individuals switching into and out of the plan. I can also focus on a population of salaried, non-union workers (and their covered spouses) in order to create a more homogeneous sample. Disaggregating the rates by demographic group also allows me to determine the correlation between the spending growth rates of different groups, and which may be larger or more variable.

Growth rates at a daily frequency also allow me to use high frequency finance data. Returns on stocks and bonds are available daily (though generally not on the weekends), which should make correlations between spending growth and asset returns easier to find and verify, should they exist. The use of daily trend also highlights the difference between medical episodes (continuous) and the trading of assets (Monday-Friday working hours, excluding holidays). In other words, this is not like a hedge where there is a market for the underlying security or commodity, but rather one where the risk may be correlated with available securities.

2 MarketScan data

The MarketScan data contains patient level claims data for private insurers surveyed by Thomson Healthcare (Adamson, Chang and Hansen 2006). Thompson uses the claims

from the employers that submit data to create HIPAA compliant, limited use data sets. I got the data from the National Bureau of Economic Research. The data contains the reimbursements for inpatient, outpatient, and pharmaceutical encounters. The data files include population and enrollment files, and benefit design files. I have access to basic demographic information: age and sex. Geographic information, such as county and three digit zip code are also available, but are restricted to private use. There is also information on the split between the total claim amount for any encounter and the split between the net amount paid by the plan and the amount paid by the employee.

2.1 Population selected

My data comes from the health insurance plan of an anonymous employer in the Manufacturing, Nondurable Goods industry. Over the seven year period 2000-2006, there was only one plan offered by this employer: a POS plan with capitation. The employer's choice to offer only one type of plan, and not to change the plan offerings, allows me to isolate medical spending growth from other changes or plan switching behavior. I do not control for changes in the benefit design, because I want the composition of the basket of available medical care to change with changes in medical practice behavior.

My sample consists of adults aged 18-64, covered as salaried, non-union employees or the spouses of such employees. I only include individuals in a given year who are covered by the plan for the entire year. The total number of covered members varies by year between eighteen and thirty thousand members. I show the membership for these

plans by age group in figure 6.1 and table 6.1, and sex within age group in figure 6.2 and table 6.2. Starting from a relative low in 2000, the population increases by 20% and then 25%, leveling out for 2003-2005 before dropping 36% in 2006. The average population in the plan is around 25,000, with no fewer than 18,000 in any given year. I used the age groups defined on the Medstat data (18-34, 35-44, 45-54, and 55-64). Within age groups, only the oldest group, ages 55-64, is ever below 1,000, with only 841 members in 2000. The split by sex is generally even, both overall and within age groups.

2.2 Spending data used

Spending data is available for drug, inpatient, and outpatient episodes. All files contain both the total payment made for each episode, as well as the net payment made by the insurer. The drug file includes more extensive information, including coinsurance, copays, and deductibles. The outpatient file includes copays and deductibles in all years I observe, but only has coinsurance beginning in 2005. The inpatient file only contains total and net payments until 2005, when coinsurance, copay, and deductible payments are recorded. All payments have dates of service and dates of payment, so I use date of service allocating a claim amount to a claim date. While the drug and outpatient experiences have a single service date attached to them, an inpatient episode can span multiple days. For that reason, I chose to use the admission date when allocating inpatient expenses for an episode of care to a claim date.

The per capita counts of episodes shows daily patterns in inpatient, outpatient, and drug

episodes, but no clear trend over time. The inpatient counts are discrete, with few claims on any given day leading to “levels” in the graph of counts per capita by claim date and the histogram of counts. The drug and outpatient claims have two, and possibly three, different claim count levels, corresponding to weekdays and weekends, with more claims on Saturday than Sunday. There is no discernible upward trend in episodes over the years, although there does appear to be a break in the trend in outpatient counts in 2003 (see figures 6.3, and 6.4). The overall count levels are mirrored in those of subgroups, such as males versus females.

3 Spending analysis

3.1 Level of spending

Total payments are rising both between and within years. Tables 6.3 and 6.4 shows rising average nominal payments per member per day in each year, both overall and for the plan net of member payments. Mean daily spending is statistically significantly different in each year. The average total spending (by the plan and individuals) rises from \$1926 in 2000 to \$3740 in 2006, a 94% increase spread over 6 years. The corresponding compound annual growth rate is 11.7%, but the annual rates of change in mean spending range from 3% in 2003 to 16% in 2004. Net spending increase is a nearly identical 93%, but this change marks even larger variation, including a 21% increase in 2002 and a 3% decrease in 2003.

The two patterns in the aggregate spending data per capita are a general rising trend in

spending and more services on weekdays than weekends. These trends are common across both age and sex. Figures 6.5 and 6.6, show the increase in spending over time for total spending and plan spending, but it is hard to tell whether there are “jumps” in spending each year. In figures 6.7 and 6.8, the total and plan spending histograms are both bimodal, corresponding to weekdays and weekends.

While the amount of spending goes largely to inpatient episodes, followed by outpatient and then drugs, the bimodal distribution of expenses is due to outpatient episodes, with some bimodality for drugs and almost none for inpatient episodes. Figures 6.9, 6.10, and 6.11 show that all three types of episodes have a clear upward trend. It is less obvious whether any of the three trends have jumps or discontinuities from year to year. Outpatient claims clearly fall into near zero and strictly positive days, while drug claims have similar, but less stark, separation. The histogram of claims by days shows similar patterns, as well as making the number of zero inpatient claims days for the population more obvious. Even with so many insured individuals, there are 161 days there are no inpatient admissions.

On a monthly level, spending in later months of the year is on average higher than those in earlier months of the year, but not significantly. In tables 6.5 and 6.6 it appears to be rising over the months, but the monthly spending differences are not statistically significant. It is likely that most of the higher plan spending in later years is based on individuals reaching their out of pocket limits. While I cannot show this conclusively, because I do not have detailed out of pocket spending data for the inpatient files for all years, the out-of-pocket deductible payments for outpatient data shows a stark difference

between beginning and end of year payments (see figure 6.12).

Unlike the month, separating between weekdays and weekends does give meaningful differences in mean spending. In table 6.7 I show that the mean spending in total and by the plan alone is much higher on weekdays than weekends. The difference leads to highly statistically significant differences between the means, both in aggregate and testing within each year. The pattern holds for inpatient and drug claims as well as for outpatient claims, which are incurred mostly on weekdays.

I use the age group categories defined on the Medstat data (18-34, 35-44, 45-54, and 55-64) and sex to break up the spending by demographic groups. Spending is higher for females than males, and the difference between weekdays and weekends is more pronounced. While spending differs significantly among all age groups, the difference is prominent on weekdays but not weekends. Overall, each age group has significantly different spending even when compared to the closest (i.e. the adjacent) age groups, with the older groups more expensive younger ones. The same is true within each year, for total and net spending. Separating weekdays from weekends, the difference remains only for weekday spending, suggesting that weekend spending, which is largely inpatient driven, is probably generated by emergencies that have nothing to do with age related medical care.

Spending differs significantly between the sexes at almost all levels of analysis. Overall, female expenditures are higher than male expenditures. The same is true within each year, for total and net spending. Separating weekdays from weekends, the difference remains for almost all year/weekday and year/weekend combinations, suggesting that male and female

spending will differ even on an emergency basis. This effect persists even for comparison of male and female spending within age groups (see tables 6.8 and 6.9).

I also test for unit roots in the total and net spending data, both overall and by demographic group. I strongly reject the hypothesis of a unit root at every level of analysis (see table 6.10). While there is a clear trend in my spending data, the data is trend stationary over the seven years of observation. As a result, I treat the increase in spending as a single overall trend over time, with variance from the trend line due to randomness.

3.2 Trend of spending

I define trend as a continuous process that increments spending on a daily basis. Daily trend is the log of the day-to-day change in per capita spending.

$$\tau = \ln(\text{Spending per capita}_{it} / \text{Spending per capita}_{it-1}) \quad (6.1)$$

For group i , for time t

As with the aggregate data from chapter 5, I am modeling trends in nominal, rather than real, medical spending.

The choice to model trend continuously has both positive and negative effects. The main positive is that medical episodes occur in continuous time, so the growth in spending could also be a continuous time phenomenon (or at least one that is best modeled on a daily basis). Aggregating at the quarterly or annual level could obscure the true time series properties of spending growth if the process is continuous. The main downside is the difficulty in interpreting the results. If the level of spending is the same on 1/1/2000 and

12/31/2006, then the average log trend will be zero even if the spending was generally increasing over time.

The groups I analyze are total population, sex groups, age groups (18-34, 35-44, 45-54, 55-64), and sex within age groups. For some age group, sex, or age group by sex groups, there were no claims on certain days, leading to missing values for the trend rate out of 2,556 possible observations (2,557 days in 1/1/2000-12/31/2006 less 1/1/2000 where I do not observe the prior day's claims). For example, in the oldest group, aged 55-64, there are two missing observations for the male group within the age group (12/25/2000 and 12/26/2000) and two missing observations for the female group within the age group (10/1/2004 and 10/2/2004). Both of these missing observations come in pairs, because zero claims on one date (12/25/2000) or negative claims on one date (10/1/2004) make the trend uncomputable on a log basis for both dates.

The absolute value of average daily log change in spending by group is smaller than 0.001 in all cases. In some cases, the average is negative, but all figures are close to zero. However, the medians are all negative whereas the skews are all positive, which is indicative of the long right tail of the daily trends. The summarized results for the change are in table 6.11. The standard deviations are large enough that none of the means are indistinguishable from zero. In addition, the standard deviation is lowest for the total population, smaller for younger than older ages, similar for males and females overall, and increasing by age groups for females, but not males. The error rates may be related to the size of the population sampled: 25,492 on average for the entire population, virtually equally split

between males and females, and larger for ages 18-34 and 35-44 than 45-54 and 55-64. (see table 6.12).

4 Analyzing trend

The trend time series is also stationary but it does exhibit autocorrelation. As with the level of spending, there is no unit root in any of the trend time series. However, correlograms of all the time series show a large autocorrelation for the seventh lag of the time series. This likely corresponds to a weekly “seasonality”, which I address using dummy variables for day of the week. For instance, figures 6.13 and 6.14 show the autocorrelation and partial autocorrelation for the total spending series; the other autocorrelation graphs are similar.

I regress the daily trend rate on the one day lagged trend rate, and dummy variables for Monday, Saturday, and 2001–2006 (2000 is the base year). I excluded month dummies because they were not correlated with significantly different trend in the summary statistics.

The regression equation is:

$$\text{trend}_t = \beta_0 + \beta_1 \text{trend}_{t-1} + \beta_2 \text{Monday} + \beta_3 \text{Saturday} + \beta_4 2001 + \dots + \beta_9 2006 \quad (6.2)$$

The results of these regressions for trend based on total spending and trend based on plan spending are in table 6.13. While lagged trend, Monday and Saturday indicator variables are significant, none of the year dummies are significant.

I also jointly tested the hypothesis that all year dummies are 0, and could not reject the hypothesis (the probability that they were non-zero was less than 0.001). The results confirm the results I got from summary statistics on the level of spending. For my data, this

result strongly suggests that the trend in spending is a continuous process without discrete jumps.

I use the regressions without the year dummies after finding them insignificant. The regression equation is:

$$\text{trend}_t = \beta_0 + \beta_1 \text{trend}_{t-1} + \beta_2 \text{Monday} + \beta_3 \text{Saturday} \quad (6.3)$$

$$e_t = \hat{\text{trend}}_t - \text{trend}_t$$

The results of these regressions for trend based on total spending and trend based on plan spending are in table 6.14. Daily trend is negatively related to prior day's trend, although the relationship is not large. The indicators for Saturday and Monday are quite large, reflecting the drop in spending on the weekend and pickup in spending when the work week starts.

The remaining errors from the regressions are now mean zero, but still have fat tails (see figures 6.15 and 6.16). The median of the residuals is strictly positive, although they are near zero skewed with remaining large kurtosis. The correlograms and residual plots of the regression confirm that I have removed the weekly seasonality (see figures 6.17 and 6.18). It is possible that these series are over-differenced, although I only applied a single differencing to generate the trend time series.

The regression of trend by demographic groups shows that daily trend is always negatively correlated with one day lagged trend. The strength of the association varies widely. For the youngest and oldest groups, the coefficient is the largest (most negative), while it is smaller for the two middle age groups. It is also consistently larger for males than for

females. This cannot be accounted for by the size of the underlying populations, which are largest for the youngest and next to youngest groups (see table 6.15). The trend is consistent for total spending and spending by the plan alone.

The control for Monday is always positive and the control for Saturday is always negative. The size of the Monday coefficient is larger for older age groups, and larger for females than males, although this pattern does not hold for each group in the breakdown by age and sex. The Saturday variable is more negative for the two middle age groups, and for females as compared to males, but the pattern also does not hold for each group within the breakdown of age by sex. There is no clear interpretation for these variables, which I used more as controls than as explanatory variables. It may indicate that emergency spending is more important for the oldest and youngest than the middle age groups, and that it is more important for females than males, leading to a smoother spending pattern for the youngest, oldest, and males. Outpatient services are more utilized by females than males (see figures 6.19 and 6.20), which is consistent with prior findings (National Center for Health Statistics 2001), so it is also possible that the low number of weekend outpatient visits is driving female spending down more on weekends.

The regression constant is negative across all groups, and is smaller than both mean and median daily trend. This means, that, all else equal, the day-to-day trend is expected to be negative. The intercept is more negative at older ages, and roughly equal between males and females. This is also mostly the case in the age/sex breakdowns. However, the overall trend in all the data is to higher spending. This reflects the long right tail of the daily

trend data, where there are relatively few large positive day-to-day changes in spending (see table 6.15).

The power of the model to explain the variation in data varies across groups. While my model accounts for around 70% of the variation in spending for all groups aggregated, the adjusted R^2 varies from below 50% to above 60% when assessing age/sex categories. The lowest adjusted R^2 are for the youngest groups, despite the fact that the count of individuals is highest. This may reflect the fact that, for the youngest group, spending growth is hardest to predict. However, the root mean squared error is highest for the oldest groups in the data. The root mean squared error is also persistently higher for males than for females. For these groups, the predicted trend is farthest from the experienced trend, so they may generate the greatest losses from deviation from a trend line.

5 Hedging trend

5.1 Returns on assets

I have daily nominal return data for several different classes of securities. I have returns on the entire market, the health care sector and subsectors, and the risk free rate from the Fama-French research factors (Fama and French 2010). I also have returns for an index of 10 year corporate bonds, created by Dow Jones and provided by Global Financial Data (Global Financial Data 2010). The mean daily returns over the 2000-2006 time horizon vary by a factor of nearly six between the highest average return (health care services) and the

smallest (the drugs subindustry). The corporate bond series is the only one with returns reported on every day; the other series are restricted to trading days only.

The residuals from the trend regressions allow me to analyze the degree of possible common shock, as well as which assets will be valuable for hedging purposes. As I showed in the regression tables, there is a significant amount of variation to explain. While there is no correlation between the residuals and asset return, the residuals of the trend regressions are correlated with each other. The correlation coefficients for any residual and asset pair are no higher than 0.05. On a pairwise basis, the correlations are all insignificant. In contrast, the asset classes are all correlated, with the exception of the risk-free rate (which is what characterizes the risk free asset), and all the asset pairwise correlations are significant. The correlation of spending residuals are all between 0.12-0.25 for total spending (the range is 0.09-0.23 for plan spending). The pairwise correlations are all significant at the 0.001 level. The ideal hedge for the medical spending growth of a specific population would be an asset that mimics the spending growth of that population. The correlations show that the second best hedges for spending growth in one subpopulation would be an asset that mimics the medical spending growth of another subpopulation.

I next evaluate the possibility of hedging prediction errors in medical spending growth with asset returns. My approach is to regress the returns in different asset classes $R(t)$ on the residuals $e(t)$ derived from equation 6.3. My regressions are of the form:

$$R(t) = \alpha + \beta e(t) \tag{6.4}$$

If there is a correlation between the assets and the residuals, then the coefficient for β is

significant. For significant relationships, a positive β suggests buying (going long) on the asset and a negative β suggests selling (shorting) the asset. To the extent that α is non-zero, the returns to market securities are explained by other factors.

The results suggest that none of these broad asset classes are appropriate hedges for growth in spending. All of the coefficients β are small and insignificant. The results do not change when I break out the trend by demographic group. They also are not sensitive to the use of the corporate bond index, which is reported on all days, and not simply weekdays. The results are also similar when I use the trend in plan spending rather than total spending. I am not able to explain any of the variation in returns with the unpredictable portion of spending.

5.2 Excess returns on assets

Another test of the ability of assets to hedge medical spending growth is the effect of spending shocks on excess return. In chapter 5, section 3 I regress excess returns to market assets on the errors from an adaptive expectations model forecasting medical spending growth (see equation 6.1). The motivation is the desire of insurers to maximize their returns in excess of the growth in spending, rather than simply their absolute returns. I use the same method with excess return calculated with daily returns on assets and daily spending growth, and with the residuals $e(t)$ derived from equation 6.3.

I use two different techniques to deal with the mismatch between observations of medical spending and observations of asset returns. I only have asset returns for weekdays, but

I have spending returns for all days of the week. One technique is to regress excess returns for weekdays against residuals for weekdays, and drop the data for weekends. The other technique is to calculate the excess return on Friday as the difference between asset returns on Friday and spending growth for Friday and the weekend. The calculation of excess returns on assets follows the equations:

$$R_e(t) = R(t) - \tau(t) \text{ for M,Tu,W,Th}$$

$$R_e(t) = R(t) - (\tau(t) + \tau(t + 1) + \tau(t + 2)) \text{ for F} \quad (6.5)$$

$$(6.6)$$

The excess returns that only include the weekdays are strongly negative because of the large weekend to weekday transition. The exception is the bond returns, which cover all days and are thus close to mean zero. The small magnitude of the excess bond returns demonstrates the difficulties health insurers may have in generating returns in excess of spending growth. The returns in excess of trend for the whole population is negative, while it is positive for the youngest age group, negative for 35-44s, 45-54s, and 55-64s (whether across or within sex groups). It is also negative for males and females.

The excess return regressions follow the same method as in the aggregate analysis, but with daily data. The regression equation is given by:

$$R_e(t) = \alpha + \beta e(t) \quad (6.7)$$

The excess return $R_e(t)$ matches the excess return concept from the aggregate analysis. Unlike the excess return calculation in the aggregate, the trend during the week swamps the

return on assets. The excess returns without considering the weekends are dominated by the magnitude of trend. When adding in the weekend trend, the magnitude of the trend still swamps the return on assets, but by a much smaller number. The results are in table 6.17. The corporate bond results, occurring on all days of the week, are of a similar magnitude to the return on assets, so the difference is indistinguishable.

The regression results are in table 6.18. All the coefficients are one or nearly 1, with small p-values. The regression shows that the test for correlated assets in the excess return regression is swamped by correlation between the trend regression errors and trend itself. The positive aspect of my finding is that insurers are not trying to hedge daily trend. Additionally, the MarketScan data does point to the heterogeneity of spending level and volatility across populations, which is unavailable in the aggregated National Health Expenditure Data.

I also ran the results with trend and stock aggregated at the weekly level, but the results did not change substantially. The oldest group has the highest standard deviation in weekly spending. The best predictor of weekly spending growth is the one week lag in spending growth. The dummy variables for year are all insignificant, so knowing the year does not improve the prediction above and beyond the most recent observation of spending growth. One difference is that the regression explains much less of the variation, with R^2 between 0.10-0.30 rather than 0.40-0.75. The regressions also produce mean zero residuals for correlation with asset returns.

There are also similar results for the correlations with weekly asset returns. The excess

returns over this period are near zero. The correlations are 0.10 or less in absolute value terms. The the coefficients of the excess returns with the residuals are one or nearly 1, with small p-values, again showing that the test for correlated assets in the excess return regression is swamped by correlation between the trend regression errors and trend itself. The regressions of total, not excess, nominal returns on the residuals are insignificant. The fact that the results are the same at the daily and weekly levels of aggregation supports my general conclusion that there is no asset that can be used to hedge.

6 Remaining differences in volatility

6.1 Importance of volatility

Volatility in trend within and across groups is important because of my difficulty in finding a hedging asset. The results from chapter 5 suggest that there is no good asset for hedging the year-to-year fluctuations in medical spending growth. Guaranteed renewable insurance is not written on an entire population, but on identifiable classes. Additional volatility in spending for subpopulations that comprise the insured classes will make the deviation from expected trend more difficult to manage.

I also failed to find a hedging assets for the subpopulations I specified in this chapter. Higher frequency spending and asset return data did not help in surmounting the problem. Therefore, insurers may have to manage not only the fluctuations in spending growth for the entire macroeconomy, but also the additional fluctuations attributable to each class of

guaranteed renewable insurance. Alternatively, insurers could ignore entirely the volatility in aggregate spending and focus only on the medical trend for each pool of insureds.

Either perspective could generate a different trend based loading factor for different insured groups. Insurers might have to reserve for the different rates of trend in different groups. In my data, I found differences in trend rates between groups. Insurers also have to load for the riskiness of trend in order to account for larger fluctuations around the trend line for different groups. I also found some differences in riskiness of medical spending growth between groups. The limitation of my data is the six year time horizon, but that would likely generate a conservative estimate of the volatility in spending over time.

6.2 Differences by sex

Trend for males and females is not more volatile than for the total population. While spending for females is always higher than for males, the growth rates in average daily spending for the two groups often differ substantially. The standard deviation and kurtosis are slightly higher, and the skewness is slightly lower. The volatility of medical spending growth differs from the volatility of average daily spending. Daily trend for males had a lower standard deviation than for the group as a whole, and therefore a higher standard deviation for females. Daily trend is less predictable for males, with a lower R^2 and higher RMSE than for females. The predictability of daily trend for females is lower than for the population as a whole.

The difference in levels of spending between males and females is the only significant

difference I observed. The level of guaranteed renewable premiums and reserves will be higher for females than males because of the higher level of spending for females. The higher per capita spending extends to the 65 and older population, as demonstrated by other studies (Cylus et al. 2011, e.g.). If the lower predictability of trend for males held up over a longer period of time, then guaranteed renewable reserves would have to be higher for males *ceteris paribus*. Males and females might differ in other ways, such as tendency to lapse, which would be more important than the small differences in spending volatility.

6.3 Differences by age

The main differences in spending by age is that trend for the middle age groups is more predictable than for the youngest and oldest group. The raw rates of change are the most variable for the oldest group, which may be due to their being the smallest subpopulation in my data. The oldest group also had the lowest overall rate of change in aggregate from 2000 to 2006. The predictability of the groups, as measured by the R^2 and RMSE, was lowest for the oldest group, followed by the youngest group, and then the two middle groups. The regression fit the 35-44 year olds better by a slightly higher margin.

If the predictability of the youngest adults continued for a longer time period, then the reserves required for younger individuals would be higher than under deterministic trend. Recognizing the risk would mean shifting some of the front loading onto individuals who purchase guaranteed renewable contracts earlier in their lives. The effect on the oldest group would have a lower impact, since the oldest group in my sample is close to Medicare

eligibility. It is also possible that, over a longer time horizon, the differences that seem large in modeling high frequency variation in trend would not be as large, or would disappear altogether. I fit the same model to all age groups, which suggests that the overall process of medical spending growth may be independent of the increase in medical spending by age.

6.4 Implications for long term health insurance

Guaranteed renewable health insurance must be written with contract reserves in the absence of medical spending growth or aging trends. The contract is underwritten to start with a healthy pool of individuals, and some insured individuals will naturally get sick over time. The insurance does not protect people against foreseeable increases due to aging or medical spending growth. The insurer uses contract reserves to recognize the liability associated with the promise to price future premiums on a class average basis.

The volatility due to unforeseen events in the effect of age and time on medical spending must also be part of the reserve. The effect of age on claims can only be predicted from cross-sectional observations of different spending by people at different ages and trends in how the aging effect has changed over time. The effect of time on claims can only be predicted from trends in how the level of medical spending has changed over time. Any prediction will be made with error because the time trends in medical spending follow a stochastic process. High frequency, short term data demonstrates that it may be appropriate to treat medical spending growth as an aggregate process that treats all population subgroups equally. In that case, the inability to hedge spending growth with financial assets

means that the insurers must still add a risk load for medical spending growth risk to their reserves. In the next chapter, I address the question of the amount of additional reserves that might be required.

7 Tables and figures

Year	All	Age group							
		18-34		35-44		45-54		55-64	
	Count	Change	Count	Change	Count	Change	Count	Change	
2000	19,310		6,996		7,741		3,732		841
2001	24,062	25%	8,838	26%	9,327	20%	4,817	29%	1,080
2002	28,789	20%	10,467	18%	11,038	18%	5,968	24%	1,316
2003	29,403	2%	10,032	-4%	11,324	3%	6,548	10%	1,499
2004	29,447	0%	9,671	-4%	11,136	-2%	6,950	6%	1,690
2005	28,845	-2%	9,142	-5%	10,608	-5%	7,252	4%	1,843
2006	18,586	-36%	5,587	-39%	6,910	-35%	4,801	-34%	1,288
Average	25,492		8,676		9,726		5,724		1,365

Table 6.1: Covered population by age group, 2000-2006

Year	All	Age group							
		18-34		35-44		45-54		55-64	
	Male	Female	Male	Female	Male	Female	Male	Female	
2000	9,643	9,667	3,400	3,596	3,868	3,873	1,905	1,827	470
2001	11,922	12,140	4,242	4,596	4,639	4,688	2,454	2,363	587
2002	14,115	14,674	4,934	5,533	5,427	5,611	3,041	2,927	713
2003	14,406	14,997	4,694	5,338	5,585	5,739	3,317	3,231	810
2004	14,409	15,038	4,502	5,169	5,461	5,675	3,550	3,400	896
2005	13,914	14,931	4,182	4,960	5,094	5,514	3,671	3,581	967
2006	8,737	9,849	2,485	3,102	3,249	3,661	2,341	2,460	662
Average	12,449	13,042	4,063	4,613	4,760	4,966	2,897	2,827	729
									636

Table 6.2: Covered population by age group and sex, 2000-2006

Year	Mean spending			SD	N
	Daily	Annualized	Change		
2000	5.28	1,926		2.62	366
2001	5.75	2,098	9%	2.87	365
2002	6.77	2,470	18%	3.35	365
2003	6.97	2,544	3%	3.37	365
2004	8.06	2,941	16%	3.91	366
2005	9.08	3,314	13%	4.58	365
2006	10.25	3,740	13%	5.21	365

Table 6.3: Total nominal spending per member per day, by year

Year	Mean spending			SD	N
	Daily	Annualized	Change		
2000	4.06	1,483		2.15	366
2001	4.70	1,716	16%	2.38	365
2002	5.68	2,074	21%	2.91	365
2003	5.50	2,006	-3%	2.77	365
2004	6.27	2,289	14%	3.23	366
2005	7.07	2,580	13%	3.62	365
2006	7.84	2,860	11%	4.23	365

Table 6.4: Plan nominal spending per member per day, by year

Year	Mean spending	SD	N
January	7.11	3.96	217
February	7.13	3.67	198
March	7.35	3.9	217
April	7.25	3.83	210
May	7.38	4.08	217
June	7.45	4.03	210
July	7.22	4.04	217
August	7.41	3.82	217
September	7.45	4.22	210
October	7.78	4.29	217
November	7.84	4.84	210
December	8.00	4.80	217
Total	7.45	4.14	2557

Table 6.5: Total nominal spending per member per day, by month

Month	Mean spending	SD	N
January	5.10	2.88	217
February	5.35	2.80	198
March	5.65	3.07	217
April	5.65	3.05	210
May	5.76	3.19	217
June	5.89	3.24	210
July	5.81	3.33	217
August	5.90	3.05	217
September	6.04	3.51	210
October	6.30	3.46	217
November	6.39	3.94	210
December	6.61	4.02	217
Total	5.87	3.34	2557

Table 6.6: Plan nominal spending per member per day, by month

Day	Mean spending	SD	N
Total spending			
Weekend	2.51	1.19	732
Weekday	9.43	3.13	1825
Total	7.45	4.14	2557
Plan spending			
Weekend	1.96	1.04	732
Weekday	7.44	2.56	1825
Total	5.87	3.34	2557

Table 6.7: Nominal spending per member per day, weekend vs. weekday

Year	Male				Female			
	Daily	Yearly	% Change	SD	Daily	Yearly	% Change	SD
2000	4.02	1,467		2.24	6.53	2,383		3.53
2001	4.47	1,630	11%	2.45	7	2,557	7%	3.81
2002	5.24	1,914	17%	3.17	8.23	3,004	18%	4.1
2003	5.2	1,897	-1%	2.73	8.67	3,165	5%	4.33
2004	5.9	2,154	13%	2.92	10.12	3,695	17%	5.41
2005	6.82	2,491	16%	4.29	11.18	4,081	10%	5.83
2006	8.09	2,952	19%	4.8	12.16	4,439	9%	6.61

Table 6.8: Total nominal spending by sex

Year	Male			Female		
	Daily	Annualized	SD	Daily	Annualized	SD
2000	3.10	1,132	1.85	5.02	1,833	3.03
2001	3.70	1,350	2.16	5.69	2,075	3.07
2002	4.41	1,609	2.86	6.91	2,521	3.53
2003	4.06	1,480	2.29	6.88	2,510	3.58
2004	4.53	1,653	2.40	7.94	2,898	4.59
2005	5.22	1,905	3.09	8.79	3,209	4.83
2006	6.21	2,267	4.15	9.28	3,387	5.35

Table 6.9: Plan nominal spending by sex

Population	Spending	
	All	Plan
All	-35	-36
Age 18-34	-42	-44
Age 35-44	-38	-38
Age 45-54	-40	-41
Age 55-64	-44	-46
All males	-35	-36
All females	-36	-37
Males		
Age 18-34	-47	-48
Age 35-44	-43	-45
Age 45-54	-44	-43
Age 55-64	-39	-39
Females		
Age 18-34	-45	-46
Age 35-44	-41	-43
Age 45-54	-48	-48
Age 55-64	-42	-46

Table 6.10: Unit root test of daily nominal spending by population

Population	Trend						
	Freq	Mean	Median	SD	Skewness	Kurtosis	
All	2,556	0	-0.07	0.86	0.63	3.62	
Age 18-34	2,556	0	-0.04	0.91	0.32	3.74	
Age 35-44	2,556	0	-0.06	0.94	0.54	3.65	
Age 45-54	2,556	0	-0.09	1.01	0.63	3.74	
Age 55-64	2,556	0	-0.11	1.19	0.39	3.98	
All males	2,556	0	-0.07	0.91	0.54	3.6	
All females	2,556	0	-0.07	0.91	0.6	3.66	
Males							
Age 18-34	2,556	0	-0.07	1.11	0.37	3.73	
Age 35-44	2,556	0	-0.08	0.99	0.43	3.67	
Age 45-54	2,556	0	-0.1	1.07	0.45	3.57	
Age 55-64	2,554	0	-0.09	1.28	0.3	3.88	
Females							
Age 18-34	2,556	0	-0.04	0.97	0.29	3.85	
Age 35-44	2,556	0	-0.07	1.04	0.5	3.66	
Age 45-54	2,556	0	-0.11	1.11	0.68	3.84	
Age 55-64	2,554	0	-0.12	1.32	0.43	4.97	

Table 6.11: Daily log change in total nominal spending by demographic group

Population	Count
All	25,492
Age 18-34	8,676
Age 35-44	9,726
Age 45-54	5,724
Age 55-64	1,365
All males	12,449
All females	13,042
Males	
Age 18-34	4,063
Age 35-44	4,760
Age 45-54	2,897
Age 55-64	729
Females	
Age 18-34	4,613
Age 35-44	4,966
Age 45-54	2,827
Age 55-64	636

Table 6.12: Average population by group

Total spending			Plan spending		
Variable	Coefficient	p-value	Variable	Coefficient	p-value
Lag trend	-0.03	0.01	Lag trend	-0.04	< 0.01
Binary variables			Binary variables		
Monday	1.64	< 0.01	Monday	1.651	< 0.01
Saturday	-1.07	< 0.01	Saturday	-1.086	< 0.01
2001	0.00	0.96	2001	0.00	0.96
2002	0.00	0.95	2002	0.00	0.96
2003	0.00	1.00	2003	0.00	0.99
2004	0.00	0.99	2004	0.00	0.99
2005	0.00	0.97	2005	0.00	0.98
2006	0.00	0.96	2006	0.00	0.97
Constant	-0.08	< 0.01	Constant	-0.08	< 0.01
Observations	2,555		Observations	2,555	
F-statistic	797.45		F-statistic	742.71	
Adj. R^2	0.74		Adj. R^2	0.72	
Root MSE	0.44		Root MSE	0.46	

Table 6.13: Nominal spending regression

Total spending			Plan spending		
Variable	Coefficient	p-value	Variable	Coefficient	p-value
Lag trend	-0.03	0.01	Lag trend	-0.04	< 0.01
Binary variables			Binary variables		
Monday	1.64	< 0.01	Monday	1.65	< 0.01
Saturday	-1.07	< 0.01	Saturday	-1.09	< 0.01
Constant	-0.08	< 0.01	Constant	-0.08	< 0.01
Observations	2,555		Observations	2,555	
F-statistic	2397.97		F-statistic	2233.36	
Adj. R^2	0.74		Adj. R^2	0.72	
Root MSE	0.44		Root MSE	0.46	

Table 6.14: Nominal spending regression

Variable	Age Groups										Sex	
	All adults	18-34		35-44		45-54		55-64		Male	Female	
		Male	Female	Male	Female	Male	Female	Male	Female			
Lag trend	-0.03	-0.16	-0.06	-0.09	-0.15	-0.09	-0.15	-0.09	-0.14	-0.09	-0.04	
(p-value)	0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Monday	1.64	1.42	1.63	1.82	1.9	1.59	1.68	1.59	1.68	1.59	1.68	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Saturday	-1.07	-1.02	-1.19	-1.11	-1.04	-0.99	-1.14	-0.99	-1.14	-0.99	-1.14	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Constant	-0.08	-0.06	-0.06	-0.1	-0.12	-0.09	-0.08	-0.09	-0.08	-0.09	-0.08	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Observations	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	
Adj. R^2	0.74	0.58	0.67	0.65	0.52	0.63	0.71	0.63	0.71	0.63	0.71	
Root MSE	0.44	0.60	0.54	0.60	0.82	0.55	0.49	0.55	0.49	0.55	0.49	
F-statistic	2397.97	1153.56	1704.40	1609.70	936.28	1449.35	2124.68	1449.35	2124.68	1449.35	2124.68	
				Age Groups								
		18-34		35-44		45-54		55-64				
Lag trend	-0.21	-0.2	-0.13	-0.09	-0.16	-0.1	-0.22	-0.14	-0.14	-0.22	-0.14	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Monday	1.56	1.41	1.55	1.74	1.72	1.96	1.82	2.08	1.82	2.08	2.08	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Saturday	-1.05	-1.04	-1.05	-1.28	-1.03	-1.21	-0.99	-1.21	-0.99	-1.21	-1.21	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Constant	-0.07	-0.05	-0.07	-0.06	-0.1	-0.11	-0.12	-0.12	-0.12	-0.12	-0.12	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Observations	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	2,555	
Adj. R^2	0.49	0.53	0.55	0.64	0.55	0.63	0.45	0.52	0.63	0.45	0.52	
Root MSE	0.79	0.67	0.66	0.63	0.72	0.68	0.95	0.68	0.95	0.68	0.92	

Table 6.15: Regression results by demographic group for total nominal spending

Variable	Age Groups						Sex	
	All adults	18-34	35-44	45-54	55-64	Male	Female	
Lag trend	-0.03	-0.19	-0.08	-0.11	-0.16	-0.11	-0.06	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Monday	1.65	1.45	1.65	1.82	1.92	1.59	1.7	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Saturday	-1.09	-1.05	-1.21	-1.14	-1.07	-1.01	-1.17	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Constant	-0.08	-0.06	-0.06	-0.1	-0.12	-0.08	-0.08	
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
Observations	2,555	2,555	2,555	2,555	2,555	2,555	2,555	
Adj. R^2	0.72	0.55	0.65	0.64	0.51	0.61	0.70	
Root MSE	0.46	0.65	0.58	0.63	0.86	0.59	0.52	

Variable	Age Groups					
	18-34	35-44	45-54	55-64	Male	Female
Lag trend	-0.24	-0.22	-0.15	-0.10	-0.18	-0.11
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Monday	1.57	1.46	1.54	1.78	1.73	1.98
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Saturday	-1.10	-1.08	-1.08	-1.32	-1.07	-1.23
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Constant	-0.07	-0.05	-0.06	-0.07	-0.09	-0.11
(p-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Observations	2,555	2,555	2,555	2,555	2,555	2,555
Adj. R^2	0.47	0.50	0.53	0.61	0.53	0.61
Root MSE	0.87	0.74	0.71	0.68	0.76	0.72
F-statistic	743.69	867.21	947.82	1357.06	973.45	1322.52
					643.85	852.71

Table 6.16: Regression results by demographic group for plan nominal spending

Population	Market	Asset class					Risk free	Corp bonds
		Health	Health svcs	Medical eq	Drugs	Drugs		
All	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
18-34	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
35-44	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0.00
45-54	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
55-64	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
M	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
F	-0.02	-0.02	-0.01	-0.02	-0.02	-0.02	-0.02	0.00
M, 18-34	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
F, 18-34	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0.00
M, 35-44	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
F, 35-44	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0.00
M, 45-54	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
F, 45-54	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
M, 55-64	-0.02	-0.02	-0.01	-0.02	-0.02	-0.02	-0.02	0.00
F, 55-64	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0.00

Table 6.17: Daily nominal return less daily nominal trend, weekend included

Asset	All	P				opulation	
		18-34	35-44	45-54	55-64	Male	Female
Market	-1.00	-1.01	-1.03	-0.98	-0.97	-0.99	-1.02
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Health	-1.00	-1.01	-1.03	-0.98	-0.97	-0.99	-1.02
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Health svcs	-1.00	-1.01	-1.03	-0.98	-0.97	-0.99	-1.02
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Medical eq	-1.00	-1.01	-1.03	-0.98	-0.97	-0.99	-1.02
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Drugs	-1.00	-1.01	-1.03	-0.98	-0.97	-0.99	-1.02
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Risk-free	-1.00	-1.01	-1.03	-0.98	-0.97	-0.99	-1.02
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Corp bonds	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Asset	Population							
	18-34		35-44		45-54		55-64	
	Male	Female	Male	Female	Male	Female	Male	Female
Market	-1.00	-1.02	-1.02	-1.03	-1.00	-0.98	-0.99	-0.97
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Health	-1.00	-1.02	-1.02	-1.03	-1.00	-0.98	-0.99	-0.97
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Health svcs	-1.00	-1.02	-1.02	-1.03	-1.00	-0.98	-0.99	-0.97
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Medical eq	-1.00	-1.02	-1.02	-1.03	-0.99	-0.98	-0.99	-0.97
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Drugs	-1.00	-1.02	-1.02	-1.03	-1.00	-0.98	-0.99	-0.97
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Risk-free	-1.00	-1.02	-1.01	-1.03	-1.00	-0.98	-0.99	-0.97
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Corp bonds	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
(p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.18: Excess nominal return regressed on residuals

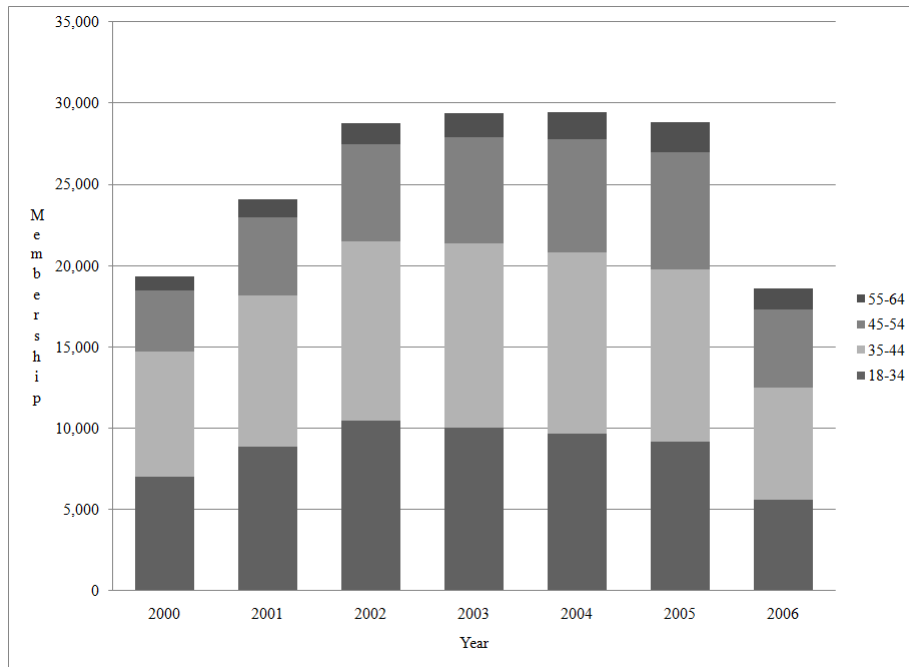


Figure 6.1: Count of members by age

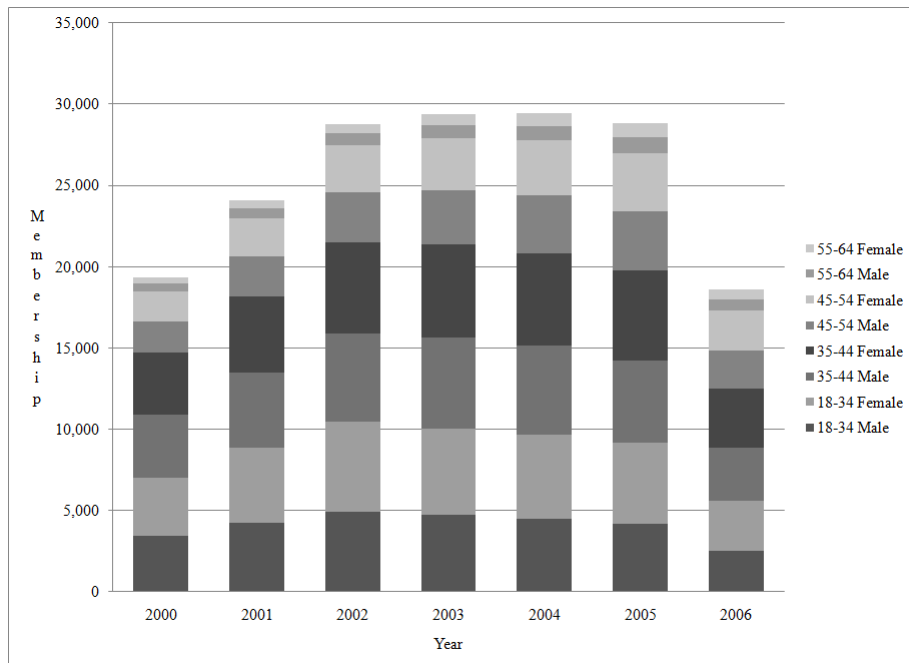


Figure 6.2: Count of members by age and sex

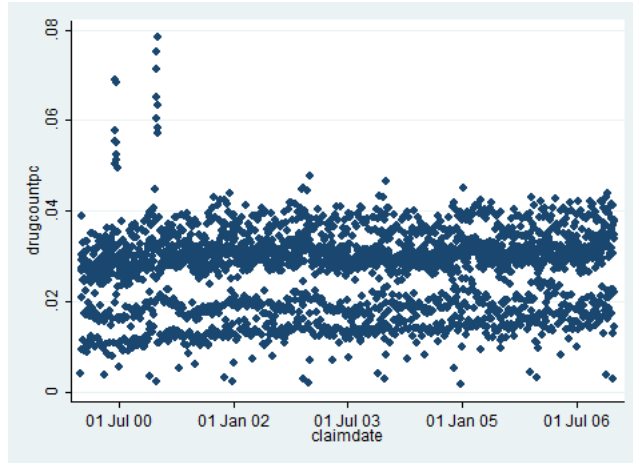


Figure 6.3: Per capita drug experiences by date

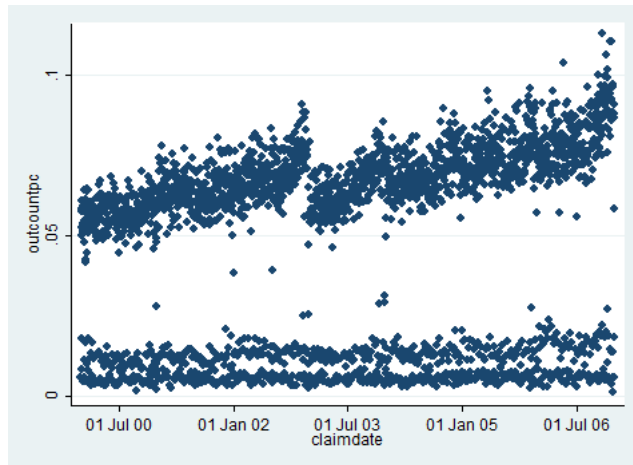


Figure 6.4: Per capita outpatient experiences by date

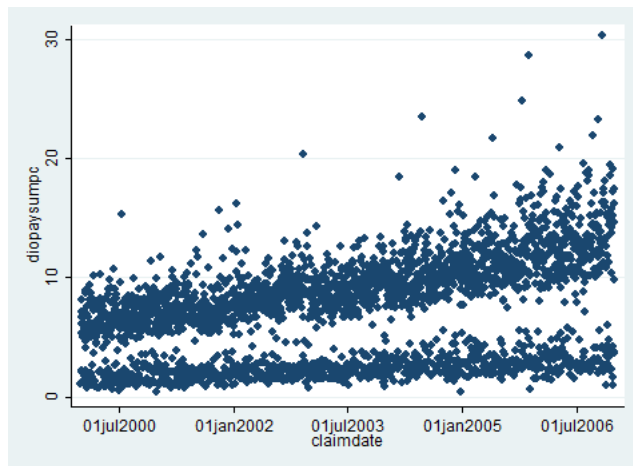


Figure 6.5: Total spending per capita

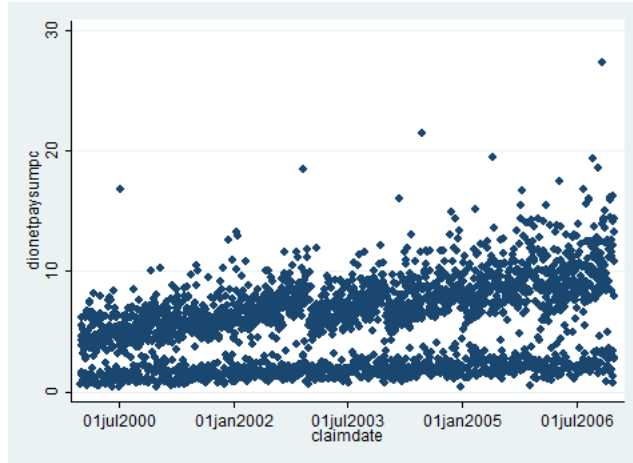


Figure 6.6: Plan spending per capita

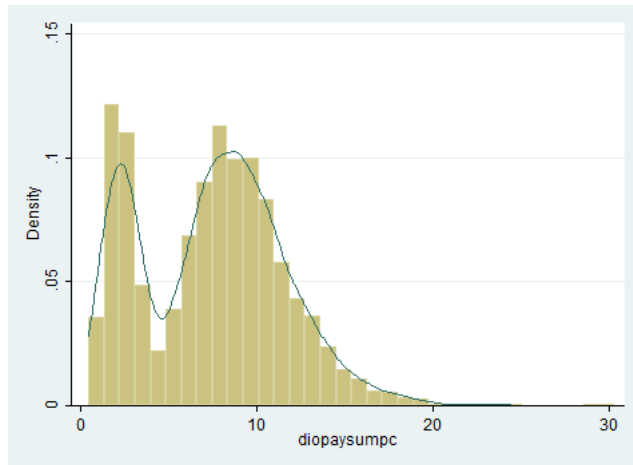


Figure 6.7: Histogram of total spending per capita

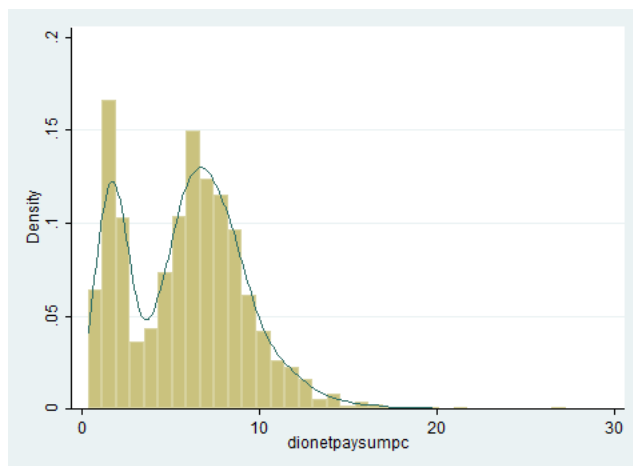


Figure 6.8: Histogram of plan spending per capita

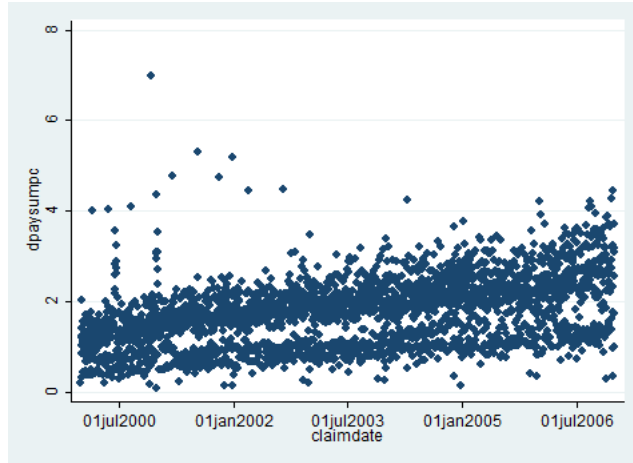


Figure 6.9: Drug spending per capita

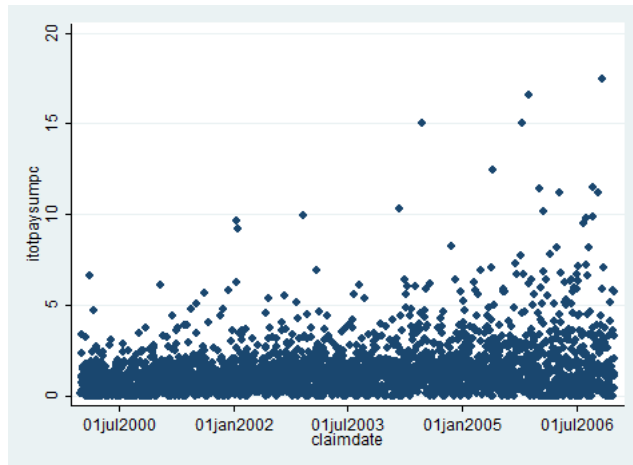


Figure 6.10: Inpatient spending per capita

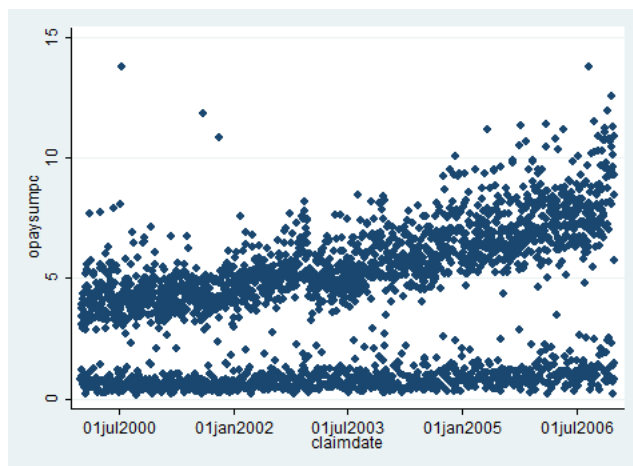


Figure 6.11: Outpatient spending per capita

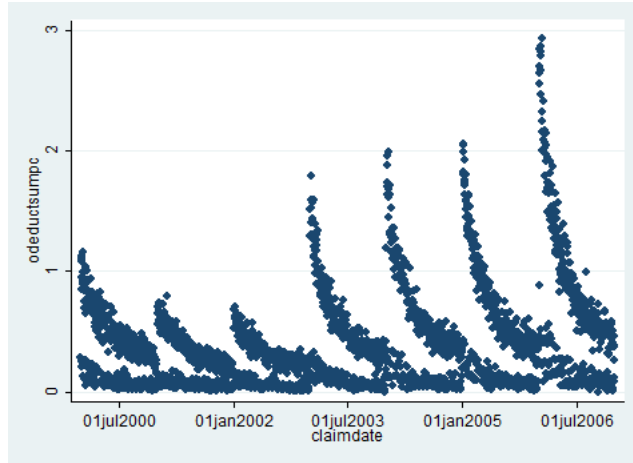


Figure 6.12: Outpatient, out-of-pocket deductible payments per member

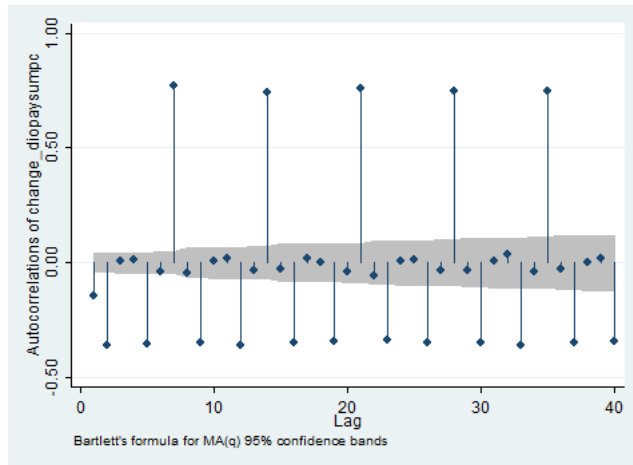


Figure 6.13: Autocorrelation for daily trend

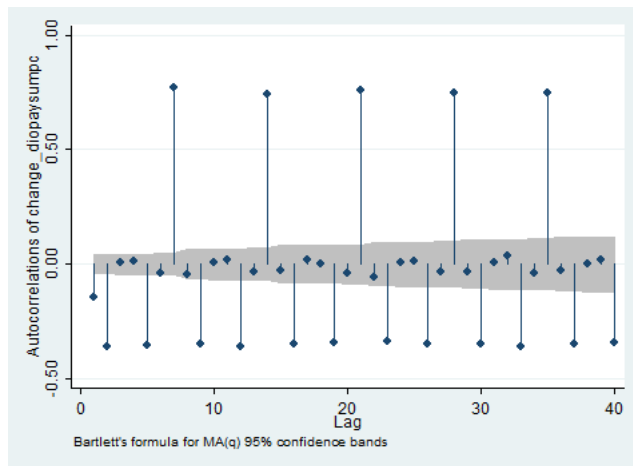


Figure 6.14: Partial autocorrelations for daily trend

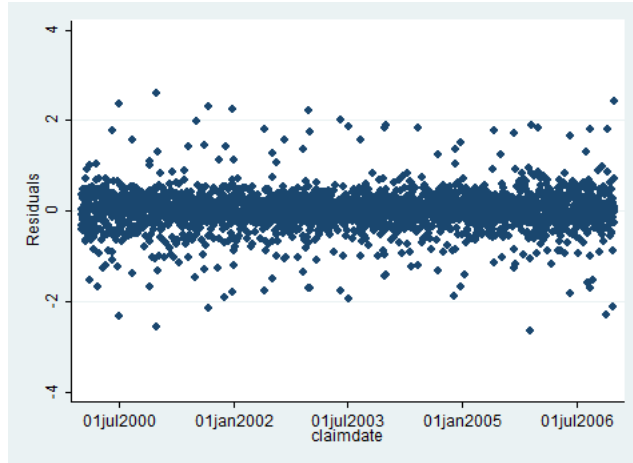


Figure 6.15: Scatter plot of residuals

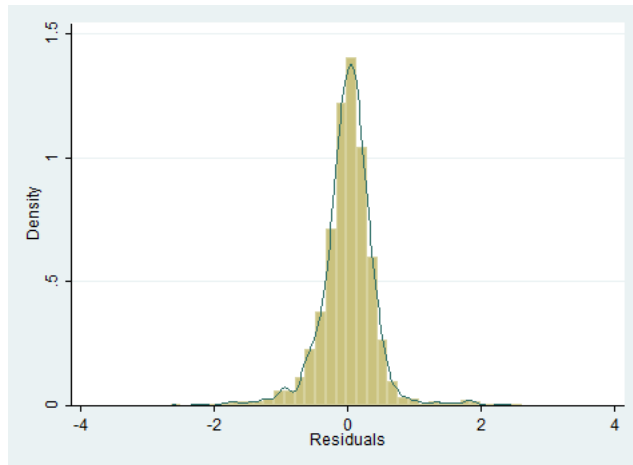


Figure 6.16: Histogram of residuals

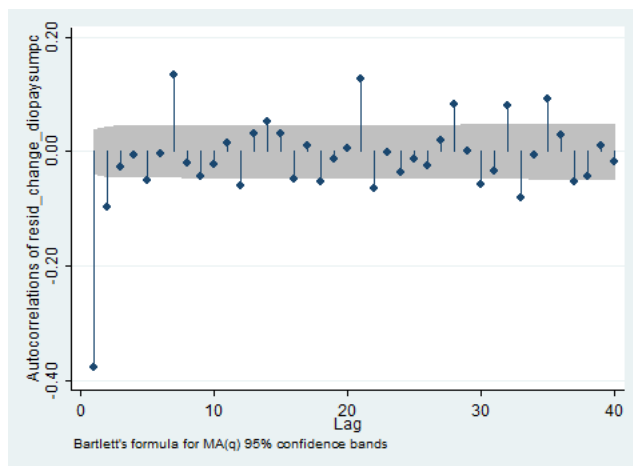


Figure 6.17: Autocorrelation for daily trend

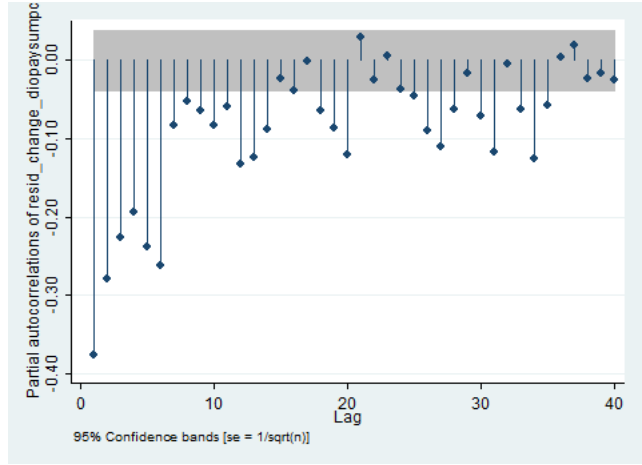


Figure 6.18: Partial autocorrelations for daily trend

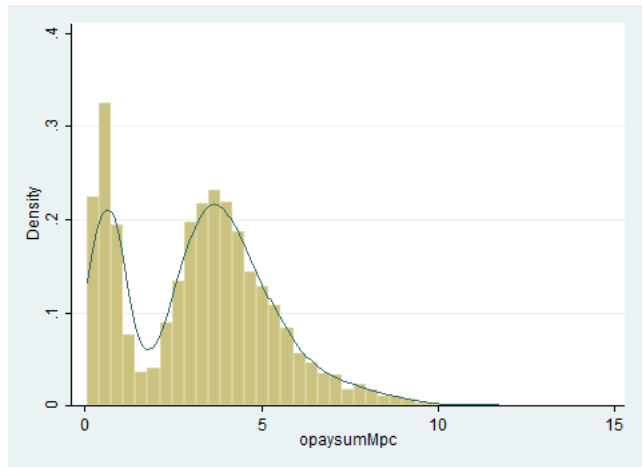


Figure 6.19: Outpatient spending per capita by males

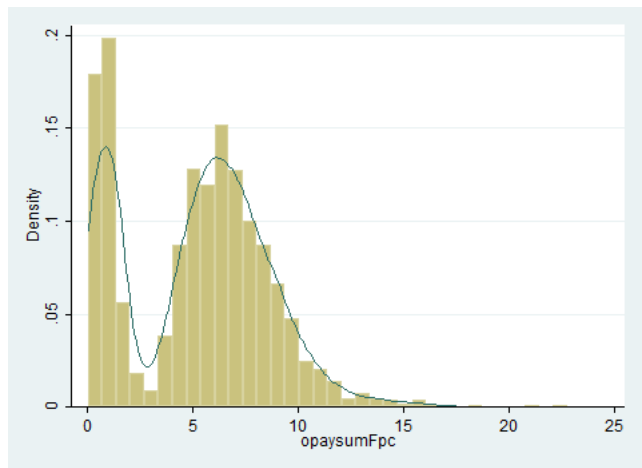


Figure 6.20: Outpatient spending per capita by females

Chapter 7

Effect of spending growth on guaranteed renewable reserves

1 Determinants of the size of reserves

1.1 Rationale for reserving against spending growth

Guaranteed renewable insurance requires a reserve for the high type insureds (see chapter 3). The reason to hold the reserve is to recognize, and hold a buffer against, the difference between the expected losses and future premiums. The expected losses are based on the probability and magnitude of losses for high type insureds, while the composition of the insured population at any time is based on the loss probabilities for high and low types. The premiums are based on the blended expectation of losses for high and low types. Class averaging across the two probability distributions leads to lower premiums for

high types relative to their expected losses, and a gap between guaranteed renewable and annual renewal premiums that must be recognized through reserves.

The size of the loss is increasing due to medical spending growth. Here, the growth and variation in L is exogenous to the particular population insured. The variability in L is uncorrelated with the mix of types, but rather is determined by the macroeconomics of health care spending. The magnitudes of reserves suggested by the results in chapters 5 and 6 are not necessarily so high as to be prohibitive, and can be “baked in” to guaranteed renewable premiums—the insurance is designed to transfer the reserves as an ex post subsidy from those who turn out to be low types to those who turn out to be high types. The spending variance, however, may cause difficulties in anticipating the required level of reserves. The insurer must dynamically adjust the level of contract reserves in response to the realized spending growth rate. The use of contract reserves will also lead to realized gains and losses over the life of the contract, possibly threatening the solvency of the insurance company.

1.2 Consequences of no hedging assets

The size of the reserves cannot be reduced by appeals to the hedging characteristics of assets purchased with reserve funds. If hedging assets did exist, firms should be borrowing an amount equal to the liability from their insureds, investing it in the hedging asset, and thereby make themselves a risk free insurance company. In the absence of the hedging assets, only the investment policy changes, not the amount of the liability. Instead, insurance

companies writing guaranteed renewable contracts must set their reserves by determining the size of losses that could be generated by above average spending growth (a negative shock, from the insurance company's point of view). The cost of the reserves will then be the economic cost of capital, which takes into account the return possibilities for invested assets, and the expected returns to a diversified portfolio.

The absence of a hedging asset raises the cost of reserves relative to their benefits. If additional reserves were available at a fixed cost of capital then insurers could solve the problem of reserving for guaranteed renewable policies by holding reserves equal to the maximum possible losses under the plan. Insurers are already required to hold contract reserves that are equal to expected losses, so in theory they are already reserving not only for losses assuming no medical trend, but also adding trend in to the pricing of policies and the size of reserves.

While the reserves required may seem "large" for a health insurer, the proper comparison is with insurers with longer duration liabilities. For example, life insurance companies must manage large reserves, often with the help of reinsurance contracts (Colquitt and Hoyt 1997). The absence of these arrangements, or their inadequacy, is a justification for the low take up of long term care insurance (Cutler 1993). A related problem in a different line of insurance is property and casualty insurance with implicit guarantees for policy renewal that face low probability, high cost risk events. Cummins and Lewis (2002) study such a situation in the light of a terrorist catastrophe, and find that it leads to a "flight to quality". Higher quality firms can be expected to take a hit to capital and then recover, tak-

ing on new investments, while lower quality firms may be wiped out and fail to pay some claims. Since guaranteed renewable contracts are more explicitly long term than property and casualty arrangements, guaranteed renewable insurers focus more on holding reserves and buying reinsurance and less on the ability to survive a one time event. The availability and cost of additional capital to respond to variations in claims is still important for guaranteed renewable insurance.

Despite the motivation to hold reserves against guaranteed renewable policies, reserves may be suboptimally low. Insurers and consumers may be locked into an equilibrium in which reserves are suboptimal, and consumers rationally believe that insurers hold suboptimal reserves. Duration rating of insurance, and the conflict in the actuarial literature regarding the “correct” level of pricing and reserving under class average guaranteed renewability, suggests that some insurers may be under-reserving and then violating the guarantee (see chapter 2). There are few examples of failure by health insurers; failure or default on contracts is more common in long term care and retiree health insurance. For individual health insurance arrangements, especially those offered by a multiproduct firm, failure to adequately reserve is more likely to manifest itself as “slippage” in the strength of the guarantee (which is hard to detect) than insurers failing to pay benefits (which is easy to detect).

It is also possible for a firm with one or two years of better than expected experience to quickly build a surplus. The amount of reserves underlying a guaranteed renewable policy will start at a high level, and it will tend to decrease over time as the time to the end of

the guarantee leads to a natural form of deleveraging (essentially, amortization of the long term liability generated by the guaranteed renewable policy). Then, the firm may choose to redeploy the surplus to writing additional guaranteed renewable contracts or other lines of insurance. The problem of too much capital chasing too few returns is the mirror image of the problem of undercapitalized insurance firms or firms that suffer unforeseen catastrophe losses. It is possible that the problem of overcapitalized firms is as bad as undercapitalized firms in that it may significantly reduce profits (Cummins and Nini 2002). Agency costs or increasing costs of capital could make higher levels of capitalization uneconomic.

1.3 Difficulties for regulators in observing reserves

Mostly, the concern is for undercapitalized insurers. Regulators must stop insurance firms from “abusing” limited liability by holding zero reserves (Munch and Smallwood 1981). Empirically, insurance commissioners may be motivated by career concerns, political pressure and other factors, in regulating premiums and reserves (Grace and Phillips 2008). Regulators are, at the very least, interested in maximizing firm solvency and minimizing the premiums charged by insurers for different insurance products. As a result, regulators in general have an interest in enforcing minimum solvency requirements on insurance, directly related to the level of reserves required by different lines of insurance. However, regulators may lack the proper incentives to monitor and penalize under-reserving that is unlikely to lead to claims default or firm ruin, especially because the goals of premium minimization and solvency maximization are in conflict.

Even in the absence of tension between the twin goals of regulators, their activity will be hampered by imperfect information. Regulators overseeing nongroup health insurance policies do have evidence on firm financial status and policy specific premiums and reserves that they can appeal to when setting minimum reserve requirements (Klein 2009). Premiums should go up by more than expected, and contract reserves should suffer relatively greater losses, in years when the expectation for future loss amounts is revised upward. In my model, these changes include increases in any of the parameters p_L, p_H, L, τ .

Observing a correct adjustment in reserves is difficult because of the ways that the loss parameters enter the reserve function. The direction of adjustments related to changes in each parameter is clear, but while some enter only as linear multipliers, others have different multiplicative effects for each year of the policy that must be summed up over the number of years remaining in the policy ¹. For example, an unexpectedly high growth in spending τ , as occurred in 2001, would linearly increase both prospective premiums and additional reserves. A deviation from the downward secular trend in p_H would increase the constant multiplier $(p_H - p_L)$ that appears in reserve equation 3.10 (see chapter 3), requiring additional reserves. However, the effect of p_H on premiums is compounded in p_L , because the derivative of the premium function with respect to p_H is $p_L L (1 + \sum_{j=0}^{N-1} (1 - p_L)^{j-1})$. So it will be much more difficult for a regulator to observe the correct adjustments to premiums with a change in p_H . The difficulties of monitoring adherence to the premium and reserving formulas could drive regulators to focus on premiums, which they can better observe.

A change in the probability of loss for low types is especially complicated as it results

¹The real world deviations from the stylized model are likely to be more complicated.

in several changes to premiums and reserves. First, the p_L term appears as a multiplier ($p_H - p_L$) outside the summation term in the reserve formula (for high types—low types always carry a zero reserve) and in the summation term itself. The extent to which p_L will effect the reserve through the multiplier and the summation depends on the number of years remaining in the guaranteed renewable contract, since the effect of the multiplier is constant while the summation term is larger for higher values of remaining years N . In addition, the reserve is based on the proportion of high types in the population, so to the extent that the shock in p_L also leads to more high types than expected there is an additional linear adjustment to the reserve. The same mathematics apply to the premium function, but in a more complicated manner, since the effect of an increase in p_L on absolute premiums is positive, but it makes guaranteed renewable premiums relatively less expensive.

Regulators have an information advantage in that they may have superior data about overall trends in losses. Medical spending growth may be independent of the particular insurer within a particular type of insurance (self-insured large group, small group, and nongroup), so the fact that regulators may need to blend data from multiple insurers should not hamper the process of calculating overall trend for the state insurance industry. Regulators therefore have strong grounds to challenge any individual insurer's choice of τ , and how it enters the reserve functions. For example, an unexpected increase in the rate of medical spending growth nationally is an increase in τ from the regulator's point of view, requiring additional reserves.

1.4 Difficulties for insurers in observing parameters

Insurers have data from internal sources when setting reserves. Insurers have access to claims data, with some lag for claims experience to flow through to observed losses (Bluhm 2007, p. 179). However, disentangling the cause of increased claims is important for correct reserving. Insurers have to distinguish between changes in quantities and prices, and the differential effect on low and high type individuals. Even with correct classification, insurers will still have to impute different types of observed changes in claims to different causes. An increase in price that equally affects the claims of both low and high types is an increase in τ . A change in quantity that equally affects the claims of both low and high types could be a simultaneous, equal increase in p_L, p_H or an increase in τ . An increase in quantity for only high types is likely an increase in p_H , but could conceivably be an increase in τ where the good is only consumed by high type individuals. Finally, the increase in τ may be continuous, while the claims arrival process is discrete, making measurement of trend more difficult (see chapter 6).

Insurers also have to blend experience from a particular pool of insured lives with other internal, and external, loss data to generate credible estimates of future losses. If experience for a given pool or state is not large enough to make accurate forecasts, then credibility is particularly important (Bluhm 2007, p. 265). Insurers can use both internal and external data sets for credibility. In the example of an increase in quantity for all types, if the increase is common across different insurance pools, it is more likely to be a change in τ than a change in probabilities. The insurer is not privy to private information regarding

τ or the rate of return on assets to counter the regulator's desire to see higher reserves. However, population variables such as the probability of a loss or the level of losses L experienced within a population is best observed by the insurer, and the regulator will have less credibility data to compare to the insurer's internal data. As a result, while the regulator could try to become more informed by requiring statements of actuarial opinion and rate justifications, the insurer always has scope to argue that its private information is superior for measuring variables other than trend. This is especially true for insurers that are larger or command a high market share in their area of operation.

2 Reserves and the problem of insurer follow through

2.1 Role of reserves in one-sided commitment

Reserves represent the potential cost of the one-sided commitment to the insurer. The contract reserves are the value of the promise, since they are calculated as the difference between the present value of future benefits and the present value of future premiums. It is not necessary to explicitly calculate the reserves in order to fulfill the commitment. It is possible, instead, to follow the rule of Cochrane (1995), and agree to "buy out" insureds who want to leave with payments that allow them to purchase insurance where the insured's portion of the premium is the "low type" rate in the future. Just because the cost of this buy out is the reserve, it is not necessary to hold this amount on the books of the firm, but merely to disburse the settlement amount when requested.

The reserve fund is crucial because of the information asymmetries found in insurance generally. “Contagion” theories of the transition of shocks from firms bearing a loss to those that are apparently well-capitalized hinge on the idea that policyholders and investors are imperfectly informed about the insurer’s financial position (Fields, Klein and Myskowski 1998). Even in annual renewal insurance, policyholders are concerned with the financial management of their insurer because of its crucial role as financial intermediary. Reserves demonstrate that the insurance company has correctly calculated the amount that it might need to pay off losses in different states of the world and that it will not be liquidity constrained. Not only does correct reserving flow into correct premium setting, but insurer’s outside option is often to exit a line of business entirely, so consumers may want to see a demonstration that the insurer is “in it for the long haul”.

Reserves are all the more crucial in one-sided commitments because insurers are not making substantial payoffs until relatively late in the contract. The big “payback” for insureds occurs when they become high loss types, which can take some time if the probability of loss (or, more generally, the probability of becoming high type) is low. The build-up, and draw down, of reserves facilitates insurance companies making a long term commitment to consumers based on the flow of funds over time. In the interim periods of the contract, proper reserving allows the firm to make moderate adjustments to recognize the gains and losses from shocks to loss amounts. The gains and losses arising from spending growth should be manageable on a year-to-year basis precisely because of the annual reserve reset process.

2.2 Reserves and the incentives to renege

The possibility that insurers will renege is one of the most important critiques of the guaranteed renewable model. In my setup, insurers are assumed not to renege but rather to act only as pass throughs for the guaranteed renewable contract. In addition, I do not allow for private information, so there is no way for an insurer to have a private signal of quality or “seriousness”. The extension to a private information situation is a natural one, and the main motivation for Phelan (1995), an early one-sided commitment model.

In a model with private information, the size of the reserve allows firms to send signals about quality and further bind themselves into their commitment. The reserve represents a floor, or minimum amount of benefits that the insurer promises to pay in the future subject to a sufficient amount of claims. The reserve mechanism then becomes an option that the insurer embeds into the guaranteed renewable insurance contract. If future claims exceed the amount of the reserve, the firm will pay out at least the amount of the reserves, and likely more, in benefits. If claims are less than the reserves, then the firm will adjust the reserve fund downward, and “refund” part of the built up reserves in lower premiums. This is not a true return of premiums—the adjustment comes in the form of truing up the future guaranteed renewable premium based on population experience—but it reduces premium payments by consumers in a similar way.

2.3 Reserves and ruin

Reserves cannot eliminate the possibility of firm ruin—there is always a stochastic element in insurance. For higher quality insurers, reserves act to minimize the losses on insurance products, though not eliminate them. For lower quality firms, proper reserving can affect the probability of ruin by encouraging capital adequacy and discouraging the firm from entering capital intensive (i.e. expensive) or volatile lines of insurance. Any capital held by the firm discourages bankruptcy, which delivers all tangible capital to creditors and wipes out intangible capital (Munch and Smallwood 1981).

The role of regulators is to examine the adequacy of provisioning for different insurance products as well as the totality of capital throughout the firm. For those firms taking too many risks, the regulatory consequences could involve strictures against new products and new customers, encouragement for more reinsurance contracts, or even, in the extreme, being wound down. In some cases, this will involve stopping insurers that are going concerns (in the sense that they are current on claims) because their reserves are inadequate, or not backed by ready money. In reality, it is more often the case that such firms, in addition to inadequate reserves, are also “a little behind” in paying claims, which could be an additional signal that the firm is getting close to the solvency/insolvency breakpoint.

3 Impact of spending growth on reserving

3.1 Size of spending growth

The primary impact of spending growth on reserves comes through the size of spending growth. Insurers will have to hold additional reserves determined by a sum of factors for each year of claims inflation. As an example is the three period guaranteed renewable premium in equation 4.12 (see chapter 4). The associated reserve for years 2 and 3 that is collected in year 1 is $[p_L L[(1 + \tau)^2(1 + (p_H - p_L))] + [(1 + \tau)^3(1 + (p_H - p_L) + (1 - p_L)(p_H - p_L))]$. The growth rates for each year come in exponentially. In addition, the trend rate τ may differ in each year.

Once the growth rates are imputed into the reserve formula, insurance companies should be holding reserves as assets. These reserves are quite large in the early years of the contract, exceeding the amount of first year premiums paid under the “duration matching” strategy (see chapter 4). If spending growth is perfectly predictable then reserves are just a time shifting loan, since the stream of premiums and benefit payments would be riskless. Insurers would only engage in financial intermediation to shift around the timing of the payments. Stable lapsation rates should not be a problem as long as they are predictable, since the effect is to lower required reserves through the channel of fewer expected high types in later years (there is no effect of low type lapsation because they carry zero reserves).

3.2 Variability of spending growth

The problem of forecasting spending growth complicates the problem of reserving. The reserves should be designed not only to handle the variability of claims due to variation in the population health of the insurance pool, but also the variability of claims due to the stochastic size of losses across time. Unlike the population variation, the size of loss problem cannot be solved through the law of large numbers, since it is a macroeconomic time series ². Instead, reserving policy must focus on the range of possibilities for losses, especially focused on the possibility of unexpectedly large losses.

One factor I observe in the time series that makes the problem easier is the persistence of spending growth. The prior year growth rate alone explains 50% of the variation in the current year growth rate using all historical data (see chapter 5). The effect of increasingly prediction power of spending growth is to reduce the fluctuations in the required reserves, and thus the need to recapitalize (or redeploy capital) on an annual basis. Additional macroeconomic variables are leading indicators for overall spending growth, and presumably there are more for spending growth within subpopulations. Alternatively, it could also be the case that for certain subpopulations, within group variation is greater than variation of aggregate spending growth.

²Even for long data series, there are many fewer data points.

3.3 Deterministic numerical example

Table 7.1 illustrates reserves for a three period model with no growth in medical spending. The loss L is \$10,000, the probability of loss for low and high types are 0.3 and 0.7, respectively. The reserve at time 0 is 0, because there are no high types. At time 1, after the payment of losses but before the payment of premiums, the average reserve per member is \$2,040, which is the premium per member at time 0 less the average loss per member at time 1 ($5040-3000$) (retrospectively) or expected future losses per member ($7200*0.7+14000*0.3$) less expected future premiums ($4200+3000$) (i.e. $9240-7200$) (prospectively). The entire reserve is allocated to high types, who have a reserve of 6800 ($2040/0.3$).

The prospective reserve for high types at time 2 after the payment of losses but before the payment of premiums is 4000 ($7000-3000$). The equivalent retrospective reserve is 2040, the sum of average premium per member at times 0 and 1 ($5040+4200$), less 7200, the sum of average losses per member at times 1 and 2 ($3000+4200$). The entire reserve is allocated to high types, who have a reserve of 4000 each ($2040/0.51$). At time 3, the reserve is back to zero both prospectively (no more premiums or losses) and retrospectively ($5040+4200+3000$)-($3000+4200+5040$)=0.

Table 7.2 illustrates reserves for a three period model with trend. The initial loss L is \$10,000, the probability of loss for low and high types are 0.3 and 0.7, respectively, and the trend rate is 8%. The reserve at time 0 is 0, because there are no high types. At time 1, after the payment of losses but before the payment of premiums, the average reserve per

member is \$2,458, which is the premium per member at time 0 less the average loss per member at time 1 (5698-3240). The entire reserve is allocated to high types, who have a reserve of \$8,193 each (2458/0.3).

At time 2, after the payment of losses but before the payment of premiums, the average reserve per member is the reserve at time 1 plus the premium per member less the average loss per member (2458+5011-(3499*0.7+8165*0.3)), or \$2,570. The entire reserve is allocated to high types, who have a reserve of \$5,039 each (2570/0.51). At time 3, the reserve after the payment of losses is zero (2570+3779-(3779*0.49+8818*0.51)). Note that at time 2, everyone pays the annual renewal premium for low types as their final guaranteed renewable premium. Also, as with the case of no trend, the reserve automatically returns to zero at time 3.

3.4 Stochastic numerical example

Now say the trend has a stochastic element. Instead of 8% trend, the trend rate is given by the results of table 5.13 (see chapter 5). The results of the raw summary statistics given in table 5.2 (see chapter 5) would imply that the upper end of the 95% confidence interval for premium growth is 14.9% because the raw moments do not include the autoregressivity in the data. Instead, the standard deviation in forecast is 2.83, which is in the middle of the range of forecast errors over the 1982-2008 period. The regression results imply an 8% forecast trend would have a 95% confidence interval of (8.00-1.96*2.83, 8.00+1.96*2.83) = (2.45%, 13.55%). In other words, the autoregressivity of spending growth shaves 1.35%

off the top end of the 95% confidence interval of forecast trend rates for the coming year.

If the firm had to set the reserve by getting the money from policyholders, it would have to impose an additional load of 10% on the guaranteed renewable premium with deterministic trend (which is already 75% higher than the annual renewal premium). If the load were still insufficient, the firm would have to recognize an additional actuarial loss. Conversely, it could book a load that was too high as a gain. Alternatively, if the insurance company could borrow the money at a fixed cost of capital which it then passed on to insureds, the load due to costs of guaranteed renewable contract reserves for trend would be a charge for the cost of capital. That would be a cost that insureds would not recover regardless of the true realization of the trend rate.

The other important inputs for reserves at time 0 are the two and three step ahead forecasts. However, two and three period lags of trend are not predictive of trend in the current period (although two period lagged growth has a significant effect in a model including one period lagged growth, see table 5.16 in chapter 5). Therefore, I assume that the forecast errors are equal to the random draw standard deviation in table 5.2 (see chapter 5), i.e. an expected mean trend of 8% and a standard deviation of 3.56. In other words, the advantage that the forecast model offers in reserving is for the one period ahead trend rate. As I have noted, the reserve in periods two and three is based on a time zero forecast of trend in those periods. So the expected value of L at times one, two, and three condition on being at time zero are 10800, 11664, and 12597. The 95% confidence intervals are (10245,11355), (9818,13669), (9258,19289).

As a result of the forecast error for trend, the insurance firm will have to use both claim reserves and contract reserves. The claim reserves will be for both low and high types, and reflect the possibility of short term losses i.e. those losses that are common to annual renewal and guaranteed renewable insurance. For example, the claim reserve for low type annual renewal insurance written at period 1 is 406, or 13.55% of 3000.

The firm will also have to set additional contract reserves i.e. reserves arising from the long term liabilities under the contract that are exacerbated by the stochastic trend rate. In addition, the ex post correct guaranteed renewable premium when the trend rate is at the upper end of the range and then returns to average is 5990. So the additional contract reserve is 596 to get to the upper end of the 95% confidence interval, which is 10.4% of the original guaranteed renewable trend.

3.5 Effect of less predictability

If the time series on trend were not autoregressive, the standard deviation of the trend rate would much higher at 3.56. The effect on claims reserves is to make them 1.2% (43) higher. In addition, contract reserves are 75 higher, for a net increase of 0.5% (32). The magnitudes of the additional reserves required if the size of spending growth is just a random draw for a known prior distribution is not significantly more than the case where the spending growth is autoregressive. However, the reason may be that the summary statistics show a statistical distribution that is not very volatile.

In essence, the great degree of predictability in the AR(1) model is not that much higher

than the summary statistics of the time series because the higher moments (especially skew and kurtosis) are so small. It is the stochastic element to the trend that drives the bulk of the additional contract reserves required to support guaranteed renewable insurance, and claims reserves for all forms of health insurance. If, for example, the spending growth rate were AR(1) when looked at as a time series but had a very high skew or kurtosis when looked at as a series of draws from a random distribution, then the time series forecasts would be much more important for the one year look ahead. For guaranteed renewability with an N well above 2, then, the most important statistics could well be the unconditional draws from the summary statistics rather than the one period ahead forecast.

4 Solutions derived from problems of reserving

4.1 Limits to reserves

The numerical examples demonstrate that the limit on reserves is the acceptable probability of nonpayment of claims. Claims are in theory infinite, so it is possible to justify any level of reserves for a given level of risk aversion. Even for limited benefit plans, the size of the fluctuation in spending growth rates above the trend line is unlimited, so the acceptable probability of nonpayment is the limiting factor on the level of reserves. The probability may be chosen by a principal (i.e. the consumer) who is willing to accept a certain level of risk, or an agent who chooses the level of risk for the consumer. If, as is likely the case, consumers are not explicitly choosing the riskiness of their insurance company, then

the level of reserves is either selected by a regulator who must balance trade-offs between low premiums, low level of insurer failure, and high degree of access to various insurance products for consumers, or by insurers maximizing the value of intangible capital. The unsatisfying element of any of the three decision makers is that the result can easily be an infinite demand for reserves.

The volatility in premiums that include loading for reserves can provide a finite ceiling for reserves independent of the decision maker. If holding reserves is costly and the cost of reserves is proportional to the level of reserves chosen ³, then at some level of reserves the dominance of guaranteed renewability over other forms of insurance will not hold. The dominance results in chapter 3 are explicitly based on the timing and amount of funds required by the insurer at each point in time. The intertemporal volatility of the policy would be increased by taking in a large reserve at time 0, even if it is largely refunded at a later time. The result is the same as a loading factor λ that is higher for guaranteed renewable policies than it is for annual renewal insurance: for individuals that are not sufficiently risk averse, the lower loaded policy with less protection may be more attractive than the higher loaded policy with more protection. An insurance company that borrows the reserves directly from policyholders and then refunds them if they are unneeded won't be able to sell this policy.

³It is likely that the cost of reserves is eventually increasing in the size of the reserves due to asymmetric taxation and agency costs.

4.2 Possible policies

The place of reserves in writing guaranteed renewable policies points to several solutions. Additional asset classes for hedging are useful in setting reserve fund investment policy if the insurer's problem of generating high enough reserves is costly capital. Long term guaranteed renewable policies (and noncancellable policies) may be unattractive in terms of the required premium. The affordability of reserves to the insurer is a greater problem than affordability of the premium to customers, since insurers should make investments equal to the total value of the contract liability when the contract is first written. The solution of partial guarantees addresses the problem of maintaining a large reserve as much as it addresses the problem of unaffordable premiums (see chapter 8).

The problem of reserves for guaranteed renewability is also generalizable to other forms of long term health insurance. The difficulty of generating and maintaining adequate reserves due to asymmetric taxation and other transactions costs will apply equally to other forms of long term insurance. The problem is less acute for guaranteed renewable plans because they are sold by multiproduct firms that can raise capital for many insurance products at once (the multiproduct firm may still be supporting uneconomic products through cross-subsidies). The same is not true for long term care or retiree medical insurance (although employers do access the capital markets for business operations, so they in theory have access to multiple sources of capital). On the other hand, the problem of reserving common to guaranteed renewable insurance and other forms of long term insurance in a multiproduct firm, such as property and casualty generalist firms, is that the existence of a reserve can

lead to inappropriate cross subsidies, wiping out the gains built up for a particular pool of guaranteed renewable insured lives.

I also speculate in the conclusion (chapter 8) about future trends that may impact the reserving for health insurance products. The main trend, that I have already mentioned, is the PPACA, which will change the entire operating environment for health insurance. Insurers may be particularly eager to blame economy-wide changes for changes in claims particular to a policy in the future, while state insurance regulators may be particularly vigilant about health insurance for the near future. These trends may shift the focus even more to premiums and solvency, and health insurance firms may have a more difficult time justifying reserves that they consider adequate to fund guaranteed renewable health insurance.

5 Conclusion

5.1 Reserves and overall uncertainty

One problem that reserves cannot solve is the global problem of parameter uncertainty in healthcare spending. The guaranteed renewable model I use, as well as the prior literature, relies on fairly stable parameters for the loss. The possibility of a regime shift in parameters underlying losses is what makes longer term guaranteed renewable contracts more expensive, and possibly prohibitive (as well as longer term contracts for noncancellable life insurance and long term care insurance (Cukierman 1980)). The level of reserves re-

quired to deal with any shift in the loss function, such as a large one-time upward shift in the level of healthcare spending, may be large. It is important to distinguish this problem from the problem of making very long term one-sided commitments under stable parameters, which I have shown to be equivalent to annual renewal insurance at the “high type” rate (see proposition 7).

The uncertainty problem is also different than the one faced by retiree health insurance, where plans simply did not calculate reserves that reflected secular trends in spending growth. Rather, the uncertainty problem is one where the probability of shifting from low to high type has nonlinearities after five or ten years, long enough that it could not be observed in the data used by any single insurer. Similar risks are those posed by large shifts in policy, like the passage of the PPACA, that then lead to legal uncertainty over the interpretation of the ongoing commitment. The best way to deal with the uncertainty problem is to limit the term of the commitment, rather than try to set aside larger and larger reserves (see conclusion).

5.2 Pricing without correlated assets as a benchmark

The value of correlated assets is not only in the ability to hedge risk, but also to price risk. A correlated asset could function as a reference price when valuing the liability under the policy. Correlated risks also allow the insurer to assess the riskiness of a line of insurance relative to other opportunities, especially the investor’s opportunity to invest their capital elsewhere. The idiosyncrasy of the risk of spending growth, and the guaranteed renewable

insurance liability generally, makes the guaranteed renewable insurer more specialized.

Unlike the case of people with idiosyncratic labor market shocks (Heaton and Lucas 1996), the insurer's problem is not one of private information. Rather, it is a problem of the pool of available capital, which may be limited if specialized investors are required. The search costs for specialized investors is a kind of transaction cost imposed on top of the rate of return demanded by such investors. However, since guaranteed renewable insurance (and retiree medical plans) are often insurance embedded in a larger firm, distinguishing the reserve requirements of the guaranteed renewable insurance line alone is difficult.

5.3 Other impediments and costs to reserving

As I have mentioned, there are several impediments to the reserving strategy I outline even if loss variables are completely predictable. The first is the issue of transactions costs associated with investment of reserves or borrowing enough to have adequate reserves in early years of the contract. Transactions costs include the fixed and proportional costs of investing, and the process of internally allocating an investment pool of all reserves to different pools of insured lives. Unless the firm is large enough to move the markets with its trading, transactions costs are part of the cost of capital.

The second issue is one of asymmetric taxation and suboptimal regulation of loss ratios. I combine these two problems because they have similar causes and similar effects. The problem of asymmetric taxation is the differential treatment of gains and losses on reserve investments. The problem of suboptimal regulation of loss ratios is the tendency

of regulators to prefer higher loss ratios to lower loss ratios without regard to the pattern of loss ratios over the life of the policy⁴. Asymmetric taxation cannot be solved through prospective premiums, since the premium setting would require a perfect forecast of asset markets—firms would need to know when markets will rise and fall in order to correctly “time” premiums for tax minimization. Company size may help, as a larger company may be able to find offsetting tax losses, but diversifications across types of insureds or lines of insurance does not help, because while the reserve is policy specific, the pool of capital is common i.e. there is no trust containing policy specific assets⁵. The loss ratio problem could be partially ameliorated with better accounting measures, but this does not address the regulator problem, which is caused by popular and political pressure on regulators. In a way, size or diversification could make the loss ratio problem worse, as it could increase the chance that some pools or lines are inappropriately classified as having loss ratios that are too low.

The signaling costs and agency costs associated with increasing reserves are a cost of reserves that become more acute as insurers wish to take on higher levels of reserves. While a healthy level of reserves may indicate an awareness of the financial obligations of the firm, write-downs or additional reserves for unanticipated losses can send a signal of firm weakness (Fenn and Cole 1994). Conversely, the separation between firm ownership

⁴At some point, regulators will also disapprove of loss ratios above 100% on solvency grounds, but low loss ratios generate substantially more scrutiny.

⁵Also, the increasing correlation across asset classes means that there will be fewer years when some asset classes lose while others gain.

and management means that managers may have the incentive to set the reserves of the firm at too low of a level since they do not lose as “enough” from firm ruin (Munch and Smallwood 1981).

5.4 Risk neutral discounting

As a result of fluctuations, the consumer with guaranteed renewable insurance, or any long term policy where the insurer takes on or shares the risk of spending growth, has a valuable option embedded in their policy. The insurer will likely not realize that the parameters or distribution of spending growth have changed, and will have to incur large actuarial losses (or gains) as it incrementally increases (or decreases) reserves in several “steps” over multiple years. The insurer could explicitly recognize the uncertainty in future growth rates through a risk neutral discount rate methodology, but the use of this method is no guarantee that the insurer will set the correct reserve level.

The implications of a risk neutral method for reserves is to discount future premiums at a higher rate than future benefits. Unexpectedly high claims may lead to less than expected lapsation rates, so there is a correlation between the deviation of losses from expectations and the required level of reserves. Discounting in this way preserves the intuition that, under a one-sided commitment, the insurer is more certain that they will pay more in claims than expected than they reverse. In addition, benefits will adjust more quickly to the new level of spending growth (perhaps instantaneously), while insurers are only able to adjust prospective premiums on a lagged basis. Delays in collecting data of sufficient credibility to

allow for rate increases will only aggravate the mismatch between premiums and benefits, and require greater reserve adjustment⁶.

One major difficulty associated with increasing reserves at the end of the year is not the “ordinary” years, when additional capital can be had at a “normal” price, but “extraordinary” years when the cost of capital for all firms is high due to an economy-wide shock. The lack of correlation between the variability in spending growth and asset returns means that the insurance firm is not subject to a correlated risk problem. In other words, if the firm faces a year of well above average spending growth, that is not likely to also be a year in which the cost of additional capital is abnormal.

The major caveat to the prediction of spending growth is the paucity of data. Reducing the sample from 1970-2008 to 1982-2008 reduces the R^2 from 0.5 to 0.4. In addition, the period 1982-2008 was a period of particularly low inflation (and, in general, little business cycle fluctuation (Bernanke 2004)), so it is unclear how relevant the recent experience is for forecasting the future. For medical care, this is all the more so true given the recent large change to the laws regulating and financing care, the Patient Protection and Affordable Care Act (U.S. Congress 2010). The tendency could be for the size of fluctuations to increase, or decrease, in the future, changing the structure of volatility. There is also the potential for a structural break in the entire time series of spending growth, with an attendant change in the level or average growth rate of spending as shown by Finkelstein (2007) for the case of the enactment of Medicare. A sufficient number of errors could even restrict the insurer’s

⁶A recent presentation about long term care insurance noted that it takes five years or more for credible data on the true cost of a block of lives (Rauch 2002).

ability to take on additional capital as investors become leery of the insurer's inability to correctly forecast growth rates. In the concluding chapter (chapter 8), I discuss possible solutions to this problem.

6 Tables

Time	0	1	2	3
Annual Renewal Premiums				
Low	3000	3000	3000	
High	7000	7000	7000	
GR premiums				
from 0	5040	4200	3000	
Total Expected claims w trend from 0				
Low	12240	7200	3000	
High	21000	14000	7000	
Total Expected claims w trend from 1				
Low		12240	7200	
High		21000	14000	
Proportion of types				
Low	1	0.7	0.49	0.343
High	0	0.3	0.51	0.657
Reserve per member				
Low	0	0	0	0
High	N/A	6800	4000	0
Total	0	2040	2040	0

Table 7.1: Three period guaranteed renewable reserve–no trend

Time	0	1	2	3
Annual Renewal Premiums				
Low	3240	3499	3779	
High	7560	8165	8818	
GR premiums				
from 0	5698	5011	3779	
Total Expected claims w trend from 0				
Low	14488	8790	3779	
High	24543	16983	8818	
Total Expected claims w trend from 1				
Low		15647	9493	
High		26506	18341	
Proportion of types				
Low	1	0.7	0.49	0.343
High	0	0.3	0.51	0.657
Reserve per member				
Low	0	0	0	0
High	N/A	8193	5039	0
Total	0	2458	2570	0

Table 7.2: Three period guaranteed renewable reserve with trend

Chapter 8

Conclusion

1 Introduction

Guaranteed renewable health insurance benefits consumers by allowing them to even out premium payments over time. The more even premium payments not only facilitate consumption smoothing, but may solve the problem that, after suffering a loss, annual renewal health insurance premiums will skyrocket. To take advantage of guaranteed renewability, the consumer must be able to confidently shift his risk onto an insurance company. That company can then spread the risk of individual losses over a group, transforming an idiosyncratic loss into a predictable loss distribution, thereby reducing the overall level of risk.

The ability of the insurance company to manage the risk of uncertain future claims and provide class average rates is reliant on the degree to which pooling individuals leads to a reduction in the level of risk. Time limits on the term of protection are a theoretically

valid way to reduce the risk, but may come at the cost of providing less protection than consumers desire. Holding reserves is another way to guarantee benefits, but reserves may be expensive, and there are limits on the willingness of companies to hold higher levels of reserves.

There are also limits to the amount of risk that insurers can manage through securities markets. Products such as health insurance futures could be used by many different types of firms, but in reality seem not to be viable. The investment returns to reserves are themselves limited by the available classes of assets, which may not support the needs of insurers in reserving for guaranteed renewable insurance. To the extent that the risk is hard or impossible for insurers to manage, alternative policies and products may facilitate some degree of guaranteed renewability for risk averse consumers. I briefly consider some of these policies in the next section.

2 Policies for improving guaranteed renewability

2.1 Making additional asset classes

The ability to hedge medical spending risk is determined by the extent of market completeness. If the variance in trend is an important risk that is not idiosyncratic, some mix of assets should exist to hedge the risk. Protection need not be perfect or perpetual, but should at least be available at some cost. If a hedge is not available, or the attendant transactions costs are high relative to the value of hedging the risk, then the risk is one that cannot be

diversified through the financial market. It is also possible that an asset, or combination of assets, that I did not consider is a good hedge for medical spending growth. It is also possible that the limited amount of data for medical spending is itself an impediment to the discovery of such assets.

If there is no hedging assets with reasonable cost, there could be some willingness by reinsurers or hedge funds to take on spending growth risks. The size and expertise of reinsurers would determine the extent of the market opportunity. If the variance in medical spending is substantially or totally uncorrelated with other assets, then underwriting this risk could generate an opportunity for hedge funds and other alternative investment vehicles because of the number of people and firms that face risks related to medical spending growth. Reasons for not writing this risk would be the unwillingness of investors to stake adequate capital for medical spending reinsurance, or sheer unpredictability of medical spending (or the difficulty of prediction over more than a short time horizon). The risk could be systematic, meaning that it can be shifted to some extent across individuals and firms but not reduced. The limited use of reinsurance by capitated providers suggests that there may not be a large market for any form of health reinsurance (Simon and Emmons 1997).

Government intervention could take the form of specialized TIPS bonds. TIPS has allowed investors to hedge inflation by indexing the bond's principal to the CPI-U index, which is partially predictive of medical spending growth. Jennings (2006) makes a case for bonds indexed only to the Medical Care portion of the CPI due to the fact that the Medical

Care time series is also hard to hedge with assets, and many liabilities are linked to medical prices rather than the overall price level. There are several problems with accurately measuring the price of medical care, leading to skepticism of the usefulness of the CPI-U, Medical Care series. Guaranteed renewable liabilities in health insurance are generally linked to total spending rather than prices, so medical inflation linked bonds would not protect against changes in quantity of care consumed. Health policy may already reflect this reality, with a new government financed long term care health insurance program in the Patient Protection and Affordable Care Act.

2.2 Partial guarantees may be possible

One solution to the cost of guarantees against reunderwriting would be partial protection. Partial protection could include term limited guaranteed renewability, and proposals for more indemnity contract type insurance. The current trends in health policy mostly run in the other direction: elimination of insurance with lifetime limits, ending the “donut hole” exposure to drug costs in Medicare Part D, and new limits on underwriting criteria for health insurance. As a result, the prospects for any kind of explicitly partial guarantee are dim.

Partial health insurance already exists as a way to finance limited amounts of health insurance. The best example is Medigap (Medicare Supplemental) insurance plans that pay for the coinsurance portion of Medicare. The success of Medigap plans is in the limitations on benefits, since Medicare is paying the majority of the premium, and Medicare premiums

are less than the amount needed to pay expected claims. Another popular, but limited, insurance plan type is the “mini-med” plan (Adamy 2010). The plans, which may offer only a few thousand dollars in total coverage are, as a result, very cheap. The best examples of a partial guarantee are high deductible and catastrophic only health plans. Guaranteed renewable plans limited to only the largest claims could be more manageable, and would provide protection against the worst episodes of illness.

2.3 More transparent accounting could help

More transparent accounting for the sources of guaranteed renewable premium increases could help all parties to insurance contracts. If insurers committed to use the increase in risk from a third party data source (such as the MarketScan database) then insured individuals could have additional confidence that premium increases were not capricious. Such data sources could also provide the credibility that insurers need to price their products, but in such a way that the business secrets inherent in their plans are not revealed. The same is true of the way that insurers account for the medical loss ratio and other accounting variables surrounding guaranteed renewable insurance. By breaking out the statistics by plan or class type, insurers could show the variance of experience from expectation, and give insureds and regulators greater confidence that the premium changes are reasonable and necessary. Additionally, more such disclosures will be required to comply with the PPACA.

3 Future research

3.1 Applications to other insurance programs

Government programs differ from private plans in that they either do not build up reserves, or cannot invest their reserves ¹. Medicare is a long term health insurance program with well known positives (less chance for use of asymmetric information) and negatives (large fiscal imbalances). New programs, such as the CLASS Act, are attempts to increase the take-up of long term care insurance through a voluntary government program. I plan to evaluate the ability of such programs to adequately fund long term care benefits given the restriction on investments imposed on these plans. Reserving for such programs is in some ways easier, because the programs can take a long term view of the funded status of the plan. Reserving is also more difficult, because the program reserve can be seen as part of the overall government balance sheet.

Another policy I plan to evaluate in future work is the prospects for retiree medical plans. GM, for example, first sold bonds to try to prefund its VEBA plan, and then gave the plan a substantial ownership interest in the firm as part of bankruptcy proceedings. A majority interest in a privately owned automobile company may or may not be the best investment for funding retiree medical programs (if anything, the retiree health plan should probably have a short position in the employer's stock). I plan to study how bringing the firm public, and allowing shares to be sold, could improve the asset/liability management

¹Sovereign wealth funds, maintained by other countries and intended to pay for social programs, are one form such investment could take.

of the GM VEBA plan. In addition, I plan to consider ways that retiree medical plans with a bankruptcy option form a kind of partial guarantee, where the insurance is conditional on the health of the insurance company.

3.2 Patient Protection and Affordable Care Act

The Patient Protection and Affordable Care Act (PPACA) will decrease the use of some long term health arrangements, while increasing others. My future research will focus on the relevant sections of the law, which include the portions affecting individual insurance, the portions affecting retiree medical care, and the portions affecting long term care insurance. I also plan to address the possibility that the law will change the overall trend for medical spending growth.

Many portions of the PPACA concerned with the individual insurance market could have the effect of reducing the availability of guaranteed renewable health insurance. The law implements rate bands for premiums, limits lifetime maximums on most plans, eliminates most pre-existing conditions exclusions along with restrictions on the use of other rating factors, and imposes floors on the medical loss ratios of insurance plans (U.S. Congress 2010). All of these policies can make guaranteed renewable plans more difficult, although to different degrees, with some evidence from state level implementation of similar policies.

The elimination of pre-existing conditions exclusions and the use of all available rating factors are the biggest impediments to guaranteed renewability. Guaranteed renewable

models are based on a homogeneous risk pool. Pre-existing conditions exclusions and rating factors both allow the insurer to set up a class of similar insureds, though in slightly different ways. Rating factors are proxies for the true statistic of interest, the expected future claims of an individual. If insurers can rate based on experience, then demographic variables may be unnecessary (although homogeneity on non-medical characteristics could still lead to lower marketing and administrative costs). Pre-existing conditions exclusions more directly address the guaranteed renewable setup where individuals start as low types and then transition at some point to becoming high types.

A more realistic model of guaranteed renewability would include the fact that people are likely to switch back and forth between the high and low types, lowering the long term cost of premiums but not the problem that insurers want to create a homogeneous pool of lives. Prior studies of community rating have found two effects on the structure of state insurance markets: fewer policies written, and higher average premium rates. There were also lower effects on the relative differences in premiums for different groups across states (Buchmueller and DiNardo 2002, Pauly and Herring 2007).

Rate bands and lifetime maximums on coverage may have less predictable effects on guaranteed renewability. The implementation may interrupt existing guaranteed renewable coverage, since insurers made guarantees based on a policy form that may have included a lifetime maximum, and was predicated on the ability to charge very different rates to different groups of insureds.

For new policies, the effect of rate bands and removal of lifetime maximum benefits

is less clear. Insurance without a lifetime maximum could be more attractive to those who could afford it, as consumers may prefer an unlimited amount of protection in the state of the world where they become a high type (as noted in chapter 7, the unlimited level of protection may interfere with reserving). Rate bands could compress insurance premiums, thus making premium setting more difficult generally. They could conceivably have the result of increasing guaranteed renewability, because one way to achieve a set ratio between the premiums of older and younger policyholders is to write higher value, guaranteed renewable insurance for the young, while writing annual renewal or shorter term guaranteed policies to those near 65. Rate bands could also discourage duration rating, which could improve adherence to guaranteed renewability but could also make insurers more wary about adverse retention.

The effect of medical loss ratios on guaranteed renewable insurance is also likely to be muted. Many states already have minimum medical loss ratio laws (AHIP 2010). More importantly, contract reserves are specifically included in the numerator of the medical loss ratio for the purposes of complying with the minimum standard (Jost 2010). Contract reserves could become an important way to comply with medical loss ratio rules, which could correct for the under-reserving as a cause of duration rating. There is some flexibility around implementation of the rule in general, so its effects may be blunted in some states (Jost 2010).

The act also establishes a new long term care insurance plan, the Community Living Assistance Services and Support (CLASS) Act. The existence of a government long term

care plan may go along with the trend of the decline in private, guaranteed renewable, long term care insurance. The continuation of previously insured customers as private companies “run out” existing lines of business may subject those policyholders to the “lock in” problem (see the literature review, section 1), because they cannot threaten exit as a way to discipline their insurers (Schultz 2010).

Consumers in the new government program may be few in number initially because they can only participate if their employers choose to offer the plan. The absence of underwriting and guaranteed issue basis for the plan could lead to a high risk profile within the small number of participants (Foster 2009). The effect of the CLASS Act itself on guaranteed renewable long term care insurance could also be small if the predicted adverse selection (Schmitz 2009, e.g.) takes place and federal subsidies are small, keeping healthier insureds in the private market.

4 Final remarks

4.1 Spending growth path may change in the future

One potential problem in the adequate funding of spending growth is the possibility of a change in underlying trends. Overall, trend has slowly been declining, but since the trend rate is still well above the growth in GDP, it is still unsustainable in the long term. The forecasts of medium to long term trend rates are especially difficult given recent legislative action that will greatly change the health care system. It is unclear when consumers will

reach a “resistance point” where they simply refuse to spend a growing portion of their budget (or, via voting, the government’s) on medical care (Getzen 2007). If medical care changes from being a superior good to being a normal good, then insurers (whether public or private) will be in a better position to implement long term premium guarantees. However, it is difficult to forecast the kind of structural changes that would lead to such a radical shift in thinking about the value of medical care.

4.2 Variance could be a problem if leverage is high

The financial distress problem stemming from spending variance is a problem of finding or creating the right assets. Better hedges could help, but it is hard to see what the best hedging assets would be. An asset whose value is linked to the level of spending for homogeneous pool of lives would help. Similar markets, such as swaps tied to specific risks, are not widely traded (Gullipalli 2007). Bonds that are tied to the level of medical care prices could also help, especially if they were issued by the U.S. government and thus low on default risk. However, medical prices are likely only a component of total spending growth, and the CPI-U medical care series is not well suited for many classes of individuals. A related problem is the scrutiny of “too low” medical loss ratios, which can also inappropriately reduce contract reserves.

4.3 Implications for other types of insurance

Long term protection in health insurance is an important goal. The future is uncertain, so a limited promise can be more valuable than a broad one if the long term risk from medical spending is a systematic risk. I focus of the cost of premium protection, and policies which could improve the strength of premium protection and the adequacy of reserves. By managing the liabilities from long term plans in a more consistent way, insurers can ensure the lowest funding costs and lowest variance in the shortfall (or surplus).

The ideas I present do not span the total set of possibilities for dealing with spending growth in long term products. Spending growth is not limited to long term plans, but premium guarantees are more sensitive to the trend and variance of medical spending growth than annual renewal insurance because of the effect on reserves. Insurers should plan ahead through the use of contract reserves and limits on insurance for future events. Regulators have a role in monitoring reserves as well as recognizing regulations and tax effects that make reserves suboptimally low. The goal of all parties should be the best possible premium guarantee that insurers can offer to provide the greatest level of protection against one of the biggest risks consumers face.

Appendix A

Theory appendix

1 Proofs

Proof of proposition 1:

Proof. The proof is by induction.

Step 1. The measure of low and high types at time 0 are 1 and 0, respectively. Therefore, the measure of individuals at this time is not driven by either probability distribution.

Step 2. Assume that at time t , the measure of types is μ_L^t and μ_H^t for low and high types, respectively. Assume that $\mu_L^t = f(p_L, t)$, $\mu_H^t = g(p_L, t)$.

At time $t+1$, the measure of types is $\mu_L^{t+1} = \mu_L^t(1-p_L)$ and $\mu_H^{t+1} = \mu_H^t + \mu_L^t p_L$. Therefore, at time $t+1$, the measure of types is driven by the probability distribution for low types only.

Therefore, the proposition is proven by induction. □

Proof of proposition 2:

Proof.

$$\begin{aligned} |p_L - 0.5| < |p_H - 0.5| &\Leftrightarrow \\ (p_L - 0.5)^2 < (p_H - 0.5)^2 &\Leftrightarrow \\ p_L^2 - p_L - 0.25 < p_H^2 - p_H - 0.25 &\Leftrightarrow \\ p_L(p_L - 1) < p_H(p_H - 1) &\Leftrightarrow \\ p_L(1 - p_L)L^2 > p_H(1 - p_H)L^2 &\Leftrightarrow \\ \text{Var}[Z_L(0)] > \text{Var}[Z_H(0)] \end{aligned}$$

□

This proof for immediate insurance extends to deferred insurance.

$$\lim \text{Var}[Z_L(k)] = \text{Var}[Z_H(k)] = \text{Var}[Z_H(0)]$$

as shown in the proof of Proposition 7.

Proof of proposition 5

Proof. Assume that a single premium is to be offered to both types.

Assume that trivial contracts are not allowed. A trivial contract is one that involves payment with certainty or overinsurance.

First, note that $G_L < G_H$. Therefore, high types certainly prefer this contract to uninsurance.

The low types under an insurance contract costing some amount A face two lotteries:

$(L, p_L; 0, 1 - p_L)$ and $(A, 1)$. In order for the latter contract to first order stochastically dominate the former, $A \leq p_L L$ (mean preserving spread).

$A < p_L L$ violates proposition 3 and actuarial fairness. Therefore, $A = p_L = G_L$. \square

Proof of proposition 7

Proof of Proposition. The proof for high types is trivial:

$${}_k G_H = G_H = p_H L$$

The proof for low types involves an infinite geometric sum:

$$\begin{aligned} {}_k G_L &= [p_L L] [1 + (p_H - p_L) + (1 - p_L)(p_H - p_L) + (1 - p_L)^2(p_H - p_L) + \\ &\quad \dots + (1 - p_L)^{N-k-1}(p_H - p_L)] \\ &= [p_L L] [1 + \sum_{j=0}^{N-k-1} (p_H - p_L)(1 - p_L)^j] \end{aligned}$$

Taking the limit as k tends toward infinity:

$$\begin{aligned} \lim_{k \rightarrow \infty} {}_k G_L &= \lim_{N \rightarrow \infty} [p_L L] [1 + \sum_{j=0}^{N-k-1} (p_H - p_L)(1 - p_L)^j] \\ &= [p_L L] [\lim_{N \rightarrow \infty} 1 + \sum_{j=0}^{N-k-1} (p_H - p_L)(1 - p_L)^j] \\ &= [p_L L] [1 + \lim_{N \rightarrow \infty} (p_H - p_L) \sum_{j=0}^{N-k-1} (1 - p_L)^j] \\ &= [p_L L] [1 + (p_H - p_L) \frac{1}{p_L}] \\ &= [p_L L] [1 + \frac{p_H}{p_L} - 1] \\ &= [p_L L] [\frac{p_H}{p_L}] \\ &= p_H L \end{aligned}$$

□

2 Maximization problems

Solution to the minimization problem in chapter 4, section 5.1:

$$\min_{\{w_u, w_d\}} \sum_{i=u,d} \left(\frac{w_i}{P_i} - iL_0 \right)^2$$

s.t. $w_u + w_d = W_0$

There are only two possibilities, so I can write out the full objective function:

$$\min_{\{w_u, w_d\}} \pi \left(\frac{w_u}{\pi} - uL_0 \right)^2 + (1 - \pi) \left(\frac{w_d}{1 - \pi} - dL_0 \right)^2$$

Substituting in the constraint expressed as a function of investment in the up claim,

$w_u = W_0 - w_d$, so that my solution is an amount invested in the up claim:

$$\min_{\{w_u, w_d\}} \pi \left(\frac{w_u}{\pi} - uL_0 \right)^2 + (1 - \pi) \left(\frac{W_0 - w_u}{1 - \pi} - dL_0 \right)^2$$

Now I maximize by taking the first order condition:

$$\frac{\partial}{\partial w_u} : 2 \left(\frac{w_u}{\pi} - uL_0 \right) - 2 \left(\frac{W_0 - w_u}{1 - \pi} - dL_0 \right) = 0$$

Then I solve for w_u :

$$\frac{w_u}{\pi} - uL_0 = \frac{W_0 - w_u}{1 - \pi} - dL_0$$

$$\frac{w_u - \pi uL_0}{\pi} = \frac{W_0 - w_u - (1 - \pi)dL_0}{1 - \pi}$$

$$(1 - \pi)w_u - \pi(1 - \pi)uL_0 = \pi W_0 - \pi w_u - \pi(1 - \pi)dL_0$$

$$w_u = \pi W_0 + (1 - \pi)(R - d)L_0$$

2.1 State contingent claims—zero losses

I show that the one period solution is optimal in a single period setting.

The investment policy for a spot contract in a complete securities market is given by the shortfall at time 1 from the two possible states of nature. In the absence of an investment policy, under fair premiums, the shortfall is:

$$p_L LR(u - 1) \quad \text{under an up move}$$

$$p_L LR(d - 1) \quad \text{under a down move}$$

The down move produces a surplus for the company under a cash investment policy. The cost at time 0 of the \$1 state contingent claims at time 1 is:

$$\pi \quad \text{under an up move}$$

$$(1 - \pi) \quad \text{under a down move}$$

The company invests the following amounts in the two assets:

$$\pi p_L L R + (1 - \pi)(R - d)p_L L \quad \text{up move claim}$$

$$(1 - \pi)p_L L R - (1 - \pi)(R - d)p_L L \quad \text{down move claim}$$

This is equivalent to the number of shares in the two assets:

$$\pi p_L u L \quad \text{up move claim}$$

$$(1 - \pi)p_L d L \quad \text{down move claim}$$

so there is 0 shortfall in all states at time 1.

2.2 Risky stock–single period solution

First note that:

$$1 - \pi = \frac{u - R}{u - d}$$

$$\pi(u - R) + (1 - \pi)(d - R) = 0$$

The first order condition of

$$\min_{\{w_S, w_B\}} \sum_{i=u,d} P_i (i w_S + R w_B - i L_0)^2$$

$$\text{s.t. } w_S + w_B = W_0$$

is

$$\frac{\partial}{\partial w_S} : 2\pi(u - R)(u w_S + R(W_0 - w_S) - u L_0) + 2(1 - \pi)(d - R)(d w_S + R(W_0 - w_S) - d L_0) = 0$$

The $R(W_0 - w_s)$ in the two parts of the equation drop out because of the relationship from above, leaving:

$$\pi(u - R)(uw_S - uL_0) + (1 - \pi)(d - R)(dw_S - dL_0) = 0$$

Using $\pi(u - R) = -(1 - \pi)(d - R)$ in the second step and $u \neq d$ in the fourth step:

$$\pi(u - R)(uw_S - uL_0) = -(1 - \pi)(d - R)(dw_S - dL_0)$$

$$(uw_S - uL_0) = (dw_S - dL_0)$$

$$(u - d)w_S = (u - d)L_0$$

$$w_S = L_0$$

In addition:

$$\mathbb{E}[X^2] = \pi u^2 + (1 - \pi)^2 d^2 = R(u + d) - ud$$

$$\text{Var}[X] = R(u + d) - ud - R^2 = R(u + d - R) - ud$$

3 Infinite period premium

For an infinite period premium in a model with discounting, the following first derivatives show how the premium responds to changes in parameters.

Denote this premium as the premium in an infinite horizon model, P_∞ . Now the derivatives and limits of this premium are:

$$\begin{aligned} \frac{\partial P_\infty}{\partial p_H} &= \beta[p_L L] \frac{\beta}{1 - \beta(1 - p_L)} \\ \frac{\partial P_\infty}{\partial p_L} &= \frac{\beta L(1 - \beta(1 - p_H))}{1 - \beta(1 - p_L)} \left[1 - \frac{\beta p_L}{1 - \beta(1 - p_L)} \right] \\ \frac{\partial P_\infty}{\partial \beta} &= \frac{p_L L}{1 - \beta(1 - p_L)} \left[1 - 2\beta(1 - p_H) + \frac{\beta(1 - \beta(1 - p_H))(1 - p_L)}{1 - \beta(1 - p_L)} \right] \end{aligned}$$

Bibliography

Abelson, Reed, “UnitedHealth to Insure the Right to Insurance,” *New York Times*, December 3 2008.

Actuarial Standards Board, “Long Term Care Insurance,” Actuarial Standard of Practice No. 18, Long-Term Care Task Force of the Actuarial Standards Board January 1999.

Adamson, David M., Stella Chang, and Leigh G. Hansen, “Health Research Data for the Real World: The MarketScan Databases,” White Paper, Thompson Healthcare January 2006.

Adamy, Janet, “McDonald’s May Drop Health Plan,” *Wall Street Journal*, September 30 2010.

AHIP, “State Mandatory Medical Loss Ratio (MLR) Requirements for Comprehensive, Major Medical Coverage: Summary of State Laws and Regulations,” naic.org/documents/committees_e_hrsi_comdoc_ahip_chart_mlr.pdf April 15 2010.

Aiyagari, S. Rao and Mark Gertler, “Asset Returns with Transactions Costs and Uninsured Individual Risk,” *Journal of Monetary Economics*, June 1991, 27 (3), 311–331.

- Akerlof, George A.**, “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics*, August 1970, 84 (3), 488–500.
- Altman, Daniel, David M. Cutler, and Richard J. Zeckhauser**, “Adverse Selection and Adverse Retention,” *American Economic Review*, May 1998, 88 (2), 122–126.
- Arrow, Kenneth J.**, “The Role of Securities in the Optimal Allocation of Risk-bearing,” *Review of Economic Studies*, April 1964, 31 (2), 91–96.
- Ball, Laurence and N.Gregory Mankiw**, “Intergenerational Risk Sharing in the Spirit of Arrow, Debreu, and Rawls, with Applications to Social Security Design,” *Journal of Political Economy*, 2007, 115 (4), 523–547.
- Bartleson, Edwin L., Allan K. Archer, Everett D. Armantrout, E. Paul Barnhart, William J. Berry, Richard W. Erdenberger, Raymond F. Killion, John S. Thompson Jr., and Charles N. Walker**, *Health Insurance Provided Through Individual Policies*, Chicago: Society of Actuaries, 1963.
- Baumol, William J.**, “Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis,” *American Economic Review*, June 1967, 57 (3), 415–426.
- Bernanke, Benjamin S.**, “Bankruptcy, Liquidity, and Recession,” *American Economic Review*, May 1981, 71 (2), 155–159.
- , “The Great Moderation,” in “Meeting of the Eastern Economic Association” Washington, D.C. February 20 2004.

- Bluhm, William F.**, “Duration-Based Policy Reserves,” *Transactions of the Society of Actuaries*, 1993, 45, 11–32.
- _____, *Individual Health Insurance*, Actex Publications, 2007.
- Bodie, Zvi**, “Common Stocks as a Hedge Against Inflation,” *Journal of Finance*, May 1976, 31 (1), 459–470.
- _____, “Shortfall Risk and Pension Fund Asset Management,” *Financial Analysts Journal*, May-June 1991, 47 (3), 57–61.
- _____, **and Victor I. Rosansky**, “Risk and Return in Commodity Futures,” *Financial Analysts Journal*, May-June 1980, 36 (3), 27–39.
- Brady, Matt**, “Goldman Establishes Swap Index For Life Settlement Market,” *National Underwriter Life & Health-Financial Services Edition*, January 7 2008.
- Buchmueller, Thomas and John DiNardo**, “Did Community Rating Induce an Adverse Selection Death Spiral? Evidence from New York, Pennsylvania, and Connecticut,” *American Economic Review*, March 2002, 92 (1), 280–294.
- Bull, Clive**, “Implicit Contracts in the Absence of Enforcement and Risk Aversion,” *American Economic Review*, 1983, 73 (4), 658–671.
- Bundorf, M. Kate, Anne Royalty, and Laurence C. Baker**, “Health Care Cost Growth Among the Privately Insured,” *Health Affairs*, September-October 2009, 28 (5), 1294–1304.

Bureau of Labor Statistics, “How BLS Measures Price Change for Medical Care Services in the Consumer Price Index,” bls.gov/cpi/cpifact4.htm April 12 2010.

_____, “Consumer Price Index (CPI),” bls.gov/cpi/ February 15 2011.

Bureau of the Public Debt, “Individual–TIPS: Rates & Terms,” treasurydirect.gov/indiv/research/indepth/tips/res_tips_rates.htm March 28 2008.

_____, “Institutional - STRIPS,” treasurydirect.gov/instit/marketablestrips/strips.htm March 28 2008.

Caldis, Todd G., “The Long-Term Projection Assumptions for Medicare and Aggregate National Health Expenditures,” Trust Fund Projections Methodology, CMS Office of the Actuary/National Health Statistics Group May 12 2009.

Campbell, Ritchie A., “The Demand for Life Insurance: An Application of the Economics of Uncertainty,” *Journal of Finance*, December 1980, 35 (5), 1155–1172.

Centers for Medicare & Medicaid Services, “National Health Expenditure Data,” cms.gov/NationalHealthExpendData/downloads/tables.pdf January 20 2011.

Cochrane, John H., “Time-Consistent Health Insurance,” *Journal of Political Economy*, June 1995, 103 (3), 445–473.

_____, “Health-Status Insurance: How Markets Can Provide Health Security,” Policy Analysis No. 633, Cato Institute February 18 2009.

- Colquitt, L. Le and Robert E. Hoyt**, “Determinants of Corporate Hedging Behavior: Evidence from the Life Insurance Industry,” *Journal of Risk and Insurance*, December 1997, 64 (4), 649–671.
- Connor, Robert, Douglas R. Wholey, Roger Feldman, and William Riley**, “The Check Is in the Mail: Determinants of Claims Payable Timing among Health Maintenance Organizations,” *Inquiry*, Spring 2004, 41 (1), 70–82.
- Cooper, Ian and Evi Kaplanis**, “Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium,” *Review of Financial Studies*, Spring 1994, 7 (1), 45–60.
- Cooper, Richard**, “The US Physician Workforce: Where Do We Stand?,” OECD Health Working Paper 37, OECD, Directorate for Employment, Labour and Social Affairs October 2008.
- Cox, Samuel H. and Robert G. Schwebach**, “Insurance Futures and Hedging Insurance Price Risk,” *Journal of Risk and Insurance*, December 1992, 59 (4), 628–644.
- Cukierman, Alex**, “The Effects of Uncertainty on Investment under Risk Neutrality with Endogenous Information,” *Journal of Political Economy*, June 1980, 88 (3), 462–475.
- Cummins, J. David**, “Insurer’s Risk: A Restatement,” *Journal of Risk and Insurance*, March 1974, 41 (1), 147–157.
- _____, “Risk-Based Premiums for Insurance Guaranty Funds,” *Journal of Finance*, September 1988, 43 (4), 823–839.

- ____ **and Christopher M. Lewis**, “Catastrophic Events, Parameter Uncertainty and the Breakdown of Implicit Long-term Contracting in the Insurance Market: The Case of Terrorism Insurance,” Center for Financial Institutions Working Papers 02-40, Wharton School Center for Financial Institutions, University of Pennsylvania July 2002.
- ____ **and** _____, “Catastrophic Events, Parameter Uncertainty and the Breakdown of Implicit Long-Term Contracting: The Case of Terrorism Insurance,” *Journal of Risk and Uncertainty*, March 2003, 26 (2), 153–178.
- ____ **and Gregory P. Nini**, “Optimal Capital Utilization by Financial Firms: Evidence from the Property-Liability Insurance Industry,” *Journal of Financial Services Research*, February 2002, 21 (1-2), 15–53.
- ____, **David Lalonde, and Richard D. Phillips**, “The Basis Risk of Catastrophic-Loss Index Securities,” *Journal of Financial Economics*, January 2004, 71 (1), 77 – 111.
- ____, **Georges Dionne, Robert Gagne, and Abdelhakim Nouira**, “The Costs and Benefits of Reinsurance,” June 2008. Available at SSRN: ssrn.com/abstract=1142954.
- Cutler, David M.**, “Why Doesn’t the Market Fully Insure Long-Term Care?,” Working Paper 4301, National Bureau of Economic Research March 1993.
- ____, *The Changing Hospital Industry: Comparing Not-For-Profit and For-Profit Institutions*, Chicago: University of Chicago Press, 2000.

Cutler, David, Mark McClellan, and Joseph Newhouse, “The Costs and Benefits of Intensive Treatment for Cardiovascular Disease,” Working Paper 6514, National Bureau of Economic Research April 1998.

Cylus, Jonathan, Micah Hartman, Benjamin Washington, Kimberly Andrews, and Aaron Catlin, “Pronounced Gender And Age Differences Are Evident In Personal Health Care Spending Per Person,” *Health Affairs*, January 2011, 30 (1), 153.

Danzon, Patricia M., “Making Sense of Drug Prices,” *Regulation*, Spring 1999, 23 (1), 56–63.

De Nardi, Mariacristina, Eric French, and John Bailey Jones, “Why do the Elderly Save? The Role of Medical Expenses,” Working Paper 15149, National Bureau of Economic Research July 2009.

Debreu, Gerard, *Theory of Value*, New Haven, CT: Yale University Press, 1959.

Doherty, Neil A. and James R. Garven, “Price Regulation in Property-Liability Insurance: A Contingent-Claims Approach,” *Journal of Finance*, December 1986, 41 (5), 1031–1050.

Dowd, Kevin, “A Proposal to End Inflation,” *Economic Journal*, July 1994, 104 (425), 828–840.

Eichner, Matthew J., Mark B. McClellan, and David A. Wise, “Insurance or Self-Insurance? Variation, Persistence, and Individual Health Accounts,” in David A.

Wise, ed., *Inquiries in the Economics of Aging*, Chicago: University of Chicago Press, 1998, pp. 19–45.

Elmendorf, Douglas W., “The Long Term Budget Outlook: Federal Debt Held by the Public Under Two Budget Scenarios,” Long Term Budgetary Issues, Congressional Budget Office June 2010.

Emmons, William R., “The Information Content of Treasury Inflation-Indexed Securities,” *The Federal Reserve Bank of St. Louis Review*, November/December 2000, pp. 25–38.

Ewing, Bradley T., Michael J. Piette, and James E. Payne, “Forecasting Medical Net Discount Rates,” *Journal of Risk and Insurance*, March 2003, 70 (1), 85–95.

Fabozzi, Frank J., Mark Pitts, and Ravi E. Dattatreya, “Price Volatility Characteristics of Fixed Income Securities,” in Frank J. Fabozzi, ed., *The Handbook of Fixed Income Securities*, Chicago: Irwin Professional Publishing, 1997.

Fama, Eugene F. and Kenneth R. French, “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, February 1993, 33 (1), 3–56.

_____ and _____, “Fama-French Factors,” mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html August 16 2010.

Fenn, George W. and Rebel A. Cole, “Announcements of Asset-Quality Problems and Contagion Effects in the Life Insurance Industry,” *Journal of Financial Economics*, April 1994, 35 (2), 181–198.

Fields, Joseph, Linda Klein, and Edward Myskowski, “Lloyd’s Financial Distress and Contagion within the US Property and Liability Insurance Industry,” *Journal of Risk and Uncertainty*, May 1998, *16* (2), 173–185.

Financial Accounting Standards Board, “Employers Accounting for Postretirement Benefits Other Than Pensions,” Statement of Financial Accounting Standards 106 December 1990.

Finkelstein, Amy, “The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare,” *Quarterly Journal of Economics*, February 2007, *122* (1), 1–37.

Ford, Ina Kay and Daniel H. Ginsburg, “Medical Care in the Consumer Price Index,” in “Medical Care Output and Productivity,” National Bureau of Economic Research, 2001, chapter 5, pp. 203–220.

Foster, Richard S., “Estimated Financial Effects of the Patient Protection and Affordable Care Act of 2009, as Proposed by the Senate Majority Leader on November 18, 2009,” Memorandum, Office of the Actuary, Center for Medicare and Medicaid Services December 10 2009.

Garven, James R., “An Exposition of the Implications of Limited Liability and Asymmetric Taxes for Property-Liability Insurance,” *Journal of Risk and Insurance*, March 1992, *59* (1), 34–56.

- Getzen, Thomas E.**, “Modeling Long Term Healthcare Cost Trends,” Research Project in Health, Society of Actuaries December 10 2007.
- Girola, James A.**, “The Long-Term Real Interest Rate for Social Security,” Research Paper 2005-02, U.S. Department of the Treasury March 30 2005.
- Global Financial Data**, “Total Return Database,” www.globalfinancialdata.com August 16 2010.
- Grace, Martin F. and Richard D. Phillips**, “Regulator Performance, Regulatory Environment and Outcomes: An Examination of Insurance Regulator Career Incentives on State Insurance Markets,” *Journal of Banking & Finance*, January 2008, 32 (1), 116–133.
- Griffith-Jones, Stephany and Krishnan Sharma**, “GDP-Indexed Bonds: Making it Happen,” DESA Working Paper 21, United Nations Department of Economic and Social Affairs April 2006.
- Gullipalli, Diya**, “Challenge for ETF Trading: Bumpy Markets Gum It Up,” *Wall Street Journal*, September 4 2007, p. C1.
- Hall, Robert E. and Charles I. Jones**, “The Value of Life and the Rise in Health Spending,” *Quarterly Journal of Economics*, February 2007, 122 (1), 39–72.
- Harrington, Scott E.**, “The Impact of Rate Regulation on Automobile Insurance Loss Ratios: Some New Empirical Evidence,” *Journal of Insurance Regulation*, December 1984, 3 (2), 182–202.

- **and Greg Niehaus**, “Capital, Corporate Income Taxes, and Catastrophe Insurance,” *Journal of Financial Intermediation*, October 2003, 12 (4), 365–389.
- **and Jack M. Nelson**, “A Regression-Based Methodology for Solvency Surveillance in the Property-Liability Insurance Industry,” *Journal of Risk and Insurance*, December 1986, 53 (4), 583–605.
- **, Steven V. Mann, and Greg Niehaus**, “Insurer Capital Structure Decisions and the Viability of Insurance Derivatives,” *Journal of Risk and Insurance*, September 1995, 62 (3), 483–508.
- Harris, Milton and Bengt Hölmstrom**, “A Theory of Wage Dynamics,” *Review of Economic Studies*, June 1982, 49 (3), 315–333.
- Heaton, John and Deborah J. Lucas**, “Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing,” *Journal of Political Economy*, June 1996, 104 (3), 443–487.
- Hendel, Igal and Alessandro Lizzeri**, “The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance,” *Quarterly Journal of Economics*, February 2003, 118 (1), 299–327.
- Herring, Bradley and Mark V. Pauly**, “Incentive-Compatible Guaranteed Renewable Health Insurance Premiums,” *Journal of Health Economics*, May 2006, 25 (3), 395–417.

- Hirshleifer, Jack**, “Efficient Allocation of Capital in an Uncertain World,” *American Economic Review*, May 1964, 54 (3), 77–85.
- Hölmstrom, Bengt**, “Moral Hazard and Observability,” *The Bell Journal of Economics*, Spring 1979, 10 (1), 74–91.
- Huang, Chi-fu and Robert H. Litzenberger**, *Foundations for Financial Economics*, Upper Saddle River, NJ: Prentice Hall, 1988.
- Irving, Susan J.**, “Debt Management: Treasury Inflation Protected Securities Should Play a Heightened Role in Addressing Debt Management Challenges,” Report to the Secretary of the Treasury, Government Accountability Office, Washington, D.C. September 2009.
- Jean-Baptiste, Eslyn L. and Anthony M. Santomero**, “The Design of Private Reinsurance Contracts,” *Journal of Financial Intermediation*, July 2000, 9 (3), 274 – 297.
- Jennings, William W.**, “Disaggregated TIPS: The Case for Disaggregating Inflation-Linked Bonds into Bonds Linked to Narrower CPI Components (Slicing and Dicing TIPS),” *Journal of Pension Economics and Finance*, November 2006, 5 (3), 325–343.
- _____, **Steve P. Fraser, and Brian C. Payne**, “Do Health Care Investments Hedge Health Care Liabilities?,” November 2007. Available at SSRN: ssrn.com/abstract=1031046.
- Jost, Timothy**, “Implementing Health Reform: Medical Loss Ratios,” healthaffairs.org/blog/2010/11/23/implementing-health-reform-medical-loss-ratios/ November 23 2010.

- Kitchen, John and Ralph Monaco**, “Real-Time Forecasting in Practice: The U.S. Treasury Staff’s Real-Time GDP Forecast System,” *Business Economics*, October 2003, 38, 10–19.
- Klein, Robert W.**, “The Insurance Industry and Its Regulation: An Overview,” in Martin F. Grace and Robert W. Klein, eds., *The Future of Insurance Regulation in the United States*, Washington, D.C.: Brookings Institution Press, 2009, chapter 2, pp. 13–52.
- Knight, Frank H.**, *Risk, Uncertainty, and Profit*, Boston, MA: Houghton Mifflin Company, 1921.
- Kolb, Robert W. and James A. Overdahl**, *Understanding Futures Markets*, sixth ed., Malden, MA: Blackwell Publishing, 2006.
- Krueger, Dirk and Harald Uhlig**, “Competitive risk sharing contracts with one-sided commitment,” *Journal of Monetary Economics*, October 2006, 53 (7), 1661–1691.
- Lamm-Tennant, Joan**, “Asset/Liability Management for the Life Insurer: Situation Analysis and Strategy Formulation,” *Journal of Risk and Insurance*, September 1989, 56 (3), 501–517.
- Larson, Erik**, “Delphi Allowed to Cancel Benefits for 15,000 Workers,” *Bloomberg*, February 24 2009.
- Leibowitz, Martin L.**, “The Dedicated Bond Portfolio in Pension Funds: Part I: Motivations and Basics,” *Financial Analysts Journal*, January-February 1986, 42 (1), 68–75.

- Levit, Katharine, Cynthia Smith, Cathy Cowan, Helen Lazenby, and Anne Martin,** “Inflation Spurs Health Spending In 2000,” *Health Affairs*, January-February 2002, 21 (1), 172–181.
- Lindh, Thomas,** “Medium-Term Forecasts of Potential GDP and Inflation using Age Structure Information,” *Journal of Forecasting*, January 2004, 23 (1), 19–49.
- Litzenberger, Robert H. and James C. Van Horne,** “Elimination of the Double Taxation of Dividends and Corporate Financial Policy,” *Journal of Finance*, June 1978, 33 (3), 737–750.
- Lucas, Deborah and Stephen P. Zeldes,** “Valuing and Hedging Defined Benefit Pension Obligations the Role of Stocks Revisited,” Working Paper 2006.
- Lutzky, Steven, John Corea, and Lisa Alecxih,** “A Survey of Employers Offering Group Long-Term Care Insurance to Their Employees,” Final Report June 20 2000.
- Mahler, Howard C. and Curtis Gary Dean,** “Credibility,” in “Foundations of Casualty Actuarial Science,” fourth ed., Arlington, VA: Casualty Actuarial Society, 2001, chapter 8.
- McDonald, Gregory J.,** “Employee Benefits: States Need Labor’s Help Regulating Multiple Employer Welfare Arrangements,” Report to Congressional Requesters B-242197, United States General Accounting Office March 1992.
- Mehra, Rajnish and Edward C. Prescott,** “The Equity Premium: A Puzzle,” *Journal of Monetary Economics*, March 1985, 15 (2), 145 – 161.

- Merton, Robert C.**, “Financial Innovation and the Management and Regulation of Financial Institutions,” *Journal of Banking & Finance*, June 1995, 19 (3-4), 461 – 481.
- Modigliani, Franco and Merton H. Miller**, “The Cost of Capital, Corporation Finance and the Theory of Investment,” *The American Economic Review*, June 1958, 48 (3), 261–297.
- Modu, Emmanuel**, “Securitization of Reinsurance Recoverables,” Best’s Rating Methodology, A.M. Best August 20 2007.
- Munch, Patricia and Dennis E. Smallwood**, “Solvency Regulation in the Property-Liability Insurance Industry: Empirical Evidence,” *Bell Journal of Economics*, Spring 1980, 11 (1), 261–279.
- **and Dennis Smallwood**, “Theory of Solvency Regulation in the Property and Casualty Insurance Industry,” in Gary Fromm, ed., *Studies in Public Regulation*, Cambridge, MA: MIT Press, 1981, chapter 3, pp. 119–180.
- Nash, John F.**, “The Bargaining Problem,” *Econometrica*, April 1950, 18 (2), 155–162.
- National Center for Health Statistics**, “Utilization of ambulatory medical care by women: United States, 199798,” Working Paper 13-149, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention July 2001.
- Novy-Marx, Robert and Joshua D. Rauh**, “The Intergenerational Transfer of Public Pension Promises,” Working Paper 14343, National Bureau of Economic Research September 2008.

Orzag, Peter, “IMAC, UBend,” whitehouse.gov/omb/blog/09/07/17/IMACUBend/ July 19 2009.

Patel, Vip and Mark V. Pauly, “Guaranteed Renewability and the Problem of Risk Variation in Individual Health Insurance Markets,” *Health Affairs*, August 28 2002, *Web Exclusives*, w280–w289.

Pauly, Mark V., “Should We Be Worried about High Real Medical Spending Growth in the United States?,” *Health Affairs*, January 8 2003, *W3*, 15–27.

— **and Bradley Herring**, “Risk Pooling And Regulation: Policy And Reality In Today’s Individual Health Insurance Market,” *Health Affairs*, May-June 2007, *26* (3), 770–779.

— **and Robert D. Lieberthal**, “How Risky is Individual Health Insurance?,” *Health Affairs*, May 2008, *27* (3), w242–w249.

— , **Howard Kunreuther, and Richard Hirth**, “Guaranteed Renewability in Insurance,” *Journal of Risk and Uncertainty*, March 1995, *10* (2), 143–156.

Pearce, Douglas K. and V. Vance Roley, “Stock Prices and Economic News,” Working Paper 1296, National Bureau of Economic Research April 1985.

Peden, Edgar A. and Mark S. Freeland, “Insurance Effects on US Medical Spending (1960-1993),” *Health Economics*, December 1998, *7* (8), 671–687.

Phelan, Christopher, “Repeated Moral Hazard and One-Sided Commitment,” *Journal of Economic Theory*, August 1995, 66 (2), 488–506.

Rauch, Mike, “Understanding The Guaranteed Renewable Aspect of Long-Term Care Insurance,” soa.org/files/pdf/02-Boston-Rauch-06.pdf October 28 2002. Society of Actuaries 2002 Annual Meeting.

Robbins, Edward L. and Peter H. Winslow, “Uncertainty Remains in Tax Reserve Assumptions for Guaranteed Renewable and Noncancellable Health Insurance,” *Taxing Times*, September 2010, 6 (3), 1,6–9.

Rothschild, Michael and Joseph Stiglitz, “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information,” *Quarterly Journal of Economics*, November 1976, 90 (4), 629–649.

Rubinstein, Mark, “Securities Market Efficiency in an Arrow-Debreu Economy,” *American Economic Review*, December 1975, 65 (5), 812–824.

Schlesinger, Harris, “The Optimal Level of Deductibility in Insurance Contracts,” *Journal of Risk and Insurance*, September 1981, 48 (3), 465–481.

Schmitz, Allen, “Adverse Selection and the CLASS Act,” Health Reform Briefing Paper, Milliman December 2009.

Schultz, Jennifer Saranow, “MetLife to Stop Offering New Long-Term Care Policies,” *New York Times*, November 11 2010.

Shiller, Robert J., *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks*, New York: Oxford University Press, USA, 1993.

Simon, Carol J. and David W. Emmons, "Physician Earnings at Risk: An Examination of Capitated Contracts," *Health Affairs*, May-June 1997, *16* (3), 120–126.

Sinclair, Sven H. and Kent A. Smetters, "Health Shocks and the Demand for Annuities," Technical Paper Series 2004-9, Congressional Budget Office July 2004.

Sisko, Andrea, Christopher Truffer, Sheila Smith, Sean Keehan, Jonathan Cylus, John A. Poisal, M. Kent Clemens, and Joseph Lizonitz, "Health Spending Projections Through 2018: Recession Effects Add Uncertainty To The Outlook," *Health Affairs*, March 2009, *28* (2), w346–w357.

Sisko, Andrea M., Christopher J. Truffer, Sean P. Keehan, John A. Poisal, M. Kent Clemens, and Andrew J. Madison, "National Health Spending Projections: The Estimated Impact of Reform Through 2019," *Health Affairs*, October 2010, *29* (10), 1–9.

Stewart, Kenneth J., "The Experimental Consumer Price Index for Elderly Americans (CPI-E): 1982-2007," *Monthly Labor Review*, April 2008, *131* (4), 19–24.

Telmer, Chris I., "Asset-Pricing Puzzles and Incomplete Markets," *Journal of Finance*, December 1993, *48* (5), 1803–1832.

The White House, “FACT SHEET: The Early Retiree Reinsurance Program,”
whitehouse.gov/the-press-office/fact-sheet-early-retiree-reinsurance-program May 4
2010.

U.S. Congress, “Patient Protection and Affordable Care Act,”
democrats.senate.gov/reform/patient-protection-affordable-care-act-as-passed.pdf
2010. Public Law 111-148.

Wachenheim, Leigh, “Variation by Duration in Individual Health Medical Insurance
Claims,” Research report, Society of Actuaries, Schaumburg, IL October 3 2006.

Young, Alison, “Blue Cross to Refund \$156 Million in N.C.,” *USA Today*, September 21
2010.

Zion, David and Amit Varshney, “You Dropped a Bomb on Me, GASB—Uncovering
\$1.5 Trillion in Hidden OPEB Liabilities for State and Local Governments,” Equity
Research, Credit Suisse March 22 2007.