Insurance Taxation and Insurance Fraud

by M.Martin Boyer

Working Paper 97-10
September 1997

ISSN : 1206-3304

The author would like to thank Keith Crocker, Georges Dionne, Neil Doherty and Sharon Tennyson for useful comments. Financial help was provided by the SSHRC of Canada and the S.S. Huebner Foundation. Usual disclaimers apply.

Insurance Taxation and Insurance Fraud

M.Martin Boyer

M.Martin Boyer is assistant professor, Department of finance, École des HEC.

Insurance Taxation and Insurance Fraud.[*]

M. Martin Boyer[†]

First Version: January 1998
This Version: November 1998

Abstract

It is common practice in the United States to impose a sales tax on insurance premiums. Another type of tax that is not used is the taxation of insurance benefits. It is typically argued that one should not tax insurance benefits because they are supposed to compensate for a loss. In this paper I present a case where the taxation of insurance benefits is preferable to the taxation of premium. When insurance fraud is present - in the form of ex post moral hazard - a tax on insurance premiums increases the number of fraudulent claims in the economy, whereas a tax on insurance benefits may reduce fraud. More importantly, however, agents are made better off with a benefit tax than a premium tax.

JEL classification: G22, H21, C72, D82

Keywords: Ex Post Moral Hazard, Insurance Fraud, Asymmetric Information, Taxation.

[*] I would like to thank Keith Crocker, Georges Dionne, Neil Doherty and Sharon Tennyson for comments. I would also like to acknowledge the useful insights of the seminar participants at l'École des Hautes Études Commerciales de Montréal and at the 1998 Public Economic Theory Seminars. Financial help was provided by the SSHRC of Canada, the S. S. Huebner Foundation at the University of Pennsylvania, and by the Risk Management Chair at l'École des Hautes Études Commerciales de Montréal. Usual disclaimers apply.

[†] Assistant Professor, Department of Finance, École des Hautes Études Commerciales, Université de Montréal. 3000 Côte-Ste-Catherine Montréal, QC H3T 2A7 CANADA. tel.:514-340-6704; fax.:514-340-5632; e-mail: martin.boyer@hec.ca

Insurance Taxation and Insurance Fraud

ABSTRACT. It is common practice in the United States to impose a sales tax on insurance premiums. Another type of tax that is not used is the taxation of insurance bene…ts. It is typically argued that one should not tax insurance bene…ts because they are supposed to compensate for a loss. In this paper I present a case where the taxation of insurance bene…ts is preferable to the taxation of premium. When insurance fraud is present - in the form of ex post moral hazard - a tax on insurance premiums increases the number of fraudulent claims in the economy, whereas a tax on insurance bene…ts may reduce fraud. More importantly, however, agents are made better o¤ with a bene…t tax than a premium tax.

JEL classi…cation: G22, H21, C72, D82

Keywords: Ex Post Moral Hazard, Insurance Fraud, Asymmetric Information, Taxation.

1 Introduction

It is common practice for many states in the United States to impose a sales tax on insurance premiums. Insurers in the United States face two types of indirect taxation. First, insurance premiums paid by a policyholder at taxed, and second, business operations are taxed. The premium tax is of the order of 2% on average for the life insurance industry,[1] and 2.5% for the property and casualty industry.[2]

Another type of tax that is not used is the taxation of insurance benefits. The reason why benefits are not taxed is that it does not seem fair to tax amounts that are designed to compensate for a loss. I argue in this paper that even though taxing benefits may seem to penalize policyholders who happen to be unlucky, some beneficial aspects of such a tax have been left unstudied. One such aspect is how the taxation of insurance benefits may in fact increase a policyholder's utility when insurance fraud - in the form of ex post moral hazard[3] - is present.

Insurance fraud in the United States is an industry. According to an 1995 estimates of the Rand Corporation[4] fifteen to eighteen billion dollars are paid every year for apparently fraudulent medical claims arising from automobile accidents. This amounts to about one hundred dollars per automobile insurance policy. Surely when one adds to this the amount spent for unnecessary medical expenses arising from workers' compensation claims and other health care frauds, the amount misallocated due to insurance fraud must be astronomical.

To combat insurance fraud, more and more resources are being devoted by insurance companies and state insurance regulators. For example in Massachusetts, private insurers formed the Massachusetts Insurance Fraud Bureau to pool insurance fraud information about policyholders. In California the state legislature has mandated the creation of a fraud crime

[1] See Castillo (1997). For domestic insurers, the premium tax rate in 1991 varied from 0% in Florida, Illinois, Kentucky and Ohio to 2.75% in Hawaii and Montana, with an average of 1.595%. For foreign insurers (insurers who are not domicile in the state) and alien insurers (insurers whose domicile is not in the U.S.) the premium tax rate is typically greater.

[2] See the 1995 NAIC Data Tapes, page 15, as provided by the Center for Risk Management and Insurance Research at Georgia State University. For domestic insurers, the effective tax rate in 1995 varied from 0.3% in Delaware and Oregon to 4.7% in Utah. As for the life insurance industry, foreign and alien insurers are typically taxed at a greater rate than domestic insurers. In the Canadian province of Quebec, the tax rate of the insurance premium is 5%.

[3] The notion of ex post moral hazard was first introduced by Spence and Zeckhauser (1971).

[4] Rand Corporation Institute for Civil Justice. Research Brief, May 1995.

bureau; similar legislation was also passed or suggested in Texas, Arizona, Florida and Nevada.[5] To finance the creation of such bodies, some have argued that the state should increase the sales tax on insurance premium. As we will see later in this paper, this may not solve the problem. Quite the contrary.

Using a simple game theoretic model, it is possible to model the insurance fraud problem as a simple two-player game where one player (the policyholder or agent) has private information as to the state of the world. The other player (the insurer or principal) can learn about this state of the world by auditing the reported state of the world. If the insurer audits, then it shall be assumed that she[6] learns for sure what the state of the world is. Auditing is costly to the insurer, and being caught cheating one's insurer is costly to the agent. This approach to the insurance fraud problem was first presented by Townsend (1979). It is known as the Costly State Verification approach.

A basic assumption of the model I develop is that the principal cannot commit credibly to a certain auditing strategy. This non-commitment does not allow to solve the principal-agent model in the sense that there is still fraud in the economy. The credibility of a pre-specified audit strategy has been studied and criticized by Picard (1996) and Boyer (1998). The basic idea behind the non-commitment assumption is that insurers cannot contractually guaranty that they will audit with the exact probability which eliminates all fraud in the economy. In the absence of commitment, the only thing the insurer can do is react in the best possible way to the policyholder's report. In Picard's opinion, the difference between being able to commit to an audit strategy and not being able to is the same difference as the one between being a leader in a Stackelberg game and being a follower. If the insurer can commit then the insurer can be viewed as the Stackelberg leader, with all the advantages that it entails.

This paper compares two economies that differ only with respect to the taxation scheme they are using. In both economies a simple proportional tax is used. In economy A, the government imposes a tax on the premium paid by the policyholder. In economy B, the government imposes a proportional tax on the benefits received by the policyholder from his

[5] See Brown (1993), National Underwriter (vol.99 no.18, p.11. 1995), and National Underwriter (vol.99 no.14, p.3. 1995).

[6] Throughout the paper the terms insurer and principal, and policyholder and agent will be used interchangeably. The feminine represents the insurer/principal and the masculine represents the policyholder/agent.

insurer.

The most interesting result of the paper is that in the presence of fraud agents would rather have their benefits taxed than their premiums. Even from the point of view of a policymaker whose only goal is to reduce insurance fraud using a tax scheme, I show that a premium tax not only fares worse than a benefit tax, but more importantly that it fares worse than no tax at all.

The remainder of the paper goes as follows. In the next section I present the basic make up of the model: the assumptions, the sequence of play and the basic game can be found there. I then present the benchmark case of an economy where agents do not engage in insurance fraud. In this benchmark case agents are indifferent, at the margin, over a benefit tax and premium tax. The core of the paper is located in section 3. When agents have the possibility to file a fraudulent claim, and when the principal cannot commit to an auditing strategy, then agents have a strict preference for a benefit tax over a premium tax. This preference is partially due the impact each tax has on the amount of fraud in the economy, as is shown in subsection 3.2. I then analyze the optimal tax scheme when both types of taxes are allowed. Finally I conclude in section 4 with a discussion of the results.

2 The Basics

2.1 Assumptions

Let's first start by stating the basic assumptions of the model.

A.1. The agent is risk averse ($U'(.) > 0; U''(.) < 0; U'(0) = \infty$) and has a von Neumann-Morgenstern utility function over final wealth. The insurer is risk neutral.

A.2. There are two states of nature: Loss and No Loss. A loss L occurs with probability π. Insurance is compulsory.

A.3. The agent and the principal play a game of asymmetric information. The agent knows whether he was involved in an accident, while the principal does not. The possible actions for the agent are file a claim (FC) and don't file (DF), while the possible actions for the principal are audit the claim (AC) and not audit (NA).

A.4. The insurance market is perfectly competitive in the sense that the insurer is making zero profits in expectation.

A.5. Auditing a claim is costly to the principal. This is a fixed cost denoted by $0 < c < L$.

A.6. An agent caught committing fraud must incur some fixed penalty $k > 0$.

A.7. The insurer cannot commit to any auditing strategy ex ante.

A.8. Initially there are no taxes in the economy.

A.9. The government wants to raise \$1 from the insurance industry. This dollar must be raised through a tax. In Economy A, this will be a sales tax on the premium, while in Economy B, it will be a proportional tax on the benefits. Let the tax rates be given by τ_A and τ_B respectively. The benefit tax is paid whenever a claim is filed, irrespective of whether or not the agent is audited, and whether or not he committed fraud.

The sequence of play goes as follow. In the first stage of the game a contract is signed between the principal and the agent. This contract specifies a coverage in case of a loss and a premium. As per A.4, this premium is set such that the principal earns zero expected profits on each contract. The premium tax is paid at this point. In the second stage of the game, Nature decides whether the policyholder is involved in an accident or not. This information is private to the policyholder. In stage three, the policyholder must decide what to report to the principal. He can either file a claim or not. The benefit tax is paid at this point. The last move belongs to the insurer who must decide whether to verify the reported state. Finally the payoffs are paid and the game ends. The sequence of the game is displayed in figure 1.[7] The extensive form of the game is shown in figure 2.

The type of fraud modeled here is one where a policyholder can claim that he suffered a loss when in fact he did not. A real world example of an insurance market where the above-mentioned game may be encountered is in the auto-theft (comprehensive) insurance market. The example is pertinent because 1-there is only one possible loss (the value of the car), 2-a policyholder whose automobile is not stolen can hide the car, claim that it was stolen, and collect from his insurer and 3-the insurer can send a claim adjuster, at a cost, to make sure that the car was indeed stolen.

The expected utility of the policyholder is given by

$$\begin{aligned} EU &= \pi(1-\nu^*)U(Y-(1+\tau_A)\alpha^*-L+(1-\tau_B)\beta^*) \\ &\quad +\pi\nu^*U(Y-(1+\tau_A)\alpha^*-L+(1-\tau_B)\beta^*) \end{aligned} \tag{EU}$$

[7] All figures, tables and graphs are in appendix 8.2.

$$
\begin{aligned}
&+(1-\pi)\eta^*(1-\nu^*)U(Y-(1+\tau_A)\alpha^*+(1-\tau_B)\beta^*) \\
&+(1-\pi)\eta^*\nu^*U(Y-(1+\tau_A)\alpha^*-\tau_B\beta^*-k) \\
&+(1-\pi)(1-\eta^*)U(Y-(1+\tau_A)\alpha^*)
\end{aligned}
$$

where Y is the policyholder's initial wealth, τ_A and τ_B are respectively the premium and the benefit tax rates, η is the probability that a policyholder defrauds his insurer, ν is the probability that an insurer audits a policyholder's claim, π is the probability that a loss occurs, L is the size of the loss, α is the premium paid, and β is the benefit received in case of a loss.

The first two lines in (EU) represents the utility the agent receives when he suffered a loss and files a claim.[8] It simplifies to $\pi U(Y-(1+\tau_A)\alpha^*-L+(1-\tau_B)\beta^*)$ since the payment to the agent is the same whether or not there is an audit. This modeling differs from that of Mookherjee and Png (1989, 1990) who let the agent receive a bonus whenever he is audited and found to have told the truth. The third and fourth lines in (EU) represent the expected utility an agent receives from filing a claim given that he did not suffer a loss. He commits fraud with probability η^*, and is caught with probability ν^*. When he is caught, he incurs monetary penalty $k > 0$. The last line represents the agents's expected utility from reporting the truth when he is not involved in an accident.

The benchmark case where agents never defraud the principal is first presented to see what happens in a perfect world. The case where an agent attempts sometimes to extract rents from the principal by misreporting his loss is then presented in section 4.

2.2 Benchmark: No Fraud

In this section of the paper I present the benchmark case where agents never commit fraud. Let subscript A (B) denote the economy where premiums (benefits) are taxed. The principal's basic problem is then to design a contract that maximizes the agent's expected utility

[8] The agent always files a claim when there is a loss. This allows to eliminate the expected payoff of the agent who does not file a claim when he suffered a loss since it never occurs. Since we have a one-period model, the agent is not concerned with the possibility that his second period premium is dependent on what he tells his insurer in the first period (no bonus-malus system). In other words, there is no reason to under-report.

subject only to a zero-profit constraint.

$$\max_{\beta_i} W_i \qquad s.t. \quad \alpha_i = \pi \beta_i \qquad i \in \{A; B\} \tag{BP}$$

The W_i's represent the policyholder's expected utility, and are equal to

$$W_A = \pi U\left(Y - (1+\tau_A)\alpha_A - L + \beta_A\right) + (1-\pi)U\left(Y - (1+\tau_A)\alpha_A\right) \tag{1}$$

$$W_B = \pi U\left(Y - \alpha_B - L + (1-\tau_B)\beta_B\right) + (1-\pi)U\left(Y - \alpha_B\right) \tag{2}$$

The first order conditions are then

$$\begin{aligned} \frac{\partial W_A}{\partial \beta_A} = \; & 0 = \pi[1 - \pi(1+\tau_A)]U'\left(Y - (1+\tau_A)\alpha_A - L + \beta_A\right) \\ & - \pi(1-\pi)(1+\tau_A)U'\left(Y - (1+\tau_A)\alpha_A\right) \end{aligned} \tag{3}$$

for Economy A, and

$$\frac{\partial W_B}{\partial \beta_B} = 0 = \pi(1-\tau_B)U'\left(Y - \alpha_B - L + (1-\tau_B)\beta_B\right) - \pi(1-\pi)U'\left(Y - \alpha_B\right) \tag{4}$$

for Economy B. Note that A.8 states that initially the two tax rates are zero ($\tau_A = \tau_B = 0$). This means that if the government does not raise any money, then the two first order conditions are the same. The implication is summarized in lemma 1.

Lemma 1 As $\tau_A \to 0$ and $\tau_B \to 0$, then $\beta_A \to \beta$ and $\beta_B \to \beta$.

Proof All the proofs are in appendix 8.1. $\square$

What lemma 1 tells us is that when tax rates tend to zero, then the coverage chosen in the two economies will be the same. A well known result in the economic literature is that if there exists a premium proportional loading factor, policyholders are better off being less than fully insured ($\beta < L$).[9] Since taxes play the same role as a proportional premium loading, lemma 1 tells us that when taxes are positive, agents choose to be partially insured in both economies. Using this lemma, we can now prove the first theorem of the paper.

Theorem 1 If initially there are no taxes ($\tau_A = \tau_B = 0$) and if agents never commit fraud then agents are indifferent between marginal increases in τ_A and τ_B.

[9] See for example see Arrow (1970).

We can infer from theorem 1 that if there is no fraud, agents are indifferent between a marginal tax that is levied on premiums and one that is levied on insurance benefits - at least when the initial tax rates are zero. The reason is that the effect of the benefit tax is reflected directly in the premium. The wealth of the agent when the benefit tax is paid is $Y - \alpha - L + (1 - \tau_B)\beta$. Since $\alpha = \pi\beta$, the agent's wealth may be rewritten as $Y - \pi\beta - L + \beta - \tau_B\beta$ and $Y - (\pi + \tau_B)\beta - L + \beta$. Letting $\tau_B = \pi\lambda$, the agent's wealth becomes $Y - (1 + \lambda)\pi\beta - L + \beta$, which is the functional form of the agent's wealth under a premium tax. It is therefore possible to rewrite the benefit tax as a premium tax when there is no fraud. It then makes intuitive sense for the agent to be indifferent between the two tax tools.

In an economy where fraud is possible, agents are not indifferent between a premium tax and a benefit tax, however. In fact they strictly prefer to have their benefits taxed than their premiums. This result is demonstrated in the next section.

3 Presence of Fraud

3.1 Optimal Contract

Suppose a policyholder suffers no disutility from committing fraud. Then the problem faced by the principal in designing the optimal contract has to take into account the behavior of each player. This means that the policyholder's optimal reporting strategy, as well as the insurer's optimal auditing strategy must be derived. In this game setup, the optimal strategies of each player leads to a Perfect Bayesian Nash Equilibrium (PBNE). This unique PBNE in mixed strategies is, by definition, a sextuplet. This PBNE is presented in lemma 2.

Lemma 2 Provided that certain existence conditions hold, the only PBNE[10] in mixed strategies of this game is such that

1-The agent always files a claim if he is involved in an accident;

[10] In this game the notions of Perfect Bayesian Nash Equilibrium and Sequential Equilibrium coincide. The PBNE in mixed strategy of this game is unique, if it exists. In a 2×2 game (two players each having two possible actions), there is at most one mixed strategy equilibrium (see Myerson, 1991). A sufficient condition for an equilibrium in mixed strategy to exist is that $\pi \leq \frac{1}{2}$.

2-The agent files a claim with probability η_i if he is not involved in an accident;

3-The principal never audits a policyholder who does not file a claim;

4-The principal audits with probability ν_i when an agent files a claim.

For each economy, the probability of committing fraud and the probability of auditing are different; η_A, η_B, ν_A and ν_B are equal to

$$\eta_A = \left(\frac{c}{\beta_A - c}\right)\left(\frac{\pi}{1-\pi}\right) \tag{RC_A}$$

$$\eta_B = \left(\frac{c}{\beta_B - c}\right)\left(\frac{\pi}{1-\pi}\right) \tag{RC_B}$$

and

$$\nu_A = \frac{U(Y - (1+\tau_A)\alpha_A + \beta_A) - U(Y - (1+\tau_A)\alpha_A)}{U(Y - (1+\tau_A)\alpha_A + \beta_A) - U(Y - (1+\tau_A)\alpha_A - k)} \tag{AC_A}$$

$$\nu_B = \frac{U(Y - \alpha_B + (1-\tau_B)\beta_B) - U(Y - \alpha_B)}{U(Y - \alpha_B + (1-\tau_B)\beta_B) - U(Y - \alpha_B - \tau_B\beta_B - k)} \tag{AC_B}$$

The nature of this PBNE is rationally anticipated by the principal when she design the contract that maximizes the agent's expected utility. It is interesting to note that the function that represents the probability of committing fraud has the same shape in both economies. In fact, if the same benefit is chosen in the two economies ($\beta_A = \beta_B$), then the probabilities of fraud are the same ($\eta_A = \eta_B$), even though the probabilities of auditing are not necessarily the same.

The problem then faced by the principal in economy A is to design a contract that specifies a benefit level and a premium that maximize the agent's expected utility given by

$$\begin{aligned} \max_{\alpha_A;\beta_A} EU_A = {} & \pi U(Y - (1+\tau_A)\alpha_A - L + \beta_A) \\ & + (1-\pi)\eta_A(1-\nu_A)U(Y - (1+\tau_A)\alpha_A + \beta_A) \\ & + (1-\pi)\eta_A\nu_A U(Y - (1+\tau_A)\alpha_A - k) \\ & + (1-\pi)(1-\eta_A)U(Y - (1+\tau_A)\alpha_A) \end{aligned} \tag{PP_A}$$

subject to

$$\alpha_A = \pi\beta_A + (1-\pi)\beta_A\eta_A(1-\nu_A) + c\nu_A[\pi + (1-\pi)\eta_A] \tag{ZP_A}$$

$$\nu_A = \frac{U(Y - (1+\tau_A)\alpha_A + \beta_A) - U(Y - (1+\tau_A)\alpha_A)}{U(Y - (1+\tau_A)\alpha_A + \beta_A) - U(Y - (1+\tau_A)\alpha_A - k)} \tag{AC_A}$$

$$\eta_A = \left(\frac{c}{\beta_A - c}\right)\left(\frac{\pi}{1-\pi}\right) \tag{RC$_A$}$$

The first constraint, ZP_A, is a constraint of zero expected profits for the insurer. The premium the insurer collects must be equal to her expected payout, which includes the benefits paid, both to those who truly had a loss ($\pi\beta_A$) and those who were not caught committing fraud ($(1-\pi)\beta_A\eta_A(1-\gamma_A)$), and the cost of the insurer's auditing strategy. The second and third constraints represent the PBNE strategies of the players. When designing the contract the principal must anticipate rationally what strategies will be played. There is no need for a participation constraint since insurance is compulsory, and is equivalent to choosing $\beta_A^* = 0$.

Substituting the constraints into the objective function, the problem becomes

$$\max_{\beta_A} EU_A = \pi U\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right) + (1-\pi) U\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right) \tag{5}$$

The first order condition of this problem is

$$\begin{aligned} 0 = {} & \pi\left[1 - \pi\frac{(1+\tau_A)\beta_A(\beta_A - 2c)}{(\beta_A - c)^2}\right] U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right) \\ & - (1-\pi)(1+\tau_A)\pi\frac{\beta_A(\beta_A - 2c)}{(\beta_A - c)^2} U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right) \end{aligned} \tag{FOC$_A$}$$

We can see that the total tax collected by the government, assuming there are N policyholders in the economy is given by

$$T_A = N\tau_A\pi\frac{\beta_A^2}{\beta_A - c} \tag{6}$$

In economy B the problem is similar. Again we have that the principal designs a contract that specifies a benefit-premium couple that maximizes the agent's expected utility subject to the same constraints. The only difference between the two problems is how the tax is collected. In Economy B benefits are taxed instead of premiums. The problem for the principal in economy B is thus

$$\begin{aligned} \max_{\alpha_B;\beta_B} EU_B = {} & \pi U(Y - \alpha_B - L + (1-\tau_B)\beta_B) \\ & + (1-\pi)\eta_B(1-\gamma_B)U(Y - \alpha_B + (1-\tau_B)\beta_B) \\ & + (1-\pi)\eta_B\gamma_B U(Y - \alpha_B - \tau_B\beta_B - k) \\ & + (1-\pi)(1-\eta_B)U(Y - \alpha_B) \end{aligned} \tag{PP$_B$}$$

subject to

$$\alpha_B = \pi\beta_B + (1-\pi)\beta_B\delta_B(1-\gamma_B) + c\gamma_B[\pi + (1-\pi)\delta_B] \qquad (ZP_B)$$

$$\gamma_B = \frac{U(Y-\alpha_B+(1-\tau_B)\beta_B) - U(Y-\alpha_B)}{U(Y-\alpha_B+(1-\tau_B)\beta_B) - U(Y-\alpha_B-\tau_B\beta_B-k)} \qquad (AC_B)$$

$$\delta_B = \left(\frac{c}{\beta_B - c}\right)\left(\frac{\pi}{1-\pi}\right) \qquad (RC_B)$$

This problem simplifies to

$$\max_{\beta_B} EU_B = \pi U\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right) + (1-\pi)U\left(Y - \pi\frac{\beta_B^2}{\beta_B - c}\right) \qquad (7)$$

The first order condition of this problem is

$$0 = \pi\left[(1-\tau_B) - \pi\frac{\beta_B(\beta_B-2c)}{(\beta_B-c)^2}\right]U'\left(Y - \pi\frac{\beta_B^2}{\beta_B-c} - L + (1-\tau_B)\beta_B\right) \qquad (FOC_B)$$
$$-(1-\pi)\pi\frac{\beta_B(\beta_B-2c)}{(\beta_B-c)^2}U'\left(Y - \pi\frac{\beta_B^2}{\beta_B-c}\right)$$

We can see that the total tax collected by the government, assuming there are N policyholders in the economy is given by

$$T_B = N\tau_B\pi\frac{\beta_B^2}{\beta_B - c} \qquad (8)$$

It is interesting to note that the amount of money collected by the government in each economy depends on the same functional form. It is clear that if the benefit chosen in the two economies is the same (i.e., $\beta_A = \beta_B$) the tax rate must be the same for the government to collect the same amount. For T_A to equal T_B, it has to be that $N\tau_B\pi\frac{\beta_B^2}{\beta_B-c} = N\tau_A\pi\frac{\beta_A^2}{\beta_A-c}$. This means that for both taxes to raise the same amount of money the premium tax rate must be related to the benefit tax as

$$\tau_A = \tau_B\left(\frac{\beta_B^2}{\beta_A^2}\right)\left(\frac{\beta_A - c}{\beta_B - c}\right) \qquad (9)$$

We see that the only determinant of whether the premium tax rate is greater than the benefit tax rate is the equilibrium benefit. If the optimal benefit under a premium tax is lower (greater) than under a benefit tax, then the premium tax rate will be greater (lower) than the benefit tax rate.

In the initial state where $\tau_A = \tau_B = 0$, the government does not raise any money, and the two first order conditions are the same. Thus if $\tau_A = \tau_B = 0$, then the benefit chosen are the same and the premium are the same. This yields lemma 3.

Lemma 3 As $\tau_A \to 0$ and $\tau_B \to 0$, then $\beta_A \to \beta$ and $\beta_B \to \beta$.

Lemma 3 tells us that if the initial tax rates are zero then the coverage chosen in either economy is the same. It is telling us the same basic conclusion as lemma 1 in the benchmark case. We recall that the only parameters to have an impact on the tax rate differences were the equilibrium benefits chosen in each economy. From this third lemma we now know that these benefits are the same as the taxes approach zero. From equation (9) it follows that $\tau_A = \tau_B$. Thus the total tax yield is the same in the two economies.

Given the PBNE found in lemma 2, it has to be that the premium that yields zero profits to the insurer cannot be equal to the expected benefits. For example part of the premium has to pay for the costly audits the insurer conducts. The question then becomes whether policyholders are still indifferent between the two forms of taxation. The answer is no. Given that all tax rates are zero initially, a policyholder strictly prefers a marginal increase in his benefit tax rather than in his premium tax. This is presented as theorem 2, which is the most interesting result of the paper.

Theorem 2 As $\tau_A \to 0$ and $\tau_B \to 0$, an agent gets greater utility from a benefit tax than a premium tax.

What theorem 2 shows is that a policyholder faced with the possibility of paying a benefit tax or a premium tax prefers the former to the latter. This result stems primarily from the fact that policyholders have the possibility of defrauding their insurance company. This contrasts with the results obtained in theorem 1 where policyholders are indifferent between a premium tax and a benefit tax. If policyholders are able to misreport their loss, they strictly prefer to have their benefits taxed rather than their premiums.

Theorem 2 does not tell us, however, what are the policyholder's tax preferences away from $\tau_A \to 0$ and $\tau_B \to 0$. Graphs 1 and 2 in appendix 8.2 show that the results presented in theorem 2 hold for a range of government needs. These graphs were computed using functional forms for the policyholder's utility function (logarithmic in graph 1 and exponential in graph 2). The value of the different parameters are displayed in tables 1 and 2. In these two graphs we see for any amount in the computed range of government needs that the policyholder's expected utility is greater under a benefit tax scheme than under a premium tax scheme. We also see that the difference between expected utilities increases as the government's needs increases.

The reason why a benefit tax is preferred is two-fold. The first reason is that a greater burden of the tax is supported by agents who are successful at committing fraud. With a premium tax, everybody in the economy bears the same tax burden. Agents who commit fraud, however, bear a greater share of the total tax bill if a benefit tax is imposed. With a benefit tax the proportion of the total tax bill borne by agents who commit fraud is $\frac{(1-\pi)\eta}{\pi+(1-\pi)\eta}$, whereas with a premium tax the proportion is $(1-\pi)\eta$. Since $\pi+(1-\pi)\eta < 1$, it follows that a greater burden of the tax falls on agents who commit fraud. It is the case, however, that agent who truly had a loss must also bear a greater proportion of the tax. Nevertheless, since agents do not pay the tax as often with a benefit tax as with a premium tax, they are willing to take that chance in order for a greater burden to be assumed by those who commit fraud.

This effect is clearer when we examine the burden of the different agents as the probability of an accident occurring (π) varies. As π decreases, fewer agents must bear the tax bill, which means that each agent who suffered a loss must pay more. The flip side to this is that an even greater burden of the tax is supported by agents who commit fraud ($\left(\frac{\partial}{\partial\pi}\frac{(1-\pi)\eta}{\pi+(1-\pi)\eta}\right) < 0$). This means that although the tax bill is paid by fewer agents who had a true loss, their total tax bill is smaller relative to the tax bill of agents who commit fraud.

The second and most important reason why a benefit tax is preferred is that it provides a better tool for reducing fraud in the economy than a premium tax, as shown in the next section. Since the principal-agent problem is not solved (i.e.: it is not always in the agent's best interest to tell the truth), there is still fraud in the economy.[11] The difference between the benefit tax and the premium tax is that the benefit tax induces the principal to lose more (pay greater benefits) if she does not audit. The principal has, therefore, greater incentives to audit more often. Greater incentives for the principal to audit implies that the agent must reduce his probability of committing fraud if he wants the principal to remain indifferent

[11] It is clear that everybody would be better off if the principal could commit to an auditing strategy. Such a commitment would induce truth telling always on the part of the agent. We know from the revelation principle that amongst the set of best contracts there exists one such that the agent always tell the truth. It is straightforward to show that the contract derived here is Pareto inferior to the contract one would derive using the revelation principle.

between auditing and not auditing.

3.2 Impact on Fraud

We know from theorem 2 that agents strictly prefer a benefit tax to a premium tax when insurance fraud is present. Why is it then that in the real world we observe taxes on premiums and not on benefits if the latter tax yields higher expected utility? One possibility could be that policymakers have another agenda than the welfare of policyholders.

A possible item on their agenda could be the reduction of the amount of fraud in the economy. This seems reasonable to presume. Since fraud is a crime, its reduction could increase the welfare of everyone in the economy in a way that cannot be captured by the use of von Neumann-Morgenstern utility functions over final wealth. It then seems plausible that a tax on premiums would reduce the amount of fraud by a greater amount than a tax on benefit. This is not the case, however.

The total number of fraudulent claims in economy i is given by $TAF_i = N(1 - \pi)\eta_i$, where N is the number of policyholders in the economy. Substituting for η_i, the total number of fraudulent claims is $\frac{N\pi c}{\beta_i - c}$. It is easy to see that the total number of fraudulent claims is smaller in the economy where benefits are taxed if and only if $\beta_B > \beta_A$.[12] It then becomes sufficient to show that an agent buys more coverage in economy B to demonstrate that premium taxation is not the best way to reduce fraud using a tax scheme.

Looking at the two first order conditions (FOC_A and FOC_B) it is not obvious that β_B is greater than β_A. I need to show that β_B is greater than β_A in a more indirect manner. For example suppose that around $\tau_A \to 0$ and $\tau_B \to 0$, either β_A decreases faster than β_B when taxes increase, or β_A increases slower than β_B when taxes increase. If this supposition is right, then β_B is greater than β_A around $\tau_A \to 0$ and $\tau_B \to 0$. All we need to prove then is that $\frac{d\beta_A}{d\tau_A}$ evaluated at $\tau_A = 0$ is smaller than $\frac{d\beta_B}{d\tau_B}$ evaluated at $\tau_B = 0$. This leads to corollary 1.

Corollary 1 As $\tau_A \to 0$ and $\tau_B \to 0$, then $\frac{d\beta_A}{d\tau_A} < \frac{d\beta_B}{d\tau_B}$ and thus $\beta_B > \beta_A$.

What corollary 1 tells us is that a tax on the premium induces policyholders to choose a benefit that is smaller than the benefit chosen when the tax is levied on the benefits

[12] Working with the total amount of fraudulent benefit paid ($\frac{N\pi c\beta_i}{\beta_i - c}$) yields the same result.

received, at least when insurance is not taxed initially. It is well known in the literature that a proportional tax on insurance premiums, which acts as a proportional loading factor, reduces the coverage chosen by policyholders. This was shown amongst others by Arrow (1970). It was never showed, however, what happened when we tax insurance benefits. According to corollary 1, the premium tax reduces the optimal level of coverage by a greater amount than the benefit tax.

The reason why this result is obtained comes from the fact that the agent only cares about his after-tax benefit. When premiums are taxed, the agent's after-tax benefit is β_A. On the other hand, when benefits are taxed the agent's after-tax benefit is $(1-\tau_B)\beta_B$. Suppose the agent wanted the exact same after tax benefit. Since $\tau_B > 0$, it then has to be that $\beta_B > \beta_A$ for the agent to receive the same exact after tax benefit.

A consequence of corollary 1 is that a premium tax induces more fraud than a benefit tax. To see why, note that $\frac{\partial \eta}{\partial \beta} = -\frac{\pi}{1-\pi}\frac{c}{(\beta-c)^2}$ is negative. This means that as the benefit increases, the amount of fraud decreases. The reason is that with a higher β, the principal has more to lose by not auditing. Therefore the agent needs to reduce his probability of committing fraud (η) in order for the principal to remain indifferent between auditing and not auditing. From corollary 1 we know that the optimal benefit is greater at the margin when a benefit tax is imposed. It therefore follows that a premium tax induces more fraud than a benefit tax.

It is interesting to notice that $\frac{d\beta_A}{d\tau_A}$ is always negative, while $\frac{d\beta_B}{d\tau_B}$ may sometimes be positive. This means that an increase in the tax rate may involve a greater before-tax coverage for the policyholder. This is shown as corollary 2.

Corollary 2 An increase in the benefit tax around $\tau_B = 0$ increases in the pre-tax optimal benefit if and only if the agent's measure of absolute risk aversion measured at wealth $W_B^L = Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + \beta_B$ is such that

$$R_A(W_B^L) \geq \frac{1}{\beta_B\left(1-\frac{\pi\beta_B(\beta_B-2c)}{(\beta_B-c)^2}\right)} \tag{10}$$

This second corollary shows that by increasing the benefit tax rate one may reduce fraud. We know that fraud decreases when benefits increase. From corollary 2 we can then conclude that if the policyholder has a high enough level of absolute risk aversion in the state where he suffers a loss, then the pre-tax optimal benefit will increase, and the amount of fraud will therefore decrease.

Why is it then that taxes are collected on premiums rather than on benefits? Perhaps because of the uneasiness of taxing losses (which are what benefits are supposed to cover), state assemblies decided that a premium tax was the only acceptable way to reduce insurance fraud using taxes. The question then becomes whether a premium tax does in fact reduce the amount of fraud in the economy compared to no tax at all. The answer is no. Around the initial zero-tax point, a premium tax increases fraud. This is what corollary 3 shows.

Corollary 3 A sufficient condition for a premium tax to increase the amount of fraud in the economy (versus no tax at all) is for the agent's utility function to display non-increasing risk aversion.

Corollary 3 says that imposing a tax on insurance premium to fund an Insurance Fraud Bureau creates a market for that government agency; there is more fraud, and thus more business for the Insurance Fraud Bureau. This result depends, however, on the assumption of non-increasing absolute risk aversion. How reasonable is this assumption? Pratt (1964) and Arrow (1970) say that it is very reasonable (see also Eeckhoudt and Gollier, 1995).

The rationale behind this result is straightforward. Since a premium tax acts as a proportional loading on the insurance premium, and since we know that a proportional load reduces the optimal benefit, we can conclude that a premium tax reduces the optimal benefit. Combined with the fact that a smaller equilibrium benefit induces more fraud, it follows that an increase in the premium tax creates more fraud.

If a tax on insurance premiums induces more fraud, why would it be that funding for insurance fraud crime units comes from a tax on insurance premiums? It would seem more efficient to tax insurance benefits since it would cause a smaller increase (and perhaps a decrease)[13] in the amount of fraud. The last possible reason why taxes on premiums are so

[13] Fraud in an economy will decrease when benefits are taxed whenever

$$-\frac{U''(Y-\alpha_B-L+(1-\tau_B)\beta_B)}{U'(Y-\alpha_B-L+(1-\tau_B)\beta_B)} \geq \frac{(\beta_B-c)^2}{\beta_B[(1-\tau_B)(\beta_B-c)^2-\pi\beta_B(\beta_B-2c)]}$$

Which means that the coefficient of absolute risk aversion has to be greater than some minimum value.

popular is that they reduce the number of successful fraud in an economy. This is not even the case.

Corollary 4 A sufficient condition for a premium tax to increase the amount of successful fraud (versus no tax at all) is for the agent's utility function to display non-increasing risk aversion.

From the analysis presented above, it is possible to conclude four things. The first is that policyholders prefer to see their benefits taxed rather than their premiums; this is stated as theorem 2. The second is that a tax on the premiums not only is not the best way to reduce insurance fraud (taxing benefits is), it actually increase fraud compared to no tax at all. On the other hand, a tax on benefits may reduce the amount of fraud in the economy. These results are stated as corollary 1, corollary 3 and corollary 2 respectively. Finally corollary 4 shows that a premium tax will increase the number of fraudulent claims that are successful in the economy.

The reason why a benefit tax is better to curb insurance fraud than a premium tax is not because the policyholder has less to gain from committing fraud. In a mixed strategy equilibrium, the policyholder's payoff only affects the principal's optimal strategy. A policyholder who has less to gain from committing fraud will only see his probability of being audited go down; it does not affect his reporting strategy. His reporting strategy is affected by the impact taxes have on the payoff of the principal. This is the real reason why a benefit tax is better than a premium tax. By taxing premiums the government reduces the equilibrium benefit received by the agent in case of an accident. This means that the insurer has a greater payoff (lower negative payoff, $\beta_A < \beta_0$) in case of an accident. To keep the insurer indifferent between auditing and not auditing the agent must perpetrate fraud more often.

Similarly, when a tax is imposed on the benefits received, then the benefit paid by the insurer may increase under certain conditions. This means that the insurer has more to lose in the case of an accident, even though the agent does not necessarily collect more. Thus to keep the insurer indifferent between auditing and not auditing the agent needs to reduce the weight he assigns to the action that calls for the filing of a claim in case of no accident. Whether the policyholder collects more or less money from the insurer is irrelevant when choosing his optimal strategy. What he needs only consider is what are the payo¤s to the insurer.

3.3 Using Both Tax Schemes

Up to now only one tax could be used. The goal of this section is to analyze the optimal combination of taxes when the government is allowed to use both taxes in the economy. If only one is available, the tax that yields greater utility to the policyholder is a bene…t tax. When the two taxes can be used it is clear that the policyholder cannot be worse o¤ than when only the bene…t tax is available. What I want to show in this section is how the two tax schemes interact.

From the previous section's results, we can anticipate that the optimal combination of taxes will be such that bene…ts are taxed more than premiums since a bene…t tax is preferred to a premium tax. We can also anticipate the possibility that the bene…t tax may be set so high as to be able to o¤er premium subsidies.

Letting $® = (1 + ¿_A)¼\frac{-2}{¡c}$, the problem may then be stated as

$$\max_{¿_A;¿_B;\_} EU = ¼U(Y ¡ ® ¡ L + (1 ¡ ¿_B)^{-}) + (1 ¡ ¼)U(Y ¡ ®) \qquad (11)$$

subject to the government's zero-pro…t constraint[14]

$$(¿_A + ¿_B)¼\frac{-2}{¡c} = \frac{G}{N} = g \qquad (12)$$

Solving this problem allows me to state the last result of the paper.

Theorem 3 If both a premium tax and a bene…t tax can be imposed, and if the government's needs are zero (i.e. $g = 0$) then the optimal tax scheme is such that the …rst best attainable (in the limit). Furthermore, if $g > 0$, bene…ts are taxed in such a way that it may be possible to subsidize premiums. This occurs if $g < ¿_B¼\frac{-2}{¡c}$.

As anticipated most of the tax is collected using the bene…t tax.[15] What is surprising,

[14] We recall that the total tax collected by the government using a bene…t tax is $T_B = N¿_B¼\frac{-2}{¡c}$, while the total tax using a premium tax is $T_B = N¿_B¼\frac{-2}{¡c}$. Combining both taxes yields the goverment's constraint.

[15] It is obvious that imposing insurance bene…ts and redistributing the excess in the form of a premium subsidy Pareto-dominates imposing insurance bene…ts only. This is straightforward since only taxing bene…ts is the same as choosing $¿_A = 0$ in the two-tax system.

however, is that by using both tax schemes the first best allocation is almost attainable (in the limit that is). This is done trough a 100% tax on insurance benefits.

The reason the first best allocation is attainable as $\tau_B \to 1^-$ is that all waste due to insurance fraud is eliminated. This is done in threes ways. First, the total cost of auditing is driven to zero by a reduction in the probability of audit. To see why, recall that the probability that a claim is audited is given by

$$\nu = \frac{U(Y - \alpha + (1 - \tau_B)\beta) - U(Y - \alpha)}{U(Y - \alpha + (1 - \tau_B)\beta) - U(Y - \alpha - \tau_B\beta - k)} \tag{13}$$

By letting $\tau_B \to 1^-$, it is clear that the numerator tends to zero ($\nu \to 0$) while the denominator remain positive. Therefore the insurer no longer has any reason to audit because the policyholder has nothing to gain by committing fraud. Thus no resources are wasted in audits. The second reason why this allocation is optimal it that the probability a fraudulent claim is filed is drive to zero. To see why, recall that the agent's probability of committing fraud is given by

$$\eta = \left(\frac{\pi}{1 - \pi}\right)\left(\frac{c}{\beta - c}\right) \tag{14}$$

Letting $\beta = \frac{L}{1-\tau_B}$ (as obtained in the first best) it is clear that $\eta \to 0$ as $\tau_B \to 1^-$.

The third reason is that the agent never has to pay the penalty for cheating. Recall that this penalty is paid if and only if an agent is caught committing fraud through an audit. Since the principal's probability of audit is driven to zero, and the agent's probability of fraud is driven to zero, the probability that the agent pays the penalty is driven to zero as well.

These arguments are similar to the ones I made for theorem 2. The goal of the tax is to reduce the waste inherent to insurance fraud. By using the two tax schemes the reduction in fraud is such that the first best is attainable. When only one tax scheme was used, the tax scheme that reduced fraud the most (the benefit tax) was optimal.

4 Discussion and Conclusion

The goal of this paper was to present the problem faced by a government who wants to raise a certain amount of money through the insurance sector. Governments need more and more money to fund their projected insurance fraud crime bureaus, and it seems natural that such monies come from the insurance sector itself. One possible way to raise money through the insurance sector is by levying a sales tax on insurance premiums. This sort of tax exists in the United States and amounts on average to two percent of the premium paid. Another tax that is not used is a tax on insurance bene…ts.

In an environment where truth is always told and where full insurance is purchased, it was shown that agents are indi¤erent between a tax levied on bene…ts and one levied on premiums. Since agents are indi¤erent between the two taxes, and a premium tax may be more marketable politically, it seems normal that premium taxes were chosen as the preferred mean of insurance taxation. Policyholders are indi¤erent however, only in an environment where they never commit fraud. If it is costly for the principal to observe the agent's true loss, then the agent will prefer to see his bene…ts taxed. This was shown as theorem 2.

The reason why agents prefer a bene…t tax to a premium tax is three-fold: 1-it implicitly penalizes those who are successful committing fraud, 2-it may reduce the policyholder's incentive to commit fraud, and 3-it may increase the insurer's incentive to audit. By taxing bene…ts the governments ends up collecting money from the fraudulent elements of the economy. Even though they were successful in extracting rents from the insurer, those agents end up paying taxes on those rents. There may also be a reduction in the incidence of fraud since the bene…ts collected by the policyholder before tax may increase with a tax on insurance bene…ts. An increase in bene…ts reduces the probability that a policyholder commits fraud. This increase in bene…t is also the rationale behind an insurer's higher probability to audit. The possible reduction in the incidence of fraud and the possible (greater) reduction in the incidence of successful fraud are elements that explain why policyholders prefer bene…t taxation to premium taxation. I showed with an example that these theoretical …ndings may hold when the government's needs are greater. This suggests that the suboptimality of the premium tax is not due in the model to the …rst order approach used.

Since taxes on insurance premiums are not optimal, why are they being used to such an extent. A possible explanation is that policymakers do not care only about policyholders well being, they also care about crime in the society. Since insurance fraud is a crime, it is not unrealistic to presume that some policy makers who would like nothing more than to reduce fraud in the society. Taxing premiums seem to be a good way to achieve that goal. This is not the case, however. A premium tax not only fares worse that a bene…t tax in curbing fraud, it fares worse than no tax at all. A premium tax induces more fraud, as opposed to a bene…t tax that may actually reduce fraud.

Since premium taxes are Pareto-dominated by bene…ts taxes, and since premium taxes increase the incidence of fraud in the economy, one has to wonder why premiums are still being taxed. A possibility may come from the fact that not all agents behave opportunistically. In the real world the psychological loss may vary from one policyholder to the next. It is then possible that although premium taxes make those who potentially commit fraud worse o¤, it makes those who do not better o¤. However, this seems unlikely since agents who never commit fraud are indi¤erent at the margin between the two types of taxes.

Another possible explanation is that premium taxes are much easier to collect. A feature I haven't examined is how costly it is to administer the di¤erent tax programs. If it is very costly to administer a bene…ts tax scheme, then it may well be in everybody's best interest to use taxes on premiums.

5 References

1. Arrow, K (1970). Essays in the Theory of Risk-Bearing, North Holland, Amsterdam and London.

2. Bond, E. W. and K. J. Crocker (1997). Hardball and the Soft Touch: The Economics of Optimal Insurance Contracts with Costly State Veri…cation and Endogenous Monitoring. Journal of Public Economics, 63:239-254.

3. Boyer, M. M. (1998). Over-Compensation as a Partial Solution to Commitment and Renegotiation Problems: The Case of Ex Post Moral Hazard. Working Paper 98-04, Risk Management Chair, HEC-Université de Montréal.

4. Brown, R. A. (1993). Worker's Compensation: State Enactments in 1992. Monthly Labor Review, 116:50-55.

5. Castillo, M. K. (1997). How the Taxed In‡uence the Determination of State Tax Policy: The Case of the Life Insurance Industry. Working Paper, MIT.

6. Crocker, K. J. and J. Morgan (1997). The Optimality of Costly State Falsi…cation: Sharecropping and Insurance Fraud. Mimeo, University of Michigan.

7. Crocker, K. J. and J. Morgan (1998). Is Honesty the Best Policy? Curtailling Insurance Fraud Through Optimal Contracts. Journal of Political Economy, 106:355-375.

8. Eeckhoudt, L. and C. Gollier (1995). Risk: Evaluation, Management and Sharing. Harvester Wheatsheaf, Hertfordshire.

9. Gibbons, R. (1992). Game Theory for Applied Economists. Princeton University Press, 1992.

10. Mookherjee, D. and I. Png (1989). Optimal Auditing, Insurance and Redistribution. Quarterly Journal of Economics, 104:205-228.

11. Mookherjee, D. and I. Png (1990). Enforcement Costs and the Optimal Progressivity of Taxes. Journal of Law, Economics and Organizations, 6:411-431.

12. Myerson, R.B. (1991). Game Theory. Harvard University Press, Cambridge, MA.

13. National Association of Insurance Commissioners. 1995 Data Tapes. Center for Risk Management and Insurance Research, Georgia State University.

14. National Underwriter, vol.99 no.18, p.11. 1995.

15. National Underwriter, vol.99 no.14, p.3. 1995.

16. Pratt, J. (1964). Risk Aversion in the Small and in the Large. Econometrica, 32:122-136.

17. Picard, P. (1996). Auditing Claims in the Insurance Market with Fraud: The Credibility Issue. Journal of Public Economics, 63:27-56.

18. Rand Corporation Institute for Civil Justice. Research Brief, May 1995.

19. Spence, A. M. and R. Zeckhauser (1971). Insurance, Information and Individual Action. American Economic Review, 61:380-387.

20. Townsend, R. M. (1979). Optimal Contracts and Competitive Markets with Costly State Veri…cation. Journal of Economic Theory, 21:265-293.

6 Appendix

6.1 Proofs

Proof of lemma 1. In economy A the first order condition as $\tau_A \to 0$ can be rewritten as

$$\begin{aligned}
0 &= \lim_{\tau_A \to 0} \pi[1 - \pi(1+\tau_A)]U'(Y - (1+\tau_A)\pi\beta_A - L + \beta_A) \\
&\quad - \lim_{\tau_A \to 0} \pi(1-\pi)(1+\tau_A)U'(Y - (1+\tau_A)\pi\beta_A) \\
&= \pi(1-\pi)U'(Y - \pi\beta_A - L + \beta_A) - \pi(1-\pi)U'(Y - \pi\beta_A) \qquad (15)
\end{aligned}$$

if and only if

$$1 = \frac{U'(Y - \pi\beta_A)}{U'(Y - \pi\beta_A - L + \beta_A)} \qquad (16)$$

While in economy B the first order condition as $\tau_B \to 0$ becomes

$$\begin{aligned}
0 &= \lim_{\tau_B \to 0} \pi[1 - \tau_B - \pi]U'(Y - \pi\beta_B - L + (1-\tau_B)\beta_B) \\
&\quad - \lim_{\tau_B \to 0} \pi(1-\pi)U'(Y - \pi\beta_B) \\
&= \pi(1-\pi)U'(Y - \pi\beta_B - L + \beta_B) - \pi(1-\pi)U'(Y - \pi\beta_B) \qquad (17)
\end{aligned}$$

if and only if

$$1 = \frac{U'(Y - \pi\beta_B)}{U'(Y - \pi\beta_B - L + \beta_B)} \qquad (18)$$

Since the first order conditions are the same when $\tau_A \to 0$ and $\tau_B \to 0$, then $\beta_A = \beta_B = \beta$. ∎

Proof of theorem 1. What we want to show here is that

$$\lim_{\tau_A \to 0} \frac{dW_A}{d\tau_A} = \lim_{\tau_B \to 0} \frac{dW_B}{d\tau_B} \iff \lim_{\tau_A \to 0} \frac{\partial W_A}{\partial \tau_A} = \lim_{\tau_B \to 0} \frac{\partial W_B}{\partial \tau_B} \qquad (19)$$

Solving for the partial derivatives[16] yields

$$\begin{aligned}
0 = &- \lim_{\tau_A \to 0} \alpha_A \pi U'(Y - (1+\tau_A)\alpha_A - L + \beta_A) \qquad (20) \\
&+ \lim_{\tau_A \to 0} \alpha_A (1-\pi)U'(Y - (1+\tau_A)\alpha_A) \\
&- \lim_{\tau_B \to 0} \pi\beta_B U'(Y - \alpha_B - L + (1-\tau_B)\beta_B)
\end{aligned}$$

[16] Using the envelope theorem:

$$\frac{dW_i}{d\tau_i} = \frac{\partial W_i}{\partial \tau_i} + \frac{\partial W_i}{\partial \beta_i}\frac{\partial \beta_i}{\partial \tau_i} = \frac{\partial W_i}{\partial \tau_i}$$

since $\frac{\partial W_i}{\partial \beta_i} = 0$ by the first order condition.

Using lemma 1 and the fact that $\alpha = \pi\beta$, equation (20) simplifies to

$$0 = [\pi U'(Y - \alpha - L + \beta) + (1 - \pi)U'(Y - \alpha)] - U'(Y - \alpha - L + \beta) \tag{21}$$

and

$$U'(Y - \alpha - L + \beta) = U'(Y - \alpha) \tag{22}$$

which is the first order condition of the problem as $\tau_A \to \tau_B \to 0$. Therefore the agent is indifferent between the two marginal tax increases. $\square$

Proof of lemma 2. Using backward induction, the proof is similar to Gibbons (1992). Looking at the left hand side of figure 2, it is clear that $\gamma(DF) = 0$. Suppose there is an accident. Then filing a claim (FC) dominates not filing (DF), whatever the principal does. By not filing the best the agent can do is get a payoff of $-L$. On the other hand by filing a claim, the payoff to the agent is $\beta - L$.

When the principal sees that the agent played DF, she knows for sure that she is not playing at the upper node of information set [1.0]. Therefore the principal knows with probability one that she is at the lower node of information set [1.0]. Consequently the only meaningful strategy for the principal when DF is played is to never audit. This is straightforward since she gets $-c$ if she audits, and 0 if she does not. We have now found three of the six elements of the sextuplet. Lets now move to the right side of the figure where things are much more interesting.

Let ν_i be the probability (in a mixed strategy sense) of auditing a filed claim. The strategy of the principal at information set [1.L] must be such that the agent is indifferent between telling the truth and committing fraud, given that he was not involved in an accident. For the agent to be indifferent between telling the truth and lying, ν_i must solve:

$$U(Y - \hat{\alpha}_i) = \nu_i U(Y - \hat{\alpha}_i - \hat{k}_i) + (1 - \nu_i)U(Y - \hat{\alpha}_i + \hat{\beta}_i) \tag{23}$$

which means that

$$\nu_i = \frac{U(Y - \hat{\alpha}_i + \hat{\beta}_i) - U(Y - \hat{\alpha}_i)}{U(Y - \hat{\alpha}_i + \hat{\beta}_i) - U(Y - \hat{\alpha}_i - \hat{k}_i)} \tag{24}$$

where

$$\left(\hat{\alpha}_i; \hat{\beta}_i; \hat{k}_i\right) = \begin{cases} ((1 + \tau_A)\alpha_A; \beta_A; k) & \text{if } i = A \\ (\alpha_B; (1 - \tau_B)\beta_B; \tau_B\beta_B + k) & \text{if } i = B \end{cases} \tag{25}$$

We now have four of the six elements of our sextuplet PBNE. All that is left to calculate is the belief of the insurer at information set [1.L] and the strategy of the agent given that there was no accident. Let η_{-} be the probability (in the mixed strategy sense) that the agent files a claim when there was no accident. In other words, η_{-} is the probability that the agent commits fraud. By Bayes' rule we can find the exact value of $\gamma(FC)$, the principal's posterior belief that there was indeed an accident given that the agent filed a claim. $\gamma(FC)$ is equal to

$$\gamma(FC) = \frac{\pi}{\pi + (1-\pi)\eta_{-}} \tag{26}$$

Only one strategy of the agent will induce the principal to be indifferent between auditing and not auditing. That strategy must be such that $\gamma(FC)$ solves

$$(-c-\beta_{-})\gamma(FC) + (-c)(1-\gamma(FC)) = -\beta_{-} \tag{27}$$

and

$$\gamma(FC) = \frac{\beta_{-} - c}{\beta_{-}} \tag{28}$$

Substituting for $\gamma(FC)$ in (26), yields that the agent's probability of committing fraud is[17]

$$\eta_{-} = \left(\frac{c}{\beta_{-} - c}\right)\left(\frac{\pi}{1-\pi}\right) \tag{29}$$

Since all six elements of our PBNE have been found, the proof is done. $\square$

Proof of lemma 3. In economy A the first order condition as $\tau_A \to 0$ becomes

$$\frac{U'(Y-\alpha_A - L + \beta_A)\frac{(\beta_A - c)^2}{(1+\lambda)\beta_A(\beta_A - 2c)}}{\pi U'(Y-\alpha_A - L + \beta_A) + (1-\pi)U'(Y-\alpha_A)} = 1 \tag{30}$$

While in economy B the first order condition as $\tau_B \to 0$ becomes

$$\frac{U'(Y-\alpha_B - L + \beta_B)\frac{(\beta_B - c)^2}{(1+\lambda)\beta_B(\beta_B - 2c)}}{\pi U'(Y-\alpha_B - L + \beta_B) + (1-\pi)U'(Y-\alpha_B)} = 1 \tag{31}$$

Since the first order conditions are the same when $\tau_A \to 0$ and $\tau_B \to 0$ then $\beta_A = \beta_B = \beta$. $\square$

Proof of theorem 2. Let V_A (V_B) represent the policyholder's expected utility when a marginal premium tax (benefit tax) is considered. Using lemma 3 it is clear that if $\tau_A =$

[17] Notice that we need to assume that $\pi < \frac{\beta_{-} - c}{\beta_{-}}$ for the reporting probability to be in the zero-one interval. If not, then the agent will always commit fraud when he has a low loss. A sufficient condition is to assume that $\pi < \frac{1}{2}$ since, as we can see in the first order condition $\beta_{-} > 2c$.

$\tau_B = 0$, then $V_A = V_B$. I want to show that $\frac{dV_A}{d\tau_A} < \frac{dV_B}{d\tau_B}$. Using the envelope theorem, it is sufficient to show that

$$\lim_{\tau_A \to 0} \frac{dV_A}{d\tau_A} < \lim_{\tau_B \to 0} \frac{dV_B}{d\tau_B} \tag{32}$$

This equation holds if and only if

$$\begin{aligned} 0 > \; & -\lim_{\tau_A \to 0} \alpha_A \pi U'(Y - (1+\tau_A)\alpha_A - L + \beta_A) \\ & -\lim_{\tau_A \to 0} \alpha_A (1-\pi) U'(Y - (1+\tau_A)\alpha_A) \\ & +\lim_{\tau_B \to 0} \pi \beta_B U'(Y - \alpha_B - L + (1-\tau_B)\beta_B) \end{aligned} \tag{33}$$

Using lemma 3, yields

$$\begin{aligned} 0 > \; & -\alpha\pi U'(Y - (1+\tau_A)\alpha_A - L + \beta_A) \\ & +\alpha_A (1-\pi) U'(Y - (1+\tau_A)\alpha_A) \\ & -\pi\beta_B U'(Y - \alpha_B - L + (1-\tau_B)\beta_B) \end{aligned} \tag{34}$$

This holds if and only if

$$\frac{\pi}{1-\pi}\left[\frac{\beta}{\alpha} - 1\right] < \frac{U'(Y-\alpha)}{U'(Y-\alpha-L+\beta)} \tag{35}$$

From the first order conditions (FOC_A) or (FOC_B), we know that

$$\frac{U'(Y-\alpha)}{U'(Y-\alpha-L+\beta)} = \frac{(\beta-c)^2 - \pi\beta(\beta-2c)}{(1-\pi)\beta(\beta-2c)} \tag{36}$$

Substituting this equation in the previous one and substituting for $\alpha = \pi\frac{\beta^2}{\beta - c}$ yields

$$\frac{\pi}{1-\pi}\left[\frac{\beta - \pi\frac{\beta^2}{\beta-c}}{\pi\frac{\beta^2}{\beta-c}}\right] < \frac{(\beta-c)^2 - \pi\beta(\beta-2c)}{(1-\pi)\beta(\beta-2c)} \tag{37}$$

Simplifying, gives us that a marginal increase in τ_A yields lower expected utility than a marginal increase in τ_B if an only if $c > 0$, which is the case by A.5. ■

Proof of corollary 1. We know that $\beta_A < \beta_B$ if and only if $\frac{d\beta_A}{d\tau_A} < \frac{d\beta_B}{d\tau_B}$ around $\tau_A = \tau_B = 0$. The first order condition in Economy A is

$$\begin{aligned} 0 = \; & (\beta_A - c)^2 U'(Y - (1+\tau_A)\alpha_A - L + \beta_A) \\ & -\pi(1+\tau_A)\beta_A(\beta_A - 2c)U'(Y - (1+\tau_A)\alpha_A - L + \beta_A) \\ & -(1-\pi)(1+\tau_A)\beta_A(\beta_A - 2c)U'(Y - (1+\tau_A)\alpha_A) \end{aligned} \tag{38}$$

Let Ψ_A represent the right-hand side of (38). Using total derivatives, the impact of a change in the tax rate on the coverage chosen is equal to $\frac{d\beta_A}{d\tau_A} = -\frac{\frac{\partial\Psi_A}{\partial\tau_A}}{\frac{\partial\Psi_A}{\partial\beta_A}}$. $\frac{\partial\Psi_A}{\partial\beta_A}$ is equal to

$$
\begin{aligned}
\frac{\partial\Psi_A}{\partial\beta_A} = {} & 2(\beta_A - c)\left[1 - \pi(1+\tau_A)\right]U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right) \qquad (39)\\
& - 2(\beta_A - c)\left[1 - \pi(1+\tau_A)\right]U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right)\\
& - 2\tau_A(\beta_A - c)U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right)\\
& + (1-\pi)\frac{\pi(1+\tau_A)\beta_A(\beta_A - 2c)}{(\beta_A - c)^2}]U''\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right)\\
& + (\beta_A - c)^2\left[1 - \frac{\pi(1+\tau_A)\beta_A(\beta_A - 2c)}{(\beta_A - c)^2}\right]^2 U''\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right)
\end{aligned}
$$

This expression is negative since $U'(.) > 0$, $U''(.) < 0$ and $\beta_A > L$ around $\tau_A = 0$. $\frac{\partial\Psi_A}{\partial\tau_A}$ equals

$$
\begin{aligned}
\frac{\partial\Psi_A}{\partial\tau_A} = {} & -\alpha_A(\beta_A - c)^2U''\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right) \qquad (40)\\
& + \alpha_A\pi(1+\tau_A)\beta_A(\beta_A - 2c)U''\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right)\\
& - \pi\beta_A(\beta_A - 2c)U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c} - L + \beta_A\right)\\
& - (1-\pi)\beta_A(\beta_A - 2c)U'\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right)\\
& + \alpha_A(1-\pi)(1+\tau_A)\beta_A(\beta_A - 2c)U''\left(Y - (1+\tau_A)\pi\frac{\beta_A^2}{\beta_A - c}\right)
\end{aligned}
$$

The same technique can be used in economy B to find the value of $\frac{d\beta_B}{d\tau_B}$. Using the first order condition, FOC_B, we get

$$
\begin{aligned}
0 = {} & (1-\tau_B)(\beta_B - c)^2U'(Y - (1+\lambda)\alpha_B - L + (1-\tau_B)\beta_B) \qquad (41)\\
& - \pi\beta_B(\beta_B - 2c)U'(Y - (1+\lambda)\alpha_B - L + (1-\tau_B)\beta_B)\\
& - (1-\pi)\beta_B(\beta_B - 2c)U'(Y - (1+\lambda)\alpha_B)
\end{aligned}
$$

Let Ψ_B equal the right-hand side of (41). Using the same technique as above, $\frac{\partial\Psi_B}{\partial\beta_B}$ is equal to

$$
\begin{aligned}
\frac{\partial \Pi_B}{\partial \beta_B} = {} & 2(\beta_B - c)(1-\pi)U'\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right) \\
& - 2(\beta_B - c)(1-\pi)U'\left(Y - \pi\frac{\beta_B^2}{\beta_B - c}\right) \\
& - 2\tau_B(\beta_B - c)U'\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right) \\
& + (1-\pi)\frac{\pi\beta_B(\beta_B - 2c)}{(\beta_B - c)^2}U''\left(Y - \pi\frac{\beta_B^2}{\beta_B - c}\right) \\
& + (\beta_B - c)^2\left[(1-\tau_B) - \frac{\pi\beta_B(\beta_B - 2c)}{(\beta_B - c)^2}\right]^2 U''\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right)
\end{aligned} \tag{42}
$$

which is negative since $U'(\cdot) > 0$, $U''(\cdot) < 0$ and $\beta_B > L$ around $\tau_B = 0$. Finally $\frac{\partial \Pi_B}{\partial \tau_B}$ is equal to

$$
\begin{aligned}
\frac{\partial \Pi_B}{\partial \tau_B} = {} & \pi\beta_B^2(\beta_B - 2c)U''\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right) \\
& - \beta_B(1-\tau_B)(\beta_B - c)^2U''\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right) \\
& - (\beta_B - c)^2U'\left(Y - \pi\frac{\beta_B^2}{\beta_B - c} - L + (1-\tau_B)\beta_B\right)
\end{aligned} \tag{43}
$$

To complete the proof, notice that (39) and (42) are very similar. In fact when $\tau_A \to 0$ and $\tau_B \to 0$ it is easy to derive that $\frac{\partial \Pi_A}{\partial \beta_A} = \frac{\partial \Pi_B}{\partial \beta_B} < 0$ (using lemma 3 in the process). Thus $\frac{d\beta_A}{d\tau_A} \leq \frac{d\beta_B}{d\tau_B}$ if and only if $\frac{\partial \Pi_A}{\partial \tau_A} \leq \frac{\partial \Pi_B}{\partial \tau_B}$. All that is left to prove is that $\frac{\partial \Pi_A}{\partial \tau_A}$ is indeed smaller or equal to $\frac{\partial \Pi_B}{\partial \tau_B}$. This occurs if and only if

$$
\begin{array}{c}
-\alpha_A(\beta_A - c)^2U''(Y - \tilde{\alpha}_A - L + \beta_A) \\
+\tilde{\alpha}_A\pi\beta_A(\beta_A - 2c)U''(Y - \tilde{\alpha}_A - L + \beta_A) \\
-\pi\beta_A(\beta_A - 2c)U'(Y - \tilde{\alpha}_A - L + \beta_A) \\
-(1-\pi)\beta_A(\beta_A - 2c)U'(Y - \tilde{\alpha}_A) \\
+\tilde{\alpha}_A(1-\pi)\beta_A(\beta_A - 2c)U''(Y - \tilde{\alpha}_A)
\end{array}
\leq
\begin{array}{c}
\pi\beta_B^2(\beta_B - 2c)U''\left(Y - \alpha_B - L + \tilde{\beta}_B\right) \\
-\tilde{\beta}_B(\beta_B - c)^2U''\left(Y - \alpha_B - L + \tilde{\beta}_B\right) \\
-(\beta_B - c)^2U'\left(Y - \alpha_B - L + \tilde{\beta}_B\right)
\end{array} \tag{44}
$$

where $\tilde{\alpha}_A = \alpha_A(1+\tau_A)$ and $\tilde{\beta}_B = \beta_B(1-\tau_B)$. Taking the limit as $\tau_A \to 0$ and $\tau_B \to 0$ (which means that $\beta_A = \beta_B = \beta$) and manipulating yields

$$
\begin{aligned}
& \left[1 - \pi\frac{\beta(\beta - 2c)}{(\beta - c)^2}\right]U'(Y - \alpha - L + \beta) \\
- & \left[1 - \pi\frac{\beta(\beta - 2c)}{(\beta - c)^2}\right](\alpha - \beta)U''(Y - \alpha - L + \beta) \leq 0 \\
- & (1-\pi)\frac{\beta(\beta - 2c)}{(\beta - c)^2}\left[U'(Y - \alpha) - \alpha U''(Y - \alpha)\right]
\end{aligned} \tag{45}
$$

Substituting the first order condition in (45) and manipulating further yields

$$-\frac{U'(Y-\alpha-L+\beta)}{U''(Y-\alpha-L+\beta)}\alpha \geq -\frac{U'(Y-\alpha)}{U''(Y-\alpha)}(\alpha-\beta) \tag{46}$$

replacing α by its value in terms of β, we get

$$R_0 \geq R_L\left(1-\frac{\beta-c}{\pi\beta}\right) \tag{47}$$

where R_L and R_0 are the measures of absolute risk aversion in the loss and no-loss state ($R_L > 0$, $R_0 > 0$). It is easily shown that $1-\frac{\beta-c}{\pi\beta} < 0$ if and only if $(1-\pi)\beta > c$, which is always the case if an equilibrium in mixed strategies to the game exists. Therefore (47) always holds.[18] $\square$

Proof of corollary 2. We want to show under what conditions will $\frac{d\beta_B}{d\tau_B}$ be greater than zero around $\tau_B = 0$. Using the same technique as in the proof of corollary 1, we have that $\frac{d\beta_B}{d\tau_B} = -\frac{\frac{\partial\Sigma_B}{\partial\tau_B}}{\frac{\partial\Sigma_B}{\partial\beta_B}}$. Since we know that $\frac{\partial\Sigma_B}{\partial\beta_B}$ is always negative, then $\text{sign}(\frac{d\beta_B}{d\tau_B}) = \text{sign}(\frac{\partial\Sigma_B}{\partial\tau_B})$. All is left to show then is under what conditions will $\frac{\partial\Sigma_B}{\partial\tau_B}$ be positive for $\frac{d\beta_B}{d\tau_B}$ to be positive as well. We recall that

$$\begin{aligned}\frac{\partial\Sigma_B}{\partial\tau_B} = {}& \pi\beta_B^2(\beta_B-2c)U''\left(Y-\pi\frac{\beta_B^2}{\beta_B-c}-L+(1-\tau_B)\beta_B\right) \\ & -\beta_B(1-\tau_B)(\beta_B-c)^2U''\left(Y-\pi\frac{\beta_B^2}{\beta_B-c}-L+(1-\tau_B)\beta_B\right) \\ & -(\beta_B-c)^2U'\left(Y-\pi\frac{\beta_B^2}{\beta_B-c}-L+(1-\tau_B)\beta_B\right)\end{aligned} \tag{48}$$

We then have that $\frac{\partial\Sigma_B}{\partial\tau_B} \geq 0$ around $\tau_B = 0$ if and only if

$$-\frac{U''\left(Y-\pi\frac{\beta_B^2}{\beta_B-c}-L+\beta_B\right)}{U'\left(Y-\pi\frac{\beta_B^2}{\beta_B-c}-L+\beta_B\right)} \geq \frac{1}{\beta_B\left(1-\frac{\pi\beta_B(\beta_B-2c)}{(\beta_B-c)^2}\right)} \tag{49}$$

Which means that the measure of absolute risk aversion measured at $Y-\pi\frac{\beta_B^2}{\beta_B-c}-L+\beta_B$ be greater than some endogenous value that goes to zero when $\beta_B \to \infty$. This completes the proof. $\square$

[18] If on the other hand we had $\lambda > 0$, we would get that $R_0 \geq R_L(1+\lambda-\frac{\beta-c}{\pi\beta})$. In this case a sufficient condition for a benefit tax to be prefered would be $\lambda < \frac{(1-\pi)\beta-c}{\pi\beta}$.

Proof of corollary 3. The total amount of fraud when there is no tax (TAF_0) and with a tax on premiums (TAF_A) are given by $\frac{\pi c}{\beta - c}$ and $\frac{\pi c}{\beta_A - c}$ respectively. $TAF_A > TAF_0$ if and only if $\beta_A < \beta$, which occurs when $\frac{d\beta_A}{d\tau_A} < 0$. Since we know from corollary 1 that $\frac{\partial \Phi_A}{\partial \beta_A} < 0$, it follows that $\frac{d\beta_A}{d\tau_A} < 0$ if and only if $\frac{\partial \Phi_A}{\partial \beta_A} < 0$. This occurs if and only if

$$\frac{\begin{array}{c}\pi U'(Y - (1+\tau_A)\alpha_A - L + \beta_A) \\ + (1-\pi) U'(Y - (1+\tau_A)\alpha_A - L + \beta_A)\end{array}}{\alpha_B (1+\tau_A)(1-\pi)} = \begin{array}{c} U''(Y - (1+\tau_A)\alpha_A) \\ - U''(Y - (1+\tau_A)\alpha_A - L + \beta_A)\Gamma \end{array} \tag{50}$$

where

$$\Gamma = \frac{(\beta_A - c)^2 - \pi(1+\tau_A)\beta_A(\beta_A - 2c)}{(1-\pi)(1+\tau_A)\beta_A(\beta_A - 2c)}$$

It is obvious that the left-hand side is always positive. It is then sufficient to show that the term on the right-hand side is non-positive to complete the proof. This occurs if and only if

$$\frac{U''(Y - (1+\tau_A)\alpha_A)}{U''(Y - (1+\tau_A)\alpha_A - L + \beta_A)} \geq \frac{(\beta_A - c)^2 - \pi(1+\tau_A)\beta_A(\beta_A - 2c)}{(1-\pi)(1+\tau_A)\beta_A(\beta_A - 2c)} \tag{51}$$

Substituting the first order condition into (51), and manipulating yields

$$-\frac{U''(Y - (1+\tau_A)\alpha_A)}{U'(Y - (1+\tau_A)\alpha_A)} \geq -\frac{U''(Y - (1+\tau_A)\alpha_A - L + \beta_A)}{U'(Y - (1+\tau_A)\alpha_A - L + \beta_A)} \tag{52}$$

which is always the case if the utility function displays non-increasing risk aversion: $R_0 \geq R_L$ since $\beta_A > L$ around $\tau_A = 0$. $\square$

Proof of corollary 4. The total number of successful fraudulent claims when there is no tax (TSF_0) and with a tax on premiums (TSF_A) are given by

$$TSF_0 = \pi \frac{c}{\beta - c}\left[\frac{U(Y-\alpha) - U(Y-\alpha-k)}{U(Y-\alpha+\beta) - U(Y-\alpha-k)}\right] \tag{53}$$

$$TSF_A = \pi \frac{c}{\beta_A - c}\left[\frac{U(Y-(1+\tau_A)\alpha_A) - U(Y-(1+\tau_A)\alpha_A-k)}{U(Y-(1+\tau_A)\alpha_A+\beta_A) - U(Y-(1+\tau_A)\alpha_A-k)}\right] \tag{54}$$

It is clear that

$$\frac{U(Y-\alpha) - U(Y-\alpha-k)}{U(Y-\alpha+\beta) - U(Y-\alpha-k)} < \frac{U(Y-(1+\tau_A)\alpha_A) - U(Y-(1+\tau_A)\alpha_A-k)}{U(Y-(1+\tau_A)\alpha_A+\beta_A) - U(Y-(1+\tau_A)\alpha_A-k)} \tag{55}$$

around $\tau_A = 0$ if and only if $\frac{\partial(1-\gamma_A)}{\partial \tau_A} > 0$. This occurs when

$$\frac{\begin{array}{c}U'(Y-(1+\tau_A)\alpha_A) \\ - U'(Y-(1+\tau_A)\alpha_A-k)\end{array}}{\begin{array}{c}U'(Y-(1+\tau_A)\alpha_A+\beta_A) \\ - U'(Y-(1+\tau_A)\alpha_A-k)\end{array}} > \frac{\begin{array}{c}U(Y-(1+\tau_A)\alpha_A) \\ - U(Y-(1+\tau_A)\alpha_A-k)\end{array}}{\begin{array}{c}U(Y-(1+\tau_A)\alpha_A+\beta_A) \\ - U(Y-(1+\tau_A)\alpha_A-k)\end{array}} \tag{56}$$

which is always the case when $U'''(\cdot) < 0$. Thus a sufficient condition for $TSF_A > TSF_0$ is that $TAF_A > TAF_0$. We know from corollary 2 that a sufficient condition for $TAF_A > TAF_0$ is that the utility displays non-increasing absolute risk aversion. $\blacksquare$

Proof of theorem 3. The first order conditions of this problem are

$$\begin{aligned}\frac{\partial EU}{\partial \beta} &= 0 = \pi\left[(1-\tau_B) - \Phi\left(\frac{\beta-2c}{\beta(\beta-c)}\right)\right]U'(Y-\Phi-L+(1-\tau_B)\beta) \\ &\quad - (1-\pi)\Phi\left(\frac{\beta-2c}{\beta(\beta-c)}\right)U'(Y-\Phi) + \mu(\tau_A+\tau_B)\pi\frac{\beta(\beta-2c)}{(\beta-c)^2}\end{aligned} \tag{57}$$

$$\frac{\partial EU}{\partial \tau_A} = 0 = \left[\mu - \pi U'(Y-\Phi-L+(1-\tau_B)\beta) - (1-\pi)U'(Y-\Phi)\right]\pi\frac{\beta^2}{\beta-c} \tag{58}$$

$$\frac{\partial EU}{\partial \tau_B} = 0 = -\pi\beta U'(Y-\Phi-L+(1-\tau_B)\beta) + \mu\pi\frac{\beta^2}{\beta-c} \tag{59}$$

$$\frac{\partial EU}{\partial \mu} = 0 = (\tau_A+\tau_B)\pi\frac{\beta^2}{\beta-c} - g \tag{60}$$

where $\Phi = (1+\tau_A)\pi\frac{\beta^2}{\beta-c}$ and μ is the Lagrange multiplier. What I want to show is that the first best is almost reached. This occurs if $L = (1-\tau_B)\beta$, and $\Phi = \pi L$. To see why, substitute $L = (1-\tau_B)\beta$ and $\Phi = \pi L$ into the function to maximize show in (11). This yields

$$EU = \pi U(Y - \pi L - (1-\tau_B)\beta + (1-\tau_B)\beta) + (1-\pi)U(Y-\pi L) \tag{61}$$

This is nothing else that the problem faced by the players when no fraud occurs: this yields the first best outcome. Suppose therefore that $\beta = \frac{L}{1-\tau_B}$, which means

$$\begin{aligned}\Phi &= (1+\tau_A)\pi\frac{\beta^2}{\beta-c} = (1+\tau_A)\pi\frac{\left(\frac{L}{1-\tau_B}\right)^2}{\frac{L}{1-\tau_B}-c} \\ &= \pi\left(\frac{1+\tau_A}{1-\tau_B}\right)\left(\frac{L^2}{L-(1-\tau_B)c}\right)\end{aligned} \tag{62}$$

From (60), it is clear that if $g = 0$, $\tau_A = -\tau_B$. We then have that $\left(\frac{1+\tau_A}{1-\tau_B}\right) = 1$ and

$$\Phi = \pi\left(\frac{L^2}{L-(1-\tau_B)c}\right) \tag{63}$$

Thus, $\Phi = \pi L$ if and only if $\tau_B \to 1^-$.[19]

[19] I shall complete the proof by showing what happens as when τ_B is exactly equal to 1. By taking $\tau_B \to 1$ we reach a maximum, while by letting $\tau_B = 1$ we reach a minimum.

Let us see if the first order conditions hold as $\tau_B \to 1^-$. From (58) it is clear that

$$\mu = U'(Y - \pi L) \tag{64}$$

Substituting in (57) and simplifying yields

$$\beta(\beta - 2c) = (\beta - c)^2 \tag{65}$$

Letting $\beta = \frac{L}{1-\tau_B}$, we get

$$\frac{L}{(1-\tau_B)^2}(L - 2(1-\tau_B)c) = \frac{(L-(1-\tau_B)c)^2}{(1-\tau_B)^2} \tag{66}$$

The two sides of this equation are equal if and only if $\tau_B \to 1^-$, which is what we wanted.

Finally, from (59) we have that

$$\frac{1}{\mu}U'\left(Y - (1+\tau_A)\pi\frac{\beta^2}{\beta - c} - L + (1-\tau_B)\beta\right) = \frac{\beta}{\beta - c} \tag{67}$$

Since $\mu = U'(Y - \pi L)$, this equation holds if $\frac{\beta}{\beta - c} = 1$. Substituting for $\beta = \frac{L}{1-\tau_B}$ yields

$$\frac{\frac{L}{1-\tau_B}}{\frac{L}{1-\tau_B} - c} = 1 \tag{68}$$

Which occurs if and only if $\tau_B \to 1^-$, which is what we wanted. Therefore as $\tau_B \to 1^-$, the first best allocation may be achieved..

The last thing I need to show is that $\tau_B \to 1^+$ is a minimum. The way I shall proceed is to show that $\tau_B \to 1^+$ results in the same allocation as in autarchy. Autarchy occurs when

$$EU = \pi U(Y - L) + (1-\pi)U(Y) \tag{69}$$

This is equivalent to choosing $\beta = 0$ in (11):

$$EU = \pi U\left(Y - (1+\tau_A)\pi\frac{\beta^2}{\beta - c} - L + (1-\tau_B)\beta\right) + (1-\pi)U\left(Y - (1+\tau_A)\pi\frac{\beta^2}{\beta - c}\right) \tag{70}$$

$\beta = 0$ solves (58), (59) and (60). The only complication is to solve (57):

$$\begin{aligned} 0 \;=\; & \pi\left[(1-\tau_B) - (1+\tau_A)\pi\frac{\beta(\beta-2c)}{(\beta-c)^2}\right]U'(Y - \alpha - L + (1-\tau_B)\beta) \\ & - (1-\pi)(1+\tau_A)\pi\frac{\beta(\beta-2c)}{(\beta-c)^2}U'(Y-\alpha) + \mu(\tau_A+\tau_B)\pi\frac{\beta(\beta-2c)}{(\beta-c)^2} \end{aligned} \tag{71}$$

Letting $\beta = 0$, we obtain

$$\pi(1 - \tau_B)U'(Y - L) = 0 \tag{72}$$

This occurs only if $\tau_B \to 1^+$. Since autarchy cannot be preferred to the first best allocation, it follows that $\tau_B \to 1^-$ is a maximum (first best allocation), while $\tau_B \to 1^+$ is a minimum (autarchy).

To prove the final point (that premiums are subsidized using the benefit tax), recall from (60) that

$$\tau_A = g\left(\frac{\beta - c}{\pi\delta}\right) - \tau_B \tag{73}$$

If $g < \pi\tau_B\frac{\delta}{\beta - c}$, then the premium tax rate is negative. This means that there is a premium subsidy. For greater government needs ($g \geq \pi\tau_B\frac{\delta}{\beta - c}$), $\tau_A \geq 0$. $\bullet$

6.2 Figures, Tables and Graphs

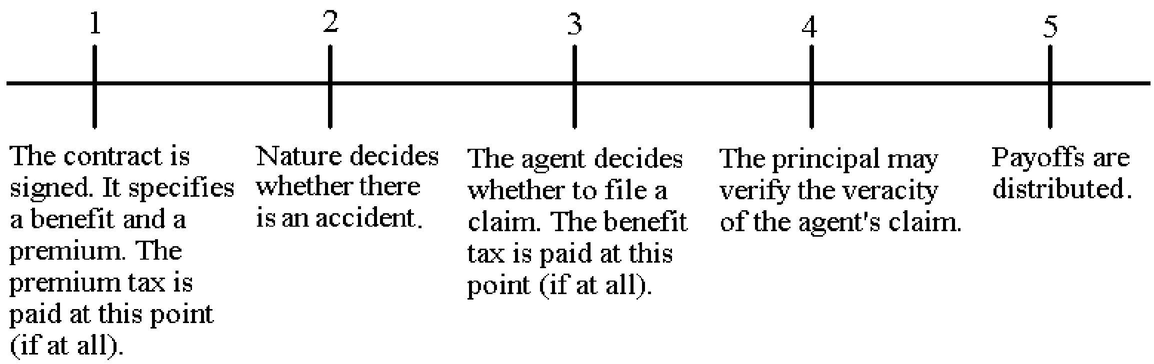

Figure 1: Sequence of play.

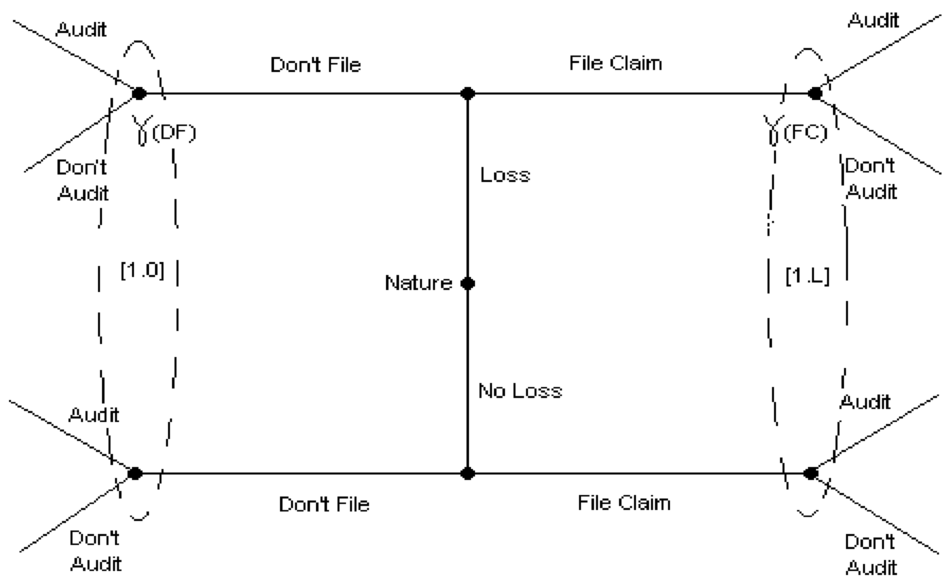

Figure 2: Extensive form of the claiming game.

Table 1

Parameter values used for the Logarithmic utility function

Parameter	Meaning	Value
U(.)	Utility function	LN(.)
Y	Initial wealth	50
L	Possible loss	25
¼	Probability of loss	0.2
c	Cost of auditing	10
k	Penalty for committing fraud	10

Table 2

Parameter values used for the exponential utility function

Parameter	Meaning	Value
U(.)	Utility function	$¡ e^{¡ a(:)}$
a	Coe¢ciant of absolute risk aversion	0.8
Y	Initial wealth	10
L	Possible loss	25
¼	Probability of loss	0.2
c	Cost of auditing	10
k	Penalty for committing fraud	10

$$\alpha = \pi\beta + (1 - \pi)\beta\eta(1 - \kappa) + c\kappa[\pi + (1 - \pi)\eta]$$

$$\eta = \left(\frac{c}{\beta - c}\right)\left(\frac{\pi}{1 - \pi}\right)$$

$$\kappa = \frac{U(Y - \alpha + \beta) - U(Y - \alpha)}{U(Y - \alpha + \beta) - U(Y - \alpha - k)}$$

$$\max_{\beta_A} EU_A = \pi U(Y - (1+\tau_A)\alpha_A - L + \beta_A) + (1-\pi) U(Y - (1+\tau_A)\alpha_A)$$

$$\alpha = \pi \frac{\beta_A^2}{\beta_A - c}$$

$$\max_{\beta_B} EU_B = \pi U(Y - \alpha_B - L + (1-\tau_B)\beta_B) + (1-\pi) U(Y - \alpha_B)$$

$$\alpha = \pi \frac{\beta_B^2}{\beta_B - c}$$