Yes, No, Perhaps? - Explaining the Demand for Risk Classification Insurance with Fuzzy Private Information

Richard Peter∗, Andreas Richter†, Petra Steinorth‡

December 9, 2010

Abstract

We examine the demand for risk classification insurance with fuzzy private information using an expected utility set-up. We first demonstrate that standard assumptions on private information cannot explain the observed demand for risk classification insurance. By allowing for fuzzy heterogeneous private information, we show that a unique equilibrium exists that can be of the form that some of the individuals buy risk classification insurance while others do not, as observed in real-life insurance markets. We derive a condition for the existence of such an equilibrium and show that this is particularly driven by the risk attitude of the insured, the volatility of outcomes without risk classification insurance and the distribution of information. We also make welfare comparisons and show that in the case of multiple equilibrium candidates the one with the highest portion of risk classification insurance buyers is Pareto superior and, thus, the unique equilibrium.

Keywords: Risk classification, insurance economics, private information

JEL-Classification: D11, D42, D82, G22

1 Introduction

Risk classification is an important instrument in insurance markets. As is well known, without risk classification markets may suffer from adverse selection if the insured possess private information on their risk type. Rothschild and Stiglitz (1976) show that there can be Nash equilibria where insured are implicitly categorized according to their risk type via a self-selection design. However, such equilibria do not always exist. Wilson (1977) and Miyazaki (1977) demonstrate, using a modified equilibrium definition, that
there are always equilibria with cross subsidization between the different risk types.\(^1\) In a multi-period setting, Cooper and Hayes (1987) show that implicit risk categorization in a competitive market is possible as low risks are more willing to enter a contract with experience rating while high risks choose short-term contracts.\(^2\) Bond and Crocker (1991) provide a model in which first-best Nash equilibria exist if insurers can use the insureds’ consumption decisions to draw conclusions about their risk types.

Yet, risk classification is not only used in cases where the insured have private information. Also, when information about risk types is symmetric, Crocker and Snow (2000) show that risk classification has to be implemented due to competitive pressure. Subsequently, risk classification or categorization is an important tool for insurers to be able to offer stable equilibrium contracts.

However, the welfare implications of risk classification are not clear. Hoy (1982) e.g. assesses welfare implications and shows that risk classification can also decrease the utility of individuals and social welfare, in particular when individuals face the risk of being classified as high risks and having to pay a higher insurance premium. Moreover, Hoy (2006) discusses the impact of governmental regulation on insurers’ classification methods and states conditions under which such restrictions increase or decrease social welfare. Harrington and Doerpinghaus (1993) confirm the ambiguous welfare effects of risk categorization for automobile insurance. They are tightened with categorization costs, as Crocker and Snow (1986) show.

As the welfare implications of risk classification are ambiguous and a threat of risk classification in the future might decrease individual utility, insuring against risk classification seems to be sensible for a risk averse person. Previous work on risk classification insurance has mostly been conducted in the field of health insurance.\(^3\) Concerning genetic testing, it has been discussed e.g. by Tabarrok (1994) whether risk classification insurance should be mandatory for those who want to take a genetic test, such that individuals do not face the risk of having to pay higher insurance premiums after the test when they turn out to be a high risk. Protection against risk classification is particularly important in the field of genetic testing as Doherty and Thistle (1996) find that individuals might refrain from undergoing a genetic test due to classification risk. In the course of their paper, they also allow for varying private information among different agents but in a discrete and less general way than in this paper.

\(^1\)However, Puelz and Snow (1994) use automobile insurance data to reject the hypothesis that some individuals subsidize others.

\(^2\)Other authors have examined optimal contracts in monopoly markets; see e.g. Townsend (1982), Dionne (1983) and Dionne and Lassere (1985, 1987).

\(^3\)Among other reasons, this might be the case as medical costs for different risk types vary immensely.
Risk classification insurance has also been discussed generally for health insurance contracts. In these settings, part of the health insurance premium or stand-alone risk classification insurance is used to insure individuals against being classified as a high risk at some future point in time, e.g. by Cochrane (1995), Pauly et al. (1995), Kifmann (2001) and Kifmann (2002). However, these papers do not assume individuals to have any kind of private information on their future health state. Accordingly, all individuals choose to insure against risk classification and the demand issue is not assessed.

There are several examples where risk classification insurance is available. Stand-alone risk classification insurance has been offered e.g. in the German private health insurance market for quite some time. One group of potential buyers of risk classification insurance are those who plan to switch from statutory health insurance, where premiums are not risk-based but income-dependent, to private health insurance where premiums are risk-based. In order to be eligible for switching to private health insurance, individuals have to meet certain criteria. They either have to wait for some following years where they meet an income threshold or they have to be self-employed or a civil servant. Risk classification insurance protects individuals from having to pay a higher insurance premium due to conditions that may arise in the period between buying classification insurance and finally being allowed to switch to private health insurance. Examining demand patterns for this risk classification insurance in this market shows that there is a significant portion among the future privately insured that buy risk classification insurance. Following the report of the German private health insurers for the accounting year 2008/2009 there were 48,900 newly insured paying full private health coverage and 20,500 purchased risk classification insurance (http://www.pkv.de/w/files/shop_zahlenberichte/pkv_zahlenbericht_2008_2009.pdf). Even though there is no data available on how many of the 48,900 newly insured in 2008/2009 had bought risk classification insurance before switching to private health insurance, it suggests that a significant portion had this insurance before as there are 20,500 new purchaser of risk classification insurance.

Also, risk classification insurance is an option in many auto insurance contracts, for example in the U.S., UK or Germany. Individuals can decide to purchase the so-called accident forgiveness option where they do not have to pay higher insurance premiums after having an accident. Nevertheless, the driver’s accident history provides the insurer with useful information to improve her evaluation of the insured’s risk type. As accident forgiveness is usually either included in some contracts or can be added at additional

---

4This is in particular sold to young drivers who are most commonly expected to be a higher risk type.
costs to the existing contract, data on the actual demand is hard to come by. However, as accident forgiveness seems to be widely offered, one can deduce that there may be substantial demand. Accordingly, there are markets where risk classification insurance is in demand. However, evidence from those markets suggests that only a fraction of the population actually purchases this insurance. Standard assumptions on private information that everyone either knows their future risk type for certain or has no private information at all do not explain the market evidence fully. If individuals already know their future risk type, risk classification insurance does not work. Nevertheless, without information on the risk type, insuring yourself against risk classification in the future at actuarially fair prices will be beneficial for every risk averse person if the risk type is revealed in the future. In such a situation it is surprising that there is risk classification in the first place as adverse selection would not be a problem and risk classification would only impose an additional income risk on the insured. Thus, the aim of this paper is to investigate the role of the distribution of information on the demand for risk classification insurance and to analyze, whether the above-mentioned market observations can be explained via a more general model of private information. In this paper, we show that allowing private information to be fuzzy and heterogeneous among agents there can be situations where only part of the insured buy risk classification insurance. Therefore, we introduce a continuous signal containing the information on future risk types in addition to the standard adverse selection approach conducted by Rothschild and Stiglitz (1976). Nevertheless, the risk types are discrete, opposing the work by Riley (1979) who shows that introducing continuous risk types leads to non-existence of any kinds of Nash equilibria. By adding a time horizon to our decision process, which allows individuals to adjust their consumption according to their insurance strategy, we are able to proof the existence of a unique Nash equilibrium in the market in any case.

The remainder of this paper is structured as follows. In the next section, the general model background is described including the timeline of actions and the insurance decision. A signal containing private information is introduced. As a benchmark, we first examine trivial signal distributions distinguishing between complete private information and no private information regarding types. Under the assumptions of this paper, we show that with perfect private information, risk classification insurance is driven out of

\footnote{By fuzzy private information we refer to a situation where the insured do not know their risk type for sure but have a better knowledge of the probability of becoming either low or high risks than the insurer.}

\footnote{The modeling of information is similar to but more general than that of Steinorth (2010), who uses this approach in the context of annuities. Crocker and Snow (1994) also allow for imperfect private knowledge but one fundamental difference from our model is that they use a discrete set-up.}
the market, while no private information induces all individuals to buy risk classification coverage. Subsequently, we examine the more realistic case of fuzzy private information, which is heterogeneous among agents. Therefore, a general model for the distribution of the informative signal is developed and an equilibrium condition is proposed. We show that a situation with demand for risk classification insurance always fulfills a necessary equilibrium condition, yet several equilibrium candidates are possible. Subsequently, we derive a sufficient condition for an interior cutoff where only part of the population buy risk classification insurance. This condition is fulfilled more easily if the degree of risk aversion is higher, the outcomes without risk classification insurance are more volatile and the curvature of the premium function for risk classification insurance at zero is steeper which is directly related to the distribution of information. Assuming the premium function to be convex, we show that the above-mentioned sufficient condition is also necessary for observing demand for risk classification insurance. In this case, there is only one cutoff apart from zero. Afterwards, we analyze the different equilibrium candidates from a welfare perspective. We show that a higher cutoff always dominates a lower one in the Pareto sense and, thus, the highest is the equilibrium. In addition, we discuss why the equilibrium must be unique. We then illustrate these observations with some examples discussing the influence of specific signal distribution and investigating the impact of different attitudes towards risk on the relative position of the market equilibrium. The paper ends with some concluding remarks.

2 Model Background

We assume that there are only two points in time and that individuals have a time-additively separable utility function $u(C)$, which is identical at both points in time and depends on the consumption level $C \geq 0$. We assume individuals to be risk averse and thus, $u'(C) > 0$ and $u''(C) < 0$. There is no time discount.

At $t_1$, individuals cannot be distinguished from the insurers’ point of view and they do not face any monetary risk. Without loss of generality, we assume that lifetime’s wealth $W$ is already available at $t_1$. Also at $t_1$, individuals receive a signal $z \in [0,1]$ on their risk type at $t_2$. The signal cannot be observed by insurers and it reflects the probability of becoming a low risk at $t_2$. Yet, the insurer is assumed to know about the signal distribution in the population. At $t_2$, all individuals become either a high or a low risk type, which can be observed by the insurers. Individuals now face the risk of losing a fixed amount $T$ with probability $p_H$ if they turn out to be a high risk and $p_L$ if they are a low risk, where $0 < p_L < p_H < 1$. In order to distinguish this risk clearly
from classification risk, we name this risk a health risk, due to the real-life examples of classification insurance described in the introduction. However, the general model set-up allows for incorporating other risks as well. We assume a perfectly competitive market framework where insurers - after the revelation of risk type at $t_2$ - offer insurance at a fair price $\alpha_H \cdot p_H \cdot T$ for high risks and $\alpha_L \cdot p_L \cdot T$ for low risks, respectively, where $\alpha_L, \alpha_H$ denote the coinsurance rates. Without classification insurance, individuals face the following decision problem:

$$
\max_{C_1,\alpha_L,\alpha_H} V = u(C_1) + z \cdot \left[ p_L \cdot u(W - C_1 - T + \alpha_L \cdot T - \alpha_L \cdot p_L \cdot T) \\
+ (1 - p_L) \cdot u(W - C_1 - \alpha_L \cdot p_L \cdot T) \right] \\
+ (1 - z) \left[ p_H \cdot u(W - C_1 - T + \alpha_H \cdot T - \alpha_H \cdot p_H \cdot T) \\
+ (1 - p_H) \cdot u(W - C_1 - \alpha_H \cdot p_H \cdot T) \right]
$$

where $C_1$ denotes consumption at $t_1$ and $V$ is the overall utility. To prevent default at $t_2$, we assume that consumption, residual loss (the part of the loss that is not covered by the insurance contract) and the premium for low and high risks respectively do not exceed the initial wealth, formally

$$
C_1 + (1 - \alpha_i) \cdot T + \alpha_i \cdot p_i \cdot T \leq W, \quad i \in \{L, H\}.
$$

In the following analysis, we assume these constraints not to be binding, for instance because the loss does not exceed initial wealth and individuals do not consume excessively, i.e. $C_1 + T < W$.$^7$ We will, thus, neglect the above-named constraints.

The following proposition summarizes the results for this set-up.

**Proposition 1.** If classification insurance is not available, individuals will fully insure against the health risk at $t_2$. They split consumption in such a way that they consume most at $t_2$ if they are a low risk and least if they are a high risk at $t_2$. Consumption at $t_1$ lies between these levels.

**Proof.** Refer to the appendix for the proof. \qed

Regarding the optimal consumption level, we can be more precise.

**Remark 1.** The first-order condition with respect to $C_1$ uniquely defines an optimal

---

$^7$Overly consuming turns out not to be optimal anyway, which alleviates the latter restriction.
consumption level depending on \( z \) in the following sense: For each \( z \in [0,1] \) there is one and only one \( C_1 \in \left[ \frac{W-p_H\cdot T}{2}, \frac{W-p_L\cdot T}{2} \right] \) that satisfies the optimality condition.

**Proof.** Refer to the appendix for the proof. \( \square \)

**Remark 2.** The optimal \( C_1 \) increases in the probability \( z \) of becoming a low risk. For CARA-utility functions \( C_1 \) is a convex function of \( z \).

**Proof.** Refer to the appendix for the proof. \( \square \)

Assume now that classification risk insurance is available at \( t_1 \). This coverage is available for a premium \( P \) that has to be paid at \( t_1 \) and that we will specify later. It refunds the difference between the premium for a high risk and the premium for a low risk at \( t_2 \).\(^8\) If this insurance coverage is purchased, the maximization problem changes to:

\[
\max_{C_1, \alpha} V = u(C_1 - P) + (z \cdot p_L + (1 - z) \cdot p_H) \cdot u(W - C_1 - T + \alpha \cdot T - \alpha \cdot p_L \cdot T) \\
+ (z \cdot (1 - p_L) + (1 - z) \cdot (1 - p_H)) \cdot u(W - C_1 - \alpha \cdot p_L \cdot T).
\]

In the following two sections, we will discuss the influence of different signal distributions on the decision on whether to insure against risk classification. To keep track of the considerations, the following timeline depicts the course of actions.

---

\(^8\)We assume here that individuals cannot buy partial coverage for risk classification. Naturally, this slightly limits the analysis. However, this limitation does not seem too severe as the authors could not find an example in insurance markets where risk classification insurance is available as partial coverage.
3 Model

3.1 Benchmark: Trivial Information Distributions

In this section, we assess the demand reaction to the offer of risk classification insurance with standard assumptions on private information. We first assume that the signal does not include any private information, but only provides information about the average probability $\eta$ of becoming a low risk in the population, which is per assumption also known to the insurer.\(^9\) In this framework, assuming risk classification insurance to be offered at an actuarially fair rate, the maximization problem (2) can be transformed into:

$$
\max_{C_1,\alpha} V = u(C_1 - (1 - \eta) \cdot \alpha \cdot (p_H - p_L) \cdot T) + (\eta \cdot p_L + (1 - \eta) \cdot p_H) \cdot u(W - C_1 - T + \alpha \cdot T - \alpha \cdot p_L \cdot T) + (\eta \cdot (1 - p_L) + (1 - \eta) \cdot (1 - p_H)) \cdot u(W - C_1 - \alpha \cdot p_L \cdot T).
$$

Comparing the situation with and without risk classification insurance, the following proposition holds.

**Proposition 2.** Without private information and if individuals are risk averse, risk classification insurance will be demanded by all individuals as it increases individual and social welfare compared to a situation without risk classification insurance.

**Proof.** Refer to the appendix for the proof.\(\square\)

As insurance does not influence the overall mean of final wealth over both points in time, it is the same with and without risk classification insurance. However, as we have seen that final wealth is equal at both points in time with risk classification insurance but varies without, the optimum with classification insurance stochastically dominates the optimum without classification insurance in the second-order sense. Accordingly, a risk averse individual prefers to buy risk classification insurance.\(^{10}\)

Now, let us assume instead that individuals already know for sure at $t_1$ what kind of risk type they are at $t_2$. Again, we assume that the risk type is revealed to the insurance company at $t_2$. Those individuals who are low risks at $t_1$ with certainty are not willing to

---

\(^9\)We assume here that $0 < \eta < 1$, as $\eta = 0$, $\eta = 1$ represent the trivial cases, in which only high or low risks exist.

\(^{10}\)This result conforms classical insurance theory that with actuarial fair insurance premium risk averse individuals choose full insurance, which is for instance covered by Mossin (1968).
pay a positive premium for risk classification insurance as they know for sure that they will not have to pay the higher health insurance premium and thus will never benefit from the coverage against classification risk. However, risk classification insurance cannot be offered costlessly as high risk types are attracted by such an offer at $t_1$, too, and risk types cannot be distinguished then. Therefore, only high risk types might be attracted by risk classification insurance. This would be anticipated by the insurers who then price the coverage offered accordingly. However, as high risks do not really face any classification risk since they know their risk type for certain, risk classification insurance cannot increase their individual utility. In this setting, the following proposition holds.

**Proposition 3.** If private information is complete at $t_1$, risk classification insurance does not enhance welfare of either low or high risk types and thus is not in demand.

To sum up, the two benchmark cases that were analyzed, fully uninformed consumers and fully informed consumers, imply that either all or no insured buy risk classification insurance. However, these results do not help us to understand what seems to be observable in real-life markets, namely that only part of the population buys risk classification insurance. We show in the next section that we can explain such equilibria by allowing for a more general modeling of private information.

### 3.2 Fuzzy Private Information

In this section, we assume that the distribution of the signal $Z$ is continuous on the interval $[0, 1]$. The signal again reflects the probability of becoming a low risk and the expected value of the signal is denoted by $\eta$, which is, thus, the expected probability of becoming a low risk among the population. Accordingly, agents receive information of different quality regarding their future risk type: a signal very close to the corners does contain significant private information while the private information is less when the signal is closer to $\eta$. This is due to the fact that the average probability of becoming a low or a high risk type is known to the insured and the insurers. The heterogeneity among agents is thus reflected in two perspectives: the probability of becoming a low risk - the individual’s signal itself - and the precision of the signal - the individual’s signal in light of the overall signal distribution.\(^{11}\)

We again consider a health risk that occurs at $t_2$. Allowing for information that differs

---

\(^{11}\)We assume the signal to be costless and to be received always. Previous work in the field of risk classification has widely focused on situations where individuals can decide whether they receive a (costless or not) signal, which, in contrast to our setting, provides perfect information on the risk type, see e.g. Doherty and Thistle (1996), Doherty and Posey (1998) and Strohmenger and Wambach (2000). By introducing the signal into our model, we assume that individuals already possess (costless) private information that is, in contrast to the mentioned papers, fuzzy.
in precision regarding the future risk type of insured persons seems consistent with real life. Some might already have quite precise information on their future risk type as e.g. they know that they have unhealthy living habits or they know about their genetic disposition due to family members or genetic tests. Those individuals receive a signal close to zero. However, others might have very healthy living habits and know that their relatives never suffered from a disease that is influenced by genetic disposition, i.e. they receive a signal closer to one. Those with average habits and disposition receive a signal close to $\eta$.

By modeling the signal in such a way, it becomes clear that someone having received a signal $z = 1$ (a definite low risk) can never be interested in buying risk classification insurance as this individual does not suffer from classification risk and would subsidize others when purchasing risk classification insurance. However, a definite high risk $z = 0$ is very interested in buying risk classification insurance as long as individuals with a lower signal buy it, too. Knowing their risk type for sure, they do not benefit from the fact that they insure against classification risk but from the subsidization from lower risk types buying classification insurance.

Those individuals with a less precise signal ($z \neq 0,1$) evaluate risk classification insurance in a different way. As they face classification risk they benefit from insuring against it. However, as they might be a lower/higher risk than average they might also have a negative/positive benefit due to subsidization in risk classification insurance. There is a trade-off between benefiting from insurance protection and subsidizing others if the personal signal is above the average signal among the risk classification buyers. Yet, it can be the case that the value of risk classification insurance is so low that negative effects of subsidization overshadow the benefits from insurance. If this holds true for all values of $z$ nobody will buy classification insurance due to adverse selection. Thus, the aim of the paper is to determine whether and under which conditions risk classification insurance is purchased by at least some of the individuals.

We examine whether a critical threshold exists such that the individual who receives this signal is indifferent between buying and not buying risk classification insurance. We refer to the threshold as the cutoff signal and to an individual receiving it as the cutoff individual. For the cutoff individual, the benefit from risk classification insurance exactly corresponds to the disutility of having to subsidize others. Accordingly, every individual with a higher signal than the cutoff signal decides against risk classification insurance, as subsidizing others is more costly to them and the benefit of insuring against risk classification is less. For reverse reasons, individuals with lower signals buy risk
classification insurance. A cutoff is, thus, formally \((z^*, P(z^*))\) defined by the cutoff signal and the cutoff premium for risk classification insurance, which we discuss later. As the cutoff individual is indifferent between the two alternatives, the cutoff signal \(z^*\) must solve the following equation:

\[
\text{utility without risk classification insurance} \quad u(C_1) + z^* \cdot u(W - C_1 - P_L) + (1 - z^*) \cdot u(W - C_1 - P_H) = \frac{2 \cdot u \left( \frac{W - P_L - P(z^*)}{2} \right)}{\text{utility with risk classification insurance}}
\]

where \(P_L\) denotes the insurance premium against the health risk for low risk types and \(P_H\) for the high risks, respectively, and \(P(z^*)\) is the premium for risk classification insurance. Accordingly, a necessary and sufficient condition for a cutoff is that the utility without risk classification must be equal to the utility with risk classification insurance for the cutoff individual. Note without classification insurance expected utility is given by proposition 1 and that for the situation with risk classification insurance, we have already inserted the optimal consumption and insurance strategy. Assuming risk classification insurance to pay the difference for the full coverage premiums it is optimal to insure fully and consume equally at both points in time for any given insurance premium \(P(z^*)\).\(^{12}\) Regarding the insurance premium for risk classification \(P(z^*)\) we assume that the insurers anticipate a cutoff when existent and, thus, it holds that

\[
P(z^*) = (1 - E(Z|Z < z^*)) \cdot (P_H - P_L).
\]

As we have mentioned before, multiple solutions to (4) are possible. Thus, the question arises of how an insurer can anticipate which cutoff will be established on the market. This issue will be solved later as we will show that one distinct cutoff is Pareto superior to all the others and will, thus, be the equilibrium in the competitive market. For the moment, let us just assume that the insurer can determine all the possible cutoffs. As a higher \(z^*\) indicates that the cutoff individual is more likely to become a low risk, it holds that

\textbf{Remark 3.} \(P(z^*)\) decreases in \(z^*\).

\textit{Proof.} Please refer to the appendix for the proof. \(\square\)

\(^{12}\) We do not provide the proof that it is optimal to consume equally at both points in time given that the individual purchases risk classification insurance and fully insures the health risk. In this case, there is no uncertainty and we neglect interest. Among others Kimball (1990) shows this result.
In the following, we will determine cutoffs that solve (4). Note again that (4) is a necessary but not sufficient condition for an equilibrium as multiple solutions to (4) are possible. We assume that cutoffs are anticipated by insurers and that they align the premium for risk classification insurance according to each cutoff. We will later discuss the implications of multiple solutions to (4) and determine the equilibrium under all possible cutoff signals. The first candidate for an equilibrium is given by the following proposition:

**Proposition 4.** A cutoff signal \( z^* = 0 \) always solves equation (4).

*Proof.* For \( z^* = 0 \) it holds that \( \Pr(Z \leq z^*) = \Pr(Z = 0) = 0 \) \(^{13}\) and thus

\[
\Pr(0) = P_H - P_L.
\]

The right-hand side of (4) simplifies to \( 2 \cdot u \left( \frac{W - P_H}{2} \right) \) while the left-hand side of (4) can be rewritten as \( u(C_1(0)) + u(W - C_1(0) - P_H) \). As there is no risk of classification at \( t_2 \), it can be derived that optimally \( C_1(0) = \frac{W - P_L}{2} \). \(^{14}\)

This proposition shows that a cutoff where only a definite high risk buys classification insurance is a possible candidate for an equilibrium. With \( z^* = 0 \), a definite high risk does not face uncertainty on risk classification in the second period and risk classification insurance is priced accordingly. Therefore, there is no benefit for the definite high risk from insuring against risk classification. Without loss of generality we therefore assume that \( z^* = 0 \) implies that nobody buys classification insurance. In the further analysis we investigate additional candidates for an equilibrium that solve (4), i.e. interior cutoffs where not only definite high risks are interested in classification insurance. After having determined all the solutions to the necessary equilibrium condition (4) we investigate which candidate can be the equilibrium cutoff signal.

**Remark 4.** \( z = 1 \) cannot be a cutoff point.

*Proof.* If \( z = 1 \) were a cutoff point, the insurer would anticipate it and calculate the fair premium using (5)

\[
P(1) = (1 - \Pr(Z \leq 1)) \cdot (P_H - P_L) = (1 - \Pr(Z))(P_H - P_L) = (1 - \eta)(P_H - P_L).
\]

\(^{13}\)One has to consider that \( \{z = 0\} \) has zero probability; but as \( \Pr(Z) = Z = \id(Z) \) one obtains that \( \Pr(Z \mid Z = 0) = \id(0) = 0 \).

\(^{14}\)The first-order condition yields \( u'(C_1(0)) = u'(W - C_1(0) - P_H) \) and the injectivity of \( u' \) entails the optimal \( C_1(0) \).
For cutoff points equation (4) holds and using $C_1(1) = \frac{W - P_L}{2}$ from equation (18), we obtain

\[
\begin{align*}
&u\left(\frac{W - P_L}{2}\right) + u\left(\frac{W - W - P_L - P_L}{2}\right) = 2u\left(\frac{W - P_L}{2}\right) \\
&\quad \leq 2u\left(\frac{W - P_L - (1 - \eta)(P_H - P_L)}{2}\right).
\end{align*}
\]

Due to the injectivity of $u$ this can only hold if $\eta = 1$. Therefore $z = 1$ almost surely and we would face a group of certain low risks only, which contradicts the assumptions.

\[\square\]

The intuition of Remark 4 is that it can never be beneficial for an individual with a signal $z = 1$ to buy risk classification insurance. There is no uncertainty about risk classification for this individual and she has to subsidize others when buying risk classification insurance. Remark 4 is, however, important for deriving a condition for an interior cutoff in the further analysis.

We define the function

\[
g(z) = u(C_1) + z \cdot u(W - C_1 - P_L) + (1 - z) \cdot u(W - C_1 - P_H) - 2 \cdot u\left(\frac{W - P(z) - P_L}{2}\right)
\]

which displays the utility difference between the cases without and with risk classification insurance. The premium for risk classification insurance is assumed to be priced according to the inserted signal. At a real cutoff, this function must be equal to zero. If the inserted value is not a cutoff $g(z) \neq 0$.

Proposition 4 implies that $g(0) = 0$ while Remark 4 and its proof show that $g(1) > 0$. Accordingly, an interior cutoff requires that there is a $z \in (0, 1]$ with $g(z) \leq 0$. We use the notation

\[
W_1 = \frac{W + P_H}{2} - P_H \quad \text{and} \quad W_2 = \frac{W + P_H}{2} - P_L.
\]

Then, the following proposition holds.

**Proposition 5.** Assuming $\lim_{z \downarrow 0} \frac{E[Z|Z \leq z]}{z}$ exists and denoting it by $\lambda$, a sufficient condition for an interior cutoff is that

\[
\frac{u(W_2) - u(W_1)}{W_2 - W_1} < \lambda \cdot u'(W_1).
\]
Proof. Using $g(0) = 0$ and $g(1) > 0$, the idea of the proof is to look at $g'(0)$, the right-hand side derivative of $g$ at 0, and check whether this is negative. This implies that $g$ has at least another null in $(0, 1)$. Therefore, we derive

$$g'(z) = C'_1(z) \cdot u'(C_1(z)) - z \cdot C'_1(z) \cdot u'(W - C_1(z) - P_L)$$
$$-(1 - z) \cdot C'_1(z) \cdot u'(W - C_1(z) - P_H) + u(W - C_1(z) - P_L)$$
$$-u(W - C_1(z) - P_H) + P'(z) \cdot u\left(\frac{W - P(z) - P_L}{2}\right)$$
$$= u(W - C_1(z) - P_L) - u(W - C_1(z) - P_H)$$
$$+ P'(z) \cdot u\left(\frac{W - P(z) - P_L}{2}\right)$$

as for an optimal consumption without risk classification insurance $u'(C_1(z)) - z u'(W - C_1(z) - P_L) - (1 - z) \cdot u'(W - C_1(z) - P_H) = 0$. Inserting $z = 0$ yields

$$g'_+ (0) = u(W - C_1(0) - P_L) - u(W - C_1(0) - P_H) \quad (7)$$

$$+ P'_+ (0) \cdot u\left(\frac{W - P(0) - P_L}{2}\right)$$
$$= u(W_2) - u(W_1) - \lambda \cdot (P_H - P_L) u'(W_1)$$

as $C_1(0) = \frac{W - P_H}{2}$, $P(0) = P_H - P_L$ and $P'_+ (0) = -\lambda \cdot (P_H - P_L)$. (7) is less than zero as long as

$$\frac{u(W_2) - u(W_1)}{W_2 - W_1} < \lambda \cdot u'(W_1)$$

due to $P_H - P_L = W_2 - W_1$. \hfill \Box$

Proposition 5 shows that the existence of an interior cutoff seems to be driven by three different features. First, it depends on the shape of the utility function, in particular the degree of risk aversion. If $h$ is an increasing concave transformation ($h' > 0$, $h'' < 0$) and $v := h(u)$, then if (6) holds for $u$, it follows that it holds for $v$. The intuition is that more risk averse individuals benefit more from the insurance protection against risk classification and, therefore, the condition for the existence of an interior cutoff is fulfilled more easily. Second, the left-hand side of (6) decreases in the difference between the health insurance premiums for low and high risks assuming $P_H$ to be constant. This is because a higher difference in premiums increases the variance of outcomes without risk classification insurance. This also makes risk classification insurance more attractive.\footnote{Note that $v$ exhibits greater absolute risk aversion than $u$. The reverse implication is not true, as one can think of situations where (6) holds for $v$, but does not for $u$.}
Third, the distribution of information also influences whether an interior cutoff exists. In condition (6), this is displayed by $\lambda$ on the left-hand side. Obviously, a higher $\lambda$ will make condition (6) to be fulfilled more easily, cet.par. This is intuitive insofar that a higher $\lambda$ determines how much the premium reduces if starting at a cutoff with $z^* = 0$ the only marginally better individual purchases RCI. It is clear this marginal individual would then be the best purchaser of RCI and it has to subsidize the certain high risk. The economic reasoning for the existence of an interior cutoff here is that if it is worthwhile for the only marginally better individual than the definite high risk to subsidize the high risk, there must be an even better signal who is just indifferent between subsidizing others within the RCI and foregoing insurance protection from RCI. This is because we know that expected utility without RCI is greater if a cutoff signal of 1 is assumed. Therefore, there must be an interior cutoff between the marginally better individual and the certain low risk. For now, condition (6) is a sufficient condition. In the following we will also discuss assumptions which will make this condition necessary and sufficient at the same time. We summarize the discussed results in the following remark.

**Remark 5.** Condition (6) of Proposition 5 is more easily fulfilled

1. the greater the degree of risk aversion, cet. par.

2. the greater the difference between low and high risks, cet. par.

3. the greater the absolute slope of the premium function at 0, cet. par.

In particular, the second statement of Remark 5 seems to be supported by market observations. As mentioned in the Introduction, forms of risk classification insurance have mostly been seen in auto and health insurance. Those two markets are known for the fact that premiums vary substantially between different insured persons.

We can derive the following proposition:

**Proposition 6.** Assuming $P''(z)$ exists and is non-negative implies that

1. condition (6) of Proposition 5 is not only sufficient but also necessary for the existence of an interior cutoff.

2. if condition (6) is fulfilled there is a unique interior cutoff.

**Proof.** We show that $g''(z) > 0$. Then $g$ does not have any inflection points and $g'(0) < 0$ becomes necessary for another null to exist. If $g$ does not have an inflection, there can
only be one extra null in addition to \( z^* = 0 \). Therefore, we compute

\[
g''(z) = C_1'(z) \left( u'(W - C_1(z) - PH) - u'(W - C_1(z) - PL) \right) > 0
\]

\[
+ P''(z) \cdot u' \left( \frac{W - P(z) - PL}{2} \right) \geq 0
\]

\[
- \left( \frac{P'(z)}{2} \right)^2 \cdot u'' \left( \frac{W - P(z) - PL}{2} \right) < 0.
\]

Assuming \( P''(z) \geq 0 \) implies that the premium is decreasing at a non-increasing absolute rate. This implies that the consumption at both points in time also increases at a non-decreasing rate. Accordingly, the expected utility with risk classification insurance is concave. Examining 8 shows that expected utility without risk classification insurance is convex in \( z \). Therefore, the expected utility with risk classification insurance must have a greater slope at 0 than without risk classification insurance to ensure that another intersection point between these expected utilities exists.

### 3.3 Equilibrium and Welfare Implications

As mentioned before (4) is only a necessary condition for an equilibrium as more than one cutoff signal is possible. The following proposition examines the welfare consequences of different cutoffs and will, thus, help to determine the equilibrium cutoff.

**Proposition 7.** Assume two cutoffs \( z_1^* \) and \( z_2^* \) with \( z_1^* < z_2^* \). The cutoff in \( z_2^* \) Pareto dominates the cutoff in \( z_1^* \).

**Proof.** Comparing the two cutoffs \( z_1^* \) and \( z_2^* \), we distinguish three groups in the population. The first group denoted by \( A \) consists of those who do not buy risk classification insurance under both cutoffs. Accordingly, being in \( A \) implies having received a signal greater than \( z_2^* \). Utility without risk classification insurance depends only on the individual signal and, thus, does not change with different cutoffs.

The second group, \( B \), includes all individuals who switch from no classification insurance under \( z_1^* \) to classification insurance under \( z_2^* \). As for any given signal \( z \in (z_1^*, z_2^*) \) it holds that the expected utility without risk classification insurance is equal under both cut-
offs, the switchers must be better off. Otherwise they would not switch to classification insurance.

The last group is denoted by C and includes all the individuals who buy risk classification insurance under both cutoffs. As $P'(z^*)$ decreases in $z^*$ individuals in C are better off as they pay a lower premium under $z^*_2$ than under $z^*_1$.

Subsequently, the utility of individuals in A and the cutoff individual stays constant when increasing the cutoff from $z^*_1$ to $z^*_2$ while all others are better off.

Regarding possible equilibria, the following proposition can be derived from Proposition 7.

**Proposition 8.** In the market exists a uniquely determined Nash-equilibrium, which is the highest cutoff.

**Proof.** The first step of the proof is to argue that a Pareto inferior cutoff cannot be the equilibrium in a competitive market. This is due to the fact that between any non-highest cutoff and the highest cutoff (including $z^* = 0$) there is a $\tilde{z}$ where $g(\tilde{z})$ is negative, i.e. the expected utility with risk classification insurance is greater than without. Offering RCI at a price $P(\tilde{z})$ attracts all purchasers of the lower cutoff contract and so many others that the contract makes a positive profit. Thus, there is always such a value of $z$ that destabilizes a cutoff if not the highest cutoff. However, due to competitive pressure, $\tilde{z}$ cannot be the equilibrium.

Accordingly, the highest cutoff is the only remaining candidate for a Nash equilibrium, which is inspected by checking the two Nash conditions adapted to an expected utility set-up by Rothschild and Stiglitz (1976):

1. No contract in the equilibrium set makes negative expected profits: this holds true as all offered contracts make zero profit due to actuarially fair pricing.

2. There is no contract outside the equilibrium set that, if offered, will make a non-negative profit: a contract for risk classification insurance aiming at a lower cutoff increases the insurance premium for all individuals and is, thus, not in demand. Offering a contract aiming for a higher cutoff attracts all former risk classification buyers but not enough better risks such that it generates negative profits. This is because at any higher $z$ than the highest cutoff, $g(z)$ is positive and the expected utility without risk classification insurance is greater than if $z$ is a cutoff. Therefore, fewer individuals additionally join a higher cutoff than required and, thus, a lower premium than with the highest cutoff induces negative profits.
Conclusively, $z^* = 0$ will only be the equilibrium if there is no interior cutoff. Otherwise, $z^* = 0$ is dominated in the Pareto sense and the (highest in the case of multiple) interior cutoff is the equilibrium.\(^{16}\) This proposition also guarantees the existence of a unique equilibrium in any case.

4 Examples

To illustrate the results further, we will discuss several examples in order to determine thresholds and more detailed conditions for interior cutoffs. We first examine different signal distributions. Figure 2 displays the three different density functions\(^{17}\) used in the following remark.

![Figure 2: Examples of densities for the distributions in Remark 6](image)

**Remark 6.** 1. Assuming $Z \sim \text{unif}[0,1]$, which implies an average probability of becoming a low risk of $\frac{1}{2}$, there is an interior cutoff if and only if

\[
\frac{u(W_2) - u(W_1)}{W_2 - W_1} < \frac{1}{2} u'(W_1).
\]

\(^{16}\)Note that in this setting, self-selecting equilibria cannot be implemented as insurers only vary the price of risk classification insurance.

\(^{17}\)Note that these density functions can be derived from a beta distribution. The beta distribution permits full flexibility in the modeling of information. However, to establish analytical tractability, we decided to only examine special beta distributions that nevertheless capture essential features of the signal.
2. Assuming $Z$ to have a density $(\alpha_1 + 1)x^{\alpha_1} \cdot 1_{[0,1]}(x)$, $\alpha_1 > 0$, which implies an average probability of becoming a low risk of $\frac{\alpha_1 + 1}{\alpha_1 + 2} \in \left(\frac{1}{2}, 1\right)$, there is an interior cutoff if and only if
\[
\frac{u(W_2) - u(W_1)}{W_2 - W_1} < \frac{\alpha_1 + 1}{\alpha_1 + 2} u'(W_1).
\] (10)

3. Assuming $Z$ to have a density $(1 - \alpha_2)x^{-\alpha_2} \cdot 1_{[0,1]}(x)$, $\alpha_2 \in (0, 1)$, which implies an average probability of becoming a low risk of $\frac{1 - \alpha_2}{2 - \alpha_2} \in \left(0, \frac{1}{2}\right)$, there is an interior cutoff if and only if
\[
\frac{u(W_2) - u(W_1)}{W_2 - W_1} < \frac{1 - \alpha_2}{2 - \alpha_2} u'(W_1).
\] (11)

Proof. Refer to the appendix for the proof.

Remark 6 demonstrates that the necessary and sufficient condition for an interior cutoff can be fulfilled more easily for the second signal distribution as $\frac{\alpha_1 + 1}{\alpha_1 + 2} > \frac{1}{2}$ for $\alpha_1 > 0$ and is more difficult to fulfill for the third signal distribution as $\frac{1 - \alpha_2}{2 - \alpha_2} < \frac{1}{2}$ for $\alpha_2 \in (0, 1)$. As Figure 2 shows, the second distribution has more probability mass on the right and the density function is strictly increasing. Therefore, the subsidization of others is less painful as there are on average fewer high risks in the population and an interior cutoff is achieved more easily. The reverse holds true for the third type of signal.

In the next step we determine the impact of different types of risk aversion for a given signal distribution. The next remark gives an overview of the technical manipulations that will be interpreted afterwards.

Remark 7. Let $Y$ denote $W_2 - W_1 = P_H - P_L$.

1. Assuming $u(x) = -e^{-\beta x}$ for $\beta > 0$ and, thus, making individuals exhibit CARA, condition (6) becomes
\[
1 - e^{-\beta Y} < \lambda \beta Y
\] (12)
and it is met for $\lambda < 1$ if and only if $\beta Y > x_1(\lambda)$ for an implicitly defined threshold $x_1 = x_1(\lambda)$.

2. Assuming $u(x) = x^\gamma$ for $\gamma \in (0, 1)$ and thus making individuals exhibit DARA and CRRA, condition (6) turns to
\[
\frac{W_1}{Y} \left(1 + \frac{Y}{W_1}\right)^\gamma < \lambda \gamma + \frac{W_1}{Y}.
\] (13)

\footnote{The same can be said about $u(x) = \ln x$, where we obtain
\[
\ln \left(1 + \frac{Y}{W_1}\right) < \lambda \frac{Y}{W_1},
\]
but there we lack the control of different levels of risk aversion.}
Assuming $\lambda < 1$ it is fulfilled if and only if $\frac{Y}{W_1} > x_2(\lambda)$ for an implicitly defined threshold $x_2 = x_2(\lambda)$.

In any case we obtain that the smaller $\lambda$, the greater the threshold.$^{19}$

**Proof.** Refer to the appendix for the proof. \hfill $\square$

Assuming $\lambda < 1$ and inserting $Y$ and $W_1$, condition $\beta Y > x_1(\lambda)$ can be rewritten as

$$(p_H - p_L)T > \frac{x_1(\lambda)}{\beta}$$

and condition $\frac{Y}{W_1} > x_2(\lambda)$ changes to

$$\left(\frac{2}{x_2(\lambda)}(p_H - p_L) + p_H\right)T > W.$$  

The interpretation is as follows: The loss $T$ must be sufficiently substantial compared to $W$ and the difference between loss probabilities of high and low risks must be sufficiently high for an interior cutoff to exist, which underlines the general findings from remark 5. For CARA utility functions the same holds true, but there is no wealth effect. Thus, the loss $T$ or the loss ratio $\frac{T}{W}$ and the difference between the loss probabilities qualitatively influences whether risk classification insurance is purchased. Their quantitative interaction is determined by the shape of the underlying utility function. Taking the influence of the degree of risk aversion into account allows further insights. A closer look at (14) reveals that the higher $\beta$ is the smaller the difference has to be between the loss probability for low and high risks in order to have an interior cutoff. Thus, a higher degree of absolute risk aversion enhances the possibility of risk classification insurance being purchased. Analyzing condition (13) further shows that the higher $\gamma$ is the higher the threshold $x_2(\lambda)$ will be.$^{20}$ As the coefficient of RRA is $1 - \gamma$, a higher RRA also lowers the bound for the difference in loss probabilities that has to be exceeded in order to find an interior cutoff. Thus, higher RRA favors the purchase of risk classification insurance as well.

So, we can confirm the general results from Remark 5, elaborate on the influence of risk aversion for well-established utility functions in a quantitative way and demonstrate that condition (5) is likely to be fulfilled under realistic assumptions on the parameters.

---

$^{19}$This statement holds true even for an arbitrary $u$ due to the implicit function theorem, given that the respective condition is met at all.

$^{20}$The analytical dependence on $\gamma$ was suppressed in the notation for matters of clearness.
5 Conclusion

In a two point in time model we demonstrate that allowing for fuzzy private information that is heterogeneous among agents can explain the observed demand for risk classification insurance while standard assumptions regarding private information cannot. Therefore, we first show that assuming all insured to be perfectly informed about their risk type, risk classification insurance does not contain any value as only high risks buy it. In the next step, we examine a situation where nobody has private information and there is only one risk type at the first point in time. In such a case, everybody favors of risk classification insurance. It is difficult to argue why there will be risk classification in the first place if individuals do not have any information on their risk type at all. Subsequently, we model fuzzy private information, which is heterogeneous among agents, by introducing a signal. Each individual receives a private signal on their risk type and can buy risk classification insurance for the second point in time. We show that a situation where no one buys risk classification insurance is always a candidate for an equilibrium while a situation where everyone buys risk classification insurance can never be an equilibrium. In addition, we derive a sufficient condition for an equilibrium candidate to exist where only part of the insured population chooses risk classification insurance. A higher degree of risk aversion, a higher volatility in outcomes without risk classification insurance and a steeper premium function at zero make this condition more easily fulfilled. We show that a concave premium function provides the above-mentioned sufficient condition also to be necessary for an interior cutoff, which is unique in this case.

The last part of the analysis examines social welfare in our model. We show that a higher cutoff always dominates a lower one in the Pareto sense. Accordingly, the equilibrium is always the highest cutoff where social welfare is also higher than without risk classification insurance. We provide several examples to illustrate the above-mentioned results.

In addition to providing a theoretical explanation for the observed demand for risk classification insurance, the paper also includes crucial pricing information for risk classification insurance. In particular, the information that rather individuals with a high likelihood of becoming a high risk purchase risk classification insurance. Furthermore, the paper analyzes the market features that lead to a demand for risk classification insurance and thus allows the identifications those markets where the offer of such insurance makes sense from an economic point of view.

The paper focused on a set up where the risk classification insurance contract was sold individually from the actual insurance contract. As discussed in the introduction, some-
times risk classification insurance is not sold separately but included in the second insurance contract as e.g. accident forgiveness in car insurance. However, the results from our paper seem to be transferable to such a context even though some features of joint contracts may then be neglected as e.g. trying to increase customer’s retention by the additional offer of risk classification protection.

Regarding future research, several additional aspects could be included in the analysis. First, we assume that individuals can never buy partial risk classification insurance coverage as all the examples in insurance markets do not offer this option. Relaxing this assumption as well as adding further points in time to the analysis may lead to very interesting theoretical results. In addition, an application of the signal distribution to the general one-period Rothschild and Stiglitz (1976) insurance set-up also seems a worthwhile expansion of our paper.

References


A Proof of Proposition 1

(1) yields the following three first-order conditions:

\[ \frac{\partial V}{\partial \alpha_L} = z \cdot (1 - p_L) \cdot p_L \cdot T \cdot (u'(W - C_1 - T + \alpha_L T - \alpha_L p_LT)) \]
\[ - u'(W - C_1 - \alpha_L p_LT)) \overset{1}{=} 0, \]

\[ \frac{\partial V}{\partial \alpha_H} = z \cdot p_H \cdot (1 - p_H) \cdot T \cdot [u'(W - C_1 - T + \alpha_H \cdot T - \alpha_H \cdot p_H \cdot T) - u'(W - C_1 - \alpha_H \cdot p_H \cdot T)] \overset{1}{=} 0, \]

and

\[ \frac{\partial V}{\partial C_1} = u'(C_1) - z \cdot [p_L \cdot u'(W - C_1 - T + \alpha_L \cdot T - \alpha_L \cdot p_L \cdot T)] \]
\[ + (1 - p_L) \cdot u'(W - C_1 - \alpha_L \cdot p_L \cdot T)] - (1 - z) \cdot [p_H \cdot u'(W - C_1 - T + \alpha_H \cdot T - \alpha_H \cdot p_H \cdot T) \]
\[ + (1 - p_H) \cdot u'(W - C_1 - \alpha_H \cdot p_H \cdot T)] \overset{1}{=} 0. \]
From (15) and (16) it follows that full insurance is optimal regardless of whether the individual turns out to be a low or a high risk. Thus, the third first-order condition (17) simplifies to

\[ u'(C_1) - z \cdot u'(W - C_1 - p_L \cdot T) - (1 - z) \cdot u'(W - C_1 - p_H \cdot T) = 0. \]  

(18)

As \( W - C_1 - p_L \cdot T > W - C_1 - p_H \cdot T \), it holds that \( W - C_1 - p_L \cdot T > C_1 > W - C_1 - p_H \cdot T \) for the optimal \( C_1 \). To assure maximality we analyze the Hessian of \( V \) for \((\alpha_H, \alpha_L) = (1, 1)\) and the optimal \( C_1 \) and obtain:

\[
\begin{pmatrix}
\frac{\partial^2 V}{\partial \alpha^2_L} & \frac{\partial^2 V}{\partial \alpha_L \partial \alpha_H} & \frac{\partial^2 V}{\partial \alpha_H \partial C_1} \\
\frac{\partial^2 V}{\partial \alpha_L \partial \alpha_H} & \frac{\partial^2 V}{\partial \alpha_H^2} & \frac{\partial^2 V}{\partial \alpha_H \partial C_1} \\
\frac{\partial^2 V}{\partial \alpha_L \partial C_1} & \frac{\partial^2 V}{\partial \alpha_H \partial C_1} & \frac{\partial^2 V}{\partial C_1^2}
\end{pmatrix}
\bigg|_{\alpha_H=1, \alpha_L=1, C_1}
= \begin{pmatrix}
-0 & 0 & 0 \\
0 & -0 & 0 \\
0 & 0 & -0
\end{pmatrix},
\]

as

\[
\frac{\partial^2 V}{\partial \alpha^2_L} \bigg|_{\alpha_L=1, C_1} = z(1 - p_L)p_L T^2 u''(W - C_1 - p_L T),
\]

\[
\frac{\partial^2 V}{\partial \alpha_H^2} \bigg|_{\alpha_H=1, C_1} = z p_H (1 - p_H) T^2 u''(W - C_1 - p_H T), \quad \text{and}
\]

\[
\frac{\partial^2 V}{\partial C_1^2} \bigg|_{\alpha_H=1, \alpha_L=1, C_1} = u''(C_1) + z u''(W - C_1 - p_L T)
+ (1 - z) u''(W - C_1 - p_H T),
\]

so the Hessian proves to be negative-definite.

**B Proof of Remark 1**

Consider the function\(^{22}\)

\[
f(C_1) := \frac{u'(C_1) - u'(W - C_1 - p_L \cdot T)}{u'(W - C_1 - p_L \cdot T) - u'(W - C_1 - p_H \cdot T)}. \]  

(19)

One can establish that

\(^{21}\)To avoid overloading the notation, the optimal consumption will not be separately labeled, but it will be clear from the context when it is addressed.

\(^{22}\)All these calculations can be done for arbitrary \( C_1 \).
\[ f(C_1) \geq 0 \iff C_1 \geq \frac{W - p_H \cdot T}{2} \]

and

\[ f(C_1) \leq 1 \iff C_1 \leq \frac{W - p_L \cdot T}{2}. \]

Differentiating \( f \) with respect to \( C_1 \) gives after some calculations

\[
f'(C_1) = \left( (u'(W - C_1 - p_L \cdot T) - u'(W - C_1 - p_H \cdot T)) u''(C_1) \right. \\
+ \left. (u'(W - C_1 - p_L \cdot T) - u'(C_1)) u''(W - C_1 - p_H \cdot T) \right. \\
+ \left. (u'(C_1) - u'(W - C_1 - p_H \cdot T)) u''(W - C_1 - p_L \cdot T) \right) \\
/ \left( u'(W - C_1 - p_L \cdot T) - u'(W - C_1 - p_H \cdot T) \right)^2,
\]

which proves to be positive for the above specified range of \( C_1 \). It can be verified that

\[ f \left( \frac{W - p_H \cdot T}{2} \right) = 0 \quad \text{and} \quad f \left( \frac{W - p_L \cdot T}{2} \right) = 1. \]

Rearranging (18) yields \( f(C_1) = z \). Let \( C_1 \) be optimal\(^{23}\), so (18) is fulfilled, in particular \( f(C_1) \in [0, 1] \). That is why \( C_1 \) lies in the above specified interval and \( f'(C_1) > 0 \) ensures uniqueness. On the other hand for any given \( z \) we can find a \( C_1 \) in the interval, that satisfies (18) due to the intermediate value theorem.

The intuition behind equation (19) is the following: The denominator represents the maximal difference in marginal utility comparing wealth today and wealth tomorrow. As the nominator ranges between 0 and the denominator, the whole fraction at the optimum represents the portion of marginal utility difference and equals the probability of becoming a low risk. For sure low risks (\( z = 1 \)) this difference can be fully exploited, for sure high risks (\( z = 0 \)) it cannot be carried out at all. All individuals with \( z \)'s between 0 and 1 thus solve the trade-off through the choice of the optimal \( C_1 \) and according to the level of uncertainty about their future risk type which is reflected in the portion of marginal utility that is attainable.

\(^{23}\)To avoid overloading the notation we drop the * to indicate the optimal \( C_1 \).
C Proof of Remark 2

For a given $z$ the optimal $C_1$ is given by equation (18), which can be rewritten as

$$F(z, C_1) := f(C_1) - z = 0.$$  

Obviously $F$ is a $C^1$-function and hence the implicit function theorem can be applied to obtain the first derivative

$$\frac{dC_1}{dz} = -\left(\frac{dF}{dC_1}(z, C_1)\right)^{-1} \cdot \frac{dF}{dz}(z, C_1) = \frac{1}{f'(C_1)} > 0.$$  

This proves the first part of the remark.

Checking for the second derivative we get

$$\frac{d^2C_1}{dz^2} = -\frac{\frac{d}{dz}f'(C_1)}{(f'(C_1))^2} = -\frac{f''(C_1)\frac{dC_1}{dz}}{(f'(C_1))^2},$$

so to determine the sign of the second derivative we have to work out $\frac{d}{dz}f'(C_1)$. After some computations we obtain

$$\frac{d}{dz}f'(C_1) = (u'(W - C_1 - p_LT) - u'(W - C_1 - p_HT)) \frac{dC_1}{dz} \frac{d}{dz} \left[ (u'(W - C_1 - p_LT) - u'(W - C_1 - p_HT)) u'''(C_1) \right. \\
- \left. (u'(W - C_1 - p_LT) - u'(C_1)) u''(W - C_1 - p_HT) \right. \\
- \left. (u'(C_1) - u'(W - C_1 - p_HT)) u''(W - C_1 - p_LT) \right] \\
+ \frac{2}{(u'(W - C_1 - p_LT) - u'(W - C_1 - p_HT))} \frac{dC_1}{dz}$$

Using $\frac{dC_1}{dz} = \frac{1}{f'(C_1)}$, this further simplifies to

$$\left[ (u'(W - C_1 - p_LT) - u'(W - C_1 - p_HT)) u'''(C_1) \right. \\
- \left. (u'(W - C_1 - p_LT) - u'(C_1)) u''(W - C_1 - p_HT) \right. \\
- \left. (u'(C_1) - u'(W - C_1 - p_HT)) u''(W - C_1 - p_LT) \right]/$$
and

\[ (u'(W - C_1 - p_L \cdot T) - u'(W - C_1 - p_H \cdot T)) u''(C_1) \]

\[ + (u'(W - C_1 - p_L \cdot T) - u'(C_1)) u''(W - C_1 - p_L \cdot T) \]

\[ + (u'(C_1) - u'(W - C_1 - p_H \cdot T)) u''(W - C_1 - p_L \cdot T) \]

\[ + 2 \cdot \frac{u''(W - C_1 - p_L T) - u''(W - C_1 - p_H T)}{u'(W - C_1 - p_L T) - u'(W - C_1 - p_H T)} \]

CARA with \( r_A(x) = -\frac{u''(x)}{u(x)} = \beta \) \(^{24}\) for instance implies \( \frac{u''(x)}{u(x)} = \beta^2 \). Thus \( \frac{d}{dx} f'(C_1) \) further simplifies and finally proves to be \(-2\beta < 0\). Hence \( \frac{d^2 C_1}{dz^2} > 0 \) which completes the proof.

## D  Proof of Proposition 2

We obtain the following first-order conditions:

\[
\frac{\partial V}{\partial \alpha} = -(1 - \eta) \cdot (p_H - p_L) \cdot T \cdot u'(C_1 - (1 - \eta) \cdot \alpha \cdot (p_H - p_L) \cdot T) \tag{20}
\]

\[ + (\eta \cdot p_L + (1 - \eta) \cdot p_H) \cdot T \cdot (1 - p_L) \cdot u'(W - C_1 - T + \alpha \cdot T - \alpha \cdot p_L \cdot T) \]

\[ - (\eta \cdot (1 - p_L) + (1 - \eta) \cdot (1 - p_H)) \cdot p_L \cdot T \cdot u'(W - C_1 - \alpha \cdot p_L \cdot T) \]

\[ = 0, \]

and

\[
\frac{\partial V}{\partial C_1} = u'(C_1 - (1 - \eta) \cdot \alpha \cdot (p_H - p_L) \cdot T) \tag{21}
\]

\[ - (\eta \cdot p_L + (1 - \eta) \cdot p_H) \cdot u'(W - C_1 - T + \alpha \cdot T - \alpha \cdot p_L \cdot T) \]

\[ - (\eta \cdot (1 - p_L) + (1 - \eta) \cdot (1 - p_H)) \cdot u'(W - C_1 - \alpha \cdot p_L \cdot T) \]

\[ = 0 \]

\( \alpha = 1 \) and \( C_1 = \frac{W + (1 - \eta) \cdot T \cdot (p_H - p_L) - p_L \cdot T}{2} \) solve the first-order conditions. Indeed, inserting \( \alpha = 1 \) into equation (20) yields

\[
\frac{\partial V}{\partial \alpha} = -(1 - \eta)(p_H - p_L)Tu'(C_1 - (1 - \eta)(p_H - p_L)T)
\]

\[ + [(1 - p_L)T(\eta p_L + (1 - \eta)p_H) - p_LT(\eta(1 - p_L) + (1 - \eta)(1 - p_H))] \]

\[ u'(W - C_1 - p_L T) \]

\[ = -(1 - \eta)(p_H - p_L)Tu'(C_1 - (1 - \eta)(p_H - p_L)T) \]

\[ + T(1 - \eta)(p_H - p_L)u'(W - C_1 - p_L T) \]

\(^{24}\) \( r_A(x) \) denotes the Arrow-Pratt measure of local absolute risk aversion.
\[
= (1 - \eta)(p_H - p_L)T \left[ u'(W - C_1 - p_L T) - u'(C_1 - (1 - \eta)(p_H - p_L)T) \right] \\
\equiv 0
\]

Using \( \eta \neq 1 \) and the injectivity of \( u' \) due to \( u'' < 0 \), we can conclude that it is fulfilled if and only if

\[
C_1 = \frac{1}{2}(W - p_L T + (1 - \eta)(p_H - p_L) T).
\]

Using \( \alpha = 1 \) for equation (21), we get

\[
\frac{\partial V}{\partial C_1} = u'(C_1 - (1 - \eta)(p_H - p_L)T) - (\eta p_L + (1 - \eta)p_H)u'(W - C_1 - p_L T) \\
- \ (\eta(1 - p_L) + (1 - \eta)(1 - p_H))u'(W - C_1 - p_L T) \overset{!}{=} 0 \\
\Rightarrow u'(C_1 - (1 - \eta)(p_H - p_L)T) = u'(W - C_1 - p_L T),
\]

and again

\[
C_1 = \frac{1}{2}(W - p_L T + (1 - \eta)(p_H - p_L) T).
\]

Checking for the second-order condition we obtain

\[
\frac{\partial^2 V}{\partial C_1^2} = u''(C_1 - (1 - \eta)\alpha(p_H - p_L) T) \\
+ \ (\eta p_L + (1 - \eta)p_H)u''(W - C_1 - T + \alpha T - \alpha p_L T) \\
+ \ (\eta(1 - p_L) + (1 - \eta)(1 - p_H))u''(W - C_1 - \alpha p_L T)
\]

and with the above-calculated values for \( \alpha \) and \( C_1 \) this resolves to

\[
\frac{\partial^2 V}{\partial C_1^2} = \left(1 + \eta p_L + (1 - \eta)p_H + \eta(1 - p_L)\right)u'' \left( \frac{W - p_L T - (1 - \eta)(p_H - p_L) T}{2} \right) \\
+ \ (1 - \eta)(1 - p_H)u'' \left( \frac{W - p_L T - (1 - \eta)(p_H - p_L) T}{2} \right) < 0.
\]

Doing so for equation (20) yields
\[
\frac{\partial^2 V}{\partial \alpha \partial^2} = (1 - \eta)^2 (p_H - p_L)^2 T^2 u'' (C_1 - (1 - \eta) (p_H - p_L) T) + (\eta (1 - p_L) + (1 - \eta) (1 - p_H)) p^2 L T^2 u'' (W - C_1 - \alpha T - \alpha p_L T).
\]

In order to determine the Hessian, we obtain

\[
\frac{\partial^2 V}{\partial \alpha \partial C_1} = (1 - \eta) (p_H - p_L) T u'' (C_1 - (1 - \eta) (p_H - p_L) T). \]

and using the \(\alpha\) and \(C_1\) from above this simplifies to

\[
-2T^2 (p_H - p_L)^2 (1 - \eta) u'' (W - (1 - \eta) (p_H - p_L) T) + p_T (1 - p_L) (p_H - p_L) + (1 - \eta) (p_H - p_L) T u'' (W - C_1 - \alpha T - \alpha p_L T).
\]

Calculating the determinant of the Hessian yields

\[
\det \left( \begin{array}{ccc}
\frac{\partial^2 V}{\partial \alpha \partial^2} & \frac{\partial^2 V}{\partial \alpha \partial C_1} & \frac{\partial^2 V}{\partial \alpha \partial C_1} \\
\frac{\partial^2 V}{\partial \alpha \partial C_1} & \frac{\partial^2 V}{\partial C_1 \partial^2} & \frac{\partial^2 V}{\partial C_1 \partial \alpha} \\
\frac{\partial^2 V}{\partial \alpha \partial C_1} & \frac{\partial^2 V}{\partial C_1 \partial \alpha} & \frac{\partial^2 V}{\partial C_1 \partial^2}
\end{array} \right) \bigg| \bigg| \begin{array}{c}
\alpha, C_1 = (1, W + (1 - \eta) (p_H - p_L)) T \\
\alpha, C_1 = (1, W + (1 - \eta) (p_H - p_L)) T
\end{array}
\]

\[
= 2T^2 ((1 - \eta)^2 (p_H - p_L)^2 + (1 - \eta) (1 - \eta) (p_H - p_L)^2 + (1 - \eta) (p_H - p_L) T u'' (W - C_1 - \alpha T - \alpha p_L T))
\]

\[
\geq \eta p_L (1 - p_L) + \eta p_H (1 - \eta) > 0.
\]
For the last step we use the fact, that $0 < \eta = E(Z) < 1$ and accordingly $\eta^2 < \eta$ and $-\eta^2(p_H - p_L)^2 > -\eta(p_H - p_L)^2$. Thus the determinant of the Hessian is positive and the first entry is negative, which proves that the Hessian is negative-definite, so we have a maximum. Accordingly, the optimal strategy when buying risk classification insurance is to consume equally at both points in time and to fully insure the health risk. The overall utility is then

$$V = 2 \cdot u \left( \frac{W - (1 - \eta) \cdot T \cdot (p_H - p_L) - p_L \cdot T}{2} \right).$$

This is greater than the optimal strategy without risk classification insurance as the expected value of outcomes is the same with and without risk classification insurance, however without risk classification insurance, there is a positive variance of outcomes.

### E Proof of Remark 3

We determine

$$E(Z|Z \leq z^*) = \frac{\int_0^1 1_{\{Z < z^*\}} \cdot z \cdot h(z) dz}{\int_0^z h(z) dz} = \frac{\int_{z^*}^{z^*} z \cdot h(z) dz}{F(z^*)}$$

with $h$ being the density of the distribution of the signal $Z$ and $F$ being its cumulative distribution function, and

$$\frac{d}{dz^*}E(Z|Z \leq z^*) = \frac{z^* \cdot h(z^*) \cdot F(z^*) - h(z^*) \cdot \int_0^{z^*} z \cdot h(z) dz}{(F(z^*))^2}.$$  

This expression is always greater than zero as $z^* \cdot F(z^*) > \int_0^{z^*} z \cdot h(z) dz$. Accordingly,

$$P'(z^*) = -\frac{z^* \cdot h(z^*) \cdot F(z^*) - h(z^*) \cdot \int_0^{z^*} z \cdot h(z) dz}{(F(z^*))^2} \cdot (p_H - p_L) < 0.$$
### F  Proof of Remark 6

The three cases can be handled simultaneously. We first check the expectation of the different distributions:

\[
E(Z) = \begin{cases} 
\int_0^1 x^2 \, dx = \frac{1}{2} \\
\int_0^1 (\alpha_1 + 1)x^{\alpha_1+1} \, dx = \frac{\alpha_1+1}{\alpha_1+2} \\
\int_0^1 (1 - \alpha_2)x^{-\alpha_2+1} \, dx = \frac{1}{2-\alpha_2} 
\end{cases}
\]

The conditional expectations resolve to

\[
E(Z|Z \leq z) = \begin{cases} 
\int_0^z x \, dx = \frac{1}{2} z \\
\int_0^z (\alpha_1 + 1)x^{\alpha_1+1} \, dx = \frac{\alpha_1+1}{\alpha_1+2} z \\
\int_0^z (1 - \alpha_2)x^{-\alpha_2+1} \, dx = \frac{1}{2-\alpha_2} z 
\end{cases}
\]

Plugging this information into (5) and differentiating yields

\[
\frac{dP}{dz}(z) = \begin{cases} 
-\frac{1}{2}(P_H - P_L) \\
-\frac{\alpha_1+1}{\alpha_1+2}(P_H - P_L) \\
-\frac{1-\alpha_2}{2-\alpha_2}(P_H - P_L)
\end{cases}
\]

Accordingly, \( \frac{d^2P}{dz^2}(z) = 0 \) and condition (6) is not only sufficient but also necessary in all three cases. Inserting \( \frac{dP}{dz}(z) \) into (6) yields the three above stated conditions.

### G  Proof of Remark 7

For the first case, we calculate

\[
\frac{-e^{-\beta W_2} + e^{-\beta W_1}}{W_2 - W_1} < \lambda \beta e^{-\beta W_1} \iff e^{-\beta W_1} - e^{-\beta (W_1 + Y)} < \lambda \beta Y e^{-\beta W_1}.
\]

Cancelling out \( e^{-\beta W_1} \), we obtain

\[
1 - e^{-\beta Z} < \lambda \beta Z.
\]

Figure 3 displays both sides of inequality (12) with \( \beta Z \) being charted on the x-axis and illustrates that for \( \lambda < 1 \) a threshold \( x_1(\lambda) \neq 0 \) is obtained which decreases in \( \lambda \).
For the second case

\[
\frac{W_2^\gamma - W_1^\gamma}{W_2 - W_1} < \lambda \gamma W_1^{\gamma-1} \iff \frac{(W_1 + Y)^\gamma - W_1^\gamma}{Y^\gamma Y^{1-\gamma}} < \lambda \gamma W_1^{\gamma-1}
\]

\[
\iff \left(1 + \frac{W_1}{Y}\right)^\gamma - \left(\frac{W_1}{Y}\right)^\gamma < \lambda \gamma \left(\frac{W_1}{Y}\right)^{\gamma-1} \iff \frac{W_1}{Y} \left(1 + \frac{Y}{W_1}\right)^\gamma < \lambda \gamma + \frac{W_1}{Y}.
\]

Figure 4: Graphical illustration of condition (13) for different values of \(\lambda\) and \(\gamma = 0.5\)

Figure 4 illustrates both sides of inequality (13) with \(\frac{W_1}{Y}\) being charted on the x-axis and shows that for \(\lambda < 1\) a threshold \(x_2(\lambda) \neq 0\) is given as a decreasing function in \(\lambda\). Alternatively, the implicit function theorem can be applied here to prove the dependence of the thresholds on \(\lambda\).