

Fels, Markus

Conference Paper

Why Do People Buy Insurance? A Modern Answer to an Old Question

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Fels, Markus (2021) : Why Do People Buy Insurance? A Modern Answer to an Old Question, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242418>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Why Do People Buy Insurance? A Modern Answer to an Old Question

Markus Fels*

28th February, 2021

Abstract

I revisit the question of which motive underlies insurance demand. I draw on the literature of state-dependent utility and on the literature of imperfectly divisible consumption to argue that the general purpose of insurance is not a risk transfer, but meeting a conditional need. In this way, insurance aligns the risk in one's financial endowment with the risk in one's financial needs. This generalization extends the classical view of insurance. I show how this extension greatly simplifies the generalization of classic results. I also discuss how the novel definition has implications for our discipline's research agenda and policy advice.

JEL Classification: D01, D81, G22

Keywords: Insurance, Risk Aversion, State Dependence, Divisibility of Consumption, Risk Preferences

*University of Dortmund (TU), Department of Economics, Vogelpothsweg 87, 44227 Dortmund, Germany; E-mail address: markus.fels@udo.edu.

1 Introduction

There is hardly a question in Economics to which members of the discipline - proverbial in their inclination to disagree - seem more united in their answer. When asked what makes people buy insurance, most economists answer “risk aversion”, a general preference of a certain amount of wealth over an uncertain one with identical expected value. Such a risk-averse person is willing to pay a positive amount of money, a risk premium, to exchange her uncertain wealth for a certain one. Buying insurance is regarded as exactly that trade. The two topics of insurance and risk aversion are inextricably intertwined in the canon of Economics. While above account suggests that one, risk aversion, is supposed to explain the other, insurance, the economic literature is fraught with examples of insurance being regarded as the ultimate evidence of risk aversion: We see a lot of people buying insurance, hence, they must be risk-averse (Arrow 1971, p. 91, Gollier 2001, p.18). In this way, the tautological circle closes: people buy insurance because they are risk-averse. And how do we know that they are risk-averse? Because they insure. It is this tautology that seems to underlie the remarkable resilience of the idea of risk aversion as a general characteristic of people’s preferences in the face of empirical evidence that questions its generality. It is further reinforced by various prominent economists equating the concept of risk aversion with the concept of diminishing marginal utility of money (Yaari 1965; Arrow 1974; Epstein and Tanny 1980; Karni 1983; Karni 1985) and the intuitive plausibility of the latter. As I will argue below, the idea of a diminishing value of money can be preserved without assuming risk aversion. It is thus the combination of an observation of bustling insurance markets and the false equation of risk aversion with diminishing marginal utility of money that keeps risk aversion alive despite growing reservations.¹

¹In a similar vein, O’Donoghue and Somerville (2018) point to non-expected utility models as alternative explanations for behavior, such as insurance, associated with risk aversion. In contrast, I seek to emphasize that insurance can be explained within an expected-utility framework without assuming risk aversion. Retaining the expected-utility framework has the advantage of emphasizing alternative motives for insurance purchase, while non-expected utility frameworks are better suited capturing the decision processes involved in expressing one’s motives. The resort to blaming the failures of the risk-aversion hypothesis to explain behavior on the expected-utility framework has a long

This paper seeks to answer the question why people buy insurance without resorting to the idea of risk aversion. Based on several theoretical developments in the literature, I present three archetypical models of insurance demand based on different motives: risk aversion, state-dependent preferences, the access motive. These stylized models allow two conclusions. First, risk aversion should be regarded as an exception rather than the rule, and, second, people have plenty of reason to buy insurance even if they are not risk-averse. In the course of this, the question of what actually constitutes insurance is revisited and answered in a different, although not entirely new, way.

I proceed to show how the novel definition of insurance can help to find simple generalizations to two classic insights of the insurance literature: the optimality of a deductible (Arrow 1963; Mossin 1968; Arrow 1971) and the effect of ex-ante moral hazard (Pauly 1974; Shavell 1979). I discuss why the novel definition allows for simple generalizations instead of resulting in the vast number of case distinctions that characterize the existing literature of insurance under state-dependent preferences (Dionne 1982; Rey 2003; Huang and Tzeng 2006).

Finally, I point out how the novel definition allows a new perspective on existing and ongoing research in several applications of (insurance) economics. The idea of insurance being based on risk aversion, and risk aversion alone, has broad implications for our discipline. It influences the research agenda that our discipline follows and the policy advice that it offers to the societies that we live in. I specify how a modern idea of insurance can help us design a new research agenda and improve our policy advice.

tradition. Friedman and Savage (1948) formulated their well-known hypothesis in response to the criticism that the simultaneous observation of insurance and gambling indicate the failure of attempting to rationalize behavior with (expected-) utility maximization. Later, the observation of a significant risk aversion over modest stakes and its implication for risk aversion over large stakes (Rabin 2000), now known as the Rabin Calibration Paradox, has been interpreted as evidence against the expected-utility framework (Rabin and Thaler 2000), overlooking the possibility that it may equally well disqualify the idea of a universally concave utility function, and, in this way, the idea of risk aversion.

2 Three simple models of insurance demand and a definition

In this section, I present three archetypes of motives for insurance demand. Starting with the classic model of insuring a wealth risk, I discuss the underlying mechanisms that make such an insurance desirable to a risk-averse individual. I then proceed to present two other motives for insurance: state-dependent preferences as a motive to insure longevity risk and the access motive as a motive to insure health risks. These models show why risk aversion is not the common denominator in insurance demand. By analyzing the assumptions underlying each model, it is possible to arrive at a general idea of what motivates insurance demand.

2.1 The classic: Risk Aversion

Let us first revisit the traditional model of insurance in order to depict its basic ingredients. It posits that people are risk-averse and that insurance is a means to rid oneself of risk. Given that insurance is basically a transfer of wealth across states of nature, the usefulness of such transfers is linked to the assumption that the ideal distribution of wealth across states of nature is an equal distribution, i.e., there is a preference for certainty over risk, or, in short, risk aversion. The desirability of insurance then requires an a-priori unequal distribution of wealth that can be equalized through transfers of wealth from states with higher to states with lower wealth. Formally, consider two states of nature s_1 and s_2 that occur with probability p and $1 - p$ respectively. Let the a-priori wealth distribution be the lottery $\tilde{w} = (w_1, p; w_2, 1 - p)$ with expected value $\bar{w} = E[w]$. Let $U(\tilde{w}) = pu(w_1) + (1 - p)u(w_2)$ be the expected utility of the lottery \tilde{w} . The individual wants to insure state 1, i.e., transfer wealth from state 2 to state 1 if there is an actuarially-fair transfer $t > 0$, such that the lottery $\tilde{w}_t = (w_1 + t, p; w_2 - \frac{p}{1-p}t, 1 - p)$ is strictly preferred to \tilde{w} . If the person is risk-averse, then $u(\bar{w}) \geq U(\tilde{w})$, with strict inequality for any lottery \tilde{w} with $w_1 \neq w_2$.

Then, if $w_1 < w_2$, the transfer $t^* = \bar{w} - w_1 > 0$ is desirable as it equalizes wealth across states, $\tilde{w}_t = (\bar{w}, p; \bar{w}, 1 - p) = (\bar{w}, 1)$, and $U(\tilde{w}_t) = u(\bar{w}) > U(\tilde{w})$. Figure 1 illustrates the motive for insurance based on risk aversion familiar from various textbooks.

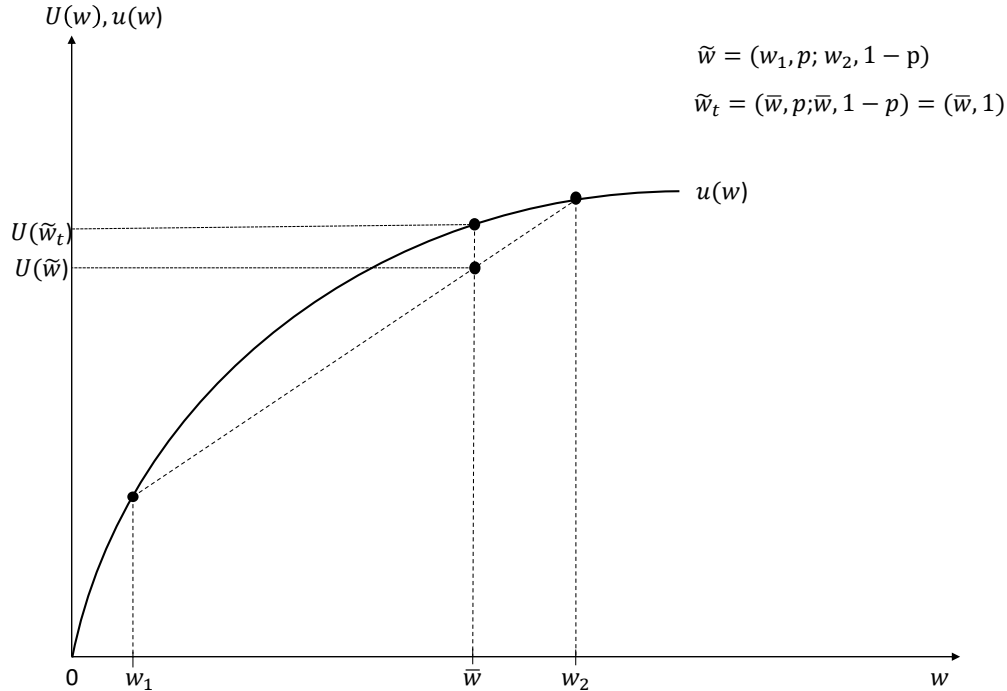


Figure 1: The utility from insuring a loss $L = w_2 - w_1$ in wealth if risk-averse

In this understanding, insurance is a means to reduce one's risk exposure. It is this view of insurance that forms the basic framework in which insurance theory operates and in which actual insurance arrangements are analyzed and evaluated. This is not remarkable given the historic origins of insurance as means to mitigate the financial consequences of major losses in wealth through shipwreck or fire. And it is still highly relevant as a framework to analyze potential losses in wealth, be it through fire, flood, or other disasters. However, the business of insurance now reaches far beyond the coverage of losses in financial assets, addressing personal risks such as disease, disability, longevity, and many more. Unfortunately, the framework of insurance has not evolved despite several opportunities for doing so.

Risk aversion, i.e., the desirability of an equal distribution of wealth across states, is typically justified with the plausibility of a diminishing marginal utility of wealth: $u''(w) < 0$. It implies that a unit of money transferred from a state with high wealth to a state with low wealth results in an increase in utility.

$$\begin{aligned} \frac{\partial U(\tilde{w}_t)}{\partial t} \Big|_{t=0} &= p [u'(w_1) - u'(w_2)] > 0 \\ &\Leftrightarrow u'(w_1) - u'(w_2) > 0 \\ &\stackrel{u'' < 0}{\Leftrightarrow} w_1 < w_2 \end{aligned}$$

It is important to recognize that it is the difference in the (marginal) value of money across state 1 and state 2, $u'(w_1) > u'(w_2)$, that ultimately motivates the wealth redistribution. However, utility is derived not from money itself but from the utility of the consumption goods that we can buy with it. This is the basic tenet on which the idea of diminishing utility of money is built. A rational individual spends money on the most important expenses first, with additional money being spent on expenditures of less and less value. Yet, if the utility of money is derived from the utility of the consumption opportunities that we can buy with it, then the optimality of an equal distribution of wealth across states of nature requires two conditions. First, the value of the consumption opportunities that money can buy must be identical across states of nature, i.e., the function $u(w)$ is state-independent. Second, an additional unit of money can only allow the purchase of less valuable items than any unit of money so far. I argue that both of these conditions are rarely met in practice, and are thus the reason why people are not risk-averse. I will also point out that both can form the basis of a motive to insure. Thus, ironically, the very reasons that make people not risk-averse make them want to buy insurance.

2.2 State-Dependent Utility of Money

It is straightforward that the value of certain consumption possibilities is state-dependent. Aggressive medical treatments or a place in a nursing home are not deemed desirable unless one finds oneself in a situation of particular need. Similarly, expenditures on various activities might have more or less value depending on the state of one's health. Given that, it is natural to assume that the (marginal) value of money depends on more than just the amount of money already available: $u(w) = u_s(w)$. The expected utility of a lottery is then given by $U(\tilde{w}) = pu_1(w_1) + (1 - p)u_2(w_2)$. Yet, if the marginal value of money is larger in some states than in others, this presents a rationale for insuring the former states.

$$\begin{aligned} \left. \frac{\partial U(\tilde{w}_t)}{\partial t} \right|_{t=0} &= p [u'_1(w_1) - u'_2(w_2)] > 0 \\ &\Leftrightarrow u'_1(w_1) - u'_2(w_2) > 0 \end{aligned}$$

Note that for insurance to be desirable in such a setting, a loss in state 1, such that $w_1 < w_2$, is no longer necessary. All that is required is a difference in the (marginal) utility of wealth $u'_s(w)$. If $w_1 = w_2 = \bar{w}$ but $u'_1(\bar{w}) > u'_2(\bar{w})$, then $\partial U(\tilde{w}_t)/\partial t|_{(t=0)} > 0$ and a desirable transfer $t > 0$ from state 2 to state 1 exists. Specifically, with $u_s(w)$ being twice continuously differentiable and $u'_s > 0, u''_s < 0$, the transfer t^* that maximizes expected utility is implicitly defined by

$$u'_1(w_1 + t^*) = u'_2(w_2 - \frac{p}{1-p}t^*). \tag{1}$$

In strong contrast to the previous setup, the optimal allocation of wealth across states can now be an unequal one. In conclusion, people are not risk-averse as they prefer particular unequal distributions

of wealth over an equal distribution.² More importantly, insurance is exactly the means to achieve these desirable distributions, and, hence, a means to acquire particular risks. These consequences of state-dependent utility are in sharp contrast to the traditional understanding of insurance and the motives that we assume to drive its purchase.

It is noteworthy that the idea of state-dependent utility is not novel (Hirshleifer 1966; Arrow 1974), but the literature seems more concerned with fitting its findings into the traditional framework than questioning the appropriateness of the framework. Cook and Graham (1977) and subsequent work define the “ransom” as the financial loss that is equivalent to the utility loss in case of an undesirable event such as disease, and then follow the traditional agenda of insurance economists to ask to what extent a rational individual would want to mitigate this loss through insurance. This line of inquiry is natural given the standard understanding of insurance. However, it also shows the basic shortcomings of the traditional framework. Conditions can be derived under which the rational individual underinsures, fully insures, or even overinsures the loss. In even stronger contrast to the standard understanding of insurance, conditions can be derived under which the rational individual wants to insure the state in which the loss does not occur.³ These findings all indicate that the traditional framework does not suffice anymore. First, even if insurance is fair, the optimal allocation of wealth across states is no longer an equal one. In short, an individual with state-dependent preferences is not risk-averse. Second, the loss itself is basically irrelevant for insurance demand. What matters is the (marginal) utility of money and how it differs across states. Finally, the way to achieve the optimal distribution of wealth across states is insurance. Yet, this means that insurance is a means to acquire the risk that is embodied in the optimal distribution.

The fact that losses have less relevance in a setting with state-dependent utility is particularly

²Note that this does not mean that they are risk-seeking as they only prefer some, and not all, unequal distributions over the equal distribution.

³To see this, simply consider the case in which $w_1 > w_2$ and $u'_1(w_1) > u'_2(w_2)$.

obvious in the case of insurance against the risk of longevity. Given that most people deem a long life more desirable than a shorter one, annuities are a prime example of an insurance that pays benefits in a state in which a gain occurs. Yaari (1965) sets up a model in which - plausibly - an individual derives utility from money only conditional on being alive. Yet, this is a model of state-dependent utility for which the optimal solution is to annuitize all wealth, or, put differently, transfer all wealth from the state in which it has no value, death, to the state in which it has positive value, life. Consider Figure 2 for a simple example in which survival (state 1) and death (state 2) are equally probable. Starting from a position of equal wealth in both states, $\tilde{w} = (\bar{w}, \frac{1}{2}; \bar{w}, \frac{1}{2})$, insurance allows to transfer all wealth from state 2, in which it has no value, to state 1, in which it has value. The optimal wealth distribution that is achieved through this transfer is $\tilde{w}_t = (2\bar{w}, \frac{1}{2}; 0, \frac{1}{2})$.

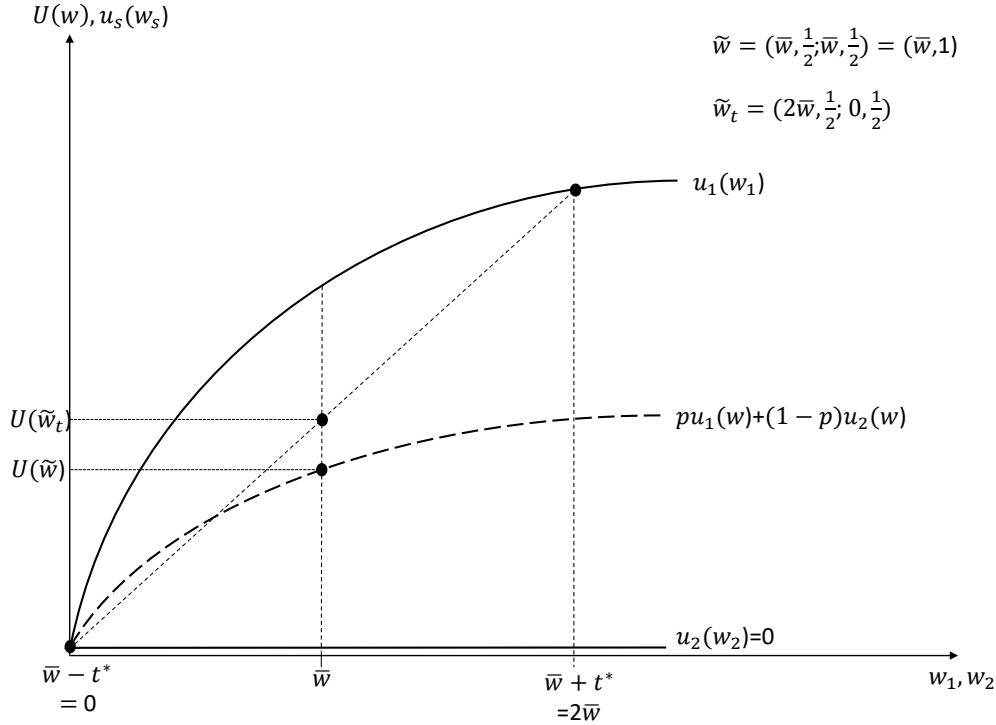


Figure 2: Insuring longevity

Is it in any way meaningful to compare the benefit size⁴ of $2\bar{w}$ to a cash-equivalent of the “loss”

⁴The optimal transfer requires an actuarial-fair insurance premium of \bar{w} in return for a benefit payment of $2\bar{w}$ in state 1.

occurring in state 1 in which the insurance pays off? Note that with $u_1(0) = u_2(w)$, the cash-equivalent of state 1 occurring is exactly the gain of \bar{w} , i.e., the individual must be compensated by an amount of $-\bar{w}$ for the “misery” of staying alive as compared to dying early. Calculating the amount of coverage of said “loss” as the relation between benefit size and ransom would then be -2 . Now, suppose the individual is a religious person believing in an afterlife. And, given that the individual is confident to be a devout person, it anticipates $u_2(w) = H$, with $H = u_1(w^h)$ for some immense amount $w^h \gg \bar{w}$. The amount of coverage would then be $2\bar{w}/(w^h - \bar{w}) \approx 0$ for $w^h \rightarrow \infty$. The cash-equivalent approach thus yields results on the amount of coverage that are highly dependent on the ex-ante assumptions on $u_1(w) - u_2(w)$.⁵

Above considerations show that the comparison of the benefit size to some monetary loss or to a cash-equivalent of a loss might seem intuitive in some contexts with $w_1 < w_2$ and $U_1(w) \leq U_2(w)$ (see the literature on irreplaceable assets). However, it cannot serve as a general framework for insurance under state-dependent preferences. In a general framework, it makes more sense to measure the amount of coverage by comparing the actual transfer t that is chosen in a given insurance model with state-dependent utility with the transfer t^* that would equalize marginal utility across states.⁶ The transfer t^* is the transfer that fully addresses the larger financial need that occurs in the insured state. Any deviation from t^* can then simply be interpreted as the amount of unmet need that results from premium loadings, asymmetric information, and other frictions that are known to decrease demand for coverage.

The optimal wealth distribution $\tilde{w}_t = (2\bar{w}, \frac{1}{2}; 0, \frac{1}{2})$ is clearly unequal, which means that the

⁵An alternative suggestion by Schlesinger (1984) to measure the level of coverage against the size of some monetary loss L that happens simultaneously with the change in the utility function is equally problematic. Again, the level of coverage now depends on the assumptions of $L = w_2 - w_1$. In our case above, $L = 0$, hence the level of coverage would be undefined. If we instead assumed that survival increases lifetime wealth due to a larger earnings potential, then $L < 0$, and the level of “loss coverage” would be negative.

⁶In case of a corner solution, as we have in the example above, t^* denotes the transfer that an individual selects if an actuarial-fair transfer is possible.

individual is not risk-averse. Yaari, unfortunately confusing diminishing marginal utility of wealth with risk aversion, wrongly concludes that his work shows that a risk-averse individual wants to annuitize its entire lifetime wealth. The appropriate conclusion is that the person with state-dependent preferences wants to fully annuitize his wealth, i.e., insure, although not being risk-averse.

The confusion of diminishing marginal utility with risk aversion is not an exception but the rule in the literature.⁷ The work of Karni (1983) and Karni (1985) is extremely valuable in showing the cost of this confusion: it shows that equating diminishing marginal utility with risk aversion requires to redefine what “risk” means.⁸ Importantly, it can no longer mean an absence of certainty. For regardless of whether one wants to follow Karni’s proposed redefinition or not, some redefinition is necessary given that the optimal distribution of wealth is typically unequal under state-dependent preferences. But this necessity allows another conclusion: if we insist on understanding “risk” as the absence of certainty and on understanding “risk aversion” as a general preference of certainty over its absence, then state-dependent preferences rule out risk aversion. In an effort to reconcile the idea of state-dependent preferences with the idea of risk aversion, Karni illustrates why the former rules out the latter.

In sum, the literature on state-dependent preferences shows that we can both retain the plausible assumption of diminishing marginal utility and retain the idea of insurance being valuable without having to assume risk aversion. The confluence of diminishing marginal utility and state dependence creates a strong incentive to reallocate wealth across states, and, hence, to insure. Neither a potential

⁷See e.g. Arrow (1974). See also Epstein and Tanny (1980) and the subsequent literature on correlation attitudes. The only exception to the confusion of diminishing marginal utility with risk aversion, that I am aware of, is Hirshleifer (1965). In a framework that allows for state-dependent utility, he recognizes that “risk aversion [...] is only a special case” (p. 534).

⁸Karni defines the “reference set” as the optimal distribution of a given wealth level across states of nature. This reference set typically incorporates an unequal distribution of wealth across states. Karni proceeds to propose an alternative definition of “risk” as stochastic deviations from this reference set, as this redefinition would allow a partial ranking of people according to their “risk aversion”. However, even that benefit of the redefinition is questionable, as rankings are only possible between people with identical reference sets. Since it is impossible to find any two people with identical reference sets in reality, the theoretical possibility of partial rankings seems irrelevant in practice.

loss nor risk aversion are necessary for the existence of an insurance motive.

2.3 Indivisibilities in Consumption

Abstracting from state-dependence for the moment, the second condition for risk aversion to result from diminishing marginal utility of wealth is that marginal utility of wealth is diminishing at *all* wealth levels. Starting with Friedman and Savage (1948), who suggested that local convexities in the utility function might help to explain why we observe gambling, several scholars have investigated potential sources of such convexities. Ng (1965) proposes that these result from the fact that some consumption opportunities are simply not perfectly divisible. If some consumption opportunities are indivisible, or imperfectly divisible, then the marginal utility of wealth exhibits jump discontinuities at wealth levels at which it becomes optimal to buy an indivisible (increment of a) consumption good. These jump discontinuities are again the result of an optimization behavior of the rational consumer.⁹ With local convexities in the utility function over wealth, some lotteries over wealth are strictly preferred to their expected value. Hence, the individual is not risk-averse as some gambles are actually desirable. The literature on indivisibilities in consumption (Ng 1965; Jones 2008; Vasquez 2017) correspondingly underlines that it can be rational for an individual to engage in both insurance and gambling.

By producing a rationale for gambling, indivisibilities in consumption seem to reduce the desirability of insurance at first glance. However, as Fels (2020a) shows, if some indivisible consumption opportunities are state-dependent, a new rationale for insurance arises as it is preferable to finance state-dependent indivisible consumption opportunities across instead of within state. Financing state-dependent indivisibilities across states is desirable for two reasons. First, if the marginal util-

⁹The argument that indivisibilities in consumption lead to local convexities does not require diminishing marginal utility of divisible consumption, however. If marginal utility of divisible consumption is constant, it is the utility function over wealth itself and not only the marginal utility function that exhibits jump discontinuities (Vasquez 2017; Fels 2020b; Fels 2020a).

ity of divisible consumption is diminishing, financing an indivisible consumption opportunity across states instead of within state reduces the opportunity cost of indivisible consumption in terms of divisible consumption. Second, if the cost of an indivisible consumption opportunity (like a medical treatment) is large, then an individual may not be able to purchase it even if he finds its value worth paying the cost. Insurance allows to overcome this affordability barrier by transferring wealth into the state in which the costly consumption opportunity arises. In this way, insurance is valuable in providing an insuree with access to an otherwise unaffordable consumption opportunity. This second advantage of financing indivisibilities across states has first been pointed out by Nyman (1999b) when suggesting an access value in health insurance. Nyman's access value is thus a special case of how the existence of a state-dependent indivisible consumption opportunity gives rise to a value in insuring.

To show the access value of insurance in a simple framework, assume that utility depends linearly on perfectly divisible non-medical consumption, the price of which we normalize to 1. State 1 signifies sickness while state 2 signifies good health. A treatment for the sickness is available at cost $c > 0$ and confers a value $v > c$ in state 1 and a value 0 in state 2. The treatment only confers its value if fully consumed (hence the indivisibility). Since $v > c$, it is optimal to seek treatment in state 1 as soon as it is affordable: $w_1 \geq c$. Hence, the state-dependent indirect utility functions over money are given by:

$$u_1(w_1) = w_1 + \mathbb{I}_{(w_1 \geq c)}(v - c),$$

$$u_2(w_2) = w_2.$$

Insuring state 1, i.e. transferring money from state 2 into state 1, is beneficial if $w_1 < c$ and $w_2 \geq \frac{p}{1-p}(c - w_1)$. In that case, the individual is unable to afford treatment without insurance, but

is able to transfer enough wealth from state 2 to state 1 to gain access to treatment. Again, this is perfectly possible with $w_1 = w_2 = \bar{w}$, i.e., a situation with a-priori certainty in wealth. Figure 3 illustrates the access value of insurance with two equiprobable states.

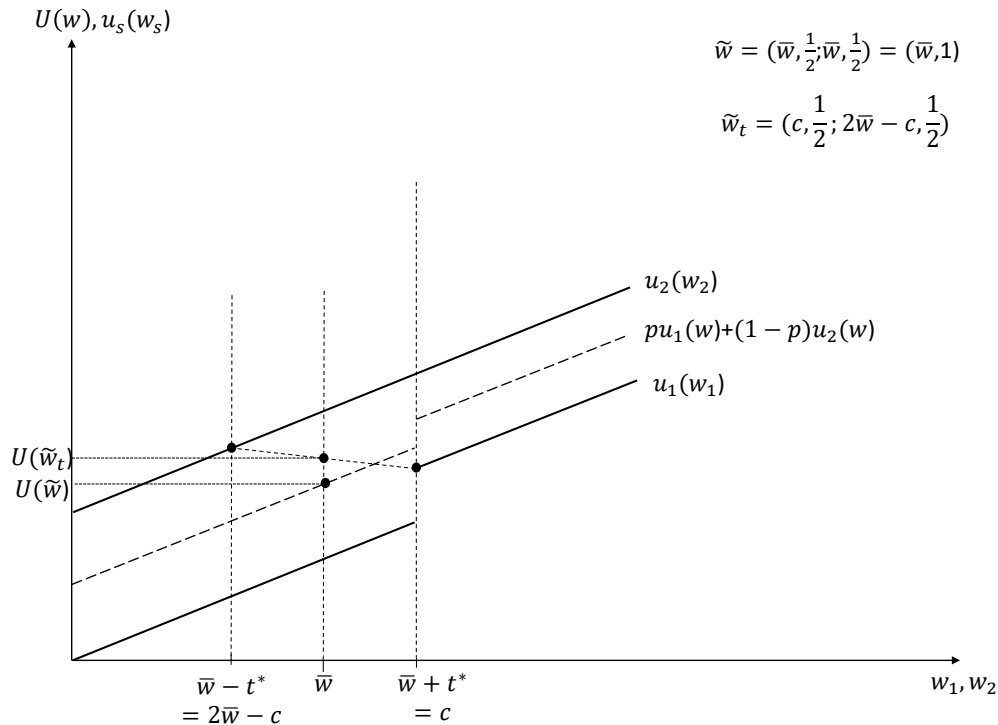


Figure 3: The Access Value of Insurance

Starting from a position of equal wealth in both states, $\tilde{w} = (\bar{w}, \frac{1}{2}; \bar{w}, \frac{1}{2})$, insurance allows to transfer enough wealth from state 2 to state 1 such that the cost of treatment c becomes affordable. Nyman (1999b) calls the utility gain $U(\tilde{w}_t) - U(\tilde{w})$ the “access value” of insurance.¹⁰ As our simple example above indicates, $U(\tilde{w}_t) - U(\tilde{w}) = p(v - c)$, and hence, the value of insurance is directly tied to the net value of the medical treatment. As first noted in Nyman (1999a) and further discussed later, this has strong implications for the evaluation of moral hazard in insurance.

Note that, with a desire to finance state-dependent indivisible consumption opportunities across states, it is again not risk aversion that leads to a motive for insurance. In fact, the person is not

¹⁰In Figure 3, a utility loss of $\epsilon = u_2(0) - u_1(0)$ is assumed. Note that the access value of insurance, $U(\tilde{w}_t) - U(\tilde{w})$, is independent of the existence, sign, and size of ϵ . It is only assumed to unclutter the figure.

risk-averse as it prefers an unequal distribution of wealth - in which more wealth is allocated to the state in which the indivisible consumption opportunity occurs - to an equal distribution of wealth. As in the case with state-dependence but perfectly divisible consumption, the optimal allocation of wealth across states being unequal rules out risk aversion. Also, as in the previous case, the very reason that makes a person not risk-averse is not reducing but increasing the desirability of insurance. But if it is not a desire for risk reduction alone that drives insurance demand, we require a novel definition of insurance.

2.4 A Definition of Insurance

Both the ideas of state-dependence of preferences and of indivisibilities in consumption rule out risk aversion as a general preference for certainty. Simultaneously, both ideas suggest a rationale for insurance purchase. However, these rationales for insurance require us to give up the idea that insurance is solely a means to rid ourselves of undesired risks. Instead, they show that insurance can also be desirable for acquiring certain risks. The following definition, first proposed in Fels (2019), seeks to capture this idea.

Definition 1. *Insurance is the directed transfer of wealth across states in order to meet conditional (financial) needs.*¹¹

In more technical terms, insurance is an alignment of the risk in one's financial assets to the risk in one's financial needs. This reformulation shows why the novel definition is broader than the traditional definition of insurance as a means to reduce one's risk exposure. It comprises the old

¹¹To my knowledge, Braess (1960, p. 14) is the first to argue that the purpose of insurance is to address a conditional need. Nyman (2003, p. 30), closest to my definition, proposes that insurance is a directed transfer of wealth across states. In contrast to my understanding of insurance, both require a loss in the state of the world that insurance targets either as the source of the conditional need (Braess 1960, pp. 11-14) or to distinguish insurance from gambling Nyman (2003, p. 133). I seek to underline that it is not the presence of risk in financial wealth per se, but the misalignment between the risk in financial wealth and the risk in financial needs that is the basic requirement for insurance to be desirable.

idea as it does not rule out that the risk of a loss in assets can be a rationale for insurance purchase, with fire, flood, or car insurance being prominent examples. In these cases, there is no variation of financial needs across states, but a loss through fire or flood induces a variation in financial assets across states. Insurance realigns the two by compensating the financial loss, thereby aligning the variation in financial resources with the (in this case, non-) variation in financial needs. In other situations, however, there is a variation in financial needs across states. In these cases, insurance redistributes wealth into states with larger financial needs, thereby aligning the variation in financial resources to the variation in financial needs. In health insurance, for example, the primary purpose of the wealth transfer is not the compensation of a loss, but the financing of (medical) expenses satisfying the needs that are exclusive to the state (of sickness). More strikingly, the insurance of longevity cannot be directed at compensating a loss, as the insured state is associated with a gain. It is only if we stay alive long enough, that we need financial resources to finance our consumption, a need that is obviously absent in the state in which we die early. In contrast to the old definition, the new definition does not restrict differences in wealth to be the sole source of a desirable wealth transfer. Exclusive consumption opportunities, or expenses that only have value in certain states, be they divisible or indivisible, are allowed as another source. The definition separates insurance from gambling in that the identity of the state in which the benefit is paid matters for insurance to be desirable. Given that some financial needs are conditional, thus dependent on particular events to occur, insurance can only help to meet these needs if the wealth transfer is directed towards the states in which the needs occur. In contrast, the desirability of gambling is not related to the exact identity of the winning state. Intuitively, it seems equally desirable to win the jackpot with one's "lucky" numbers as with one's "unlucky" numbers.

3 Two Applications and two generalizations

In this section, I show how the novel definition of insurance may guide us towards simple generalizations of results that have been derived in the context of state-independent preferences. First, I show how the result of the optimality of a deductible in case of a loaded insurance premium (Arrow 1963; Mossin 1968; Arrow 1971) generalizes to a setting of state-dependent preferences. Second, I show how the result of incomplete coverage under ex-ante moral hazard (Pauly 1974; Shavell 1979) generalizes to a setting of state-dependent preferences. Both questions have already been investigated in the literature.¹² However, lacking the framework of the general definition of insurance that I suggest in this paper, the previous literature is loaded with case distinctions and derivations for special cases. The purpose of this section is to carve out how the novel definition of insurance yields simple generalizations of the classic results, from which the classic results directly and intuitively follow.

3.1 Loading and Optimal Insurance

Consider a model with two states $s = 1, 2$. A decision-maker can buy insurance that pays an indemnity q in state 1 that occurs with probability $p \in (0, 1)$. Following a common notation in the literature, the utility function over wealth w is given by $V(w)$ in state 1, and by $U(w)$ in state 2. I assume that $V'(w_1) \geq U'(w_2)$ where w_s denotes the initial wealth in state $s = 1, 2$.¹³ Insurance is available at the price of a loaded insurance premium $\pi = (1 + l)pq$ that needs to be paid in both

¹²See e.g. Rey (2003) and Huang and Tzeng (2006) on the optimality of a deductible and Dionne (1982) on ex-post moral hazard under state-dependent preferences.

¹³Note that this assumption is actually without loss of generality. I simply choose to label state 1 and state 2 in such a way that it is met. The prior analysis showed that it is desirable to use insurance to transfer money into the state with larger marginal utility. Hence, if $V' < U'$ we will find that the optimal level of q will be negative. In that case, state 2 would be the insured state, and we would need to reinterpret $pi - q$ as the premium payment, and $-q_i$ as the benefit payment in state 2. In order to avoid this, I maintain $V'(w_1) \geq U'(w_2)$.

states. The DM chooses q to maximize the expected utility

$$Z(q) = pV(w_1 - \pi + q) + (1 - p)U(w_2 - \pi)$$

The FOC, given an inner solution, yields

$$p(1 - p) [V' - U'] - lp [(1 - p)U' + pV'] = 0 \tag{2}$$

We see directly that loading leads to incomplete coverage in the sense that it reduces the incentive to transfer money from state 2 into state 1. With $l = 0$, the DM chooses a level of coverage q^* that equalizes marginal utility across states. In contrast, if l is too large and/or the difference $V' - U'$ is too small, we have a corner solution of $\tilde{q} = 0$.¹⁴

Proposition 1. *With a loaded premium, the optimal level of coverage $0 \leq \tilde{q} < q^*$.*

The classic result by Arrow (1963, Arrow (1971) and Mossin (1968) can thus be generalized that a proportional loading reduces the incentive to transfer wealth into the state with larger marginal utility. Importantly, there will be no wealth transfer between states of nature with sufficiently similar marginal utility. On the one hand, this directly implies the optimality of a deductible in the special case of state-independent marginal utility, $V'(w) = U'(w)$, of which state-independence of the entire utility function is itself a special case. For, in case of state-*independent* marginal utility, small differences in marginal utility can only result from small wealth losses.

On the other hand, this underlines the limited importance of monetary or utility losses in case of state-*dependent* marginal utility, for state 1 could involve a heavy utility loss ($V \ll U$) or monetary

¹⁴ $\tilde{q} < 0$ is never an optimal solution. Effectively, this would mean that the DM wants to transfer money into state 2. Consider such a setup by simply relabeling state 1 and state 2 such that $V' \leq U'$ holds. Quick inspection of the first derivative of $Z(q)$ in this case reveals that it is strictly negative at $q = 0$ for all $l \geq 0$ showing that a premium loading, however large, never makes the DM want to insure the state with lower marginal utility.

loss ($w_1 \ll w_2$) in comparison to state 1, yet still remain uninsured if the difference in marginal utility is small. Similarly, state 1 could involve a small wealth loss, no loss at all, or even a monetary gain $w_1 > w_2$, yet, the DM still seeks to transfer a large amount of money into state 1 if $V' \gg U'$. In short, the states that remain uninsured are characterized by small gains in marginal utility, and not necessarily by small losses compared to the uninsured state.¹⁵

This finding offers a novel rationale for value-based insurance design that does not require any information asymmetry. State 2 is associated with a larger marginal utility compared to state 1 if expenditures (such as health expenditures) have larger marginal utility in the former state. Proposition 1 then implies that health insurance with a loaded premium should only cover states in which health expenses have a large benefit-to-cost ratio as only these lead to a large increase in marginal utility as compared to the healthy state. In contrast, health expenses with a lower benefit-to-cost ratio only slightly increase the marginal utility of a state, thus should not be covered according to the proposition.

More importantly, by adopting the novel definition, we can see how the special-case result, that it is not optimal to insure small losses in wealth with a loaded premium, generalizes to a framework that does not restrict preferences to be state-independent. In general, the result means that it is not optimal to insure states that are characterized by only small increases in marginal utility. The failure of the existing literature to realize this simple generalization stems from an insistence on a framework that is suitable in a special case (insurance is about covering losses) to a case where this framework fails (insurance is about covering differences in needs that may or may not result from losses). The point, that I seek to convey here, is not that the results of the previous literature

¹⁵Huang and Tzeng (2006) claim that the optimal insurance policy involves a deductible if $U' > V'$. This is a misinterpretation of their finding. By restricting q to be non-negative, they fail to realize that, in this setting, the optimal coverage q^* is strictly negative. With their restriction on q , they misinterpret their solution of $q = 0$ as a deductible. For a similar misinterpretation, see Rey (2003).

are false, but that they seem to be derived as if state-dependence is an extension of the classic model in the sense that an additional assumption is added to this classic model. As a result, it is investigated whether and under what circumstances classic results, such as the optimality of a deductible, still hold under that extension. Instead, it is the other way around: the classic case of state-independence requires additional assumptions as compared to the more general case. To generalize the findings from the classic setup, we thus need to ask what, in general, characterizes the states that remain uncovered although there is an underlying motive to seek coverage, instead of asking whether small losses still remain uncovered.

3.2 Ex-ante Moral Hazard

Consider a model with two states $s = 1, 2$. A decision-maker can buy insurance that pays an indemnity q in state 1 at the expense of an actuarially fair premium payment π in both states.¹⁶ The probability $p(x)$ of state 1 is a function of effort x of the decision-maker. The DM can influence the probability p in both directions incurring cost $C(x)$. I make the following assumptions. Both $p(x)$ and $C(x)$ are twice continuously differentiable. A positive value of x can be interpreted as effort to prevent state 1, a negative value of x can be interpreted as effort to promote state 1, where both efforts show diminishing returns: $p_x < 0 \forall x$ and $p_{xx} > 0, \forall x > 0$; $p_{xx}(0) = 0$; $p_{xx} < 0, \forall x < 0$. To capture that any effort is costly, and its marginal cost is increasing, I assume that $C(0) = 0$; $C_x > 0, \forall x > 0$; $C_x < 0, \forall x < 0$; $C_{xx} > 0$.

The DM maximizes the expected utility

$$Z(x, q) = p(x)V(w_1 - \pi + q) + (1 - p(x))U(w_2 - \pi) - C(x)$$

¹⁶Again, I label states such that $V'(w_1) \geq U'(w_2)$ holds.

where $V(\cdot)$ is the utility function in state 1 and $U(\cdot)$ is the utility function in state 2. w_s is the wealth of the individual in state $s = 1, 2$.

The DM chooses x after an insurance contract is signed. We consider two scenarios. Under symmetric information, x is observable. Under asymmetric information, x is not observable by the insurer.

Scenario 1: Symmetric information

The insurer charges an actuarially fair premium π that depends on both effort x and coverage q :

$$\pi(x, q) = p(x)q.$$

In that case, q is given at the moment of choice of x , yet π may vary with x . In a competitive market environment, we will have $\pi(x, q) = p(x)q$. Then, the DM chooses x in order to maximize

$$Z(x, q) = p(x)V(w_1 - \pi(x, q) + q) + (1 - p(x))U(w_2 - \pi(x, q)) - C(x)$$

The FOC for a maximum is given by

$$p_x(V - U) - \pi_x [(1 - p)U' + pV'] = C_x \tag{3}$$

$$\Leftrightarrow -p_x [U - V + q(pV' + (1 - p)U')] = C_x \tag{4}$$

Given that $p_x < 0$, this condition is satisfied for $U > V$ only if $x^* > 0$. If the DM has a preference for state 2 (despite the wealth transfer from state 2 to state 1), effort is exerted to prevent state 1.¹⁷

Note, however, that even if $V = U$ at q^* , i.e. without any state preference after the wealth transfer, effort is exerted to prevent state 1. The intuition is as follows. As we will derive below, $q^* > 0$, i.e., it is optimal to transfer money from state 2 into state 1. If $C'(0) = 0$, then it is optimal to

¹⁷Consider health insurance or long-term care insurance as possible examples.

exert some effort to prevent the insured state, as it reduces the opportunity cost of this transfer.¹⁸ Note that this implies that no effort, $x^* = 0$, is optimal only if the DM has a state preference for the insured state 1, $V > U$, after insurance. Finally, if $V \gg U$ after the wealth transfer, in the sense that the DM has a state preference for state 1 that dominates the incentive to reduce the opportunity cost of insurance, then $x^* < 0$. The DM has an incentive to promote the insured state 1. This does not necessarily imply deceptive activities, such as arson, but may simply reflect a natural state preference.¹⁹ Anticipating its own effort choice $x^*(q)$, the DM chooses the level of coverage q in order to maximize

$$Z(x^*(q), q) = p(x^*(q))V(w_1 - \pi(x^*(q), q) + q) + (1 - p(x^*(q)))U(w_2 - \pi(x^*(q), q)) - C(x^*(q))$$

The FOC yields:

$$\left\{ \underbrace{p_x(V - U) - \pi_x [(1 - p)U' + pV']}_{=0} - C_x \right\} \frac{\partial x^*}{\partial q} - p(1 - p)(U' - V') = 0 \quad (5)$$

$$\Leftrightarrow p(1 - p)(V' - U') = 0 \quad (6)$$

which means that the DM chooses the level of coverage $q^* \geq 0$ that equalizes marginal utility across states.²⁰

Scenario 2: Asymmetric information

Under asymmetric information, the premium π and coverage q are both given at the moment of effort choice.

¹⁸With actuarially fair insurance, an additional unit of wealth in state 1 costs $p/(1 - p)$ units of wealth in state 2.

¹⁹Consider longevity insurance as an obvious examples. Annuities pay in case of survival of the recipient. The insured person's interest in promoting survival is hardly deceptive, and probably not significantly altered by purchasing insurance.

²⁰Here, it becomes obvious that, if $V' < U'$, then it must hold that $q^* < 0$, and, hence, the DM wants to insure state 2. Anticipating this, we chose to label state 1 and state 2 in such a way that $V' \geq U'$.

Hence, the DM chooses effort \hat{x} to maximize

$$Z(x) = p(x)V(w_2 - \pi + q) + (1 - p(x))U(w_1 - \pi) - C(x)$$

The FOC for a maximum is given by

$$-p_x(U - V) = C_x \tag{7}$$

Given that, $p_x < 0$, $\hat{x} > 0$ if and only if $U > V$. That is, the DM only exerts preventive effort if there remains a state preference for state 2 after the wealth transfer through insurance. The DM seeks to promote the insured state whenever there is a state preference for state 1. The reason for the lower incentive for preventive effort (higher incentive for promotive effort) is the fact that under asymmetric information the reduced opportunity cost of the wealth transfer that result from increased preventive effort cannot be passed on to the insuree as effort is not observable.

The insights regarding the impact of asymmetric effort on preventive effort can thus be simply generalized as follows.

Proposition 2. *If actions that influence the occurrence of the insured state are not observable/contractible, then this reduces incentives to prevent and increases incentives to promote the insured state.*

We see, again, that the major difference that state dependence makes is the *identity* of the insured state. It means that it is always the state with larger marginal utility that is insured. It is not necessarily the one with lower wealth w_s or lower utility U, V . Hence, it is not the occurrence of a loss in wealth and/or utility that define the insured state, but the occurrence of larger marginal utility.

While utility losses do not determine the identity of the state that the DM wants to insure, they

determine the state that the DM wants to promote or prevent. If there is an a-priori state preference for state 2 (1), this produces an incentive to prevent (promote) the insured state. Buying insurance always makes the insured state comparatively more favorable, i.e., it either mitigates an a-priori preference for the uninsured state or strengthens an a-priori preference for the insured state. This translates into lower effort preventing/ higher effort promoting state 1, respectively. This effect of insurance is stronger under asymmetric information.

We can determine the influence of insurance on effort by defining $g(x, q) = -p_x(U - V) - C_x$ and using the implicit function theorem:

$$\frac{\partial \hat{x}}{\partial q} = -\frac{\partial g / \partial q}{\partial g / \partial x} \quad (8)$$

Given that $\frac{\partial g}{\partial x} > 0$ given our assumptions on $p(x)$ and $C(x)$, and $p_x < 0$, the sign of $\frac{\partial \hat{x}}{\partial q}$ is the opposite of the sign of $\frac{\partial g}{\partial q}$.

$$\frac{\partial g}{\partial q} = -p_x [U'(-\pi_q) - V'(1 - \pi_q)] \quad (9)$$

$$= -p_x [U'\pi_q + V'(1 - \pi_q)] \quad (10)$$

At (\hat{x}, \hat{q}) , it must hold that $0 < \pi_q < 1$. Hence, we can conclude that $\frac{\partial g}{\partial q} > 0$, and, thus $\frac{\partial \hat{x}}{\partial q} < 0$. Under asymmetric information, larger coverage unambiguously reduces the effort to prevent/increases the effort to promote the insured state.

When selecting insurance coverage \hat{q} under asymmetric information, both the actuarial premiums $\pi(q)$, that the insurer charges, and the anticipated probability $p(x)$ need to reflect the effort choice $\hat{x}(q)$ that the DM anticipates to make given the coverage choice. Hence, $\pi(q) = p(\hat{x}(q))q$ and

$p(x) = p(\hat{x}(q))$. Given that, the DM chooses \hat{q} to maximize

$$Z(q) = p(\hat{x}(q))V(w_1 - \pi(q) + q) + (1 - p(\hat{x}(q)))U(w_2 - \pi(q)) - C(\hat{x}(q)).$$

The FOC yields

$$\left\{ \underbrace{p_x(V - U) - C_x}_{=0} \right\} \frac{\partial \hat{x}}{\partial q} - p_x q [(1 - p)U' + pV'] \frac{\partial \hat{x}}{\partial q} - p(1 - p)(U' - V') = 0 \quad (11)$$

$$\Leftrightarrow p(1 - p)(V' - U') - p_x q [(1 - p)U' + pV'] \frac{\partial \hat{x}}{\partial q} = 0 \quad (12)$$

Given that $p_x < 0$, the influence of asymmetric information on the level of coverage is determined by the sign of $\frac{\partial \hat{x}}{\partial q}$. As we have determined the sign to be negative, we can conclude that asymmetric information unambiguously reduces insurance coverage. However, it never reduces optimal coverage to zero, since $\partial Z / \partial q > 0$ at $q = 0$.

Proposition 3. *If actions that influence the occurrence of the insured state are not observable/contractible, then this reduces insurance coverage: $0 < \hat{q} < q^*$.*²¹

Asymmetric information on actions that influence the probability of the insured state reduce the level of coverage that an individual seeks.

Again, this generalization is intuitive and extremely simple compared to the extant literature. The complication and the necessity to make several case distinctions in the existing literature, again, only arise because the authors apply the classic framework to a setting in which it no longer applies. Insisting that insurance is about covering losses, the different levels of coverage, \hat{q} and q^* , are compared to different loss measures. This can be a classic monetary loss $L = w_2 - w_1$,

²¹While looking similar at first glance, this result is in strong contrast to Dionne (1982) who claims that $0 < \hat{q} < q^*$ irrespective of which state has larger marginal utility. A quick inspection of equations 6 and 12 reveals that, without our assumption of $V'(w_1) \geq U'(w_2)$, we would have $q^* < \hat{q} < 0$. In this case, state 2 is the insured state and $-q$ constitutes the size of the insurance benefit in state 2.

or a monetary loss-equivalent like the “ransom” R that leaves the DM indifferent between the two states $U(w_1 - R) = V(w_2)$ (Cook and Graham 1977). Schlesinger (1984) e.g. proposes to talk of full insurance if $q = R$ and to talk of complete insurance if $q = L$. The results from above are, however, completely independent of L and R . Indeed, the model places no restrictions on L and R : they can be positive, negative, zero, of opposite or similar sign; L can be larger, identical, or smaller than R . Depending on what assumptions are made on L and R , we can thus analyze an abundance of different cases. In some cases, e.g. the case $R \geq L > 0$ that primarily motivates the literature on irreplaceable commodities (Cook and Graham 1977; Schlesinger 1984), this seems to make intuitive sense. The limits of this approach become obvious in other cases, such as longevity insurance, where $R < L \leq 0$ seems more plausible. In general, how should we call the cases with $q/L < 0$ and or $q/R < 0$? It is hardly incomplete coverage of a loss, if the DM seeks to insure the state in which a monetary/utility *gain* occurs.

The realization that insurance is intended to transfer money into states with larger (marginal) utility makes these comparisons obsolete. It suggests another, quite simple, point of comparison: to compare the transfer, that an individual chooses, to the transfer that fully eliminates this underlying motive. In other words, coverage q needs to be compared to the level of coverage q^* that equalizes marginal utility across states. With this point of reference, we can conclude that ex-ante moral hazard always results in incomplete coverage.

4 Why a novel definition matters

Apart from the possibility to offer a simple way to generalize classic results for state-independent utility to broader settings, the novel definition of insurance puts into perspective and offers some guidance with regard to several strands of existing and ongoing research. The novel definition

differs from the previous understanding of insurance as a risk-reducing mechanism in two important regards. First, it shifts the focus of optimal insurance from the mitigation of a potential loss to the meeting of a potential need. Second, it allows risk preferences to have far more nuance than the traditional trichotomy of risk aversion, risk neutrality, and risk love. Both have important consequences for the questions that our discipline should ask (research agenda) and for the advice that we deduce from our research (policy implications). I want to discuss some of these consequences here.

4.1 What questions should we ask?

First, given that the mitigation of a loss is no longer the sole (or sometimes even an) underlying rationale for insurance, the standard question to what extent insurance should cover a given loss can be misdirected. It is justified in a context in which the difference between states is solely a difference in wealth (fire, flood). Here, the misalignment between (the risk in) financial needs and (the risk in) financial means stems from the former being certain while the latter are uncertain. The focus on losses is misleading, however, in a context in which the main difference between states is due to needs that are specific to a state of nature (health, longevity). The optimal level of coverage must then be derived in relation to the state-specific need, not in relation to some correlated loss as is commonly done in the literature on the insurance of irreplaceable commodities (Cook and Graham 1977; Shioshansi 1982; Schlesinger 1984; Huang and Tzeng 2006). In the previous section, I illustrate why deriving the optimal coverage of a loss is misleading in the context of state-dependent preferences. It first requires translating a utility loss into a wealth loss, the ransom, only to find that optimal insurance coverage only depends on the difference in marginal utility across states (Arrow 1974; Schlesinger 1984), an information for which the calculation of the ransom is often not helpful. As illustrated in the example on longevity insurance, it is not particularly informative to calculate the

optimal level of coverage in relation to some loss equivalent (or correlated monetary loss) if a given difference in marginal utility can be associated with utility (and/or monetary) losses of different sizes, or, more strikingly, with a utility/monetary *gain*.²² The focus on the mitigation of losses, that our traditional theory of insurance encourages, obfuscates the more important question in the setting of state-dependent preferences: how much coverage is needed to address the conditional needs that are implicit in the difference in marginal utility across states. Instead of asking “How much of the loss should be covered?”, the question “To what extent can the conditional need be met?” seems more appropriate in settings in which state-dependent preferences are a natural assumption.

Following the traditional view, the optimal design of insurance is typically analyzed as a trade-off between incentive provision and risk transfer (Pauly 1968; Shavell 1979). If the purpose of insurance is not to transfer risk, but to meet a conditional need, then the optimal level of incentives needs to recognize how these incentives can undermine the capability of insurance to address a conditional need. Cost-sharing requirements may have little effect on the consumption-smoothing value of insurance, but can have a decisive impact on an insurance’s capability to address a conditional need if these requirements impose access barriers (Fels 2020c). A novel question then arises: “How much unmet needs are acceptable in exchange for incentives?”. Answering this question may not only give rise to a different answer on the optimal extent, but also on the optimal design of incentives in insurance contracts.

The traditional understanding of insurance as risk transfer and gambling as risk acquisition frame these behaviors as natural opposites leading to repeated attempts to reconcile the two behaviors given evidence of their simultaneous prevalence (Friedman and Savage 1948; Ng 1965; Conlisk 1993; Hartley and Farrell 2002; Chetty and Szeidl 2007; Jones 2008; Vasquez 2017). Acknowledging that both behaviors can be different types of risk acquisition implies that there is no need for

²²This shortcoming of the “cash-equivalent model” seems to have been anticipated by Arrow (1974, pp. 4-5).

reconciliation in the first place. Instead, the relation between the two behaviors is far more complex even allowing for a complementary role (Fels 2020a). This gives rise to the question “How do these behaviors interact?”. Understanding under what circumstances the two behaviors reinforce or substitute each other may allow deriving testable predictions on their joint occurrence.²³

Once we give up the simplification of a world in which all individuals either dislike all risk (risk-averse), love all risk (risk-seeking), or never care (risk-neutral), the question of the stability of risk preferences appears odd. If risk preferences derive from state-dependent needs, then there is no reason to expect them to be stable across context or time (Barseghyan, Prince, and Teitelbaum 2011; Schildberg-Hörisch 2018). Needs, both conditional and unconditional ones, change over time. Similarly, risk preferences derived in one insurance context, say health insurance, should be different from risk preferences in another insurance context, say home insurance. This is no sign of instability of preferences, but a mere consequence of these insurances addressing different conditional needs. Instead of worrying about an instability of preferences, we could ask “What does the different willingness to reallocate resources across states tell us about the difference in conditional needs?”. Answering this question might help us informing optimal insurance design and policy. In addition, it might help us to understand differences in insurance take-up across markets.

Finally, the foundation of insurance demand on conditional needs implies that we should exert great caution when linking risk preferences elicited in the laboratory to questions of optimal insurance because the former abstract from the conditional needs that need to inform the latter. Laboratory evidence on decision-making under risk is crucial to understand the choice procedures that individuals apply, which are highly relevant in an insurance context. Yet, laboratory settings abstract from the conditional needs relevant in each insurance context. This means that the risk

²³In a recent paper, Amentier et al. (2018) find that individuals with higher wealth have both more insurance and hold more risky assets. Even after controlling for wealth, the positive correlation remains robust.

preferences elicited in a laboratory context differ from the risk preferences that govern insurance choices.²⁴ Instead of trying to recover risk preferences from laboratory data that may then end up having little relevance in actual insurance decisions, we could focus our attention on the question “What can laboratory evidence tell us about decision-making under risk?”. It is exactly this question that laboratory evidence seems most valuable to answer, and the growing literature on non-expected utility models that is driven by this evidence is a testament to this value.

4.2 What advice can we give?

Beyond suggesting new perspectives for our research agenda, the different understanding of insurance and its purpose has major implications for the policy implications that we can derive from our theories.

One of the most influential predictions of insurance theory is that full insurance will be accompanied by moral hazard (Arrow 1963), an increased consumption of the insured service, leading to a decrease in welfare (Pauly 1968). The evaluation of this behavioral response of the insured as welfare-decreasing has been criticized by De Meza (1983) and Nyman (1999a) in the context of health insurance. These works show that an increase in consumption can be the underlying purpose of insurance instead of an undesired side effect. This interpretation of the behavioral response becomes more obvious if one understands the purpose of insurance (in this case health insurance) of addressing a conditional need (in this case medical need). If insurance is intended to reallocate resources into a state that is characterized by a specific need, an increase in the consumption of those services that address the need is a desired consequence of insurance. The increase of consumption by the insured is then a sign of insurance being successful in achieving its original purpose,

²⁴Indeed, several studies have found risk preferences elicited in the laboratory to have little (if any) predictive power with respect to actual insurance take-up (Delavande et al. 2018; Jaspersen et al. 2019; Charness et al. 2020).

and not a sign of misaligned incentives. Policy advice that seeks to reduce this behavioral response could actually undermine the value of insurance instead of strengthening it (compare Nyman 2003, pp. 145-149). A classic article by Feldstein (1973) illustrates this mistake as it argues for stronger cost-sharing requirements in health insurance to achieve a better compromise between risk transfer and incentives. Recognizing that the utilization response is, at least in part, a benefit and not a cost of insurance alters the evaluation of how much insurance individuals should seek. The different understanding of the purpose of insurance can also fundamentally change our evaluation of existing insurance systems. In a prominent example, the public insurance program of Medicaid is criticized for offering only a very limited insurance value (Brown and Finkelstein 2008; Brown and Finkelstein 2011). This criticism is perfectly valid if insurance is seen solely as a mechanism for risk transfer. However, if insurance is intended to address a conditional need, in this case the need of basic services of long-term care, then Medicaid is achieving its intended goal: ensuring that everyone, regardless of financial status, can meet the conditional need of long-term care at least on a basic level (Fels 2020d). The evaluation of the public insurance scheme thus fundamentally changes depending on what we deem the purpose of insurance.

Even in cases in which adopting the new definition does not change our policy advice, maintaining the traditional understanding of insurance as a risk transfer can undermine the effectiveness of our advice. The “annuitization puzzle” refers to the observed reluctance to annuitize a major part of one’s retirement wealth despite a large insurance value in doing so. Several explanations have been put forward. According to one of them, behavioral biases lead consumers to misperceive annuities as risky investment (Brown 2007; Hu and Scott 2007; Benartzi et al. 2011). This suggestion of mistaken consumer perceptions is straightforwardly false, for annuities are risky investments (as any insurance). The claim of a misperception follows from the traditional characterization of insurance as a risk transfer as the risk embodied in insurance is supposed to hedge against an ex-ante risk in

endowment. It is exactly the context of annuity markets where this traditional framework breaks down as annuities pay off in a state that is not characterized by a loss in wealth resulting in insurance increasing the variation of wealth across states. It is also annuity markets where the danger of maintaining the traditional framework becomes obvious. Brown (2007) points out that lay person might follow the misleading heuristic that “Insurance is for bad events” in their refusal to annuitize. Ironically, Brown fails to mention that this misleading heuristic is exactly what Economics teaches.²⁵ Worse, economists are perpetuating this misleading heuristic in the context of annuities by referring to longevity risk as “the risk of being unable to sustain [one’s] consumption should [one] live longer than expected” (Brown et al. 2008) or “the risk of outliving one’s retirement wealth” (Benartzi et al. 2011). On a technical level, the problem with these definitions is that they confuse the terms of risk, in the insurance context usually reserved for a variation in (wealth) endowment across states of nature, with the possibility of needing additional financial resources, a variation in preference across states of nature.²⁶ On a heuristic level, it deliberately confuses the possibility of a longer life with a negative event, thereby framing annuities according to exactly the heuristic that Brown calls misleading. This confusion is necessary if one insists on the traditional framework which requires to frame annuities as the “safe alternative” against a “risk”. What these definitions of longevity risk actually describe is not a risk in wealth, but an uncertain need. By adopting the new definition of insurance, annuities can simply be described as provisions for a potential need, or, more technically, an alignment of the risk in endowment with the risk in needs. Increasing the salience of these potential needs is associated with larger annuitization (Brown et al. 2008). People thus seem to understand the value of insurance in providing for uncertain needs. Economists might help

²⁵Brown (2007, p. 24), in contrast, claims that “[a]n economist’s view of insurance is that it is a mechanism for transferring resources from states of low marginal utility of income [...] to states of high marginal utility of income.” There is no Economics textbook that I am aware of in which insurance is taught in this way without associating the difference in marginal utility with a loss, i.e., a “bad event”.

²⁶Note, however, that annuitization does nothing to reduce the risk in preferences. It simply helps to accommodate these preferences by introducing a risk in one’s financial wealth that matches the risk in preferences.

them by stopping to teach the traditional view of insurance as risk transfer that creates misleading heuristics. In addition, instead of denying the risk involved in annuitization, we could try to explain the necessity of taking a calculated risk in providing for an uncertain need.

5 Conclusion

This paper seeks to show that the literature has, over many years, produced several theoretical reasons to doubt that people are risk-averse and to question that risk aversion is the sole motive underlying insurance purchase. This does not mean that risk aversion needs to vanish from theoretical models. On the contrary, in many settings, it captures behavior and its underlying motives reasonably well to serve as a useful modeling assumption. However, there are also settings in which the deviations from risk aversion are crucial for our behavioral predictions and for our policy advice. In this paper, I argue that insurance is one of these settings. It is high time that our discipline takes its own findings seriously and moves beyond the traditional understanding of insurance as a mechanism for risk transfer. Abandoning the traditional view of insurance presents several opportunities for novel research questions and has the potential to greatly improve our policy advice.

References

- Amentier, O., J. Foncel, and N. Treich (2018). Insurance and portfolio decisions: A wealth effect puzzle. Working Paper.
- Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review* 53(5), 941–973.
- Arrow, K. (1971). *Essays in the theory of risk-bearing*. North-Holland Publishing Company.

- Arrow, K. (1974). Optimal insurance and generalized deductibles. *Scandinavian Actuarial Journal* 1, 1–42.
- Barseghyan, L., J. Prince, and J. Teitelbaum (2011). The risk preferences stable across contexts? evidence from insurance data. *American Economic Journal* 101(2), 591–631.
- Benartzi, S., A. Previtiero, and R. Thaler (2011). Annuitization puzzles. *Journal of Economic Perspectives* 25(4), 143–164.
- Braess, P. (1960). *Versicherung und Risiko*. Betriebswirtschaftlicher Verlag Dr. Th. Gabler.
- Brown, J. (2007). Annuitization puzzles. NBER Working Paper 13537.
- Brown, J. and A. Finkelstein (2008). The interaction of public and private insurance: Medicaid and the long-term care insurance market. *American Economic Review* 98(3), 1083–1102.
- Brown, J. and A. Finkelstein (2011). Insuring long-term care in the united states. *Journal of Economic Perspectives* 25(4), 119–142.
- Brown, J., J. Kling, S. Mullainathan, and M. Wrobel (2008). Why don't people insure late-life consumption? a framing explanation of the under-annuitization puzzle. *American Economic Review: Papers & Proceedings* 98(2), 304–309.
- Charness, G., T. Garcia, T. Offerman, and M. Villeval (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty* 60, 99–123.
- Chetty, R. and A. Szeidl (2007). Consumption commitments and risk preferences. *The Quarterly Journal of Economics* 122(2), 831–877.
- Conlisk, J. (1993). The utility of gambling. *Journal of Risk and Uncertainty* 6, 279–304.
- Cook, P. and D. Graham (1977). The demand for insurance and protection: The case of irreplaceable commodities. *The Quarterly Journal of Economics* 91(1), 143–156.

- De Meza, D. (1983). Health insurance and the demand for medical care. *Journal of Health Economics* 2, 47–54.
- Delavande, A., J. Ganguli, and F. Mengel (2018). Attitudes to uncertainty and household decisions. Working Paper.
- Dionne, G. (1982). Moral hazard and state-dependent utility function. *The Journal of Risk and Insurance* 49(3), 405–422.
- Epstein, L. and S. Tanny (1980). Increasing generalized correlation: A definition and some economic consequences. *Canadian Journal of Economics* 13, 16–34.
- Feldstein, M. (1973). The welfare loss of excess health insurance. *The Journal of Political Economy* 82(2), 251–280.
- Fels, M. (2019). Risk attitudes with state-dependent indivisibilities in consumption. Ruhr Economic Papers No. 805.
- Fels, M. (2020a). Complex risk attitudes with piece-wise linear utility functions. Working Paper.
- Fels, M. (2020b). Financing state-dependent indivisible expenses with insurance. Working Paper.
Available at <https://ssrn.com/abstract=3587649>.
- Fels, M. (2020c). Mental accounting, access motives, and overinsurance. *Scandinavian Journal of Economics* 122(2), 675–701.
- Fels, M. (2020d). On the value of medicaid in providing access to long-term care. *Journal of Public Economic Theory* 22(4), 933–948.
- Friedman, M. and L. Savage (1948). The utility analysis of choices involving risk. *Journal of Political Economy* 56(4), 279–304.
- Gollier, C. (2001). *The Economics of Risk and Time*. The MIT Press.

- Hartley, R. and L. Farrell (2002). Can expected utility theory explain gambling. *The American Economic Review* 92(3), 613–624.
- Hirshleifer, J. (1965). Investment decision under uncertainty: Choice-theoretic approaches. *The Quarterly Journal of Economics* 79(4), 509–536.
- Hirshleifer, J. (1966). Investment decision under uncertainty: Applications of the state-preference approach. *The Quarterly Journal of Economics* 80(2), 252–277.
- Hu, W.-Y. and J. Scott (2007). Behavioral obstacles in annuity markets. *Financial Analysts Journal* 63(6), 71–82.
- Huang, R. and L. Tzeng (2006). The design of an optimal insurance contract for irreplaceable commodities. *Geneva Risk and Insurance* 31, 11–21.
- Jaspersen, J., M. Ragin, and J. Sydnor (2019). Predicting insurance demand from risk attitudes. NBER Working Paper 26508.
- Jones, L. (2008). A note on the joint occurrence of insurance and gambling. *Macroeconomic Dynamics* 12, 97–111.
- Karni, E. (1983). Risk aversion for state-dependent utility functions: Measurement and applications. *International Economic Review* 24(2), 637–647.
- Karni, E. (1985). *Decision-Making Under Uncertainty: The Case of State-Dependent Preference*. Harvard University Press.
- Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy* 76(4), 553–568.
- Ng, Y.-K. (1965). Why do people buy lottery tickets? Choices involving risk and the indivisibility of expenditure. *Journal of Political Economy* 73(5), 530–535.

- Nyman, J. (1999a). The economics of moral hazard revisited. *Journal of Health Economics* 18, 811–824.
- Nyman, J. (1999b). The value of health insurance: the access motive. *Journal of Health Economics* 18, 141–152.
- Nyman, J. (2003). *The Theory of Demand for Health Insurance*. Stanford University Press.
- O’Donoghue, T. and J. Somerville (2018). Modeling risk aversion in economics. *Journal of Economic Perspectives* 32(2), 91–114.
- Pauly, M. (1968). The economics of moral hazard: Comment. *American Economic Review* 58(3), 531–537.
- Pauly, M. (1974). Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection. *The Quarterly Journal of Economics* 88(1), 44–62.
- Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* 68(5), 1281–1292.
- Rabin, M. and R. Thaler (2000). Anomalies: Risk aversion. *The Journal of Economic Perspectives* 15(1), 219–232.
- Rey, B. (2003). A note on optimal insurance in the presence of a nonpecuniary background risk. *Theory and Decision* 54, 73–83.
- Schildberg-Hörisch, H. (2018). Are risk preferences stable? *The Journal of Economic Perspectives* 32(2), 135–154.
- Schlesinger, H. (1984). Optimal insurance for irreplaceable commodities. *The Journal of Risk and Insurance* 51(1), 131–137.
- Shavell, S. (1979). On moral hazard and insurance. *The Quarterly Journal of Economics* 93(4), 541–562.

Shioshansi, F. (1982). Insurance for irreplaceable commodities. *The Journal of Risk and Insurance* 49(2), 309–320.

Vasquez, M. (2017). Utility of wealth with many indivisibles. *Journal of Mathematical Economics* 71, 20–27.

Yaari, M. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies* 32(2), 137–150.