Classification risk in health insurance: The interaction of genetics, prevention, and insurance*

Julia Holzapfel†

December 14, 2023

Abstract

This paper studies regulatory regimes for the use of genetic and behavioral information in health insurance pricing. Explicitly considering both genetic and behavioral information allows to account for the different nature of the underlying risk factors and to carve out how they interact with each other. In particular, I show that banning the use of behavioral information exacerbates adverse selection on genetic differences. On the other hand, banning the use of genetic information can impact the capability of the use of behavioral information to mitigate moral hazard. As a consequence, banning the use of genetic information and promoting the use of behavioral information maximizes ex ante welfare if the distribution of genetic risk is not too dispersed and prevention is equally productive for all individuals. Otherwise, regulators face an equity-efficiency trade-off and only a social planner can resolve this trade-off.

Keywords: Classification Risk · Long-Term Health Insurance · Classification Bans · Adverse Selection · Moral Hazard

JEL Classifications: D82 · D86 · G22 · G28 · G52 · I13

*I thank Michael Hoy, Tobias Huber, Johannes Jaspersen, Lu Li, Richard Peter, Andreas Richter, Casey Rothschild, Jörg Schiller, and participants of the 2021 DVfVW Annual Conference, the 2021 Huebner/ARIA Doctoral Colloquium, the 2021 ARIA Meeting, the 2021 EGRIE Seminar, the 2022 SCOR-TSE Workshop on Risk Markets and Value Creation, the 2023 DGF Annual Meeting, and seminar participants at Coburg University, Wellesley College, and UMR SMART Rennes for their helpful comments and suggestions.

†Munich Risk and Insurance Center, LMU Munich, Schackstr. 4, 80539 Munich, Germany. Email: holzapfel@lmu.de.

1 Introduction

For many diseases, such as diabetes, cancer, or cardiovascular disease, the risk of disease onset depends on the interaction of genetic and behavioral risk factors. Medical studies suggest that modifiable risk factors, such as a high body-mass index or tobacco smoking, account for more than a quarter of healthcare spending in the US (Bolnick et al., 2020). Technological progress has significantly extended insurers' possibilities to use behavioral data for risk assessment. For example, insurers can use smartphone apps and digital stamp cards to collect information about preventive examinations or gym attendance and wearable devices to track individuals' physical activity. Most of the existing risk classification literature, however, focuses on the use of immutable characteristics, such as age, ethnicity, or gender (e.g., Hoy, 1982; Crocker and Snow, 1986) or genetic information (e.g., Tabarrok, 1994; Doherty and Thistle, 1996), in insurance pricing.

In this paper, I discuss the use of both genetic and behavioral information in health insurance pricing to analyze which information insurers should be allowed to use from an economic point of view. I show how behavioral risk factors can create complications especially when the productivity of prevention depends on individuals' genetic disposition. Information asymmetries with regard to genetic risk factors relate to problems of adverse selection, information asymmetries with regard to behavioral risk factors relate to problems of moral hazard. Starting with the seminal works of Akerlof (1970) for adverse selection as well as Shavell (1979) for moral hazard, the two types of information asymmetries have typically been studied separately in the literature.[1] This paper therefore contributes to the existing literature by jointly considering both types of information asymmetries and carving out how they interact with each other.

Conditions such as diabetes, cancer, or cardiovascular disease develop gradually over time and lead to high expected healthcare spending over a longer period of time. To account for this long-term character of the development of health risks, I consider a two-period model with short- and long-term health insurance contracts. The model of the insurance market is based on the Einav, Finkelstein, and Cullen (EFC) framework, in which insurers offer exogenously fixed contracts and only compete on premiums (see, Einav et al., 2010; Einav and Finkelstein, 2011). Insurers adjust the premium of short-term contracts when new information about an individual's risk type is revealed. Therefore, individuals face a classification (or premium) risk which effectively reduces the extent to which health risks are insured. In contrast, long-term contracts promise a premium stream that does not depend on the individual's future health condition. In the spirit of Hendel and Lizzeri (2003), insurers can commit to future premiums whereas individuals cannot commit to keep their contract when new information about their risk type is revealed. Therefore, to keep individuals who turn out to be low-risk types in the contract, long-term contracts feature a pre-

[1] Notable exceptions are Stewart (1994) and Jack (2002), who discuss effects on welfare and equilibrium existence, respectively, when both problems are present simultaneously, as well as Einav et al. (2013), who empirically investigate the possibility that individuals select insurance coverage based on their anticipated moral hazard response to insurance.

payment (or front-loading). Depending on the regulation in place, this prepayment may depend on individuals' genetic disposition and prevention behavior.

My results show that using behavioral information in insurance pricing improves welfare in a Pareto-sense whereas deciding about the use of genetic information results in an equity-efficiency trade-off. Hence, standard results about the welfare effects of risk classification from models of pure adverse selection or pure moral hazard are robust to the introduction of the other dimension of information. Regarding the interaction of adverse selection and moral hazard, this paper shows that, on the one hand, banning the use of behavioral information exacerbates adverse selection on genetic differences. If insurers do not use behavioral information in pricing, insurance discourages prevention which increases long-term health costs. As a consequence, the price of insurance increases and individuals at low genetic risk leave the pool of insured. On the other hand, banning the use of genetic information impacts the capability of the use of behavioral information to mitigate moral hazard if the productivity of prevention depends on individuals' genetic disposition. If insurers cannot tailor incentives for prevention to genetic risk, individuals may choose inefficient levels of prevention even if insurers use behavioral information in pricing. As a consequence, if the distribution of individuals' genetic disposition is not too dispersed and prevention is equally productive for all individuals, banning the use of genetic information and promoting the use of behavioral information can achieve the ex ante welfare maximizing outcome in a private market. Otherwise, regulators face an equity-efficiency trade-off and only a social planner intervention can resolve this trade-off.

Currently, medical research is only at an early stage of understanding the development of multifactorial diseases. Genetic effects that have been identified so far are largely of moderate size and the interaction between genetic and behavioral risk factors is still far from being perfectly understood (Qi et al., 2008; Pomeroy et al., 2009). Medical studies suggest that individuals at high genetic risk benefit the most from a healthy lifestyle but behavioral factors have a strong effect on the probability of disease for all genetic risk groups (e.g., Said et al., 2018). Therefore, given the current state of medical knowledge, banning the use of genetic information and promoting the use of behavioral information has the potential to eliminate classification risk in a private health insurance market with long-term contracts and to reduce healthcare spending due to modifiable risk factors. In the future, however, increased medical knowledge may lead to a more dispersed distribution of genetic risk and offer more opportunities to individualize prevention measures based on genetic information. In this case, the intervention of a social planner may be necessary to avoid inefficient prevention behavior and tackle classification risk in health insurance.

This paper contributes to two strands of literature. First, it contributes to the risk classification literature by discussing the use of genetic and behavioral information in insurance pricing. Although the use of behavioral information constitutes a promising avenue not only to assess risks more precisely but also to encourage preventive behavior, it has received little attention in the

literature so far. Similar to this paper, Hoy (1989) considers individuals who possess different prevention technologies. He analyzes the welfare implications of screening mechanisms matching individuals to their prevention technologies. However, he does not discuss the welfare implications of screening mechanisms that capture preventive behavior. Bond and Crocker (1991) and Polborn (2008) study risk classification based on behavior that is causally or statistically correlated with loss propensities and discuss effects on consumer choices. They consider individuals with different preferences for a hazardous good such that the consumption of the hazardous good serves as a signal about individuals' exogenous risk type. I complement their work as preventive behavior does not serve as a signal about genetic risk in my model but I explicitly disentangle genetic and behavioral factors to discuss the use of either type of information. By doing so, I show how incorporating the relationship between genes and behavior has important implications for appropriate regulation of the use of information in insurance pricing.

Second, the paper contributes to the literature about classification risk and long-term insurance contracts by introducing preventive behavior as a risk factor affecting the development of diseases. Long-term health insurance contracts incorporating prepayments to insure classification risk are common and have been empirically studied in Germany (Hofmann and Browne, 2013) and Chile (Atal, 2019). Recently, they have also received some attention as a potentially welfare-improving alternative to the current regulatory framework in the US because they avoid the problem of adverse selection created by community rating (Atal et al., 2021; Ghili et al., 2022). Pauly et al. (1995) and Cochrane (1995) show that long-term contracts can eliminate classification risk in an ideal setting. If individuals have some private information about their future risk type, however, Peter et al. (2016) demonstrate that only some individuals or none will cover themselves against classification risk with a long-term contract. While Peter et al. (2016) only consider private information regarding exogenous risk factors, I extend their work by disentangling exogenous genetic and endogenous behavioral risk factors. Explicitly considering both types of information allows to account for the different nature of the two types of risk factors and to carve out how they interact with each other. This paper therefore contributes to the literature by showing how introducing prevention affects the effectiveness of long-term contracts for eliminating classification risk.

This paper proceeds as follows. The next section sets up the formal model. Section three determines the optimal level of prevention in a general setting. Section four studies market outcomes under different regulatory regimes for the use of genetic and behavioral information and discusses the resulting welfare implications. Section five illustrates how a social planner can resolve the equity-efficiency trade-off which regulators face when they decide about the use of information. Section six relates the model more closely to the existing literature on risk classification as well as on long-term and fixed contracts, and discusses the underlying model assumptions. The final section concludes.

2 The model

2.1 Individuals

Individuals maximize their expected utility in a two-period setting. Their utility function is time-additively separable. The felicity functions u and v of final wealth in the first and in the second period, respectively, are twice differentiable.[2] Both felicity functions are increasing and concave ($u' > 0$, $u'' < 0$ and $v' > 0$, $v'' < 0$) representing non-satiation and risk aversion. Initial wealth is given by w_1 in the first period and by w_2 in the second period.

In the first period t_1, individuals know their genetic disposition $z_0 \in [0, 1]$ and decide about their prevention expenditures $e \geq 0$ and risk-free savings s. I introduce savings to limit the role of long-term contracts to the insurance of classification risk rather than confounding it with intertemporal consumption smoothing.[3] For the sake of simplicity, the risk-free interest rate is set to zero and individuals do not face any risk in the first period. At the beginning of the second period t_2, each individual becomes either a high- or a low-risk type. Expected medical expenses in the second period equal P_H for high-risk types and P_L for low-risk types, with $0 < P_L < P_H$.[4] Once risk types have developed, individuals can no longer hide their risk type or influence their medical expenses.[5]

The prevention technology $z(z_0, e)$ characterizes the development of risk types. It yields the probability of becoming a high-risk type for an individual with genetic disposition z_0 and prevention expenditures e. The prevention technology has the following properties, where subscripts denote partial derivatives: $z(z_0, 0) = z_0$, $z_{z_0} > 0$, $z_e < 0$, and $z_{ee} > 0$.[6] The genetic disposition z_0 denotes the probability of becoming a high-risk type if the individual does not engage in prevention and a smaller z_0 generally represents lower genetic risk. Furthermore, prevention decreases the probability of becoming a high-risk type and the marginal productivity of prevention is decreasing. Finally, to incorporate the interaction of genetic and behavioral risk factors in the model, I adapt the terminology introduced by Hoy (1989) and say that the prevention technology exhibits

[2] The special case $v = \beta u$ yields the discounted expected utility model with $\beta \in (0, 1]$ being the rate of pure preference for the present.

[3] Frick (1998) notes that borrowing constraints may pose a problem to long-term contracts because premiums at young ages can be unaffordably high. However, Herring and Pauly (2006) show based on US medical expenditure data that in practice premiums would rather be increasing with age since both the probability of becoming a high-risk type and expected medical expenses of all risk types increase with age. Hence, borrowing constraints plausibly would not affect the attractiveness of long-term health insurance.

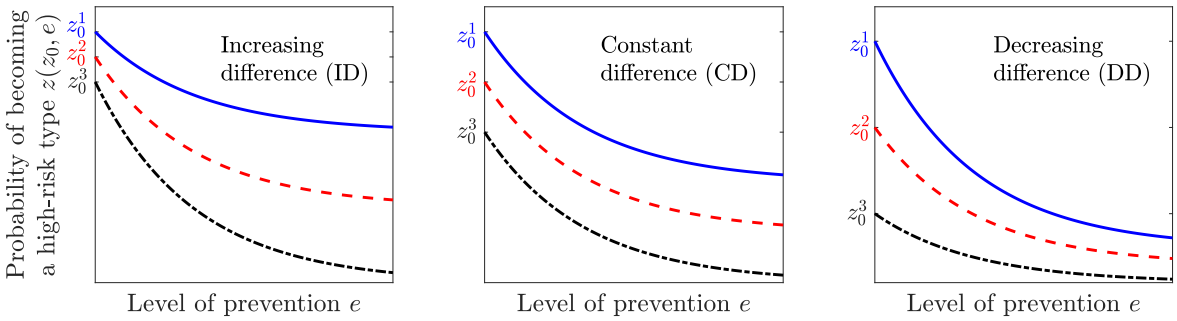
[4] This general form of expected medical expenses includes the case that both risk types face the same loss risk of size l but with different probabilities of loss, $p_L < p_H$, which yields $P_L = p_L l$ and $P_H = p_H l$. For instance, both diabetics and non-diabetics may suffer from kidney failure and incur treatment expenses for renal dialysis but the probability of such a health loss is higher for diabetics. The general form of expected medical expenses also includes the case that high-risk types have higher expenses with certainty such as when a diabetic patient needs regular insulin injections.

[5] That is, I abstract from situations in which individuals do not undertake medical examinations in order to avoid being classified as a high-risk type as well as from problems of moral hazard after the revelation of risk types.

[6] For $z_0 = 0$, the latter assumptions are relaxed to $z_e \leq 0$ and $z_{ee} \geq 0$, which yields $z(0, e) = 0$ for all $e \geq 0$.

increasing difference (ID) if $z_{ez_0} > 0$, constant difference (CD) if $z_{ez_0} = 0$, and decreasing difference (DD) if $z_{ez_0} < 0$. Figure 1 depicts the three cases.

Figure 1: ID, CD, DD



Notes: The figure displays the probability of becoming a high-risk type, $z(z_0, e)$, as a function of the expenditures on prevention e for different prevention technologies and genetic dispositions $z_0^1 > z_0^2 > z_0^3$. Since $z_{z_0} > 0$, individuals at lower genetic risk have a smaller probability of becoming a high-risk type, $z(z_0^1, e) > z(z_0^2, e) > z(z_0^3, e)$, for all $e \geq 0$. The terminology of ID, CD, and DD is motivated by the properties of the difference function $\delta(e) = z(z_0^1, e) - z(z_0^2, e) > 0$, for $z_0^1 > z_0^2$. $\delta(e)$ describes the difference between the probabilities of becoming a high-risk type for individuals with different genetic dispositions z_0^1 and z_0^2. Considering the derivative of the difference function, $\delta'(e) = z_e(z_0^1, e) - z_e(z_0^2, e) > (=, <)\ 0$ for all $e \geq 0$ if $z_{ez_0} > (=, <)\ 0$.

Intuitively, when individuals are young, they have an idea about their future risk type but do not know it perfectly. For example, individuals have some information about their genetic disposition because they know which diseases their family members suffered from in the past. Knowing their genetic disposition, they can engage in prevention in a targeted way in order to improve their future health prospects. For example, they can join a fitness class or undertake preventive medical examinations. Later in life, individuals may suffer from diseases, such as diabetes, cancer, or cardiovascular disease, that lead to high expected medical expenses over a longer period of time and therefore classify them as high-risk types. Their probability to contract such a disease depends on their genetic endowment and their engagement in prevention earlier in life. For many cardiovascular diseases and diabetes, medical studies suggest that prevention is particularly productive for individuals at high genetic risk.[7] In this case, the prevention technology exhibits DD.

2.2 Insurers

To cover medical expenses, individuals can purchase health insurance. In a competitive market, risk-neutral insurers offer short- and long-term contracts. The model of the insurance market is based on the EFC framework. That is, insurers offer exogenously fixed contracts and only compete

[7]For example, Shook et al. (2012) find that the risk-reducing effect of a higher level of fitness due to regular physical activity is stronger for individuals with a parental history of hypertension than for those without. Using polygenic risk scores and overall lifestyle categories, Said et al. (2018) also find that prevention for diabetes and various cardiovascular diseases is particularly productive for individuals at high genetic risk.

on premiums (see, Einav et al., 2010; Einav and Finkelstein, 2011). For reasons of tractability, I focus on full insurance contracts.[8] Insurers offer zero-profit contracts and individuals choose their most preferred contract from the set of available contracts.

Short-term contracts can be purchased after the revelation of risk types at the beginning of the second period. They cover medical expenses at the fair premium depending on an individual's risk type.[9] The fair type-dependent premium is equal to the expected medical expenses of P_L for low-risk types and P_H for high-risk types. Short-term contracts cover medical expenses at t_2 but they expose individuals to classification risk due to uncertain future insurance premiums.

In contrast, long-term contracts offer a premium stream that does not depend on any revealed risk type. They insure both the health risk at t_2 and the classification risk due to the evolution of risk types over time. In the spirit of Hendel and Lizzeri (2003), I model long-term contracts as contracts with one-sided commitment. Insurers can commit to future premiums whereas individuals freely choose between staying with their long-term contract and switching to a short-term contract at t_2. Therefore, to obtain insurance against classification risk, individuals must prepay some of the premiums such that the prepayment locks them into the contract. In my simple two-period model, individuals purchasing a long-term contract make a prepayment P at t_1 and pay the low-risk premium P_L to insure the health risk at t_2 regardless of their risk type. The guaranteed premium of P_L at t_2 keeps low-risk types in the contract and the fair prepayment P covers the expected medical expenses of high-risk types in excess of the low-risk premium. When insurers are allowed to use genetic or behavioral information in pricing, they can condition their offers of long-term contracts on z_0 or e (or both). Figure 2 illustrates the sequence of play.

3 Optimal prevention

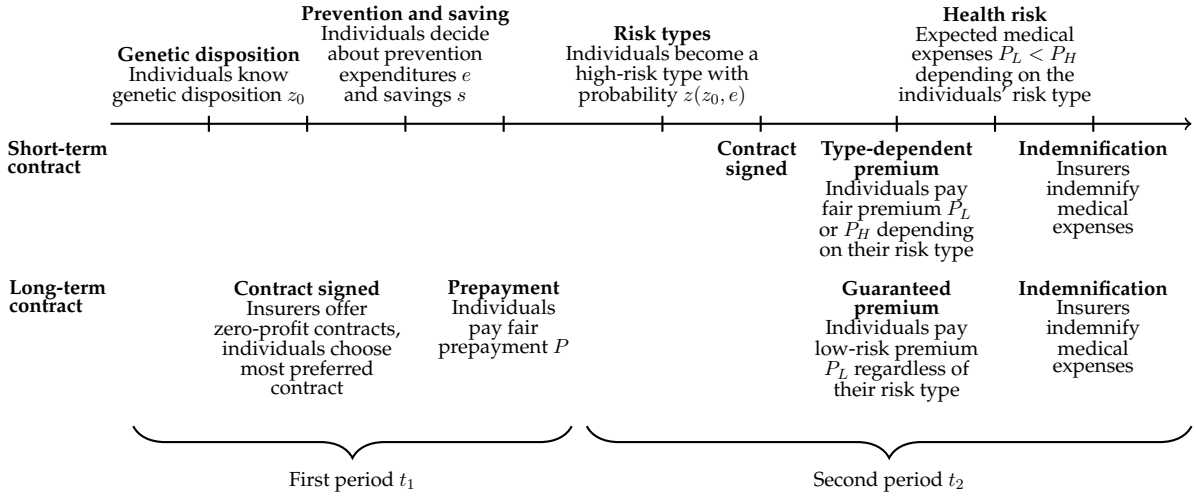
3.1 Ex ante welfare maximizing outcome

Before studying prevention behavior and market outcomes in a competitive market with short- and long-term health insurance contracts, I determine the outcome which maximizes ex ante wel-

[8] The EFC framework is increasingly used for modeling insurance markets. It is particularly well suited to model health insurance markets since these are often highly regulated and products therefore highly standardized. The restriction to fixed full insurance contracts keeps the analysis tractable and allows to highlight key trade-offs when it comes to the use of genetic and behavioral information in health insurance pricing. For a detailed discussion of the applicability of the EFC framework, see Section 6.

[9] We can think of a one-year health insurance contract here whose premium depends on the individual's health condition when entering the contract at the beginning of the year. Insurers can distinguish high- and low-risk types, for example, by checking individuals' medical history for pre-existing conditions in the underwriting process. Recently, such short-term contracts have gained some attention because they are heavily on the rise in China (Swiss Re, 2021).

Figure 2: Sequence of play



Notes: The figure displays the timing of the evolution of risk types and the insurance payments. In the first period, individuals have some information about their genetic disposition and decide about their expenditures on prevention and savings. Individuals purchasing a long-term contract also make a prepayment to insure classification risk. At the beginning of the second period, risk types are revealed. Short-term contracts cover medical expenses in the second period at the fair premium depending on an individual's risk type. Long-term contracts cover medical expenses at the fair low-risk premium even if the individual has become a high-risk type.

fare as a benchmark. That is, I look for the most preferred outcome from a perspective where individuals do not know their genetic disposition yet.[10]

Let

$$EH(e; z_0) = z(z_0, e)P_H + (1 - z(z_0, e))P_L + e$$

denote the expected lifetime health expenditures of individuals with genetic disposition z_0 and prevention expenditures e. Expected lifetime health expenditures consist of the medical expenses at t_2 and the expenditures on prevention at t_1. The ex ante welfare maximizing outcome is characterized by the optimization problem

$$\begin{aligned} \max_{\substack{e\geq 0,\\ C_1, C_2^H, C_2^L \geq 0}} \quad & \mathbb{E}\left[u(C_1) + z(z_0, e)v(C_2^H) + (1 - z(z_0, e))v(C_2^L)\right] \\ \text{s.t. } & \mathbb{E}\left[C_1 + z(z_0, e)C_2^H + (1 - z(z_0, e))C_2^L\right] \leq w_1 + w_2 - \mathbb{E}\left[EH(e; z_0)\right], \end{aligned}$$

where e, C_1, C_2^H, and C_2^L may depend on individuals' genetic disposition z_0 and expectations are taken with respect to the distribution of z_0. Due to risk aversion, the welfare-maximizing consumption stream neither depends on individuals' risk type nor on their genetic disposition, i.e. C_1 and $C_2^H = C_2^L$ do not depend on z_0. Moreover, due to non-satiation, the welfare-maximizing level of prevention maximizes total expected wealth, which holds if and only if it minimizes each indi-

[10] This approach is based on Harsanyi's (1953; 1955) veil of ignorance and has been adapted to discuss the welfare consequences of classification bans by Hoy (2006).

vidual's expected lifetime health expenditures $EH(e; z_0)$. Hence, the efficient level of prevention e^* is characterized by the first-order condition

$$EH_e = z_e(z_0, e^*)(P_H - P_L) + 1 = 0. \tag{1}$$

The second-order condition is globally fulfilled since $z_{ee} > 0$. I assume that an interior solution exists for all $z_0 \in (0, 1]$, which holds if $-z_e(z_0, 0) > \frac{1}{P_H - P_L}$, i.e. if an infinitesimal expenditure on prevention has a sufficiently large effect on the probability of becoming a high-risk type.[11] Applying the implicit function theorem to analyze how the efficient level of prevention depends on an individual's genetic disposition yields the following proposition, for which Appendix A.1 provides a proof.

Proposition 1. *The efficient level of prevention minimizes each individual's expected lifetime health expenditures. This level of prevention is lower (the same, higher) for individuals at high genetic risk than for individuals at low genetic risk if the prevention technology exhibits ID (CD, DD). The welfare-maximizing consumption stream does not depend on an individual's genetic disposition or health risk type.*

The genetic information encoded in z_0 characterizes not only individuals' probability of becoming a high-risk type but also the productivity of their prevention technology. Since $z_e < 0$, prevention is more productive at the margin if z_e is smaller, i.e. more negative. This implies that the marginal productivity of prevention decreases (is constant, increases) in z_0 if the prevention technology exhibits ID (CD, DD). Proposition 1 shows that the efficient level of prevention is higher, the more productive an individual's prevention technology.

3.2 Short-term contracts

With a short-term contract, the expected utility of an individual with genetic disposition z_0 equals

$$EU^{ST}(e, s; z_0) = u(w_1 - e - s) + z(z_0, e)\, v(w_2 - P_H + s) + (1 - z(z_0, e))\, v(w_2 - P_L + s).$$

In the first period, individuals receive the income w_1 and have prevention expenditures of e. Savings s transfer wealth between the two periods. In the second period, individuals receive the income w_2 and insure the health risk at the fair premium P_H or P_L depending on their risk type.

Interior solutions $\left(e^{ST}, s^{ST}\right)$ for the optimal levels of prevention and saving are characterized by the first-order conditions

[11] Individuals with genetic disposition $z_0 = 0$ definitely become low risks even if they do not engage in prevention. Hence, their efficient level of prevention is equal to zero.

$$
\begin{aligned}
EU_e^{ST} &= -u'\left(w_1 - e^{ST} - s^{ST}\right) - z_e\left(z_0, e^{ST}\right)\left(v\left(w_2 - P_L + s^{ST}\right) - v\left(w_2 - P_H + s^{ST}\right)\right) = 0, \\
EU_s^{ST} &= -u'\left(w_1 - e^{ST} - s^{ST}\right) + \quad z\left(z_0, e^{ST}\right)\; v'\left(w_2 - P_H + s^{ST}\right) \\
&\qquad\qquad + \left(1 - z\left(z_0, e^{ST}\right)\right)\; v'\left(w_2 - P_L + s^{ST}\right) = 0.
\end{aligned}
$$

I assume that the second-order conditions are satisfied and an interior solution exists for all $z_0 \in (0, 1]$. Optimal prevention equalizes the marginal utility cost of prevention expenditures at t_1 and their expected marginal utility benefit resulting from the decrease of the probability of paying the high-risk premium at t_2. Optimal savings smooth consumption across the two periods. Since there are no information asymmetries after the revelation of risk types and insurance is offered at the fair premium, all individuals prefer purchasing a short-term full insurance contract at t_2 over staying uninsured according to Mossin's (1968) Theorem.

3.3 Long-term contracts

Depending on the regulatory regime in place, individuals' opportunities on the insurance market may depend on their genetic disposition z_0 and prevention expenditures e. I therefore denote the prepayment of the most preferred long-term contract for an individual with genetic disposition z_0 and prevention expenditures e by $P(e, z_0)$. With a long-term contract, the expected utility of an individual with genetic disposition z_0 equals

$$
EU^{LT}\left(e, s; z_0\right) = u\left(w_1 - e - P\left(e, z_0\right) - s\right) + v\left(w_2 - P_L + s\right).
$$

Compared to individuals purchasing a short-term contract, individuals make the additional prepayment $P(e, z_0)$ in the first period which allows them to insure the health risk in the second period at the low-risk premium P_L regardless of their risk type. When purchasing a long-term contract, the individuals' consumption in both periods is given with certainty. Consumption neither depends on the individuals' risk type nor on their actual medical expenses.

Individuals choose the optimal levels of prevention and saving maximizing expected utility. Interior solutions $\left(e^{LT}, s^{LT}\right)$ are characterized by the first-order conditions

$$
\begin{aligned}
EU_e^{LT} &= u'\left(w_1 - e^{LT} - P\left(e^{LT}, z_0\right) - s^{LT}\right)\left(-1 - P_e\left(e^{LT}, z_0\right)\right) = 0, \\
EU_s^{LT} &= -u'\left(w_1 - e^{LT} - P\left(e^{LT}, z_0\right) - s^{LT}\right) + v'\left(w_2 - P_L + s^{LT}\right) = 0.
\end{aligned} \tag{2}
$$

In Appendix B.1, I show that the second-order conditions are globally fulfilled if $P_{ee} > 0$. Since $u' > 0$, the first-order condition with respect to e is equivalent to

$$
P_e\left(e^{LT}, z_0\right) = -1. \tag{3}
$$

Optimal prevention equalizes the marginal benefit from a reduced prepayment and the marginal cost of prevention which is constantly equal to 1. Since long-term health insurance removes all the risk from the consumption stream, individuals choose the level of prevention which maximizes their wealth. They then smooth their consumption across the two periods by choosing the optimal level of saving according to (2).

4 Regulatory regimes

Depending on the regulation in place, insurers can condition their offers of long-term contracts on individuals' genetic disposition z_0 and prevention expenditures e. Individuals either purchase their most preferred long-term contract or they leave classification risk uninsured and purchase a short-term contract later in life.

I now discuss how restrictions on the use of information affect individuals' prevention behavior and their demand for long-term health insurance. The resulting welfare implications, which I derive at the end of this section, can help to decide in which cases regulatory restrictions are desirable and in which cases policy-makers and insurers should promote individual underwriting. Table 1 summarizes the main results for each regime.

Table 1: Prevention behavior and demand for long-term health insurance under different regulatory regimes for the use of genetic and behavioral information

Regulatory regime	**Information used**	**Prevention behavior**	**Demand for long-term health insurance**
Full information	e and z_0	Individuals minimize their expected lifetime health expenditures $\rightarrow$ Efficient level of prevention	All individuals purchase long-term insurance But: New classification risk based on genetic risk
No individual underwriting	-	Insurance eliminates incentives for prevention $\rightarrow$ High future medical expenses	Nobody or only individuals at high genetic risk purchase long-term insurance $\rightarrow$ Complete market unraveling possible
Only behavioral information allowed	e	Individuals minimize the expected lifetime health expenditures of an "average insured" $\rightarrow$ Efficient level of prevention if productivity of prevention independent of genetic disposition	Individuals at high genetic risk purchase long-term insurance; More individuals than with no underwriting purchase long-term insurance $\rightarrow$ Efficient risk allocation if the distribution of genetic risk is not too dispersed
Only genetic information allowed	z_0	Insurance eliminates incentives for prevention $\rightarrow$ High future medical expenses	Nobody or only some individuals purchase long-term insurance; Individuals at high genetic risk never purchase long-term insurance $\rightarrow$ Complete market unraveling possible

Notes: The table provides an overview of the main results for each regulatory regime.

4.1 Full information

Zero-profit contracts

In a full information regime, insurers may use both genetic and behavioral information. Hence, they can condition their offer of a long-term contract on both z_0 and e. Individuals with genetic disposition z_0 and prevention expenditures e become a high-risk type with probability $z(z_0, e)$. Since a high risk's expected medical expenses are given by P_H but all individuals pay the low-risk premium P_L in the second period, the expected excess medical expenses of each high risk equal $P_H - P_L$. Hence, a contract offered to individuals with genetic disposition z_0 and prevention expenditures e makes zero profits if and only if the prepayment equals

$$P^{fu}(e, z_0) = z(z_0, e)(P_H - P_L).$$

For individuals with genetic disposition z_0, the set of zero-profit contracts is thus given by

$$\mathcal{C}^{fu}(z_0) = \{(P^{fu}(e, z_0), e) \mid e \geq 0\}.$$

Optimal prevention

From the set of available contracts $\mathcal{C}^{fu}(z_0)$, individuals choose the one which yields the highest expected utility. Using the general first-order condition (3), we see that this contract, and thus the optimal level of prevention e^{fu}, is characterized by

$$P_e^{fu}(e^{fu}, z_0) = z_e(z_0, e^{fu})(P_H - P_L) = -1. \tag{4}$$

Comparing this first-order condition with (1) shows that in a full information regime long-term health insurance implements the efficient level of prevention that minimizes expected lifetime health expenditures. We can also relate this level of prevention to the level of prevention that individuals choose with a short-term contract leaving classification risk uninsured, which yields the following proposition. I provide a proof in Appendix A.2.

Proposition 2. *When insurers have full information about individuals' genetic disposition and prevention expenditures, individuals with a long-term health insurance contract choose the efficient level of prevention that minimizes their expected lifetime health expenditures. Long-term contracts raise (do not change, reduce) the level of prevention compared to short-term contracts if the individual's probability of becoming a high-risk type with short-term contracts, $z(z_0, e^{ST})$, is greater than (equal to, less than) an endogenously determined threshold z^c.*

On the one hand, long-term insurance eliminates the incentive to engage in prevention to avoid the expensive high-risk premium later in life. On the other hand, expenditures on prevention allow individuals to purchase a long-term contract with a lower prepayment, which provides a new incentive for prevention. With a long-term contract, only the marginal effect of prevention on expected lifetime health expenditures matters when individuals decide about their level of prevention. With a short-term contract, high-risk types have to pay a higher premium later in life which reduces their total wealth. Hence, prevention expenditures ex post have a higher utility cost for high-risk types than for low-risk types. Therefore, the expected marginal utility cost of prevention is higher if an individual is more likely to become a high-risk type later in life. As a result, the optimal level of prevention with a short-term contract is lower (higher) than the optimal level of prevention with a long-term contract for individuals at high (low) genetic risk.

Demand for long-term health insurance

Individuals can either only insure the health risk at t_2 by purchasing a short-term contract or also insure the classification risk due to the evolution of health risk types by purchasing a long-term contract. When insurers have full information about individuals' genetic disposition and prevention expenditures, individuals can insure classification risk at their personal fair premium, which is desirable since they are risk-averse. Therefore, we obtain the following proposition, for which Appendix A.3 provides a formal proof.

Proposition 3. *When insurers have full information about individuals' genetic disposition and prevention expenditures, all individuals purchase long-term health insurance.*

The market outcome that long-term contracts attract all individuals is the same as the one obtained by Cochrane (1995) and Pauly et al. (1995) in the absence of prevention and genetic heterogeneity. The reason is that the long-term contract characterized by a particular prepayment $P^{fu}(e, z_0)$ is only offered to a homogeneous group of individuals with genetic disposition z_0 and prevention expenditures e.

Discussion

When insurers have full information about individuals' genetic disposition and prevention expenditures, long-term health insurance encourages insured individuals to choose the efficient level of prevention that minimizes their lifetime health expenditures. Hence, long-term health insurance can help to reduce medical expenses associated with diseases for which the probability of disease depends to some extent on individuals' behavior. Medical studies suggest that prevention exhibits DD for many cardiovascular diseases and diabetes. In this case, the efficient level of prevention

is higher for individuals at high genetic risk (see Proposition 1). Proposition 2 implies that individuals at high genetic risk increase their expenditures on prevention if they purchase a long-term instead of a short-term contract. Even though these individuals know that prevention is particularly productive for them, they choose an inefficiently low level of prevention with a short-term contract because they prefer to save money in order to be able to pay for the expensive high-risk contract if they become a high-risk type later in life. Long-term health insurance eliminates this inefficiency because individuals' wealth later in life no longer depends on their risk type.

Regarding the demand for long-term health insurance, all individuals purchase a long-term contract. Hence, at first glance, risk allocation is efficient since all individuals insure the classification risk due to the evolution of health risk types. However, premium discrimination based on genetic differences creates a new classification risk based on genetic instead of health risk types since individuals at high genetic risk have to make a higher prepayment than individuals at low genetic risk. Therefore, risk allocation is not efficient from behind a veil of ignorance, where genetic risk types are still to be assigned.

4.2 No individual underwriting

Optimal prevention

When insurers are not allowed to use genetic or behavioral information, no individual underwriting of long-term health insurance takes place. All individuals are offered the same long-term contracts with the same prepayments P^{no}, regardless of their genetic disposition z_0 and prevention expenditures e. Hence, the expected utility of all individuals purchasing a long-term contract with prepayment P^{no} is given by

$$EU^{no}(e,s) = u(w_1 - e - P^{no} - s) + v(w_2 - P_L + s).$$

Since the price of insurance neither depends on the level of prevention nor on an individual's risk type, individuals holding a long-term contract are not rewarded for expenditures on prevention. Therefore, the optimal expenditure is given by $e^{no} = 0$ and insurance discourages prevention. This yields the following proposition.

Proposition 4. *When no individual underwriting of long-term health insurance takes place, insurance eliminates incentives for prevention.*

Demand for long-term health insurance

Individuals choose between purchasing long-term insurance in the first period or leaving classification risk uninsured and purchasing short-term insurance in the second period. With a long-term contract, individuals' utility does not depend on their genetic disposition. In contrast, expected utility with a short-term contract is decreasing in the genetic disposition z_0. Consequently, individuals at high genetic risk are more interested in long-term health insurance than individuals at low genetic risk. For any given prepayment P^{no}, either no one purchases long-term health insurance or there exists a cutoff $z_0^* \in [0,1]$ such that individuals with genetic disposition $z_0 \geq z_0^*$ purchase long-term insurance whereas individuals with genetic disposition $z_0 < z_0^*$ prefer not to insure classification risk.

Zero-profit contracts

Since individuals who purchase a long-term contract do not engage in prevention, a fraction $\mathbb{E}\left[z_0 \mid z_0 \geq z_0^*\right]$ becomes a high-risk type in the second period if the cutoff is given by z_0^*. Hence, insurers make zero profits if and only if the prepayment equals

$$P^{no}(z_0^*) = \mathbb{E}\left[z_0 \mid z_0 \geq z_0^*\right](P_H - P_L).$$

In equilibrium, contracts must be informationally consistent. That is, the cutoff which forms under the long-term contract with the prepayment $P^{no}(z_0^*)$ indeed has to be equal to z_0^*. When there is no individual underwriting, the set of zero-profit contracts is thus given by

$$\mathcal{C}^{no} = \{P^{no}(z_0^*) \mid z_0^* \text{ informationally consistent cutoff}\}.$$

In order to analyze whether informationally consistent cutoffs exist, I first consider the extreme cases $z_0^* = 0$ and $z_0^* = 1$. The corresponding formal calculations are provided in Appendix B.2. Individuals with genetic disposition $z_0 = 0$ will definitely become a low risk later in life. They do not face classification risk but would have to subsidize worse genetic risks if they purchased long-term health insurance. They therefore prefer to purchase a short-term contract later in life. Hence, $z_0^* = 0$ cannot be an informationally consistent cutoff and, in contrast to the full information regime, long-term health insurance is not in demand by all individuals. Concerning the case $z_0^* = 1$, when insurers cannot assess prevention behavior, individuals holding a long-term contract cannot credibly commit to engage in prevention. Consequently, a cutoff of $z_0^* = 1$ would imply that all individuals purchasing long-term health insurance become a high-risk type with certainty and a zero-profit long-term contract would be priced accordingly. On the other hand, if individuals with genetic disposition $z_0 = 1$ engage in prevention, there is a chance that they can purchase

a short-term contract at the low-risk premium later in life. Therefore, individuals with genetic disposition $z_0 = 1$ prefer not to purchase a long-term contract but to engage in prevention instead and $z_0^* = 1$ cannot be an informationally consistent cutoff either.

Individuals with an intermediate genetic disposition $z_0 \in (0, 1)$ face the following trade-off when deciding whether to purchase long-term health insurance or not. On the one hand, long-term insurance is attractive because it removes classification risk. On the other hand, expenditures on prevention do not pay off with a long-term contract. Moreover, the prepayment is based on the average probability of becoming a high-risk type across the pool of insured. Consequently, individuals at low genetic risk have to subsidize individuals at high genetic risk. Therefore, two types of market outcome are possible, which I summarize in the following proposition. Appendix A.4 provides a proof.

Proposition 5. *Assume that no individual underwriting of long-term health insurance takes place. If no informationally consistent cutoff exists, long-term health insurance is not in demand. If at least one informationally consistent cutoff exists, the uniquely determined Nash equilibrium in the market is characterized by the lowest informationally consistent cutoff* $z_0^{*,no} \in (0, 1)$. *Individuals with genetic disposition* $z_0 \geq z_0^{*,no}$ *purchase long-term health insurance whereas individuals with genetic disposition* $z_0 < z_0^{*,no}$ *leave classification risk uninsured and purchase a short-term contract later in life.*

If no informationally consistent cutoff exists, there is complete unraveling of the market for long-term insurance. Otherwise, individuals at high genetic risk $z_0 \geq z_0^{*,no}$ purchase long-term insurance and do not engage in prevention whereas individuals at lower genetic risk $z_0 < z_0^{*,no}$ prefer not to insure classification risk and choose the level of prevention which maximizes their expected utility with a short-term contract. If several informationally consistent cutoffs exist, competitive forces drive the equilibrium to the lowest cutoff because all individuals prefer a lower over a higher cutoff. A lower cutoff yields a pool of insured that has on average better risk type prospects. Hence, long-term insurance is cheaper and individuals' expected utility is higher, the lower the cutoff is.

Discussion

Under a ban on the use of both genetic and behavioral information, all individuals face the same premium stream regardless of their genetic disposition. Therefore, long-term health insurance tackles classification risk from both an interim and and ex ante point of view. In contrast, long-term health insurance in the full information regime only replaced interim classification based on health types by ex ante classification based on genetic types.

On the other hand, long-term insurance discourages prevention when insurers do not use behavioral information in pricing. Under DD, prevention is particularly productive for individuals

at high genetic risk. These individuals also form the group of individuals who purchase the long-term contract if it is in demand. Thus, long-term contracts discourage preventive behavior of those individuals who should actually choose the highest level of prevention. Moreover, individuals who consider themselves not very likely to become a high-risk type consider long-term contracts too expensive when there is no individual underwriting. In the extreme case of complete market unraveling, all individuals leave classification risk uninsured and everyone is worse off than in the full information regime.

Prohibiting the use of genetic and behavioral information without further regulatory interventions would create an environment similar to the health insurance market in the US before the Affordable Care Act (ACA) came into force in 2014. In those days, the Genetic Information and Nondiscrimination Act (GINA) of 2008 prohibited the use of genetic information. Since mobile devices were not as ubiquitous as they are today, the assessment of prevention behavior was not technologically feasible. Indeed, markets failed to provide comprehensive long-term health insurance coverage and many people were not able to renew their health insurance contracts when they were diagnosed with diseases that lead to high expected healthcare spending over a longer period of time. Such market failure is in line with the result that the long-term health insurance market may unravel when no individual underwriting takes place.

4.3 Only behavioral information

Demand for long-term health insurance

When insurers may only use behavioral information, they can condition their offer of a long-term contract on e but not on z_0. Hence, long-term contracts are characterized by a prepayment P^{beh} and an expenditure on prevention e and all individuals are offered the same contracts regardless of their genetic disposition. Therefore, individuals' utility with long-term contracts does not depend on their genetic disposition z_0 and all individuals unanimously prefer the same long-term contract when they choose between different contracts (P^{beh}, e). Expected utility with short-term contracts, in contrast, is decreasing in the genetic disposition z_0. Therefore, for any long-term contract (P^{beh}, e), either no one purchases this contract or there exists a cutoff $z_0^* \in [0, 1]$ such that, when deciding between this contract and not insuring classification risk, individuals with genetic disposition $z_0 \geq z_0^*$ prefer the long-term contract whereas individuals with genetic disposition $z_0 < z_0^*$ prefer not to insure classification risk. Individuals once more face a trade-off when they decide about their insurance purchases. On the one hand, they can get rid of classification risk by purchasing long-term health insurance. On the other hand, they may have to subsidize individuals at higher genetic risk. The benefit of insuring classification risk prevails for individuals at

high genetic risk ($z_0 \geq z_0^*$) whereas the disadvantage of subsidizing worse genetic risks prevails for individuals at low genetic risk ($z_0 < z_0^*$).

Zero-profit contracts

If the level of prevention is given by e and the cutoff is given by z_0^*, a fraction $\mathbb{E}\left[z(z_0, e) \mid z_0 \geq z_0^*\right]$ becomes a high-risk type in the second period. Hence, insurers make zero profits if and only if the prepayment equals

$$P^{beh}(e; z_0^*) = \mathbb{E}\left[z(z_0, e) \mid z_0 \geq z_0^*\right](P_H - P_L).$$

Individuals' utility thus equals

$$EU^{beh}(e, s; z_0^*) = u(w_1 - e - P^{beh}(e; z_0^*) - s) + v(w_2 - P_L + s).$$

In equilibrium, contracts must again be informationally consistent for insurers to make zero-profits. That is, when the prepayment is given by $P^{beh}(e; z_0^*)$ and the long-term contract requires prevention expenditures of e, the cutoff that separates the purchasers of long- and short-term contracts must indeed equal z_0^*. The set of zero-profit contracts is thus given by

$$\mathcal{C}^{beh} = \{(P^{beh}(e; z_0^*), e) \mid e \geq 0, z_0^* \text{ informationally consistent cutoff}\}.$$

When there is no individual underwriting, we saw that the set of zero-profit contracts $\mathcal{C}^{no}$ may be empty, which implies complete unraveling of the long-term health insurance market. In Appendix B.3, I show that there always exists an informationally consistent cutoff $z_0^* \in (0, 1)$ for the level of prevention $e^*(1)$, which is efficient for individuals with genetic disposition $z_0 = 1$ (see Section 3.1). Hence, the set of zero-profit contracts $\mathcal{C}^{beh}$ is never empty and long-term health insurance is always in demand when insurers use behavioral information.

Optimal prevention

The set of zero-profit contracts $\mathcal{C}^{beh}$ comprises the candidates for Nash equilibrium. In equilibrium, only the contract which yields the highest expected utility is in demand. We obtain the following propositions, for which Appendix A.5 provides a proof.

Proposition 6. *Assume that insurers use only behavioral information in pricing. The uniquely determined Nash equilibrium in the market is characterized by the lowest informationally consistent cutoff $z_0^{*,beh} \in (0, 1)$. Individuals with genetic disposition $z_0 \geq z_0^{*,beh}$ purchase long-term health insurance whereas individuals with genetic disposition $z_0 < z_0^{*,beh}$ leave classification risk uninsured and purchase a*

short-term contract later in life. The lowest informationally consistent cutoff is smaller than when insurers use neither behavioral nor genetic information in pricing, i.e. $z_0^{*,beh} < z_0^{*,no}$.

Proposition 7. *When insurers use only behavioral information in pricing, all individuals with a long-term contract choose the level of prevention that would minimize the expected lifetime health expenditures of an individual with the average prevention technology* $z^{avg}(e) = \mathbb{E}\left[z\left(z_0, e\right) \mid z_0 \geq z_0^{*,beh}\right]$.

The cutoff result is similar to the Nash equilibrium that we obtained when no individual underwriting of long-term health insurance takes place but there are two differences which must not be overlooked. First, when insurers use behavioral information in pricing, the existence of an informationally consistent cutoff is guaranteed. Therefore, there are always some individuals who purchase long-term insurance which is not the case when no individual underwriting takes place. Second, since prevention reduces individuals' expected medical expenses, long-term health insurance is cheaper, and hence, also attractive for individuals at lower genetic risk when insurers use behavioral information. Therefore, the pool of insured in the long-term contract is larger than without any individual underwriting.

Discussion

When insurers use only behavioral information in pricing, insured individuals minimize the health expenditures of an "average insured" instead of their personal health expenditures in the full information regime. Under CD, the marginal productivity of prevention $z_e(z_0, e)$ is the same for all individuals regardless of their genetic disposition z_0. Therefore, insured individuals still choose the efficient level of prevention that minimizes their personal expected lifetime health expenditures. If individuals have differently productive prevention technologies, however, those whose prevention technology is more (less) productive than the average prevention technology $z^{avg}(e)$ choose a level of prevention that is lower (higher) than their efficient level of prevention. This inefficiency occurs because insurers must offer a one-size-fits-all contract if they cannot differentiate their offerings based on individuals' genetic disposition. Therefore, banning the use of genetic information impacts the capability of the use of behavioral information to mitigate moral hazard if the productivity of prevention depends on individuals' genetic disposition.

The use of behavioral information reduces the inefficiency in risk allocation compared to the regime without any individual underwriting because more individuals cover classification risk by purchasing a long-term contract. Using behavioral information therefore not only tackles moral hazard but also mitigates adverse selection. Nevertheless, individuals at very low genetic risk may still prefer to bear classification risk themselves in order not to subsidize worse genetic risks. In the real-world, even individuals at very low genetic risk plausibly will become a high-risk type with positive probability which implies that the support of z_0 is a strict subset of the interval $[0, 1]$.

If the cutoff $z_0^{*,beh}$ is small enough such that all individuals have a genetic disposition $z_0 \geq z_0^{*,beh}$, all individuals purchase the long-term contract and there is no inefficiency in risk allocation.

So far, medical research is only at an early stage of understanding the interaction between genetic and behavioral risk factors in the development of multifactorial diseases (Qi et al., 2008; Pomeroy et al., 2009). Therefore, the inefficiency in prevention is likely of moderate size given the current state of medical knowledge. In the future, however, increased medical knowledge may offer more opportunities to individualize prevention measures based on genetic information and the inefficiencies in prevention when insurers use only behavioral information may become more relevant. Moreover, genetic effects that have been identified are largely of moderate size for common multifactorial diseases such as diabetes or cardiovascular disease (Qi et al., 2008; Said et al., 2018). Therefore, even individuals at very low genetic risk may purchase the long-term contract because the subsidy to individuals at high genetic risk is only small and hence may be outweighed by the benefit of insuring classification risk. In this case, long-term contracts using behavioral information have the potential to eliminate classification risk in health insurance.

4.4 Only genetic information

Optimal prevention

When insurers may use genetic but no behavioral information, they can condition their offer of a long-term contract on individuals' genetic disposition z_0 but not on their expenditures e on prevention. Consequently, the consumption stream of individuals who purchase long-term health insurance neither depends on the level of prevention nor on their risk type later in life. As in the regime without any individual underwriting, long-term insurance therefore discourages prevention yielding the following proposition.

Proposition 8. *When insurers use only genetic information in pricing, long-term health insurance eliminates incentives for prevention.*

Zero-profit contracts

Since individuals holding a long-term contract do not engage in prevention, insurers make zero-profits on individuals with genetic disposition z_0 if and only if the prepayment equals

$$P^{gen}(z_0) = z_0(P_H - P_L).$$

The set of zero-profit contracts for individuals with genetic disposition z_0 therefore contains exactly one element:

$$\mathcal{C}^{gen}(z_0) = \{P^{gen}(z_0)\}.$$

Demand for long-term health insurance

To analyze who purchases a long-term contract, I once more start with the extreme cases $z_0 \in \{0, 1\}$. For individuals with genetic disposition $z_0 = 0$, who definitely become a low-risk type later in life, the prepayment equals $P^{gen}(0) = 0$. These individuals could obtain the guaranteed premium in the second period without making a prepayment. However, they do not face classification risk because they will be able to get the low-risk short-term contract with certainty. Hence, there is no need for them to enter the long-term contract and they are indifferent between short- and long-term insurance. For individuals with genetic disposition $z_0 = 1$, it holds that $P^{gen}(1) = P_H - P_L$. If these individuals purchase the long-term contract, they will be treated as definite high risks. If they engage in prevention and purchase the short-term contract later in life, however, there will at least be a small chance that they get the cheaper low-risk short-term contract. Therefore, they prefer to leave classification risk uninsured.

Individuals with intermediate genetic disposition $z_0 \in (0, 1)$ face the following trade-off when deciding which type of insurance to purchase. On the one hand, a long-term contract enables them to get rid of the undesirable classification risk. On the other hand, it does not reward prevention. In contrast to the regimes in which insurers do not use genetic information, long-term health insurance is now more expensive the higher is someone's genetic risk z_0. Therefore, the market outcome is no longer described by a cutoff. Instead, the possible market outcomes are as follows. Appendix A.6 provides a proof.

Proposition 9. *Assume that insurers use only genetic information in pricing. Either long-term health insurance is not in demand or some individuals purchase long-term health insurance and some leave classification risk uninsured and purchase a short-term contract later in life. Individuals at high genetic risk (z_0 close to 1) never purchase long-term health insurance.*

Similar to the regime without any individual underwriting, there will always be some individuals who bear the classification risk themselves and even complete market unraveling is possible. This time, however, long-term health insurance is particularly unattractive for individuals at high genetic risk whereas we cannot make a definite statement for individuals at low genetic risk.

Discussion

When insurers use only genetic information in pricing, long-term health insurance discourages prevention. Consequently, expected long-term medical expenses are high. In the previous sections, we have seen that, although being ethically questionable, the use of genetic information may be beneficial for market efficiency because it allows insurers to individualize incentives for prevention and because individuals at low genetic risk may not purchase long-term health insurance otherwise. However, Proposition 9 shows that individuals at high genetic risk can not afford

insurance when insurers use genetic information in pricing but do not classify risks based on behavior. In addition, individuals at low genetic risk may also prefer not to insure classification risk when insurers use only genetic information meaning that complete market unraveling is still possible.

4.5 Welfare comparison

The previous analysis shows that the use of information in health insurance critically influences prevention behavior and insurance demand. Having the results from the previous sections in mind, we can now discuss the resulting welfare effects. For this purpose, I distinguish between genetic and behavioral information.

Behavioral information

The following proposition shows that using behavioral information in pricing always improves social welfare in a Pareto-sense regardless of whether insurers use genetic information or not. Appendix A.7 provides a proof.

Proposition 10. *The regulatory regimes can be Pareto-ranked with regard to the use of behavioral information in pricing. Only behavioral information Pareto-dominates no individual underwriting. Full information Pareto-dominates only genetic information.*

Using behavioral information in pricing is welfare-enhancing because it incentivizes insured individuals to engage in prevention, which reduces expected future medical expenses. Individuals benefit from reduced medical expenses because they can cover classification risk at a lower price. As a consequence, more individuals insure classification risk when insurers use behavioral information. Thus, both the price reduction and the efficiency improvement in risk allocation enhance social welfare.

Genetic information

The welfare implications of the use of genetic information are less clear-cut. It is not possible to Pareto-rank regimes depending on whether insurers use genetic information or not but the use of genetic information makes some individuals better and others worse off. We obtain the following proposition, for which Appendix A.8 provides a proof.

Proposition 11. *If the long-term health insurance market unravels when there is no individual underwriting, all individuals are equally well or better off when insurers use only genetic information than when there is no individual underwriting of long-term contracts. If the market outcome without any individual*

underwriting is characterized by a cutoff $z_0^{*,no}$, there exists a critical genetic disposition $z_0^{c,1} \geq z_0^{*,no}$ such that individuals with genetic disposition $z_0 < (=,>)\ z_0^{c,1}$ are equally well or better (equally well, worse) off when insurers use only genetic information than when there is no individual underwriting of long-term contracts.

Let $z_0^{*,beh}$ denote the lowest informationally consistent cutoff when insurers use only behavioral information in pricing. There exists a critical genetic disposition $z_0^{c,2} > z_0^{*,beh}$ such that individuals with genetic disposition $z_0 < (=,>)\ z_0^{c,2}$ are better (equally well, worse) off under full information than when insurers use only behavioral information.

When insurers do not use genetic information in pricing, everybody can insure classification risk at the same price and among the purchasers of the long-term contract individuals at low genetic risk subsidize individuals at high genetic risk. The subsidy makes long-term insurance unattractive for individuals at particularly low genetic risk. These individuals may benefit from the use of genetic information because it may allow them to insure classification risk at an attractive price. Among the individuals who purchase long-term insurance regardless of whether insurers use genetic information, the ones at relatively low genetic risk benefit from the use of genetic information because it reduces their prepayment while the ones at high genetic risk suffer from a higher price of insurance.

Under DD, prevention allows individuals at high genetic risk to partly offset their genetic disadvantage and expected future medical expenses are not necessarily monotonically increasing in z_0 when insurers use behavioral information in pricing. Nevertheless, there is always a critical genetic disposition that separates the profiteers and the sufferers of the use of genetic information because individuals at high genetic risk z_0 need higher expenditures on prevention to offset their genetic disadvantage.

Discussion

The use of behavioral information improves welfare because it incentivizes prevention and thus reduces the price of insurance. The use of genetic information has ambiguous welfare effects because it makes individuals at low genetic risk better and individuals at high genetic risk worse off. General welfare effects from models of pure moral hazard or pure adverse selection are therefore robust to the introduction of the other dimension of information.

Having a closer look at the underlying mechanisms, we see, however, that the two dimensions of information also interact with each other. On the one hand, when insurers do not use genetic information in pricing, banning the use of behavioral information not only leads to moral hazard but it also exacerbates adverse selection on genetic differences. Indeed, if insurers do not use behavioral information in pricing, insurance discourages prevention. As a consequence, long-

term contracts become more expensive and, hence, unattractive to individuals at low genetic risk. As these individuals leave the pool of insured, the price of long-term contracts increases further. Therefore, a ban on behavioral information makes both individuals who leave the pool of insured and individuals who purchase long-term insurance anyway worse off and such a ban may even cause market unraveling. This interaction effect implies that the use of behavioral information attenuates the inefficiency in risk allocation that results from a ban on the use of genetic information. It therefore reinforces the positive welfare effect of the use of behavioral information.

On the other hand, when insurers use behavioral information in pricing, a ban on the use of genetic information can not only lead to adverse selection but it may also limit the capability of the use of behavioral information to mitigate moral hazard. If the productivity of prevention depends on individuals' genetic disposition and insurers cannot tailor incentives for prevention to genetic risk, individuals may choose inefficient levels of prevention even if insurers use behavioral information. A ban on the use of genetic information may thus not only result in inefficient risk allocation but also in inefficient prevention behavior. This interaction effect attenuates (but never eliminates) the positive welfare effect of the use of behavioral information.

5 Resolving the equity-efficiency trade-off

The previous section has shown that the use of behavioral information is welfare-enhancing while regulation concerning the use of genetic information has ambiguous welfare effects. Since a ban on the use of genetic information redistributes expected wealth from individuals with higher expected wealth to individuals with lower expected wealth, it has desirable distributional consequences. However, banning the use of genetic information in a private health insurance market also creates inefficiencies in prevention behavior and risk allocation. In this section, I show how a social planner can resolve this equity-efficiency trade-off. Instead of only making a binary decision about whether to use genetic and behavioral information in pricing, a social planner can design a more nuanced contract that takes particular dimensions of either type of information into account and ignores others.

I now construct such a contract and relate it to the contracts offered in a regulated private health insurance market. Consider the prepayment

$$P^{soc}(e, z_0) = P^{fu}(e, z_0) + \mathbb{E}\left[EH(e^*(\tilde{z}_0); \tilde{z}_0)\right] - EH(e^*(z_0); z_0),$$

where the expectation is taken over the genetic disposition across the population. The genetic disposition over which the expectation is taken is denoted by $\tilde{z}_0$ to avoid confusion with the genetic disposition z_0 of the particular individual whose prepayment we consider. The first part of this prepayment is the fair prepayment under full information which covers the expected excess expenses

of individuals with genetic disposition z_0 who become a high-risk type later in life. The second part is a tax-subsidy scheme which offsets immutable differences in genetic risk. The tax/subsidy is the difference between the average expenditures across the population and the expected expenditures of individuals with genetic disposition z_0 when everybody chooses the level of prevention which minimizes their expected lifetime health expenditures. Individuals at low genetic risk subsidize individuals at high genetic risk which yields a form of "ex post" genetic insurance.

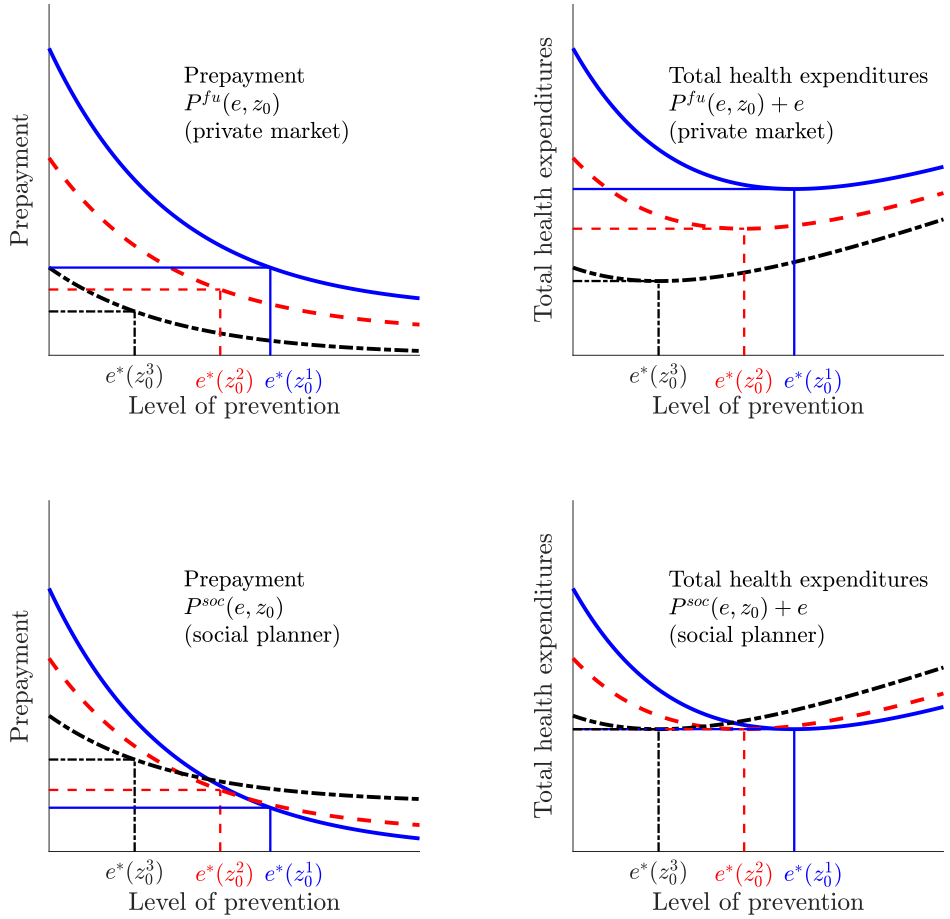
Figure 3 illustrates the differences between the contracts that private insurers offer under full information (first row) and the social planner contract (second row). The figure displays the prepayment (first column) and the total health expenditures at t_1 (consisting of the prepayment and the prevention expenditures, second column) of individuals with long-term health insurance as a function of their prevention expenditures. Since medical studies suggest that prevention technologies for many diseases, such as diabetes or cardiovascular disease, exhibit DD, I focus on this case. Under DD, individuals at high genetic risk can partly offset their genetic disadvantage because prevention is more productive for them. Hence, under full information, the difference between the prepayments for individuals with different genetic dispositions $z_0^1 > z_0^2 > z_0^3$ is smaller in equilibrium than it would be if they did not engage in prevention. If the difference in the productivity of prevention is large compared to the difference in exogenous genetic risk, the ordering of the prepayments may even be reversed in equilibrium. Nevertheless, individuals at high genetic risk always have higher total health expenditures than individuals at low genetic risk because they have higher expenditures on prevention.[12]

In the social planner contract, the tax-subsidy scheme works as if the social planner partially uses genetic information in pricing. If individuals do not engage in prevention, individuals at high genetic risk still have to make a higher prepayment than individuals at low genetic risk. However, the more productive prevention technology of individuals at high genetic risk reverses the ordering of the prepayments in equilibrium. As a consequence, in equilibrium, the total health expenditures consisting of the prepayment and the expenditures on prevention are the same for all individuals regardless of their genetic disposition.[13] These observations help to establish the following proposition, for which Appendix A.9 provides a proof.

[12]Formally, $\frac{d}{dz_0}P^{fu}(e^*, z_0) = \left(z_{z_0}(z_0, e^*) + z_e(z_0, e^*)\frac{de^*}{dz_0}\right)(P_H - P_L)$ which may be negative under DD since $z_{z_0} > 0$, $z_e < 0$, and $\frac{de^*}{dz_0} > 0$ in this case. For the total health expenditures at t_1, however, we obtain $\frac{d}{dz_0}\left(P^{fu}(e^*, z_0) + e^*\right) = z_{z_0}(z_0, e^*)(P_H - P_L) + (z_e(z_0, e^*)(P_H - P_L) + 1)\frac{de^*}{dz_0} = z_{z_0}(z_0, e^*)(P_H - P_L) > 0$ due to the first-order condition (4).

[13]In practice, such a contract could be implemented as follows: All individuals make the prepayment $P^{soc}(0, z_0)$ upfront which means that individuals at high genetic risk initially make a higher prepayment than individuals at low genetic risk although the difference is smaller than in the full information regime. When individuals engage in prevention, however, they get a reimbursement of $P^{soc}(0, z_0) - P^{soc}(e, z_0)$ which is higher for individuals at high genetic risk. For example, individuals at high genetic risk for breast cancer may more frequently be eligible for a reimbursement if they do a mammography screening than individuals at low genetic risk.

Figure 3: Private market with full information versus social planner



Notes: The figure displays the prepayment (first column) and the total health expenditures at t_1 (second column) as a function of individuals' prevention expenditures for different genetic dispositions $z_0^1 > z_0^2 > z_0^3$ in the DD case. The first row considers the contract that private insurers offer under full information, the second row considers the contract with the prepayment $P^{soc}(e, z_0)$ that a social planner can offer. With both contracts, individuals choose the efficient level of prevention. With the contract that private insurers offer under full information, individuals at high genetic risk have higher prevention expenditures and also have to make a higher prepayment. With the social planner contract, individuals at low genetic risk subsidize individuals at high genetic risk such that in equilibrium the total health expenditures of all insured individuals are the same regardless of their genetic disposition.

Proposition 12. *A social planner offering long-term health insurance with the prepayment $P^{soc}(e, z_0)$ together with a mandate to purchase this contract implements the welfare maximizing level of prevention and consumption stream at a balanced budget.*

Let $z_0^{*,beh}$ denote the lowest informationally consistent cutoff when insurers use only behavioral information in pricing. There exists some $z_0^{*,c} \leq z_0^{*,beh}$ such that the mandate is not necessary if the support of the genetic disposition is such that $z_0 \geq z_0^{*,c}$ for all individuals and private insurers are not allowed to use genetic information in pricing. If the prevention technology additionally exhibits CD, a ban on the use genetic information in pricing alone implements the welfare maximizing outcome in the private market.

Since the tax/subsidy in the prepayment P^{soc} does not depend on an individual's prevention expenditures, individuals purchasing the social planner contract choose the efficient level of prevention just as in the full information regime. As a consequence, if all individuals purchase the social planner contract, the contract breaks even and the tax/subsidy indeed offsets immutable differences in genetic risk and yields a consumption stream that neither depends on z_0 nor on an individual's risk type later in life. Hence, the social planner contract implements the welfare maximizing outcome at a balanced budget if all individuals purchase this contract (by choice or due to a mandate).

If the distribution of the genetic disposition z_0 is not too dispersed, a mandate is not necessary because the difference between the expected medical expenses of individuals at low genetic risk and the average expenses across the population is small. Therefore, even individuals at low genetic risk prefer the pooled long-term contract over leaving classification risk uninsured and purchasing a short-term contract later in life. If the prevention technology additionally exhibits CD, the prepayment does not need to be tailored to each individual's personal prevention technology because the productivity of prevention does not depend on an individual's genetic disposition. Hence, the contract offered by private insurers who must not use genetic information in pricing is the same contract as the one offered by a social planner. In conclusion, banning the use of genetic information in pricing is sufficient in this case to achieve a market outcome in which all individuals choose the efficient level of prevention, insure classification risk and bear the same share of the population's health expenditures.

Discussion

While banning or permitting the use of genetic information in insurance pricing has opposing equity and efficiency effects in general, a social planner can achieve the best of both worlds. A social planner can disentangle genetic factors determining insurmountable differences in expected medical expenses and genetic factors determining the productivity of prevention. Therefore, a long-term contract offered by a social planner can encourage insured individuals to choose the efficient level of prevention and still provide insurance against classification risk without disadvantaging individuals at high genetic risk.

For common multifactorial diseases such as diabetes or cardiovascular disease, genetic effects that have been identified so far are largely of moderate size (Qi et al., 2008; Said et al., 2018). In particular, individuals at low genetic risk still face a positive probability of disease meaning that the distribution of z_0 is bounded from below. Hence, a mandate may not be necessary for the social planner contract to be purchased by all individuals. Moreover, the interaction between genetic and behavioral risk factors is still far from being perfectly understood (Qi et al., 2008; Pomeroy et al., 2009). Therefore, known differences in the productivity of prevention are only small. In

conclusion, a ban on the use of genetic information may be sufficient to eliminate classification risk in the private market and to encourage individuals to choose the efficient level of prevention given the current state of medical knowledge. In the future, however, increased medical knowledge may improve the understanding of genetic risk factors and their interaction with behavioral risk factors. As a consequence, the distribution of genetic risk may be more dispersed and there may be more opportunities to individualize prevention measures based on genetic information. In this case, the intervention of a social planner may be necessary to avoid inefficient prevention behavior and tackle classification risk in health insurance.

6 Related literature and discussion

6.1 Risk classification

By discussing the pros and cons of using genetic and behavioral information in insurance pricing, this paper contributes to the risk classification literature. Hoy (1982) and Crocker and Snow (1986) study risk classification based on immutable characteristics like age, ethnicity, or gender, which are imperfectly correlated with risk. Hoy (1989) includes the possibility of prevention into the analysis. He analyzes the welfare implications of screening mechanisms matching individuals to their exogenous prevention technologies. However, he does not discuss the welfare implications of screening mechanisms that capture preventive behavior. So far, however, only few contributions explicitly discuss the use of behavioral information in pricing. Bond and Crocker (1991) as well as Polborn (2008) study the classification of risks based on the insured's consumption of products that are causally or statistically correlated with loss propensities and discuss the effects on consumer choices. In their settings, individuals have different preferences for a hazardous good and their consumption of the hazardous good serves as a signal about their exogenous risk type. I complement their work as prevention behavior does not serve as a signal about genetic risk in my model but I explicitly disentangle genetic and behavioral risk factors to discuss the use of either type of information in insurance pricing. Explicitly considering both genetic and behavioral risk factors allows to carve out the different nature of these two types of risk factors, to highlight how they interact with each other, and to derive important policy implications for insurance regulation.

Starting with the seminal works of Tabarrok (1994) and Doherty and Thistle (1996), the increasing availability of genetic tests has lead to an ongoing debate about the use of genetic information in insurance pricing. Similar to my setting, Polborn et al. (2006) discuss regulatory regimes for the use of genetic information when individuals learn their risk type over time. In their setting, however, individuals cannot influence their future risk type. Several papers compare regulatory regimes for the use of genetic test results in pricing when prevention is possible. Barigozzi and Henriet (2011) as well as Crainich (2017) assume that insurers observe preventive activities such

as medical checkups. Peter et al. (2017) assume that prevention is not observable which can be the case for lifestyle factors. Doherty and Posey (1998) as well as Hoel and Iversen (2002) compare testing and prevention behavior with symmetric information and when no underwriting takes place but do not discuss the effect of the exclusive use of either genetic or behavioral information. Bardey and De Donder (2013) study testing and prevention decisions depending on whether insurers observe the insured's prevention effort and discuss the resulting welfare implications. In general, the genetic testing literature assumes that risk types are only revealed if individuals take a genetic test. I complement this literature as I assume that individuals initially have some idea about their risk type and that their risk type develops over time and will be revealed later in life in any case. My model is in particular suited to describe the development of multifactorial diseases, such as diabetes, cancer, or cardiovascular disease. I contribute to the literature about the use of genetic information by paying particular attention to the interaction of genetic and behavioral risk factors.

6.2 Long-term health insurance

I add to the theoretical literature on long-term health insurance and classification risk by introducing prevention which allows individuals to improve their future health prospects. Long-term health insurance contracts have been proposed by Cochrane (1995) and Pauly et al. (1995) as a means to tackle classification risk. They show that long-term contracts fully eliminate classification risk in an ideal setting. However, dropping some of the assumptions for an ideal insurance market results in incomplete protection against classification risk (see, e.g., Frick, 1998; Peter et al., 2016; Hoy et al., 2021). In particular, Peter et al. (2016) demonstrate that long-term contracts insuring classification risk are no longer in demand by all individuals if they have some private information about their future risk type. To the best of my knowledge, the literature has assumed so far that the probability of becoming a high-risk type is exogenously given. For many diseases that lead to high expected healthcare spending over a longer period time, such as diabetes, cancer, or cardiovascular disease, however, the probability of disease onset depends on both genetic and behavioral risk factors. I extend the model of Peter et al. (2016) by disentangling genetic and behavioral factors and analyze how new technologies may help insurers overcome information asymmetries. By doing so, I show how the interaction of genetic and behavioral risk factors creates complications that may limit the effectiveness of long-term insurance in eliminating classification risk.

Long-term health insurance contracts are common and have been empirically investigated in Germany (Hofmann and Browne, 2013; Atal et al., 2021) and Chile (Atal, 2019). In the US, long-term contracts with prepayments exist in life insurance (Hendel and Lizzeri, 2003) and long-term care insurance (Finkelstein et al., 2005). Regarding health insurance in the US, Herring and Pauly

(2006) calculate the optimal insurance premium path that eliminates classification risk based on US medical expenditure data and show that the estimated path is close to the actual premium path in individual health insurance prior to the ACA. With the ACA coming effective in 2014, the US has taken a different approach to contend with classification risk in health insurance with managed competition in health exchanges featuring community rating and guaranteed issuance. Nevertheless, long-term contracts are still discussed as a potentially welfare-improving alternative to managed competition (see, Atal et al., 2021; Ghili et al., 2022).

6.3 Fixed contracts

My model of the health insurance market is based on the EFC framework developed by Einav et al. (2010). This supply-demand framework is in the spirit of the seminal work by Akerlof (1970). Insurers offer exogenously fixed full insurance contracts and only compete on price. The EFC framework has been widely used in analyzing selection markets, in particular for health insurance, and has also been extended by Mahoney and Weyl (2017) and Rothschild and Thistle (2022) to be applicable to a wide range of markets.[14]

Most of the risk classification literature is based on models assuming price and quantity competition in which insurers tackle adverse selection by offering self-selecting contract menus in the spirit of Rothschild and Stiglitz (1976).[15] Health insurance markets are often highly regulated and insurance contracts therefore highly standardized. Hence, the EFC framework is particularly well-suited to model health insurance markets. Assuming only price competition to study the role of asymmetric information in the context of classification risk and long-term contracts is in line with the approach of Peter et al. (2016), who also limit their analysis to premium schedules providing full insurance against classification risk. Handel et al. (2015) also study the trade-off between adverse selection and classification risk in a model based on the EFC framework when comparing regulatory regimes for health exchanges. The restriction to fixed full insurance contracts keeps the analysis tractable and allows to highlight key trade-offs when it comes to the use of genetic and behavioral information in health insurance pricing.

7 Conclusion

Technological and scientific progress over the past years has extended insurers' possibilities to gather and analyze large amounts of data for risk assessment. While most of the existing risk

[14] See Einav and Finkelstein (2023) for a recent survey of applications of the EFC framework.

[15] The underlying equilibrium concepts are the ones developed by Rothschild and Stiglitz (1976), Wilson (1977), Miyazaki (1977), and Spence (1978). For a survey of the risk classification literature based on self-selecting contract menus, see e.g. Dionne and Rothschild (2014).

classification literature focuses on the use of immutable characteristics, this paper explicitly disentangles exogenous genetic and endogenous behavioral risk factors and discusses the use of either type of information in health insurance pricing. Explicitly considering both genetic and behavioral information allows to account for the different nature of the two types of risk factors and to carve out how they interact with each other depending on the regulation in place.

I show that the use of behavioral information is always welfare-enhancing whereas the use of genetic information has ambiguous welfare effects. If the distribution of genetic risk is not too dispersed and the productivity of prevention does not depend on individuals' genetic disposition, banning the use of genetic information and promoting the use of behavioral information implements the welfare maximizing outcome. In this case, the use of behavioral information in insurance pricing eliminates both moral hazard and adverse selection. If the productivity of prevention varies between individuals, however, regulators face an equity-efficiency trade-off. A social planner can resolve the trade-off and offer a contract that takes the productivity of each individual's prevention technology into account and simultaneously establishes a tax-subsidy scheme to offset immutable genetic differences.

To implement contracts that only use behavioral information, insurers and regulators need to disentangle genetic and behavioral risk factors. In practice, it may be rather difficult to decide whether behavior results from someone's genetic disposition or from their voluntary choice. For example, it may be more difficult for individuals predisposed to disease to do sports in order to improve their health prospects. Therefore, it is important to appropriately define expenditures on prevention and identify prevention technologies based on the current state of medical research. There are already some cautious approaches to include lifestyle factors in insurance pricing. In the US, the ACA allows to impose a surcharge on tobacco users' premiums. Moreover, some employer-sponsored health insurance contracts offer workplace wellness programs that reward gym visits with insurance premium discounts. Bonus programs in German statutory health insurance reward health-promoting activities, such as joining a fitness class or attending medical checkups, by subsidizing course fees or paying out a cash bonus at the end of the year. Although empirical evidence is mixed, several studies suggest that financial incentives have the potential to encourage healthy behaviors (see, e.g., Sutherland et al., 2008; Charness and Gneezy, 2009; Royer et al., 2015). Up until now, however, contracts using behavioral information are not much in demand.[16] This paper shows that the use of behavioral information in health insurance pricing can help to reduce healthcare spending due to modifiable risk factors and to provide affordable health insurance coverage. Therefore, an interesting avenue for future research is to further study why demand for contracts based on behavioral information is low in order find ways to promote such contracts.

[16]Such a lack of acceptance not only plays a role in health insurance but also in other lines of insurance. For example, take-up rates of usage-based auto insurance are quite low although risk-based pricing is widely accepted in auto insurance in general. Potential explanations that have been discussed in the literature are privacy concerns (Gemmo et al., 2019; Biener et al., 2020) or insufficient classification accuracy (Holzapfel et al., 2023).

This paper focuses on the financial consequences of health losses which is in line with most of the existing literature on both risk classification and long-term health insurance. Diseased individuals, however, may not only suffer from the monetary losses resulting from treatment costs but also from the disease itself (Cook and Graham, 1977). Therefore, a state-dependent utility framework may be an interesting extension of the model discussed in this paper. Moreover, loading factors and heterogeneous preferences may yield ambiguous welfare implications of an insurance mandate and make public policy implications less clear-cut (Einav and Finkelstein, 2011). These extensions may affect the welfare implications of a social planner intervention and, hence, rather speak in favor of less severe market interventions in form of the regulatory restrictions on the use of information in a private market. Accounting for such effects therefore constitutes another interesting avenue for future research.

References

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics 84(3), 488–500.

Atal, J. P. (2019). Lock-in in dynamic health insurance contracts: Evidence from Chile. Available at SSRN 3485007.

Atal, J. P., H. Fang, M. Karlsson, and N. R. Ziebarth (2021). Long-term health insurance: Theory meets evidence. Available at SSRN 3559159.

Bardey, D. and P. De Donder (2013). Genetic testing with primary prevention and moral hazard. Journal of Health Economics 32(5), 768–779.

Barigozzi, F. and D. Henriet (2011). Genetic information: Comparing alternative regulatory approaches when prevention matters. Journal of Public Economic Theory 13(1), 23–46.

Biener, C., M. Eling, and M. Lehmann (2020). Balancing the desire for privacy against the desire to hedge risk. Journal of Economic Behavior & Organization 180, 608–620.

Bolnick, H. J., A. L. Bui, A. Bulchis, C. Chen, A. Chapin, L. Lomsadze, A. H. Mokdad, F. Millard, and J. L. Dieleman (2020). Health-care spending attributable to modifiable risk factors in the USA: An economic attribution analysis. The Lancet Public Health 5(10), e525–e535.

Bond, E. W. and K. J. Crocker (1991). Smoking, skydiving, and knitting: The endogenous categorization of risks in insurance markets with asymmetric information. Journal of Political Economy 99(1), 177–200.

Charness, G. and U. Gneezy (2009). Incentives to exercise. Econometrica 77(3), 909–931.

Cochrane, J. H. (1995). Time-consistent health insurance. Journal of Political Economy 103(3), 445–473.

Cook, P. J. and D. A. Graham (1977). The demand for insurance and protection: The case of irreplaceable commodities. The Quarterly Journal of Economics 91(1), 143–156.

Crainich, D. (2017). Self-insurance with genetic testing tools. Journal of Risk and Insurance 84(1), 73–94.

Crocker, K. J. and A. Snow (1986). The efficiency effects of categorical discrimination in the insurance industry. Journal of Political Economy 94(2), 321–344.

Dionne, G. and C. Rothschild (2014). Economic effects of risk classification bans. The Geneva Risk and Insurance Review 39(2), 184–221.

Doherty, N. A. and L. L. Posey (1998). On the value of a checkup: Adverse selection, moral hazard and the value of information. Journal of Risk and Insurance 65(2), 189–211.

Doherty, N. A. and P. D. Thistle (1996). Adverse selection with endogenous information in insurance markets. Journal of Public Economics 63(1), 83–102.

Einav, L. and A. Finkelstein (2011). Selection in insurance markets: Theory and empirics in pictures. Journal of Economic Perspectives 25(1), 115–138.

Einav, L. and A. Finkelstein (2023). Empirical analyses of selection and welfare in insurance markets: a self-indulgent survey. The Geneva Risk and Insurance Review 48(2), 167–191.

Einav, L., A. Finkelstein, and M. R. Cullen (2010). Estimating welfare in insurance markets using variation in prices. The Quarterly Journal of Economics 125(3), 877–921.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on moral hazard in health insurance. American Economic Review 103(1), 178–219.

Finkelstein, A., K. McGarry, and A. Sufi (2005). Dynamic inefficiencies in insurance markets: Evidence from long-term care insurance. American Economic Review 95(2), 224–228.

Frick, K. (1998). Consumer capital market constraints and guaranteed renewable insurance. Journal of Risk and Uncertainty 16(3), 271–278.

Gemmo, I., M. J. Browne, and H. Gründl (2019). Privacy concerns in insurance markets: Implications for market equilibria and social welfare. Available at SSRN 3480629.

Ghili, S., B. Handel, I. Hendel, and M. D. Whinston (2022). Optimal long-term health insurance contracts: characterization, computation, and welfare effects. Review of Economic Studies (forthcoming).

Gollier, C. (2001). The economics of risk and time. Cambridge, MA: The MIT Press.

Handel, B., I. Hendel, and M. D. Whinston (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. Econometrica 83(4), 1261–1313.

Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. Journal of Political Economy 61(5), 434–435.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. Journal of Political Economy 63(4), 309–321.

Hendel, I. and A. Lizzeri (2003). The role of commitment in dynamic contracts: Evidence from life insurance. The Quarterly Journal of Economics 118(1), 299–328.

Herring, B. and M. V. Pauly (2006). Incentive-compatible guaranteed renewable health insurance premiums. Journal of Health Economics 25(3), 395–417.

Hoel, M. and T. Iversen (2002). Genetic testing when there is a mix of compulsory and voluntary health insurance. Journal of Health Economics 21(2), 253–270.

Hofmann, A. and M. Browne (2013). One-sided commitment in dynamic insurance contracts: Evidence from private health insurance in Germany. Journal of Risk and Uncertainty 46(1), 81–112.

Holzapfel, J., R. Peter, and A. Richter (2023). Mitigating moral hazard with usage-based insurance. Journal of Risk and Insurance (forthcoming).

Hoy, M. (1982). Categorizing risks in the insurance industry. The Quarterly Journal of Economics 97(2), 321–336.

Hoy, M. (1989). The value of screening mechanisms under alternative insurance possibilities. Journal of Public Economics 39(2), 177–206.

Hoy, M. (2006). Risk classification and social welfare. The Geneva Papers on Risk and Insurance - Issues and Practice 31(2), 245–269.

Hoy, M., A. Mirza, and A. Sadanand (2021). Guaranteed renewable life insurance under demand uncertainty. Journal of Risk and Insurance 88(1), 131–159.

Jack, W. (2002). Equilibrium in competitive insurance markets with ex ante adverse selection and ex post moral hazard. Journal of Public Economics 84(2), 251–278.

Mahoney, N. and E. G. Weyl (2017). Imperfect competition in selection markets. Review of Economics and Statistics 99(4), 637–651.

Miyazaki, H. (1977). The rat race and internal labor markets. The Bell Journal of Economics 8(2), 394–418.

Mossin, J. (1968). Aspects of rational insurance purchasing. Journal of Political Economy 76(4, Part 1), 553–568.

Pauly, M. V., H. Kunreuther, and R. Hirth (1995). Guaranteed renewability in insurance. Journal of Risk and Uncertainty 10(2), 143–156.

Peter, R., A. Richter, and P. Steinorth (2016). Yes, no, perhaps? Premium risk and guaranteed renewable insurance contracts with heterogeneous incomplete private information. Journal of Risk and Insurance 83(2), 363–385.

Peter, R., A. Richter, and P. D. Thistle (2017). Endogenous information, adverse selection, and prevention: Implications for genetic testing policy. Journal of Health Economics 55, 95–107.

Polborn, M. K. (2008). Endogenous categorization in insurance. Journal of Public Economic Theory 10(6), 1095–1113.

Polborn, M. K., M. Hoy, and A. Sadanand (2006). Advantageous effects of regulatory adverse selection in the life insurance market. The Economic Journal 116(508), 327–354.

Pomeroy, J., A. M. Söderberg, and P. W. Franks (2009). Gene-lifestyle interactions and their consequences on human health. In Genetics and Sports, pp. 110–135. Karger Publishers.

Qi, L., F. B. Hu, and G. Hu (2008). Genes, environment, and interactions in prevention of type 2 diabetes: A focus on physical activity and lifestyle changes. Current Molecular Medicine 8(6), 519–532.

Rothschild, C. and P. D. Thistle (2022). Supply, demand, and selection in insurance markets: Theory and applications in pictures. Risk Management and Insurance Review 25(4), 419–444.

Rothschild, M. and J. Stiglitz (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. The Quarterly Journal of Economics 90(4), 629–649.

Royer, H., M. Stehr, and J. Sydnor (2015). Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a fortune-500 company. American Economic Journal: Applied Economics 7(3), 51–84.

Said, M. A., N. Verweij, and P. van der Harst (2018). Associations of combined genetic and lifestyle risks with incident cardiovascular disease and diabetes in the UK Biobank Study. JAMA Cardiology 3(8), 693–702.

Shavell, S. (1979). On moral hazard and insurance. The Quarterly Journal of Economics 93(4), 541–562.

Shook, R. P., D.-c. Lee, X. Sui, V. Prasad, S. P. Hooker, T. S. Church, and S. N. Blair (2012). Cardiorespiratory fitness reduces the risk of incident hypertension associated with a parental history of hypertension. Hypertension 59(6), 1220–1224.

Spence, M. (1978). Product differentiation and performance in insurance markets. Journal of Public Economics 10(3), 427–447.

Stewart, J. (1994). The welfare implications of moral hazard and adverse selection in competitive insurance markets. Economic Inquiry 32(2), 193–208.

Sutherland, K., J. B. Christianson, and S. Leatherman (2008). Impact of targeted financial incentives on personal health behavior. Medical Care Research and Review 65(6_suppl), 36S–78S.

Swiss Re (2021). Short-term health insurance: a growth engine for China's P&C insurers. Available online at: https://www.swissre.com/institute/research/sigma-research/Economic-Insights/short-term-health-insurance-china.html.

Tabarrok, A. (1994). Genetic testing: an economic and contractarian analysis. Journal of Health Economics 13(1), 75–91.

Wilson, C. (1977). A model of insurance markets with incomplete information. Journal of Economic Theory 16(2), 167–207.

A Mathematical proofs

A.1 Proof of Proposition 1

I have already shown that the efficient level of prevention minimizes each individual's expected lifetime health expenditures. To compare the prevention expenditures of different individuals depending on their genetic disposition, I apply the implicit function theorem on the first-order condition (1) which yields

$$\frac{de^*}{dz_0} = -\frac{z_{ez_0}(z_0, e^*)}{z_{ee}(z_0, e^*)}.$$

Since $z_{ee} > 0$, the sign of this expression solely depends on the sign of the cross-derivative z_{ez_0} which yields the second statement. Finally, the statement about the welfare-maximizing consumption stream follows directly from individuals' risk aversion.

A.2 Proof of Proposition 2

To compare the optimal levels of prevention with long- and short-term contracts, I use the following Lemma from Gollier (2001, p. 151).

Lemma 1. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a concave function in the variables (e, s), that is, $f_{ee} < 0$ and $f_{ee}f_{ss} - f_{es}^2 > 0$, which is maximal at (e^*, s^*). Let $\bar{e} \in \mathbb{R}$ be a value we want to compare e^* with. Then, $e^* > \bar{e}$ if and only if $f_e(\bar{e}, \hat{s}) > 0$, where $\hat{s}$ is the value that maximizes $f(\bar{e}, s)$.*

When insurers use both genetic and behavioral information,

$$P_{ee}^{fu}(e, z_0) = z_{ee}(e, z_0)(P_H - P_L) > 0.$$

Hence, the second-order conditions are globally satisfied for expected utility with long-term insurance under full information which I denote by $EU^{fu}(e, s; z_0)$ (see Appendix B.1). Therefore, I can apply Lemma 1 to compare optimal prevention with long- and short-term insurance. Let $\hat{s}$ be the level of saving that maximizes $EU^{fu}\left(e^{ST}, s; z_0\right)$. $\hat{s}$ solves the first-order condition

$$EU_s^{fu}\left(e^{ST}, \hat{s}; z_0\right) = -u'\left(w_1 - e^{ST} - P^{fu}\left(e^{ST}, z_0\right) - \hat{s}\right) + v'\left(w_2 - P_L + \hat{s}\right) = 0.$$

According to Lemma 1, $e^{fu} > e^{ST}$ if and only if $EU_e^{fu}\left(e^{ST}, \hat{s}; z_0\right) > 0$, which holds if and only if $P_e^{fu}\left(e^{ST}, z_0\right) + 1 < 0$. We have

$$
\begin{aligned}
P_e^{fu}\left(e^{ST}, z_0\right) &= z_e(z_0, e^{ST})(P_H - P_L) \\
&= \frac{-u'(w_1 - e^{ST} - s^{ST})}{v(w_2 - P_L + s^{ST}) - v(w_2 - P_H + s^{ST})}(P_H - P_L) \\
&= -\frac{z(z_0, e^{ST})v'(w_2 - P_H + s^{ST}) + (1 - z(z_0, e^{ST}))v'(w_2 - P_L + s^{ST})}{v(w_2 - P_L + s^{ST}) - v(w_2 - P_H + s^{ST})}(P_H - P_L),
\end{aligned}
$$

where I used $EU_e^{ST}\left(e^{ST}, s^{ST}; z_0\right) = 0$ to obtain the second line and $EU_s^{ST}\left(e^{ST}, s^{ST}; z_0\right) = 0$ to obtain the third line. Hence, $P_e^{fu}\left(e^{ST}, z_0\right) + 1 < 0$ if and only if

$$
\begin{aligned}
z\left(z_0, e^{ST}\right) v'\left(w_2 - P_H + s^{ST}\right) + \left(1 - z\left(z_0, e^{ST}\right)\right) v'\left(w_2 - P_L + s^{ST}\right) \\
> \frac{v\left(w_2 - P_L + s^{ST}\right) - v\left(w_2 - P_H + s^{ST}\right)}{P_H - P_L}.
\end{aligned}
$$

The left-hand side is a convex combination of the slope of v at the wealth levels $w_2 - P_H + s^{ST}$ and $w_2 - P_L + s^{ST}$. The right-hand side represents the slope of the secant line between these two wealth levels. Due to the concavity of v, $v'\left(w_2 - P_H + s^{ST}\right) > \frac{v\left(w_2 - P_L + s^{ST}\right) - v\left(w_2 - P_H + s^{ST}\right)}{P_H - P_L} > v'\left(w_2 - P_L + s^{ST}\right)$. Let $z^c \in (0,1)$ such that $z^c v'\left(w_2 - P_H + s^{ST}\right) + (1 - z^c)\, v'\left(w_2 - P_L + s^{ST}\right) = \frac{v\left(w_2 - P_L + s^{ST}\right) - v\left(w_2 - P_H + s^{ST}\right)}{P_H - P_L}$. Then, $P_e^{fu}\left(e^{ST}, z_0\right) + 1 < 0$ if and only if $z\left(z_0, e^{ST}\right) > z^c$.

A.3 Proof of Proposition 3

For an individual with genetic disposition z_0, it holds that

$$
\begin{aligned}
EU^{ST}\left(e^{ST}, s^{ST}; z_0\right) &= u\left(w_1 - e^{ST} - s^{ST}\right) + \quad z\left(z_0, e^{ST}\right)\ v\left(w_2 - P_H + s^{ST}\right) \\
&\qquad + \left(1 - z\left(z_0, e^{ST}\right)\right)\ v\left(w_2 - P_L + s^{ST}\right) \\
&\leq u\left(w_1 - e^{ST} - s^{ST}\right) + v\left(w_2 - z\left(z_0, e^{ST}\right) P_H - \left(1 - z\left(z_0, e^{ST}\right)\right) P_L + s^{ST}\right) \\
&= u\left(w_1 - e^{ST} - P^{fu}\left(e^{ST}, z_0\right) - \tilde{s}\right) + v\left(w_2 - P_L + \tilde{s}\right) \\
&= EU^{fu}\left(e^{ST}, \tilde{s}; z_0\right),
\end{aligned}
$$

where I used the concavity of v to obtain the inequality in the second line and defined $\tilde{s} := s^{ST} - z(z_0, e^{ST})(P_H - P_L) = s^{ST} - P^{fu}(e^{ST}, z_0)$ to obtain the third line. Since $\left(e^{fu}, s^{fu}\right)$ maximizes $EU^{fu}(e, s; z_0)$, this yields $EU^{ST}\left(e^{ST}, s^{ST}; z_0\right) \leq EU^{fu}\left(e^{fu}, s^{fu}; z_0\right)$ for all z_0 and all individuals (weakly) prefer long- over short-term insurance. As I assume an interior solution $e^{ST} > 0$ for all $z_0 \in (0, 1]$, it holds that $z(z_0, e^{ST}) \in (0, 1)$ and the inequality in the second line is strict for all $z_0 \in (0, 1]$. That is, only certain low risks with genetic disposition $z_0 = 0$ are indifferent between short- and long-term insurance and all others strictly prefer long-term insurance over short-term insurance.

A.4 Proof of Proposition 5

We first apply the envelope theorem to formally show that expected utility with short-term contracts is decreasing in the genetic disposition z_0. Indeed,

$$\frac{dEU^{ST}}{dz_0} = z_{z_0}(z_0, e^{ST}) \left(v(w_2 - P_H + s^{ST}) - v(w_2 - P_L + s^{ST}) \right) < 0.$$

Since expected utility with a long-term contract and no individual underwriting does not depend on z_0, this implies that for every long-term contract there is a cutoff z_0^* such that individuals with genetic disposition $z_0 \geq z_0^*$ prefer the long-term contract and individuals with genetic disposition $z_0 < z_0^*$ prefer short-term contracts.

In the following, I denote expected utility under the long-term contract with the prepayment $P^{no}(z_0^*)$ by $EU^{no}(e, s; z_0^*)$. We know that $EU^{ST}(e^{ST}, s^{ST}; z_0 = z_0^*) > EU^{no}(e^{no}, s^{no}; z_0^*)$ for the potential cutoffs $z_0^* \in \{0, 1\}$. If $EU^{ST}(e^{ST}, s^{ST}; z_0 = z_0^*) > EU^{no}(e^{no}, s^{no}; z_0^*)$ for all potential cutoffs $z_0^* \in [0, 1]$, no long-term contract that would be in demand makes non-negative profits. Hence, long-term insurance is not in demand. If there exists at least one informationally consistent cutoff $z_0^* \in (0, 1)$ such that $EU^{ST}(e^{ST}, s^{ST}; z_0 = z_0^*) = EU^{no}(e^{no}, s^{no}; z_0^*)$, individuals with genetic disposition $z_0 \geq z_0^*$ prefer the long-term contract with the prepayment $P^{no}(z_0^*)$ over short-term insurance and long-term insurance is in demand.

To show that the long-term contract with the lowest informationally consistent cutoff constitutes the Nash equilibrium, I show that it fulfills the following two criteria.

1. The contract makes non-negative expected profits.
2. There is no contract outside the equilibrium set that, if offered, will make a non-negative profit.

A long-term contract based on an informationally consistent cutoff makes zero profits by definition. Moreover, long-term contracts which are priced based on an informationally consistent cutoff are the only long-term contracts which make zero profit. Hence, no other long-term contract can be in the equilibrium set. Indeed, for a contract that would yield positive profits, there always exists a contract with a lower prepayment that makes zero profit (by continuity since a contract with a prepayment equal to zero always yields negative profits). Since all individuals unanimously prefer a contract with a lower prepayment, the contract with positive profits cannot be in demand in equilibrium.

In general, there may be several informationally consistent cutoffs $z_0^* \in (0, 1)$ fulfilling $EU^{ST}(e^{ST}, s^{ST}; z_0 = z_0^*) = EU^{no}(e^{no}, s^{no}; z_0^*)$. The envelope theorem yields

$$\frac{dEU^{no}}{dz_0^*} = u'\left(w_1 - P^{no}\left(z_0^*\right) - s^{no}\right)\left(-P_{z_0^*}^{no}\left(z_0^*\right)\right).$$

Moreover, denoting the density of the distribution of the genetic disposition z_0 by f and its cumulative distribution function by F,

$$\begin{aligned} P_{z_0^*}^{no}(z_0^*) &= \frac{d}{dz_0^*}\left(\frac{\int_{z_0^*}^1 z_0 f(z_0)\, dz_0}{1 - F(z_0^*)}\right)(P_H - P_L) \\ &= \frac{-z_0^* f(z_0^*)(1 - F(z_0^*)) + f(z_0^*)\int_{z_0^*}^1 z_0 f(z_0)\, dz_0}{(1 - F(z_0^*))^2}(P_H - P_L). \end{aligned}$$

Since $\int_{z_0^*}^1 z_0 f(z_0)\, dz_0 > z_0^*(1 - F(z_0^*))$, it holds that $P_{z_0^*}^{no} > 0$ and, thus, $\frac{dEU^{no}}{dz_0^*} < 0$. Hence, if several informationally consistent cutoffs exist and the corresponding long-term contracts are offered simultaneously in the market, all individuals choose the contract with the lowest cutoff because it yields the highest expected utility. Consequently, the long-term contract based on the lowest informationally consistent cutoff is the only contract fulfilling 1. and 2. and, hence, constitutes the Nash equilibrium.

A.5 Proof of Propositions 6 and 7

To show that the long-term contract with the lowest informationally consistent cutoff constitutes the Nash equilibrium, I show that it fulfills the following two criteria.

1. The contract makes non-negative expected profits.
2. There is no contract outside the equilibrium set that, if offered, will make a non-negative profit.

A long-term contract based on an informationally consistent cutoff makes zero profits by definition. Moreover, contracts which are priced based on an informationally consistent cutoff are the only long-term contracts which make zero profit. Hence, no other long-term contract can be in the equilibrium set. Indeed, for a contract that would yield positive profits, there always exists a contract with a lower prepayment and the same prevention expenditure that makes zero profit (by continuity since a contract with a prepayment equal to zero always yields negative profits). Since all individuals unanimously prefer a contract with a lower prepayment, the contract with positive profits cannot be in demand in equilibrium.

In general, there may be several contracts $(P^{beh}(e; z_0^*), e)$ with informationally consistent cutoffs $z_0^* \in (0, 1)$. By definition of the cutoff z_0^*, expected utility with the long-term contract $(P^{beh}(e; z_0^*), e)$ (which is the same for all individuals) equals expected utility with a short-term contract of indi-

viduals with genetic disposition z_0^*. Therefore, since expected utility with a short-term contract is decreasing in z_0, the contract with the smallest cutoff z_0^* maximizes expected utility. Formally,

$$\frac{dEU^{beh}}{dz_0^*} = \frac{dEU^{ST}}{dz_0}\bigg|_{z_0=z_0^*} = z_{z_0}(z_0^*, e^{ST})\left(v(w_2 - P_H + s^{ST}) - v(w_2 - P_L + s^{ST})\right) < 0,$$

where I used the envelope theorem to calculate $\frac{dEU^{ST}}{dz_0}$. Hence, if several informationally consistent cutoffs exist and the corresponding long-term contracts are offered simultaneously in the market, all individuals choose the contract with the lowest cutoff because it yields the highest expected utility (regardless of the levels of prevention associated with the different contracts). Consequently, the long-term contract based on the lowest informationally consistent cutoff is the only contract fulfilling 1. and 2. and, hence, constitutes the Nash equilibrium.

To characterize the associated level of prevention, let $z_0^*(e)$ denote the smallest informationally consistent cutoff for each level of prevention e and set $z_0^*(e) = 1$ if no such cutoff exists. Then, the optimal contract from the set of zero-profit contracts $\mathcal{C}^{beh}$ is the contract maximizing $EU^{beh}(e, s; z_0^*(e))$. We can restrict our attention to the smallest cutoff for each level of prevention without loss of generality since the globally smallest cutoff must also be the smallest cutoff for the associated level of prevention. Since the cutoff $z_0^*(e)$ depends on the level of prevention e, the general first-order condition (3) translates to

$$P_e^{beh}(e; z_0^*(e)) + P_{z_0^*}^{beh}(e; z_0^*(e))\,\frac{\mathrm{d}z_0^*}{\mathrm{d}e} = -1.$$

The equilibrium cutoff $z_0^{*,beh}$ is the smallest informationally consistent cutoff in $\mathcal{C}^{beh}$. Hence, $z_0^{*,beh}$ is a minimum of $z_0^*(e)$ and, thus, $\frac{\mathrm{d}z_0^*}{\mathrm{d}e} = 0$ when $z_0^*(e) = z_0^{*,beh}$. Therefore, the first-order condition is fulfilled if and only if $P_e^{beh} = -1$, which yields

$$P_e^{beh}(e^{beh}; z_0^{*,beh}) = \mathbb{E}\left[\, z_e\left(z_0, e^{beh}\right) \;\middle|\; z_0 \geq z_0^{*,beh}\right](P_H - P_L) = -1. \tag{A.1}$$

In the full information regime, individuals with a long-term contract minimize their personal expected health expenditures. Comparing (A.1) to (1), we see that optimal prevention minimizes the expected health expenditures of an "average insured" when insurers use only behavioral information.

For the comparison of the cutoffs with no individual underwriting and when insurers use behavioral information, we first fix the cutoff $z_0^{*,no}$ and level of saving s^{no} with no underwriting. Let $\hat{e}$ be the level of prevention that minimizes $P^{beh}(e; z_0^{*,no}) + e$. That is, $\hat{e}$ fulfills the first-order condition

$$P_e^{beh}(\hat{e}; z_0^{*,no}) = \mathbb{E}\left[\, z_e\left(z_0, \hat{e}\right) \mid z_0 \geq z_0^{*,no}\right](P_H - P_L) = -1.^{17}$$

[17] The second-order condition is globally fulfilled since $P_{ee}^{beh}(e; z_0^{*,no}) = \mathbb{E}\left[\, z_{ee}\left(z_0, e\right) \mid z_0 \geq z_0^{*,no}\right](P_H - P_L) > 0$.

Since $-z_e(z_0, 0) > \frac{1}{P_H - P_L}$ for all $z_0 \in (0, 1]$, it holds that $\hat{e} > 0$. Then,

$$\begin{aligned} EU^{beh}\left(\hat{e}, s^{no}; z_0^{*,no}\right) &= u\left(w_1 - \hat{e} - P^{beh}(\hat{e}; z_0^{*,no}) - s^{no}\right) + v\left(w_2 - P_L + s^{no}\right) \\ &> u\left(w_1 - P^{no}(z_0^{*,no}) - s^{no}\right) + v\left(w_2 - P_L + s^{no}\right) \\ &= EU^{no}\left(e^{no}, s^{no}; z_0^{*,no}\right), \end{aligned} \tag{A.2}$$

where I used $P^{beh}(e = 0; z_0^{*,no}) = P^{no}(z_0^{*,no})$ as well as $e^{no} = 0$. Since a cutoff individual with genetic disposition $z_0^{*,no}$ is indifferent between purchasing long-term insurance without any individual underwriting and not insuring classification risk, above inequality implies that this individual strictly prefers long-term insurance requiring prevention expenditures of $\hat{e}$ and aiming at the cutoff $z_0^{*,no}$ over short-term insurance. Hence, the lowest informationally consistent cutoff $z_0^*(\hat{e})$ with this level of prevention, and a fortiori the smallest informationally cutoff overall $z_0^{*,beh}$ must be smaller than $z_0^{*,no}$.

A.6 Proof of Proposition 9

We have already seen that individuals with genetic disposition $z_0 = 0$ are indifferent between short- and long-term insurance whereas individuals with genetic disposition $z_0 = 1$ strictly prefer short- over long-term insurance. The formal argument for the genetic disposition $z_0 = 1$ is the same as in the regime without any individual underwriting (see Appendix B.2).

Since expected utility with both short- and long-term insurance is continuous in the genetic disposition z_0, $EU^{ST} > EU^{LT}$ for $z_0 = 1$ implies that there exists some $\hat{z}_0 < 1$ such that the same inequality holds for all $z_0 > \hat{z}_0$. Hence, individuals at high genetic risk (z_0 close to 1) never purchase long-term health insurance. If $EU^{ST} \geq EU^{LT}$ for all $z_0 \in [0, 1]$, all individuals prefer short- over long-term insurance and long-term health insurance is not in demand. If there exist some $z_0 \in (0, 1)$ such that $EU^{ST} < EU^{LT}$, individuals with these genetic dispositions purchase a long-term contract meaning that long-term health insurance is in demand.

A.7 Proof of Proposition 10

We start with the welfare comparison when insurers do not use genetic information in pricing. That is, we compare the regime without any individual underwriting to the regime in which insurers use only behavioral information. We know from Proposition 6 that the lowest informationally consistent cutoff when insurers use only behavioral information, $z_0^{*,beh}$, is smaller than the one without any individual underwriting, $z_0^{*,no}$. We can now distinguish three groups of individuals. First, the purchasers of both types of long-term contracts are better off when insurers use behavioral information. Indeed, since $z_0^{*,beh} < z_0^{*,no}$ and $\frac{\mathrm{d}EU^{ST}}{\mathrm{d}z_0} < 0$,

$$EU^{beh} = EU^{ST}\left(e^{ST}, s^{ST}; z_0^{*,beh}\right) > EU^{ST}\left(e^{ST}, s^{ST}; z_0^{*,no}\right) = EU^{no},$$

where I used that individuals with genetic disposition equal to the cutoff are indifferent between short- and long-term insurance. Second, since $z_0^{*,beh} < z_0^{*,no}$, there are some individuals who switch from short-term insurance to long-term insurance when insurers start to use behavioral information. Since these switchers could have stayed with the same short-term contract, they must also be better off when insurers use behavioral information. Finally, the expected utility of those who purchase short-term insurance anyway does not change as the long-term contract changes. In conclusion, some individuals are better off and nobody is worse off when insurers use behavioral information.

To analyze the welfare effect of the use of behavioral information when insurers use genetic information in pricing, we compare the regime in which insurers use only genetic information to the regime with full information. Since all individuals purchase the long-term contract under full information while some or all individuals purchase short-term insurance when insurers use only genetic information in pricing, we can distinguish two groups of individuals. The ones who purchase long-term insurance anyway are better of under full information. Indeed, it holds that $P^{gen}(z_0) = P^{fu}(e = 0, z_0)$. Hence,

$$EU^{gen}(e^{gen}, s^{gen}; z_0) = EU^{fu}(0, s^{gen}; z_0) < EU^{fu}(e^{fu}, s^{fu}; z_0)$$

since $e^{fu} > 0$ for all $z_0 \in (0, 1]$. Individuals who only purchase long-term insurance when insurers use behavioral information must also benefit from the use of behavioral information because they could have stayed with the same short-term contract. In conclusion, when insurers use behavioral information, individuals with genetic disposition $z_0 \in (0, 1]$ are better off and individuals with genetic disposition $z_0 = 0$ are equally well off because they do not engage in prevention anyway.

A.8 Proof of Proposition 11

I start with the welfare comparison when insurers do not use behavioral information in pricing. That is, we compare the regime without any individual underwriting to the regime in which insurers use only genetic information. If the market unravels when there is no individual underwriting, all individuals purchase short-term insurance in this case. They may stay with the same short-term contract or switch to a long-term contract when insurers use genetic information. Hence, all individuals are equally well or better of when insurers use genetic information.

If the equilibrium without any individual underwriting is characterized by the cutoff $z_0^{*,no}$, define $z_0^{c,1} := \mathbb{E}\left[z_0 \mid z_0 \geq z_0^{*,no}\right]$. Then,

$$P^{no}(z_0^{*,no}) = z_0^{c,1}(P_H - P_L) > (=, <)\ z_0(P_H - P_L) = P^{gen}(z_0) \quad \text{for} \quad z_0 < (=, >)\ z_0^{c,1}.$$

We can now distinguish four groups. Individuals with genetic disposition $z_0 < z_0^{*,no}$ purchase short-term insurance when there is no individual underwriting. These individuals can stay with the same short-term contract or switch to a long-term contract when insurers use only genetic information. Therefore, they are equally well or better off when insurers use genetic information. Individuals with genetic disposition $z_0^{*,no} \leq z_0 < z_0^{c,1}$ purchase long-term insurance when there is no individual underwriting. Long-term insurance gets cheaper for them when insurers use genetic information ($P^{no}(z_0^{*,no}) > P^{gen}(z_0)$). Hence, they also purchase long-term insurance but at a lower price when insurers use genetic information and are therefore better off with this regime than without any individual underwriting. For individuals with genetic disposition $z_0 = z_0^{c,1}$, long-term insurance is available with the same prepayment in both regimes. Hence, the use of genetic information does not affect their welfare. Finally, individuals with genetic disposition $z_0 > z_0^{c,1}$ purchase long-term insurance when there is no individual underwriting and it gets more expensive for them when insurers use genetic information ($P^{no}(z_0^{*,no}) < P^{gen}(z_0)$). These individuals either stay with long-term insurance but have to make a higher prepayment or they switch to short-term insurance. Therefore, they are worse off when insurers use genetic information.

To analyze the welfare effect of the use of genetic information when insurers use behavioral information in pricing, we compare the regime in which insurers use only behavioral information to the regime with full information. By definition of the cutoff $z_0^{*,beh}$, individuals with genetic disposition $z_0^{*,beh}$ are indifferent between long- and short-term insurance when insurers use only behavioral information. Under full information, they prefer long- over short-term insurance since long-term insurance eliminates classification risk at the fair premium (see Proposition 3). Hence,

$$EU^{beh}(e^{beh}, s^{beh}; z_0^* = z_0^{*,beh}) = EU^{ST}(e^{ST}, s^{ST}; z_0 = z_0^{*,beh}) < EU^{fu}(e^{fu}, s^{fu}; z_0 = z_0^{*,beh})$$

for cutoff individuals with genetic disposition $z_0^{*,beh}$. On the other hand, since $z_{z_0} > 0$, we obtain for $z_0 = 1$ that

$$P^{fu}(e,1) = z(1,e)(P_H - P_L) > \mathbb{E}\left[z(z_0,e) \mid z_0 \geq z_0^{*,beh}\right](P_H - P_L) = P^{beh}(e; z_0^{*,beh})$$

and, hence, $EU^{fu}(e, s; z_0 = 1) < EU^{beh}(e, s; z_0^* = z_0^{*,beh})$ for all e and s. Therefore,

$$EU^{fu}(e^{fu}, s^{fu}; z_0 = 1) < EU^{beh}(e^{beh}, s^{beh}; z_0^* = z_0^{*,beh})$$

for individuals with genetic disposition $z_0 = 1$. Taking above inequalities together,

$$EU^{fu}\left(e^{fu}, s^{fu}; z_0 = z_0^{*,beh}\right) > EU^{beh}\left(e^{beh}, s^{beh}; z_0^* = z_0^{*,beh}\right) > EU^{fu}\left(e^{fu}, s^{fu}; z_0 = 1\right). \quad \text{(A.3)}$$

When insurers use only behavioral information in pricing, individuals with genetic disposition $z_0 \geq z_0^{*,beh}$ purchase long-term insurance and the their welfare, EU^{beh}, does not depend on their genetic endowment z_0. Under full information, all individuals purchase the long-term contract and expected utility of individuals with this contract decreases in the genetic disposition z_0. Indeed, the envelope theorem yields

$$\frac{d}{dz_0} EU^{fu} = u' \left(w_1 - e^{fu} - P^{fu} \left(e^{fu}, z_0 \right) - s^{fu} \right) \left(-P_{z_0}^{fu} \left(e^{fu}, z_0 \right) \right) < 0,$$

where I used $P_{z_0}^{fu} \left(e^{fu}, z_0 \right) = z_{z_0} \left(z_0, e^{fu} \right) (p_H - p_L) l > 0$. Together with (A.3), this implies the existence of a unique $z_0^{c,2} \in \left(z_0^{*,beh}, 1 \right)$ such that

$$EU^{beh}(e^{beh}, s^{beh}; z_0^* = z_0^{*,beh}) < (=, >)\ EU^{fu}(e^{fu}, s^{fu}; z_0) \text{ for } z_0 < (=, >)\ z_0^{c,2}.$$

Hence, for $z_0 \geq z_0^{*,beh}$, the use of genetic information makes individuals with genetic disposition $z_0 < (=, >)\ z_0^{c,2}$ better (equally well, worse) off. Individuals with genetic disposition $z_0 < z_0^{*,beh}$ purchase short-term insurance when insurers use only behavioral information and switch to long-term insurance when insurers also use genetic information. Since they could have stayed with the same short-term contract, they must be better off when insurers use both behavioral and genetic information. In conclusion, individuals with genetic disposition $z_0 < (=, >)\ z_0^{c,2}$ are better (equally well, worse) off when insurers also use genetic information.

A.9 Proof of Proposition 12

I first check that the proposed social planner contract induces individuals to choose the efficient level of prevention. Inserting $P^{soc}(e, z_0)$, the first-order condition (3) characterizes individuals' expenditures on prevention. Since the tax/subsidy in P^{soc} does not depend on the individuals' expenditures on prevention, the first-order condition with the prepayment P^{soc} is the same as with the prepayment P^{fu} in the full information regime. Therefore, individuals indeed choose the efficient level of prevention $e^{fu} = e^*$ with the social planner contract.

I now show that the proposed contract yields the same consumption stream regardless of an individual's genetic disposition or health risk type. Since $P^{fu}(e, z_0) + e = z\left(z_0, e\right)(P_H - P_L) + e = EH(e; z_0) - P_L$, the expected utility under the social planner contract of an individual with genetic disposition z_0 and prevention expenditures $e^*(z_0)$ equals

$$\begin{aligned} EU^{soc}\left(e^*(z_0), s; z_0\right) &= u\left(w_1 - e^*(z_0) - P^{soc}\left(e^*(z_0), z_0\right) - s\right) + v\left(w_2 - P_L + s\right) \\ &= u\left(w_1 - \mathbb{E}\left[EH(e^*(\tilde{z}_0); \tilde{z}_0)\right] + P_L - s\right) + v\left(w_2 - P_L + s\right), \end{aligned}$$

and the consumption stream does not depend on z_0 or the individual's health risk type.

I now check that all individuals indeed purchase the contract offered by the social planner and that the contract yields a balanced budget. If all individuals purchase the social planner contract, it yields a balanced budget since P^{fu} generates zero profits and $\mathbb{E}\left[P^{soc}(e, z_0)\right] = \mathbb{E}\left[P^{fu}(e, z_0)\right]$, where I again take expectations over the genetic disposition z_0 across the population. With a mandate to purchase the social planner contract, the claim that all individuals purchase the contract holds automatically and the social planner contract indeed implements the welfare maximizing outcome at a balanced budget.

Without a mandate to purchase the social planner contract, individuals can choose between the social planner contract, a long-term contract offered by private insurers and short-term insurance. When private insurers are not allowed to use genetic information in pricing, they can only offer long-term contracts whose prepayment does not depend on z_0. Hence, the expected utility of individuals who purchase such a contract does not depend on z_0. Since expected utility with the social planner contract also does not depend on z_0, all individuals who purchase a long-term contract unanimously choose either the one offered by the social planner or the one offered by private insurers. Therefore, if the long-term health insurance contract offered by private insurers is in demand, the existence of the social planner contract does not affect the market outcome and private insurers offer the same long-term contract as in section 4.3. When individuals decide between the two long-term contracts,

$$\begin{aligned} &EU^{soc}\left(e^*(z_0), s; z_0\right) - EU^{beh}\left(e^{beh}, s; z_0^{*,beh}\right) \\ &\quad = u\left(w_1 - e^*(z_0) - P^{soc}\left(e^*(z_0), z_0\right) - s\right) - u\left(w_1 - e^{beh} - P^{beh}\left(e^{beh}; z_0^{*,beh}\right) - s\right). \end{aligned}$$

It holds that

$$\begin{aligned} e^*(z_0) + P^{soc}\left(e^*(z_0), z_0\right) &= \mathbb{E}\left[EH(e^*(\tilde{z}_0); \tilde{z}_0)\right] - P_L = \mathbb{E}\left[\min_e EH(e; \tilde{z}_0)\right] - P_L \\ &\leq \min_e \mathbb{E}\left[EH(e; \tilde{z}_0)\right] - P_L \leq \min_e \mathbb{E}\left[EH(e; \tilde{z}_0) \mid \tilde{z}_0 \geq z_0^{*,beh}\right] - P_L \\ &= \mathbb{E}\left[EH(e^{beh}; \tilde{z}_0) \mid \tilde{z}_0 \geq z_0^{*,beh}\right] - P_L = e^{beh} + P^{beh}\left(e^{beh}; z_0^{*,beh}\right), \end{aligned} \tag{A.4}$$

where I used that the efficient level of prevention minimizes each individual's personal health expenditures whereas, when insurers use only behavioral information, individuals with a long-term contract minimize the health expenditures of an average insured, and that expected health expenditures for a fixed level of prevention e are increasing in z_0. Hence, $EU^{soc}\left(e^*(z_0), s; z_0\right) - EU^{beh}\left(e^{beh}, s; z_0^{*,beh}\right) \geq 0$ and individuals prefer the long-term contract offered by the social planner over the long-term contract offered by private insurers for any level of saving s. In conclusion, they also prefer the social planner contract when they choose the optimal level of saving with each contract. Without a mandate for long-term health insurance, individuals at very low genetic risk may prefer to leave classification risk uninsured and purchase short-term insurance. We have just

seen that expected utility with the social planner contract does not depend on individuals' genetic disposition z_0. Hence, there will again be a cutoff, which I denote by $z_0^{*,c}$ such that individuals with genetic disposition $z_0 < z_0^{*,c}$ prefer short-term insurance whereas individuals with genetic disposition $z_0 \geq z_0^{*,c}$ prefer the social planner contract. If now all individuals have a genetic disposition $z_0 \geq z_0^{*,c}$, everybody prefers the social planner contract over short-term insurance. Using inequality (A.4), $z_0^{*,c} \leq z_0^{*,beh}$ follows analogously to the proof of $z_0^{*,beh} < z_0^{*,no}$ in Proposition 6. In conclusion, the social planner contract indeed implements the efficient level of prevention and consumption stream if $z_0 \geq z_0^{*,c}$ for all individuals even if there is no mandate to purchase this contract.

Finally, if the prevention technology additionally exhibits CD, the first inequality in (A.4) becomes an equality because the level of prevention that minimizes lifetime health expenditures does not depend on an individual's genetic disposition. Moreover, the second inequality also becomes an equality for $z_0^{*,beh} = z_0^{*,c}$. Hence, when insurers use only behavioral information a contract with the prepayment $P^{beh}(e; z_0^{*,c})$ yields the same expected utility as the contract offered by a social planner. Since the social planner contract breaks even, $z_0^{*,c}$ must also be an informationally consistent cutoff when insurers use only behavioral information in pricing. In conclusion, the market outcome in a private market with a ban on the use of genetic information is the same as with the social planner contract.

B Auxiliary calculations

B.1 Second-order conditions with long-term health insurance

To increase readability, I omit arguments and define $u_{LT} := u(w_1 - e - P(e, z_0) - s)$ and $v_{LT} := v(w_2 - P_L + s)$. The second partial derivatives of EU^{LT} are given by

$$\begin{aligned} EU_{ee}^{LT} &= u''_{LT}(-1 - P_e)^2 + u'_{LT}(-P_{ee}), \\ EU_{ss}^{LT} &= u''_{LT} + v''_{LT} < 0, \\ EU_{es}^{LT} &= u''_{LT}(1 + P_e). \end{aligned}$$

Hence, the determinant of the associated Hessian matrix equals

$$\begin{aligned} D = EU_{ee}^{LT} EU_{ss}^{LT} - (EU_{es}^{LT})^2 &= \left[u''_{LT}(1 + P_e)^2 - u'_{LT} P_{ee}\right]\left[u''_{LT} + v''_{LT}\right] - \left[u''_{LT}(1 + P_e)\right]^2 \\ &= u''_{LT}(1 + P_e)^2 v''_{LT} - u'_{LT} P_{ee}\left[u''_{LT} + v''_{LT}\right]. \end{aligned}$$

It holds that $EU_{ee}^{LT} < 0$ and $D > 0$ if $P_{ee} > 0$. Hence, the second-order conditions are globally fulfilled if $P_{ee} > 0$.

B.2 Potential cutoffs $z_0^* = 0$ and $z_0^* = 1$ with no individual underwriting

To clarify which cutoff the prepayment is based on, I denote expected utility with a long-term contract with prepayment $P^{no}(z_0^*)$ by $EU^{no}(e, s; z_0^*)$. When there is no individual underwriting, all individuals would be offered a long-term contract with the prepayment $P^{no}(0) = \mathbb{E}[z_0](P_H - P_L) > 0$ if the cutoff were given by $z_0^* = 0$. Individuals with genetic disposition $z_0 = 0$ definitely become a low-risk type even if they do not engage in prevention. Since

$$\begin{aligned} EU^{ST}(e^{ST}, s^{ST}; z_0 = 0) &= \max_s u(w_1 - s) + v(w_2 - P_L + s) \\ &> u(w_1 - P^{no}(0) - s^{no}) + v(w_2 - P_L + s^{no}) \\ &= EU^{no}(e^{no}, s^{no}; z_0^* = 0), \end{aligned}$$

individuals with genetic disposition $z_0 = 0$ prefer short- over long-term insurance and $z_0^* = 0$ cannot be an informationally consistent cutoff.

If the cutoff were given by $z_0^* = 1$ and there is no individual underwriting, the prepayment would equal $P^{no}(1) = P_H - P_L$. For individuals with genetic disposition $z_0 = 1$, this yields

$$\begin{aligned} EU^{ST}(e^{ST}, s^{ST}; z_0 = 1) &= \max_{e,s} u(w_1 - e - s) + \quad z(1,e)\, v(w_2 - P_H + s) \\ &\qquad + (1 - z(1,e))\, v(w_2 - P_L + s) \\ &> \max_s u(w_1 - s) + v(w_2 - P_H + s) \\ &= \max_{\tilde{s}} u(w_1 - P^{no}(1) - \tilde{s}) + v(w_2 - P_L + \tilde{s}) \\ &= EU^{no}(e^{no}, s^{no}; z_0^* = 1), \end{aligned}$$

where I used $e^{ST} > 0$ and $z(1,0) = 1$ to obtain the inequality in the second line, and defined $\tilde{s} := s - (P_H - P_L) = s - P^{no}(1)$ in the third line. This implies that $z_0^* = 1$ cannot be an informationally consistent cutoff either.

B.3 Potential cutoffs $z_0^* = 0$ and $z_0^* = 1$ with only behavioral information

To clarify which cutoff the prepayment is based on, I denote expected utility with a long-term contract with prepayment $P^{beh}(e; z_0^*)$ by $EU^{beh}(e, s; z_0^*)$. For the potential cutoff $z_0^* = 0$, the same argument as with no individual underwriting implies $EU^{ST}(e^{ST}, s^{ST}; z_0 = 0) > EU^{beh}(e, s; z_0^* = 0)$ for all $e \geq 0$ and $s \geq 0$ (see Appendix B.2). Hence, $z_0^* = 0$ cannot be an informationally consistent cutoff.

If the cutoff were given by $z_0^* = 1$ and the level of prevention by $e^*(1)$, on the other hand, $P^{beh}(e^*(1), 1) = \mathbb{E}[z(z_0, e^*(1)) \mid z_0 \geq 1](P_H - P_L) = z(1, e^*(1))(P_H - P_L) = P^{fu}(e^{fu}, 1)$. We

therefore know from Proposition 3 that with the optimal level of saving $\hat{s}$, $EU^{beh}(e^*(1), \hat{s}; z_0^* = 1) = EU^{fu}(e^{fu}, s^{fu}; z_0 = 1) > EU^{ST}(e^{ST}, s^{ST}; z_0 = 1)$. Moreover, the previous paragraph has shown that $EU^{beh}(e^*(1), \hat{s}; z_0^* = 0) < EU^{ST}(e^{ST}, s^{ST}; z_0 = 0)$. Hence, by continuity, there must exist $z_0^* \in (0, 1)$ such that $EU^{beh}(e^*(1), \hat{s}; z_0^*) = EU^{ST}(e^{ST}, s^{ST}; z_0 = z_0^*)$, i.e. such that z_0^* is an informationally consistent cutoff.