THE RISK PROTECTION VALUE OF "MORAL HAZARD"

By

Angie Acquatella and Victoria Marone

July 2026

COWLES FOUNDATION DISCUSSION PAPER NO. 2549

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

The Risk Protection Value of "Moral Hazard"*

Angie Acquatella[§] Victoria Marone[†]

July 2026

Abstract

Health insurance lowers the out-of-pocket price of healthcare, and it is well-established that this leads to higher utilization of care. This manifestation of "moral hazard" is typically viewed as a social cost of insurance. Within a standard model, this paper shows that a consumer's ability to change her behavior in response to insurance can also play a central role in the ability of insurance to protect her from risk. We provide a theoretical characterization of this channel and quantify its importance empirically. Under standard parameterizations and estimates in the literature, we find that insurance-induced healthcare utilization can account for more than half of the total value of risk protection derived from insurance. Preventing consumers from changing their behavior would lower healthcare spending, but also result in a major loss of risk protection, on-net *reducing* social welfare in some cases. Our results suggest that under-utilization of healthcare may thus be an equally important threat to welfare as over-utilization.

Keywords: moral hazard, risk protection, optimal insurance
JEL Codes: I13, D81, I18

*We are grateful to Jason Abaluck, Felix Bierbrauer, Marika Cabral, Hector Chade, Raj Chetty, David Cutler, Eduardo Dávila, Liran Einav, Randy Ellis, Amy Finkelstein, Marty Gaynor, Francois Gerard, Christian Gollier, Gautam Gowrisankaran, Igal Hendel, Nathan Hendren, Bruno Jullien, Zi Yang Kang, Elliot Lipnowski, Lee Lockwood, Erzo Luttmer, Albert Ma, Marius Ring, Larry Samuelson, Amanda Starc, Pietro Tebaldi, Nicholas Tilipman, Bob Town, Ben Vatter, and Andre Veiga for helpful comments and discussion. We also thank seminar participants at numerous departments and conferences for comments and suggestions that greatly benefited this research. Dario Gori and Junrui Lin provided excellent research assistance. [§]Department of Economics, Toulouse School of Economics (email: angelique.acquatella@tse-fr.eu) [†]Department of Economics, Yale University (email: victoria.marone@yale.edu) and NBER

I Introduction

The possibility that insurance changes consumer behavior after the realization of risk is typically viewed as a social cost. Across settings such as unemployment insurance, disability insurance, long-term care insurance, and health insurance, the manifestation of this type of "ex-post moral hazard"—or "behavioral response" to insurance—has been widely studied and broadly influential on policy debates around the world. In the realm of health insurance, in particular, controlling moral hazard has loomed large in the design of public health insurance schemes, since Pauly (1968) first described its potential to render some types of contracts non-optimal. The subsequent literature on optimal insurance emphasized that the value derived from risk protection must be weighed against the social cost from distorted incentives (Zeckhauser, 1970). What was perhaps never equally emphasized is that insurance-induced changes in behavior may add weight to this balance on both sides. While the empirical record now shows clearly that gaining health insurance increases healthcare utilization, it remains unclear under what theoretical conditions, and thus to what empirical extent, these increases are socially undesirable. This paper aims to fill that gap. In doing so, we demonstrate the possibility that preventing insurance-induced increases in healthcare utilization may do more harm than good.

The basic intuition is as follows. Absent insurance, a consumer must make choices that respect her budget constraint, regardless of what ills befall her. In a perfect-information world, she could write contracts to help her transfer resources from good states of the world to bad. In the real world, the realized state might be difficult to verify, and thus to contract on. Second-best contracts might emerge that help her achieve the same type of resource transfer, imperfectly. Changes in her behavior in response to a second best contract might then reflect one of two things: (i) choices she would also have made in a perfect world, under the resource transfers arranged by a state-contingent contract, but not absent insurance, or else (ii) choices she would only have made under imperfect conditions. Both reflect changes in behavior in response to insurance, but only the latter reflects distorted incentives.

Formalizing this story, the single market failure in our model is ex-post asymmetric information between the insurer and the insured. For clarity, we distinguish terminologically between the phenomenon of asymmetric information—which leads to incomplete internalization of costs under feasible contracts and thus to *moral hazard*—and the distinct phenomenon of changes in healthcare utilization that are induced by insurance—what we call *moral hazard utilization*.[1] As will become clear, these two concepts play distinct roles in our analysis. While

[1]This terminology departs from the prevailing norm in the health economics literature, where measurements of insurance-induced changes in behavior are commonly referred to as measurements of *moral hazard* (e.g.,

moral hazard is unambiguously undesirable, moral hazard utilization may not be. Indeed, our empirical findings suggests that the ability to adjust behavior after the realization of risk is an important tool for risk-mitigation, and should thus be restrained with caution.

The basic elements of our model are as follows. A consumer has a fixed initial income and a known distribution over health states. She has state-dependent preferences over two goods: healthcare utilization and non-health consumption. Absent insurance, she will typically have more valuable uses of income in sicker states. A health insurance contract lowers the consumer's point-of-sale ("out-of-pocket") cost of healthcare below its market price, but places otherwise no restrictions on her behavior. The realization of the health state is the consumer's private information, and if insured, she may act fully in accordance with her ex-post incentives under the contract, enjoying reimbursement for the difference between the out-of-pocket cost of care and its full cost. All consumer characteristics relevant to the insurer's ex ante expected cost are commonly known. There is therefore no ex-ante asymmetric information (selection) in the model, only ex-post information asymmetry (ex post moral hazard). Premiums are set to be actuarially fair, such that the private and social value of insurance coincide. The singular source of value from insurance is risk protection; the insurer has the technology to harness the Law of Large Numbers, while the individual does not.

Without the possibility to contract on realized health, the ability of health insurance to shift the consumer's resources across health states derives from her differential tastes for healthcare across states. Making healthcare cheaper raises her real income in all states (gross of the premium), but crucially, it does so by *more* in states in which healthcare is more valuable to her. Though health insurance cannot transfer nominal income across states, it can thus still transfer real income across states (Marshall, 1976; Newhouse, 1978; de Meza, 1983; Nyman, 1999a, 2003). The matter of moral hazard utilization—the consumer's resulting change in behavior—is then simply a question of how these changes in purchasing power are used. The principal focus of this paper is to show that resulting changes in the consumption of healthcare are not necessarily welfare-decreasing, and that their contribution to the risk protection value of insurance can be economically large.

An essential element of our theoretical approach is to treat the consumer's fundamental risk as being over her marginal indirect utility of income, rather than over consumption directly.[2]

Pauly, 1968; Einav and Finkelstein, 2018). In our terminology, we would say these are measurements of *moral hazard utilization*. We also note that the convention in this literature is for the information asymmetry giving rise to moral hazard to be about a hidden *type* (the realized state), rather than a hidden *action*. See footnote 18 for further discussion.

[2]These two approaches are generally interchangeable, in the sense that the marginal utility of the numeraire consumption good coincides with the marginal indirect utility of income at any interior optimum. They diverge when the consumer is constrained from reaching an interior optimum, as in the counterfactual of

We define the notion of *equivalent income* as the amount of money that, in a given health state under a given insurance contract, would leave the consumer just as well off had she instead faced uninsured (and undistorted) prices. This formal notion of real income allows us to express the consumer's valuation of any insurance contract as a state-contingent vector of equivalent incomes. It also permits a novel graphical analysis of the welfare generated by insurance, extending the classic income-income diagram of Rothschild and Stiglitz (1976) to a setting with moral hazard. Nothing in this construction restricts the functional form of utility; we require no separability between health and non-health consumption and no particular parameterization of risk preferences.

A second key element is the first-best insurance contract, which we use as a benchmark for efficient behavior. This contract would arise if asymmetric information were eliminated and the health state became contractible (Zeckhauser, 1970; Cutler and Zeckhauser, 2000). Rather than distorting prices, it reallocates the consumer's initial income across states in an actuarially fair way, equating her marginal indirect utility of income across states. Moral hazard utilization under a feasible contract can then be decomposed relative to the level of utilization that would arise under this benchmark. Absent insurance, equilibrium utilization typically falls below the first-best level, resulting in *under-utilization*.[3] With insurance, utilization typically exceeds the first-best level, resulting in *over-utilization*. An insurance-induced increase in utilization thus combines a reduction in under-utilization (attributable to efficient income transfers), with some amount of over-utilization (attributable to distorted incentives). Preventing the consumer from changing her behavior in response to insurance eliminates both.

Our model offers a new perspective on a number of central questions in the health insurance literature. First, our analysis clarifies that gaining "access" (à la Nyman, 1999b) to additional healthcare utilization is still very much a risk protection motive for insurance. Once the model is cast in the light of smoothing the marginal indirect utility of income, the risk-protection interpretation of income's uses follows directly. Second, our analysis highlights that the (ex-post) demand curve for healthcare cannot alone be used to make ex ante welfare statements about the consumption of care. Since welfare from insurance must invariably be valued ex ante, the normative interpretation of its effects on consumer behavior must also be. Healthcare utilization that looks inefficient from an ex-post perspective may instead look efficient from an

interest in this paper.

[3]For an example of under-utilization, suppose there is a 0.1 percent chance of becoming sick, meaning income could be fairly transferred from the healthy to the sick state at a rate of 1 to 999. Suppose further that in a first-best world, a consumer with initial income of $100,000 would happily sacrifice $1,000 when healthy in order to have $1,000,000 when sick. Take the sickness to be cancer and a potential treatment (immunotherapy) to cost $500,000. Suppose that with a million dollars, the sick consumer would buy immunotherapy. But without those additional resources, she clearly could not do so, and would instead settle with chemotherapy for just $50,000. The gap between $50,000 to $500,000 represents under-utilization.

ex-ante perspective.[4] Third, our decomposition of moral hazard utilization can be thought of as a price-theoretic counterpart to the idea of "high-value" versus "low-value" care (Schwartz et al., 2014; Baicker, Mullainathan and Schwartzstein, 2015). "High value" care might be interpreted as anything up to the consumer's first-best level of utilization (care she would consume in a first-best world, but perhaps could not afford absent insurance) and "low-value" care as over-utilization (care she would only consume due to the price distortion). Finally, our results underscore that the empirical (uncompensated) elasticity of demand for healthcare with respect to insurance coverage is not a sufficient statistic for optimal insurance design. Considering the reason for changes in behavior is essential, and in some cases (as we demonstrate in our empirical analysis), doing so can reverse the intuition that greater demand response implies higher optimal cost-sharing.

The second part of the paper takes the model to data. Our aim is to evaluate the quantitative importance of these forces in the context of empirical estimates of consumer preferences. To that end, we simulate four populations of consumers using parameterizations and estimates from Einav et al. (2013), Marone and Sabety (2022), and Ho and Lee (2023). We find that the risk protection value of moral hazard utilization can be economically large. Under a full insurance plan, it accounts on average for as much as 10 percent of welfare. Under-utilization drives almost the entirety of this amount, despite accounting for less than 1 percent of average expected healthcare spending. The corresponding negative contribution to welfare from the social cost of moral hazard, driven by over-utilization, ranges from between 2 to 22 percent of welfare. Taken together, eliminating moral hazard utilization—both under-utilization and over-utilization combined—would have a net negative effect on welfare in two of our simulated populations. In other words, somehow preventing consumers from changing their behavior in response to insurance would *reduce* the social value that insurance provides.[5]

There is substantial heterogeneity across consumers. One important dimension is income. Consumers with low income relative to their potential health outcomes are especially likely

[4] The standard revealed preference argument is that an uninsured consumer who does not purchase a unit of care at its full cost must value it at less, and hence that the units of utilization induced by insurance must be valued below cost (Pauly, 1968). But note that the value revealed by this argument is an ex post value; the observed choices are made *after* the health state is realized, out of the income available in that state. The ex ante counterpart would ask whether the consumer would like to pay a higher premium for insurance to fund the expected cost of the induced care, or else would prefer a lower premium and a commitment not to change her behavior in response to insurance. The ex post demand response alone does not reveal the answer to this question.

[5] We emphasize that the comparison in question is not between insurance and no insurance, but rather between insurance and no-behavioral-response-insurance. Healthcare remains cheaper to the consumer either way. While there is no formal mechanism in our model to prevent moral hazard utilization, there are many mechanisms that try to do so in the real world; for example, prior authorization, utilization review, referral requirements, and step therapy protocols.

to under-utilize healthcare absent insurance. For these consumers, the risk protection value of moral hazard utilization accounts for more than half of the welfare generated by insurance (while the social cost of moral hazard amounts to only 1 percent). Eliminating moral hazard utilization in this group would thus be enormously detrimental. Among consumers with the highest incomes, in contrast, eliminating moral hazard utilization would instead increase welfare, consistent with the standard intuition. We also find that consumers who change their behavior most in response to insurance do not necessarily have lower efficient coverage levels. A large utilization response can indicate that the consumer was under-utilizing healthcare absent insurance, or that over-utilization is especially valuable to her in sicker states; in either case, moral hazard utilization will tend to be worth its cost, yielding a higher efficient coverage level.

Related Literature. The study of moral hazard utilization in health insurance has been an active area of economic research since at least the 1960's (Cutler and Zeckhauser, 2000, p. 580). Quantifying its magnitude and evaluating its contribution to the social cost of health insurance has been the focus of a now vast literature in health economics (see foundational work by Feldstein (1973), Manning et al. (1987), Manning and Marquis (1996), and Einav and Finkelstein (2018) for a recent review).[6] Our focus is on how moral hazard utilization can also provide a social benefit, and on whether in some cases, this benefit can exceed its cost. Of course, many prior authors have noted that insurance-induced utilization can be socially beneficial in the presence of additional frictions. For example, there may be supra-competitive provider prices (Crew, 1969; Frick and Chernew, 2009), costly dynamic externalities (Cawley and Ruhm, 2011; Newhouse, 2006), liquidity constraints (Ericson, Jaspersen and Sydnor, 2025), or "behavioral hazard" under which misperceived benefits result in under-use (Baicker, Mullainathan and Schwartzstein, 2015; Chandra, Flack and Obermeyer, 2024). Our arguments for social benefits from moral hazard utilization are distinct. In our setting, the benefit arises from fully rational and informed behavior in a perfectly competitive and static environment. The key necessary condition is risk, and the sole mechanism of welfare benefit is risk protection.[7]

Our analysis also relates closely to the body of work of John Nyman, who has been a long-time advocate for a more positive interpretation of insurance-induced increases in healthcare utilization (Nyman, 1999a,b, 2003, 2007, 2008; Nyman et al., 2018). Nyman provides a decomposition of moral hazard utilization into parts which are conjectured to be either welfare

[6]The social cost of insurance has also been an important object of study in disability insurance (Deshpande and Lockwood, 2022), unemployment insurance (Chetty and Finkelstein, 2013), and long-term care insurance (Lieber and Lockwood, 2019; Kesternich et al., 2025).

[7]In contrast, a social benefit from moral hazard utilization could arise in all the papers cited above even if the consumer faced no uncertainty about her health. In our model, that would be impossible.

enhancing or welfare decreasing. Our decomposition of utilization is related but distinct. In particular, Nyman defines as *efficient* the level of utilization that would arise if the consumer faced undistorted prices but were given an income transfer equal to the amount of healthcare demanded if insured, less the fair premium (Nyman, 2003, p. 82–83). This definition omits involvement of the consumer's risk preferences, which are central in ours.[8] To our knowledge, ours is the first decomposition of the welfare generated by insurance into the value that would arise absent moral hazard utilization and the value attributable to the consumer's ability to change her behavior.[9]

Beyond healthcare, the risk protection value of insurance-induced changes in behavior may arise in a number of other settings. Most immediately, the model we study is formally equivalent to one of optimal income taxation for redistribution (Mirrlees, 1971). As in our setting, the realized state (ability) is unobserved by the insurer (government), leaving price distortion (taxes) as the second-best instrument for risk-spreading (redistribution). Our analysis suggests that taxation-induced labor supply distortions need not represent a pure social cost. These same ideas also appear in the context of social insurance programs that use in-kind transfers, which reduce the eligible consumer's price of some goods below their social cost (Lieber and Lockwood, 2019). Our model parallels this structure closely. Insurance-induced increases in the use of covered benefits—for example, an increase in the use of formal care in response to long-term care insurance (Mommaerts, 2025)—may thus not necessarily represent a social welfare loss.

The paper proceeds as follows. Section II presents our model and introduces the key concepts for welfare analysis. Section III presents the empirical analysis of the model. Section IV discusses the implications of our results. Section V concludes.

[8]Critically, the (nominal) income transfers that arise under the first-best contract need not—and in all likelihood, will not—correspond to the (real) income transfers effected by traditional health insurance. Traditional health insurance may either transfer too much or too little real income towards sick states, which will drive a wedge between Nyman's definition of the efficient level of utilization and our own.

[9]Note that under the marginal approach to insurance valuation (Baily, 1978; Chetty, 2006; Finkelstein, Hendren and Luttmer, 2019), no such decomposition arises because at any single margin of coverage, the envelope theorem guarantees that the behavioral response to insurance represents a pure cost. The risk protection value of moral hazard utilization is thus an intrinsically non-marginal phenomenon, arising only for discrete changes in insurance coverage. Appendix A.4 presents our welfare decomposition in derivative form and reconciles these two perspectives.

II Model

An expected-utility-maximizing consumer faces uncertainty about her health state $l \in \mathbb{R}$, which is distributed according to F. The consumer values two goods: healthcare utilization $m \in \mathbb{R}_+$ and non-health consumption $y \in \mathbb{R}_+$. The market price (and social cost) of both goods is fixed at one throughout the analysis, such that one unit of m represents the amount of healthcare which could otherwise be transformed into one unit of the composite non-health consumption good y. The consumer is endowed with income $w_0 \in \mathbb{R}_+$. Her preferences are represented by utility function $u(y, m; l)$, where u is strictly increasing in y, strictly concave in (y, m), and twice continuously differentiable in (y, m, l).[10] Utility features a (state-dependent) bliss point with respect to m, reflecting the intuition that the returns to being poked and prodded eventually become negative. Healthcare is relatively more valuable when the consumer is in worse health, and higher l indicates worse health; in other words, $\frac{u_m}{u_y}$ is strictly increasing in l. Finally, both goods are normal (demand increasing in income).[11]

A health insurance contract x is characterized by an out-of-pocket cost function $c(m; x)$, where $c(m; x) \in [0, m]$ represents the consumer's point-of-service cost of healthcare utilization m. We restrict attention to the set of contracts $\mathcal{X}$ for which $c(\cdot; x)$ is increasing and concave. An uninsured consumer is effectively enrolled in the null contract x_0, under which $c(m; x_0) = m$. Insurance thus lowers the consumer's point-of-service price of healthcare, according to the price schedule $c(m; x)$. The out-of-pocket cost function cannot depend directly on the health state l due to ex-post asymmetric information between the consumer and the insurer about which state has realized.

When the consumer is enrolled in contract x and realizes health state l, she makes an optimal choice of healthcare utilization m by trading off its benefit against the cost of foregone non-health consumption y. The optimal choice of healthcare utilization is given by:

$$m^*(l, x, w_0 - p) = \underset{m \geq 0}{\arg\max} \;\; u(y, m; l) \quad s.t. \quad 0 \leq y = w_0 - p - c(m; x). \tag{1}$$

where $p \in \mathbb{R}_+$ represents the contract premium. The corresponding optimal amount of non-health consumption is then $y^*(l, x, w_0 - p) = w_0 - p - c(m^*(l, x, w_0 - p); x)$. Prior to the

[10] Throughout the paper we use increasing and decreasing in the weak sense of non-decreasing and non-increasing, adding "strictly" when needed, and similarly with positive and negative, and concave and convex.

[11] This model has a number of antecedents in the health economics literature. Recent work has largely built directly from Cardon and Hendel (2001), but that model is in turn closely related to the Grossman (1972) framework. The notation has varied across papers, despite the model remaining essentially unchanged.

realization of the health state, expected utility from the lottery induced by insurance is

$$U(x, w_0 - p) = \mathbb{E}_l\big[u(y^*(l, x, w_0 - p), m^*(l, x, w_0 - p); l)\big].$$

We restrict attention to actuarially fair pricing, which has the useful implication that the private and social value of insurance coincide and thus obviates the need for interpersonal comparisons.[12] The fair premium for contract x is defined (implicitly) by

$$\bar{p}(x) = \mathbb{E}_l\big[m^*(l, x, w_0 - \bar{p}(x)) - c(m^*(l, x, w_0 - \bar{p}(x)); x)\big].$$

To simplify notation going forward, let $m^*(l, x) = m^*(l, x, w_0 - \bar{p}(x))$ denote the consumer's optimal healthcare utilization under the actuarially fair premium, and similarly for $y^*(l, x)$.

We measure welfare as the dollar increase in initial income that the consumer would deem equivalent to obtaining the contract at the fair premium. It can of course be negative, as it will be if the consumer would not be willing to pay the fair premium. The welfare generated by the provision of insurance contract x, relative to the status quo of having the null contract x_0, is thus equal to $V(x)$ that solves $U(x_0, w_0 + V(x)) = U(x, w_0 - \bar{p}(x))$.[13]

II.A Decomposition of Healthcare Utilization

Our goal is to understand the role of moral hazard utilization in driving social benefit from insurance, in addition to its traditional role in driving social cost. To that end, we begin by decomposing healthcare utilization under a given insurance contract into two parts: *moral hazard utilization* (i.e., causally resulting from insurance) and *non-moral hazard utilization* (i.e., would have arisen even absent insurance). Non-moral hazard utilization $m^N(l)$ in a given health state is the amount of healthcare a consumer would find optimal absent insurance:

$$m^N(l) \equiv m^*(l, x_0, w_0) = \underset{m \geq 0}{\arg\max}\ \ u(y, m; l) \quad s.t. \quad 0 \leq y \leq w_0 - m. \tag{2}$$

[12]Such a regime could be micro-founded as a perfectly competitive insurance market with no ex-ante asymmetric information and no loading.

[13]This measure of welfare corresponds to the *equivalent variation* (for an actuarially fairly compensated price decrease on healthcare). An alternative measure is the *compensating variation*, which would be equal to $V^c(x)$ that solves $U(x_0, w_0) = U(x, w_0 - \bar{p}(x) - V^c(x))$. If the consumer's willingness to pay for insurance is invariant to her income—that is to say, if there are not income effects in the demand for insurance—the two welfare measures will coincide. Many empirical health insurance papers use utility specifications embedding this assumption. Given the general utility specification studied here, however, we use equivalent variation because it provides a consistent money-metric basis for comparing welfare across contracts (recall that the compensating variation need not rank alternatives correctly; see Chipman and Moore, 1980; Mas-Colell, Whinston and Green, 1995, Ch. 3.I).

It is invariant to the insurance contract the consumer ultimately enrolls in, but depends on both the realized health state as well as the consumer's initial income.

Moral hazard utilization—under a given insurance contract x—is then simply all remaining spending that the consumer finds optimal: $m^*(l,x) - m^N(l)$. As long as the marginal propensity to spend on healthcare is not substantially higher in healthier states or at lower income levels, expected moral hazard utilization will be positive, as is consistent with the empirical evidence (see Lemma 1 in Appendix A.1). As we explore next, however, this utilization is not entirely wasteful.

Decomposition of Moral Hazard Utilization. Consider a hypothetical, perfect world with no asymmetric information between the consumer and the insurer, so that the health state is contractible. The consumer can now freely reallocate her income across states (by, say, trading Arrow-Debreu securities) to maximize her expected utility. The state-contingent income vector that does so is given by:

$$\left(w^{FB}(l)\right) = \underset{(w(l))}{\arg\max}\ \mathbb{E}_l\left[v(w(l);l)\right] \quad \text{s.t.} \quad w(l) \geq 0 \ \text{ and } \ \mathbb{E}_l\left[w(l)\right] \leq w_0, \tag{3}$$

where $v(w;l) = u(y^*(w;l), m^*(w;l);l)$ is the consumer's indirect utility function at undistorted prices, with optimal bundles given by $m^*(w;l) = \arg\max_{0 \leq m \leq w} u(w-m, m; l)$ and $y^*(w;l) = w - m^*(w;l)$.

We denote the first-best levels of healthcare utilization and non-health consumption by $m^{FB}(l) \equiv m^*(w^{FB}(l);l)$ and $y^{FB}(l) \equiv y^*(w^{FB}(l);l)$, the consumer's optimal bundles in a first-best world.[14] At these bundles, within each state, she equates the marginal utility of a dollar spent on each of the two goods at undistorted prices, so that the marginal benefit of m equals the marginal cost of foregone y, just as she would absent insurance. But in addition, at the first-best bundles, she equates her marginal indirect utility of income v_w *across* states.[15] Her resources may thus exceed w_0 in some states and fall short in others, whereas absent insurance they equal w_0 in every state. In this precise sense, the sole reason the consumer demands insurance is to smooth her marginal indirect utility of income across health states, i.e., to protect herself from risk.

This formulation of the first-best benchmark highlights that the consumer in this model fun-

[14]Note that dependence on initial income w_0 is present but suppressed, as w_0 remains fixed throughout our analysis.

[15]Unless, of course, she finds it preferable to allocate zero income to some state(s), in which case v_w would only be equated across states with non-zero income.

damentally faces preference risk as opposed to income risk.[16] Absent insurance, her marginal indirect utility of income $v_w(w_0; l)$ varies across states—just like in a model with income risk and a single consumption good—but does so because of differential "needs" (or "wants"), rather than because of differential resources. Canonically, $v_w(w_0; l)$ is high when the consumer is sick, and low when she is healthy. The consumer would therefore benefit from reallocating her income towards sick states. And naturally, changing her income in a given state may also change her optimal bundle in that state.

For a given health state realization, moral hazard utilization $m^*(l,x) - m^N(l)$ can now be decomposed into two parts: an *under-utilization* amount $m^{FB}(l) - m^N(l)$, and an *over-utilization* amount $m^*(l,x) - m^{FB}(l)$. Under mild conditions, both components will be positive in expectation across states (See Lemmas 2 and 3 in Appendix A.1).[17] Under-utilization represents units of healthcare that the consumer would inefficiently forgo absent insurance. It is induced by insurance—and thus meets our definition of moral hazard utilization—but it is not the result of "moral hazard." Indeed, it arises under first-best insurance, in which there is no moral hazard.[18] Over-utilization, in contrast, represents additional utilization induced by the price distortion inherent in feasible insurance.

Example with Two Health States. Figure 1 illustrates the decomposition of healthcare

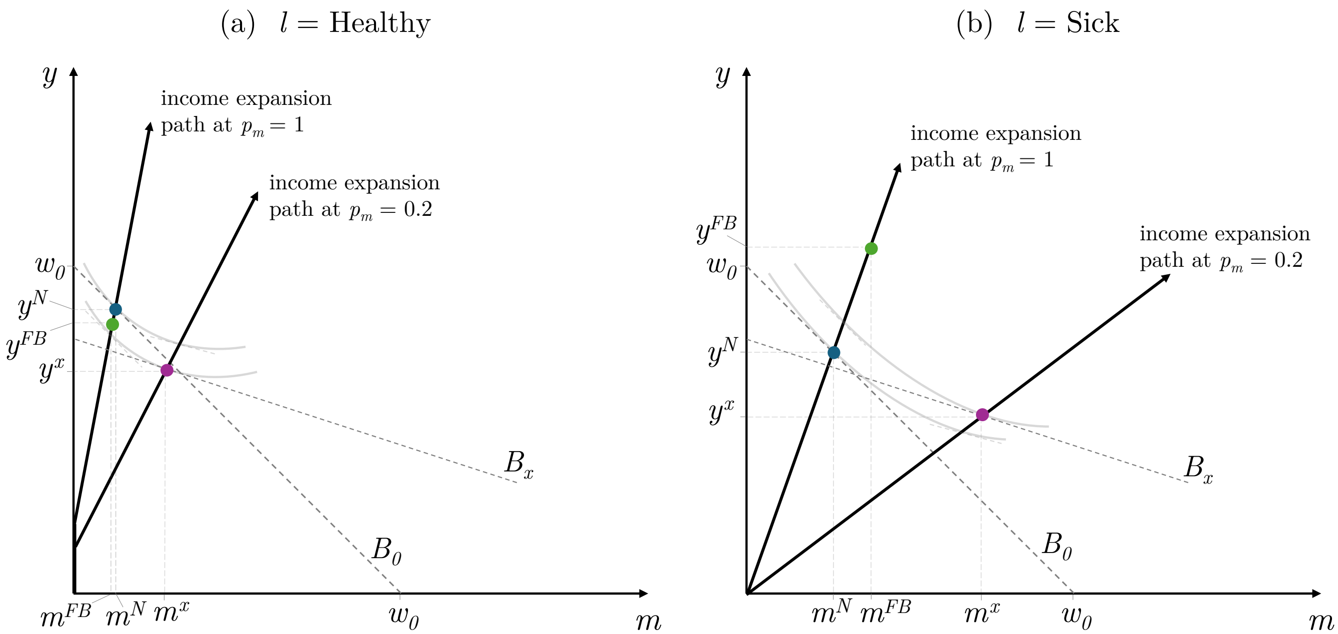
[16] This distinction clarifies much of the intuition from Nyman's work, who (rightly) took issue with models in which the only motivation for purchasing health insurance was to reduce risk of non-health consumption loss: "[T]here is a fundamental difference between insurance against a loss of property... and insurance against a 'loss' represented by the quid pro quo purchase of medical care... [W]hile the loss of income or wealth as medical expenditures is captured by conventional theory, the other side of the transaction is not, and the access to medical care can represent an extremely valuable gain. This suggests that conventional theory may be inappropriately applied to health insurance and that an alternative theory is needed to explain better the demand for this important commodity." (Nyman et al., 2018). Our analysis shows that all that is needed to explain the "access motive" for purchasing insurance is a model with budget allocation between two consumption goods, and clarifies that this motive is still predicated directly on risk protection.

[17] For under-utilization, the conditions are that the consumer's marginal propensity to consume healthcare is higher in sicker states and at higher incomes, and that the consumer prefers reallocating income from healthier to sicker states. For over-utilization, they are that the coverage provided by x is sufficiently generous, i.e., not too close to x_0 (along with some regularity conditions).

[18] It is worth re-emphasizing at this point that the use of the term "moral hazard" has been a matter of debate in the health economics literature. As noted in the introduction, we use *moral hazard* to refer to the manifestation of asymmetric information that prevents first-best contracts, and *moral hazard utilization* to refer to the behavioral response to insurance. The prevailing convention in health economics does not distinguish between these two concepts, using "moral hazard" for both (e.g., Einav and Finkelstein, 2018, p. 959). It would be reasonable to ask whether the term "moral hazard utilization" should really include *all* insurance-induced utilization, or only its "inefficient" component. Marshall (1976), de Meza (1983), Nyman (1999a), and Cutler and Zeckhauser (2000) do use this narrower definition, which corresponds loosely to what we call *over-utilization*. Our broader usage follows the prevailing norm in the health economics literature, as reflected in Pauly (1968) and Einav and Finkelstein (2018), as well as the insurance industry's own practice of using the term to mean any behavioral change by the insured—including both under- and over-utilization—that affects insurer costs (Pauly, 1983, p. 81).

utilization in a two-state example. Panel (a) corresponds to the healthy state of the world, and panel (b) to the sick state. Each panel depicts the consumer's (ordinal) preferences over (y, m) bundles in the respective state. Indifference curves on each diagram are convex to the origin, representing level sets of $u(y, m; l)$ in the given state l. On each panel, two key objects are drawn. The first is the consumer's income expansion path at undistorted prices ($p_y = p_m = 1$). This curve traces out all points in the (y, m) plane through which an indifference curve has slope -1. For any income level, this curve can thus tell us the consumer's optimal bundle at undistorted prices. Where it is upward sloping, there is a strictly positive income elasticity of demand for both goods. The second object is the income expansion path that prevails under an insurance contract x, when the consumer price of healthcare has been lowered below its market price. In the example shown, the insurance contract is simply a 20 percent coinsurance rate, such that the consumer price of m has fallen to 0.2. This income expansion path contains all the information needed for determining the consumer's choices under contract x, at any income level.

Figure 1. Decomposition of Healthcare Utilization: Two-State Example



Notes: This figure illustrates the decomposition of healthcare utilization in an example with two health states, where there is a 10 percent probability of becoming sick, and where the consumer has an insurance contract x characterized by a 20 percent coinsurance rate. Each panel depicts the consumer's (ordinal) preferences over non-health consumption y and healthcare utilization m in a given state, along with two income expansion paths (at the undistorted price $p_m = 1$, and at the insured price $p_m = 0.2$). B_0 is the budget constraint absent insurance, while B_x is the budget constraint with actuarially fair insurance.

Three different bundles are marked in each panel. The first is the consumer's optimal bundle absent insurance, (m^N, y^N) in each panel (shown in blue). This bundle is pinned down by the intersection of the no-insurance income expansion path ($p_m = 1$) and the no-insurance budget line B_0 (slope $= -1$, y-intercept $= w_0$). Because the consumer values healthcare more when sick, m^N is higher when sick than when healthy.[19] The second marked bundle in each panel is the consumer's first-best bundle (m^{FB}, y^{FB}) (shown in green), corresponding to the allocations derived from Equation 3. Since first-best insurance reallocates income across states without distorting relative prices, these bundles too lie on the income expansion path at $p_m = 1$. As drawn, the consumer's first-best allocation involves transferring income from the healthy to the sick state: starting from her status quo allocation, she would like to walk *down* her income expansion path in the healthy state in exchange for walking *up* the income expansion path in the sick state. This choice reflects cardinal information in her utility function (her risk preferences), rather than from information contained in her indifference curves over y and m (her ordinal preferences) alone. (The first-best bundles therefore cannot be derived from any tangencies on these diagrams.)

Finally, the third marked bundle is the equilibrium outcome under insurance contract x, (m^x, y^x) in each panel (shown in pink). This bundle lies at the intersection of the insurance budget line B_x (slope = -0.2, y-intercept $= w_0 - \bar{p}(x)$) and the income expansion path with insurance ($p_m = 0.2$), reflecting the tangency between an indifference curve and the budget line at the reduced out-of-pocket price. So that it is an actuarially fair shift in the budget set, the y-intercept of B_x is lowered below w_0 by the amount of the fair premium.[20] As with the no-insurance allocations (m^N, y^N), the allocations under insurance (m^x, y^x) reflect only information about the consumer's ordinal preferences (hence the ability to derive them from tangencies on these plots).

In each panel, the decomposition of moral hazard utilization $(m^x - m^N)$ can then be seen directly. Under-utilization $(m^{FB} - m^N)$ is negative in the healthy state, positive in the sick state, and positive in expectation. Over-utilization $(m^x - m^{FB})$ is positive in both states. It will be useful later to notice here that under-utilization represents movements along the income expansion path at undistorted prices, and thus represents the income effect on healthcare utilization from the first-best redistribution of income. Over-utilization, in contrast, represents the effect of the price distortion. It does not, however, correspond to a substitution effect.

---

[19] Our assumption that healthcare is relatively more valuable in sicker states ($\frac{u_m}{u_y}$ strictly increasing in l) implies that at any given price ratio, the income expansion path in the sick state will lie everywhere to the right of the one in the healthy state.

[20] The exact amount by which the insurance budget line B_x needs to be shifted down so that it corresponds to an actuarially fairly compensated price change can be seen as follows. Start by rotating B_0 around the y-intercept w_0 so that its slope corresponds to the new prices under insurance, -0.2. Then shift this line down vertically until the intersection in the healthy state falls inside the original budget set by an amount that exactly offsets the amount by which the intersection in the sick state falls outside, weighted by the respective state probabilities. Now the marked pink points in the diagram have been reached.

Alternative Decomposition of Moral Hazard Utilization: Substitution vs. Income Effects. The decomposition of moral hazard utilization between under-utilization and over-utilization does not correspond to a standard (Slutsky or Hicksian) decomposition between an income effect and a substitution effect. The waypoints separating the two components of each decomposition are simply different. In a standard Hicksian decomposition, the waypoint separating the substitution effect from the income effect is the bundle of goods which would both leave the consumer indifferent to her current situation as well as be optimal at the new prices under some income level.[21] This bundle can be seen directly in Figure 1, at the intersection between the indifference curve passing through (m^N, y^N) and the income expansion path with insurance ($p_m = 0.2$).[22] In our under-/over-utilization decomposition, in contrast, the waypoint separating under-utilization from over-utilization is the first-best bundle of goods. The two pairs of components—under-utilization/income effect and over-utilization/substitution effect—will coincide only in the particular case where consumer preferences are quasilinear in non-health consumption.

Appendix Figure A.1 shows the comparison between these alternative decompositions of the movement from the optimal bundle under no-insurance (m^N, y^N) to the optimal bundle under insurance (m^x, y^x), taking the trip along the path of either over-/under-utilization (as depicted in Figure 1) or instead along the path of income/substitution effects. Appendix Figure A.2 presents them on price–quantity axes instead of commodity-commodity axes, providing the analogs to Pauly's classic demand-based moral hazard diagram (Pauly, 1968, Figure 1). We will return to these two alternative decompositions of moral hazard utilization in our discussion of the decomposition of welfare, below.

II.B Decomposition of Welfare

Quantifying moral hazard utilization—and its contribution to a social cost of insurance—has been the focus of a now vast literature in health economics. Our focus is on how it can also provide social benefit, and on whether in some cases, this benefit can exceed its cost. We

[21] Or if traveling along the route of equivalent variation, the bundle which would both leave the consumer indifferent to her new situation as well as be optimal at the original prices under some income level (Hicks, 1943). We will say Hicksian rather than Slutsky going forward because Slutsky's original decomposition applied only to marginal price changes, obviating the need to specify a waypoint (Slutsky, 1915).

[22] Note that the decomposition presented here is not quite textbook standard. A classic Hicksian decomposition is applicable in the context of an *uncompensated* price change. With actuarially fair insurance, however, the price change is compensated in the form of the premium. A "Hicksian-type" decomposition in this setting thus has three components instead of two: (i)–(ii) the standard Hicksian income and substitution effects resulting from an uncompensated price change, and (iii) a separate income effect resulting from the premium. It is natural to collapse the two income effects, as we have done in the text.

approach this question by evaluating how the social welfare derived from insurance would change if insurance could still be given (the price decrease on healthcare), but moral hazard utilization could be prevented.

Our approach relies on a decomposition of welfare along two dimensions. The first is the classic breakdown between the value of risk protection and the social cost of moral hazard. The second is a novel decomposition between the welfare that would arise under a given insurance contract if the consumer were prevented from changing her behavior, and its complement, the incremental welfare attributable to the consumer's ability to adjust her behavior. In the latter case, we can then decompose further between changes in behavior attributable to under-utilization versus over-utilization of healthcare, as defined above.

The one remaining building block is to restate the consumer's well-being under a particular contract x in terms of indirect utility from real income, rather than in terms of direct utility from consumption. Consider the consumer's level of well-being under contract x in health state l, at her equilibrium bundle $(y^*(l,x), m^*(l,x))$. Define her *equivalent income* to be the income level $\tilde{w}(l,x)$ that, at undistorted prices, would allow her to attain this same utility level.[23] That is,

$$v(\tilde{w}(l,x);l) = u(y^*(l,x), m^*(l,x);l), \tag{4}$$

where $v(w;l)$ is again her indirect utility at undistorted prices.[24] Equivalent income translates the consumer's well-being in each state under contract x into the common currency of dollars spendable at undistorted prices. Absent insurance, the consumer has equivalent income w_0 in every state. With insurance, because healthcare is more valuable to her when sicker, her equivalent income is higher in sicker states.

Proposition 1 (Health Insurance Transfers Real Income Towards Sicker States). *For any $x \in \mathcal{X}$, equivalent income $\tilde{w}(l,x)$ is increasing in l.*

The proof is provided in Appendix A.2, but the intuition is simple. A price decrease on one good in all states feels more beneficial in states in which that good was relatively more valuable. Health insurance thus moves the consumer from the uninsured regime of fixed income but varying marginal indirect utility of income, to the insured regime of varying (equivalent)

[23]The concept of equivalent income dates back to King (1983), who developed it for the purpose of aggregating utilities of heterogeneous individuals. Somewhat surprisingly, it has, to the best of our knowledge, never been applied in the context of an insurance problem. Just as comparable dollar units are useful for welfare analysis when preferences differ across individuals, they are useful here where preferences differ across states.

[24]Equivalently, $\tilde{w}(l,x) = e(\hat{u}(l,x);l)$, where $e(\cdot;l)$ is the expenditure function at undistorted prices and $\hat{u}(l,x) = u(y^*(l,x), m^*(l,x);l)$ is the consumer's utility level under contract x. $\tilde{w}(l,x)$ could also be called the consumer's "money-metric utility" (McKenzie, 1957; Samuelson, 1974).

income but, possibly, smoother marginal indirect utility of income. Whether the consumer finds the efficiency cost of achieving this smoothing worthwhile is the question of welfare.

Welfare can now be expressed as $V(x)$ that solves:

$$\mathbb{E}_l[v(w_0 + V(x); l)] = \mathbb{E}_l[v(\tilde{w}(l,x); l)]. \tag{5}$$

Clearly, $V(x)$ will be positive only so long as the consumer in fact wanted to transfer income from healthier to sicker states, which is to say, so long as her marginal indirect utility of income $v_w(w_0; l)$ is higher when she is sicker. If not, she would have no use for this type of insurance.

The Value of Risk Protection and the Social Cost of Moral Hazard. With insurance, the consumer in state l uses total economic resources equal to $y^*(l,x) + m^*(l,x)$. Because she faces distorted prices, however, this bundle is not the one she would freely choose at undistorted prices given the same resources. The *deadweight loss* in state l is the dollar cost of this distortion:

$$D(l,x) \equiv \left[y^*(l,x) + m^*(l,x)\right] - \tilde{w}(l,x) \geq 0, \tag{6}$$

representing the gap between resources consumed and the associated equivalent income.[25] Reflecting the standard revealed preference argument (cf. footnote 4), the consumer's (ex post) valuation of her consumption bundle under insurance must fall below its resource cost, such that $D(l,x)$ is generically positive.

Taking expectations of (6), expected equivalent income under contract x can now be written

$$\mathbb{E}_l[\tilde{w}(l,x)] = w_0 - \underbrace{\mathbb{E}_l[D(l,x)]}_{\equiv\, SCMH(x)}, \tag{7}$$

where we have used the actuarial fairness condition that guarantees total expected resources consumed must equal initial income ($\mathbb{E}_l[y^*(l,x) + m^*(l,x)] = w_0$). We define $SCMH(x) \equiv \mathbb{E}_l[D(l,x)]$ to be the *social cost of moral hazard*—the expected deadweight loss resulting from the price distortion.

Equation 7 represents the fundamental decomposition of value from insurance. On one hand, insurance destroys wealth, wastefully reallocating consumption towards healthcare and away

[25]Note that $D(l,x) = 0$ if and only if the consumer's equilibrium bundle coincides with what she would freely choose at undistorted prices given the same total resources, as would occur if demand for healthcare were perfectly inelastic with respect to price (Pauly, 1968, p. 533). More generally, the magnitude of $D(l,x)$ depends not only on the price elasticity of demand for healthcare but also on the curvature of the (Hicksian) demand curve (Harberger, 1971; Hausman, 1981; Kang and Vasserman, 2025).

from non-health consumption, leaving expected equivalent income $\mathbb{E}_l[\tilde{w}(l,x)]$ to fall below the consumer's actual income w_0. On the other hand, insurance redistributes wealth, yielding the state-contingent equivalent income vector $[\,\tilde{w}(l,x)\,]$, rather than the status quo state-invariant income vector $\left[\,w_0\,\right]$. The value of risk protection reflects the consumer's valuation of this reallocation of resources, gross of its efficiency cost. Taken together, the welfare generated by insurance can be decomposed as:

$$V(x) = \underbrace{\Psi(x)}_{\substack{\textit{Value of}\\ \textit{risk protection}}} - \underbrace{SCMH(x)}_{\substack{\textit{Social cost of}\\ \textit{moral hazard}}}, \tag{8}$$

where $V(x)$ is defined implicitly in Equation 5, $SCMH(x)$ is defined in Equation 7, and $\Psi(x)$ is the residual.[26]

If there were no uncertainty about health, or if the consumer's marginal indirect utility of income $v_w(w;l)$ were constant across all health states and income levels, $\Psi(x)$ would equal zero. Such a consumer has no more valuable use for a dollar when sick than when healthy, and would not be willing to pay anything to redistribute resources across states. The provision of insurance would unambiguously reduce welfare. If, instead, the consumer has more valuable uses of income in sicker states, then the transfer of resources from healthy to sick states that can be achieved by contract x will be valuable to her. As long as the contract does not transfer *too much* income towards sick states relative to the consumer's desired (first-best) allocation of income, $\Psi(x)$ will be positive.[27]

Proposition 2 (Value of Risk Protection is Positive). *For any $x \in \mathcal{X}$, if $v_w(\tilde{w}(l,x);l)$ is increasing in l, then $\Psi(x) \geq 0$.*

The proof is provided in Appendix A.2, but the logic is simple. The consumer's desired (first-best) allocation of income equalizes her marginal indirect utility of income across states, leaving no valuable transfers unmade. Contract x transfers equivalent income toward sicker states (Proposition 1), whether or not the consumer wanted such a transfer. Such a transfer lowers v_w in sicker states, where income arrives, and raises it in healthier states, where income departs. For a consumer whose marginal indirect utility of income was higher in sicker states to begin with, the transfer therefore works to compress the differences. The hypothesis of Proposition 2 says that even at the allocation the contract delivers, marginal indirect utility

[26] This decomposition is closely related to the partition between *aggregate efficiency* and *redistribution* studied by Dávila and Schaab (2025), using initial income as welfare numeraire.

[27] Otherwise, if the consumer really only wanted a little bit of extra income when sicker, the income transfer effected by insurance could leave her worse off—in terms of smoothing her marginal indirect utility of income—than she was before.

of income *remains* higher in sicker states; the consumer would be willing to transfer *even more* income toward sick states than the contract achieved. Such a contract has moved the consumer's marginal indirect utility of income toward equalization without overshooting it. Her v_w is smoother under contract x than under x_0, so $\Psi(x) \geq 0$.[28]

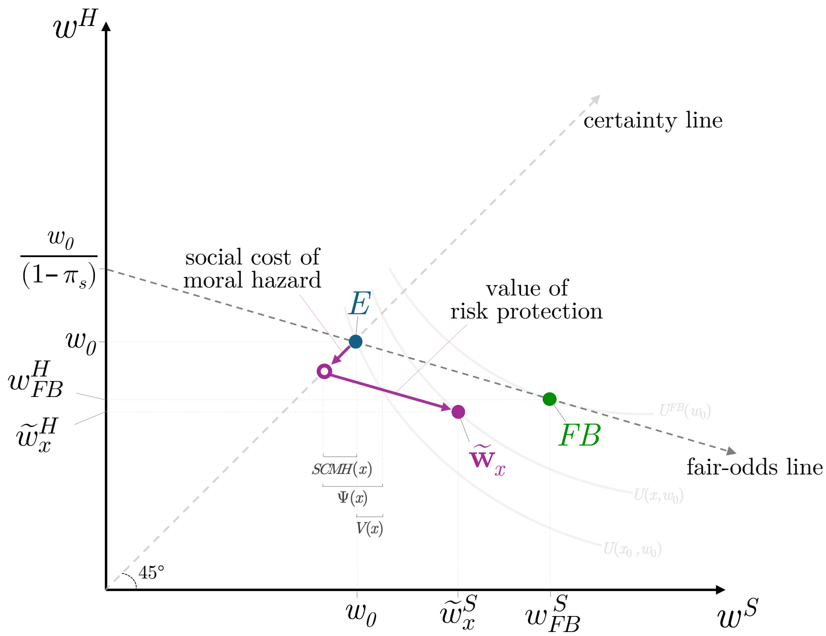
Two-State Example. The decomposition of welfare between the value of risk protection and the cost of moral hazard has a transparent geometric interpretation, illustrated in Figure 2 for the two-state example. The horizontal axis represents income spendable at undistorted prices (equivalent income) in the sick state, and the vertical axis for the healthy state. The consumer's endowment point $E = (w_0, w_0)$ (marked in blue) lies on the 45-degree line (the certainty line). Through E passes the fair-odds line, representing the set of all state-contingent income vectors with expected value w_0 (an iso-mean line). It has slope $-\pi_S/(1-\pi_S)$, where $\pi_S \in [0,1]$ is the probability of becoming sick. Indifference curves on this diagram represent level sets of expected utility $\mathbb{E}_l[v(w(l);l)]$ derived from various state-contingent income bundles. The consumer's most-preferred bundle within her budget set is the point $FB = (w^S_{FB}, w^H_{FB})$ (marked in green), indicating graphically the solution to the first-best problem in Equation 3.[29] This diagram represents the moral-hazard analog to the income-income diagram with contractible health state introduced by Rothschild and Stiglitz (1976, Figure I).

Under insurance contract x, the consumer's state-contingent equivalent income vector is represented by the point $\tilde{\mathbf{w}}_x = (\tilde{w}^S_x, \tilde{w}^H_x)$, marked in pink. Since equivalent income is higher when sick (Proposition 1), $\tilde{\mathbf{w}}_x$ lies to the right of the certainty line. And, since mean equivalent income falls below w_0 by the amount $SCMH(x)$ (Equation 7), $\tilde{\mathbf{w}}_x$ lies below the fair-odds line. As drawn, $\tilde{\mathbf{w}}_x$ lies on a higher indifference level than point E, meaning contract x is preferred to remaining uninsured. Given actuarial fairness, we can conclude that the movement from E to $\tilde{\mathbf{w}}_x$ generates positive social welfare. The magnitude of welfare $V(x)$ is equal to the certain income gain that would move the consumer from E, up the certainty line, to the intersection with the indifference curve passing through $\tilde{\mathbf{w}}_x$.

[28] The hypothesis in Proposition 2 embeds the condition, discussed above, that the consumer has more valuable uses of income in sicker states in the first place, i.e., that $v_w(w_0; l)$ is increasing in l. This is the classic condition that there is a positive covariance between the marginal indirect utility of income and the severity of the health state (Arrow, 1973; Besley, 1988). It is necessary. Without it, any health insurance contract would transfer income opposite to the desired direction, and $\Psi(x)$ would be negative. It is not sufficient, however, precisely because a contract can overshoot the income transfer. Requiring that v_w be increasing in l at the equivalent income vector the contract delivers, rather than at the consumer's initial income, is what rules out such overshooting.

[29] Note that the first-best income vector corresponds to the cost of the first-best bundles (m^{FB}, y^{FB}) in each health state in Figure 1. That is, $m^S_{FB} + y^S_{FB} = w^S_{FB}$ and $m^H_{FB} + y^H_{FB} = w^H_{FB}$. (The diagrams are not drawn exactly to scale.)

Figure 2. Decomposition of Welfare: Two-State Example



Notes: This figure depicts the welfare decomposition from Equation 8 in a two-state example. The horizontal axis is equivalent income in the sick state; the vertical axis is equivalent income in the healthy state. $E = (w_0, w_0)$ is the endowment point (no insurance). $FB = (w^S_{FB}, w^H_{FB})$ is the first-best allocation from Equation 3. $\tilde{\mathbf{w}}_x = (\tilde{w}^S_x, \tilde{w}^H_x)$ is the consumer's state-contingent equivalent income under contract x. The movement from E to $\tilde{\mathbf{w}}_x$ decomposes into (i) a movement down the certainty line (the social cost of moral hazard) and (ii) a mean-preserving redistribution of income toward sicker states (risk protection).

The movement from E to $\tilde{\mathbf{w}}_x$ can be decomposed into two orthogonal parts, mirroring Equation 8. The first is a movement southwest along the certainty line—a uniform reduction of equivalent income by amount $SCMH(x)$ in every state—representing the social cost of moral hazard. It carries the consumer from the endowment point (w_0, w_0) to the intermediate point $(w_0 - SCMH(x),\, w_0 - SCMH(x))$, marked with an open pink circle.[30] Naturally, the consumer would be indifferent between avoiding this movement and losing $SCMH(x)$ dollars from her initial income. The second is a movement southeast along the iso-mean line passing through the intermediate point—a mean-preserving redistribution of income toward the sick state. The consumer, when standing at the intermediate point, would be indifferent between moving to $\tilde{\mathbf{w}}_x$ and gaining $\Psi(x)$ dollars for sure. The welfare generated by the contract, $V(x)$, reflects the consumer's net valuation of these two movements.[31]

[30] By Equation 7, $w_0 - SCMH(x) = \mathbb{E}_l[\tilde{w}(l, x)]$, so the intermediate point is also the mean of the consumer's equivalent income vector under contract x. It is the point where the iso-mean line passing through $\tilde{\mathbf{w}}_x$ crosses the certainty line.

[31] These three magnitudes are indicated in Figure 2. $SCMH(x)$ is the certain-income distance from the intermediate point to E; $\Psi(x)$ is the certain-income distance from the intermediate point to the point where the indifference curve passing through $\tilde{\mathbf{w}}_x$ crosses the certainty line (the consumer's certainty equivalent of

Figure 2 also illustrates Proposition 2. In the two-state example, the requirement that $v_w(\tilde{w}(l,x);l)$ be increasing in l says that the indifference curve passing through $\tilde{\mathbf{w}}_x$ is at least as steep as the fair-odds line. Equivalently, $\tilde{\mathbf{w}}_x$ must lie on the segment of the iso-mean line between the intermediate point and its point of tangency with an indifference curve, where the consumer's desired redistribution along that line would be complete. By inspection, this would fail if the tangency point were to lie northwest of the certainty line (indicating the consumer did not wish to transfer income towards the sick state at all), or if $\tilde{\mathbf{w}}_x$ were to lie southeast of the tangency point (indicating that contract x transferred more income than the consumer wanted). As drawn, $\tilde{\mathbf{w}}_x$ lies on this segment, and $\Psi(x)$ is positive.

These two orthogonal movements in the space of equivalent income correspond to the Hicksian decomposition of moral hazard utilization discussed above. In each state, there is an income effect associated with moving income from w_0 to $\tilde{w}(l,x)$ at undistorted prices, and then a substitution effect of moving along the new indifference curve to the income expansion path associated with insured prices.[32] The expected excess cost of the substitution effects across all states is exactly the social cost of moral hazard. It is a pure efficiency cost, pulling $\tilde{\mathbf{w}}_x$ into the interior of the Pareto frontier. The cross-state pattern of both effects, and the associated rearrangement of equivalent income, is what delivers risk protection.

Consistent with the classic intuition, the income effects represent transfers with no efficiency cost (movement along an iso-mean line), while it is the substitution effects that create deadweight loss. What must be held in mind in the present setting (as well as in others where an ethical observer might view transfers across heterogeneous agents as beneficial) is that efficiency loss does not imply welfare loss. As Zeckhauser (1970) first emphasized, the value of risk protection derived from the price distortion may more than make up for its deleterious effects. The indifference curves on Figure 2 show this clearly.

Components of Healthcare Utilization. To isolate the specific contribution of the consumer's ability to change her behavior in response to insurance, welfare can be further decomposed as follows. Consider two hypothetical situations. In the first, imagine the consumer was provided insurance contract x—the price decrease on healthcare—but was somehow constrained to consume only $m^N(l)$ in each health state. In other words, suppose the consumer were given insurance, but was prevented from changing her behavior. Under this hypothetical, the consumer would pay actuarially fair premium $\bar{p}^N(x) = \mathbb{E}_l[m^N(l) - c(m^N(l);x)]$ and receive non-health consumption $y^N(l,x) = w_0 - \bar{p}^N(x) - c(m^N(l);x)$ in each state. Insurance

$\tilde{\mathbf{w}}_x$); and $V(x) = \Psi(x) - SCMH(x)$ is the difference.

[32] These two movements can be traced in each panel of Figure 1. Appendix Figure A.1 reproduces the same diagram with more detailed annotations. The path described corresponds to the movement from point [1] to [2] to [5].

would thus still transfer resources across states, but do so exclusively through non-health consumption. The associated equivalent income, representing the consumer's valuation of this arrangement in a given state, would be given by $\tilde{w}^N(l,x)$ that solves $v(\tilde{w}^N(l,x);l) = u(y^N(l,x), m^N(l);l)$. The welfare generated could be decomposed exactly as above, into a value of risk protection and a social cost of moral hazard.[33]

Now consider instead a second hypothetical situation, in which the consumer were constrained to use $m^{FB}(l)$, the first best amount of healthcare in each health state, rather than the no-insurance amount $m^N(l)$ or the equilibrium amount $m^*(l,x)$. Here, she analogously pays the actuarially fair premium associated with this level of healthcare utilization and the terms of the contract: $\bar{p}^{FB}(x) = \mathbb{E}_l[m^{FB}(l) - c(m^{FB}(l);x)]$, leaving her with non-health consumption $y^{FB}(l,x) = w_0 - \bar{p}^{FB}(x) - c(m^{FB}(l);x)$.[34] Her associated equivalent income would now be given by $\tilde{w}^{FB}(l,x)$ that solves $v(\tilde{w}^{FB}(l,x);l) = u(y^{FB}(l,x), m^{FB}(l);l)$.

The two hypothetical utilization schedules $m^N(l)$ and $m^{FB}(l)$ thus define two intermediate points in the space of state-contingent equivalent incomes for a given contract x. Starting from the endowment point, there is the no-behavioral-response point $\tilde{\mathbf{w}}^N(x)$, the efficient-utilization point $\tilde{\mathbf{w}}^{FB}(x)$, and finally the equilibrium point $\tilde{\mathbf{w}}(x)$. The movement from the endowment E to the equilibrium $\tilde{\mathbf{w}}(x)$, with associated welfare effect of $V(x) = \Psi(x) - SCMH(x)$, can thus be decomposed into three steps, each of which in turn can be decomposed into an incremental

[33]Note that the social cost of moral hazard (expected deadweight loss) is in general still nonzero under this hypothetical: preventing the consumer from re-optimizing her consumption bundle in response to insurance results in allocations that do not lie on the income expansion path at undistorted prices—which is to say, that have a deadweight loss. Given a transfer of resources in a given state, the consumer may have preferred to increase *both* health and non-health consumption, but here she is permitted only to increase her non-health consumption.

[34]Note that $y^{FB}(l,x)$ is not the same object as $y^{FB}(l)$. The latter is the level of non-health consumption that would arise under the first-best contract, as defined in Section II.A (p. 9). $y^{FB}(l,x)$, in contrast, is the non-health consumption that would result from enforcing the first best level of healthcare utilization, but otherwise implementing the cost sharing features of contract x. There may thus still be deadweight loss in this case.

value of risk protection and an incremental social cost of moral hazard:

$$V(x) = V^N(x) \ + \ \overbrace{\Delta V^{under}(x) \ + \Delta V^{over}(x)}^{\substack{\textit{Social value of} \\ \textit{moral hazard utilization} \\ \Delta V^{MH}(x)}} \tag{9}$$

$$\begin{aligned} &= \Psi^N(x) \ + \ \overbrace{\Delta \Psi^{under}(x) \ + \Delta \Psi^{over}(x)}^{\substack{\textit{Risk protection value} \\ \textit{of moral hazard utilization,} \\ \Delta \Psi^{MH}(x)}} \\ &\quad - SCMH^N(x) - \underbrace{[\Delta SCMH^{under}(x) + \Delta SCMH^{over}(x)]}_{\textit{Social cost of moral hazard utilization}}. \end{aligned} \tag{10}$$

In Equation 9, $V^N(x)$ represents the welfare achieved under the no-behavioral-response benchmark, $\Delta V^{under}(x)$ represents the incremental welfare obtained by moving to the first-best-utilization benchmark (incorporating care that was under-utilization absent insurance), and finally, $\Delta V^{over}(x)$ represents the incremental welfare obtained by moving to the true equilibrium (incorporating over-utilization).[35] Together, the second and third term represent the social value of moral hazard utilization.[36]

In Equation 10, each term is then broken into its two orthogonal components. The term $\Psi^N(x)$ represents the risk protection value of insurance absent the consumer's ability to change her behavior. The terms $\Delta\Psi^{under}(x) + \Delta\Psi^{over}(x)$ together represent the risk protection value of moral hazard utilization.[37]

The Risk Protection Value of Moral Hazard Utilization. The risk protection value of moral hazard utilization reflects the value of allowing the consumer to optimally exploit resources gained in bad states. To see this concretely, recall that in a given state, the no-behavioral-response hypothetical forces the consumer to allocate any additional resources she may have gained (or lost) to non-health consumption. Even if the resulting allocation left the marginal benefit of an extra unit of healthcare utilization *above* the marginal cost of losing a unit of non-health consumption, the consumer is unable to re-optimize. Deadweight loss

[35]Formally, incremental welfare is $\Delta V^{under}(x) = V^{FB}(x) - V^N(x)$, and $\Delta V^{over}(x) = V(x) - V^{FB}(x)$, where $V^N(x)$ is defined implicitly by $\mathbb{E}_l[v(w_0 + V^N(x); l)] = \mathbb{E}_l[v(\tilde{w}^N(l,x); l)]$, and similarly for $V^{FB}(x)$.

[36]Appendix A.4 presents this decomposition in derivative form, reconciling it with the marginal approach to insurance valuation.

[37]As with welfare overall, $\Delta\Psi^{under}(x) \equiv \Psi^{FB}(x) - \Psi^N(x)$ and $\Delta\Psi^{over}(x) \equiv \Psi(x) - \Psi^{FB}(x)$, where $\Psi^N(x) = V^N(x) + SCMH^N(x)$ and $\Psi^{FB}(x) = V^{FB}(x) + SCMH^{FB}(x)$. And similarly for $\Delta SCMH^{under}(x)$ and $\Delta SCMH^{over}(x)$, where $SCMH^N(x) = w_0 - \mathbb{E}_l[\tilde{w}^N(l,x)]$ and $SCMH^{FB}(x) = w_0 - \mathbb{E}_l[\tilde{w}^{FB}(l,x)]$.

follows directly, corresponding in expectation to the term $SCMH^N(x)$. The key question, from the perspective of risk protection, is whether more real income is left on the table in sicker or healthier states. If there is more to gain from re-optimizing in sicker states, then the relaxation of the no-behavioral-response constraint would lead to a relative redistribution of equivalent income towards these states.

The problem, of course, is that allowing the consumer to change her behavior comes in two parts. First, it allows her to align her utilization with what she would have chosen in a first best world, contributing $\Delta\Psi^{under}$ to the risk protection value of insurance. Second, facing a subsidized marginal price of healthcare, she cannot help but consume beyond this level. The resulting over-utilization once again results in deadweight loss, in aggregate corresponding to the full social cost of moral hazard from insurance $SCMH(x)$. Each stage shifts equivalent income across states. Whether the shift goes toward sicker states, where it may be most valuable, is the key question for signing the incremental value of risk protection. Appendix A.2 provides sufficient conditions for $\Delta\Psi^{under}$ and $\Delta\Psi^{over}$ to be positive (Propositions 3 and 4). The central requirements in each case are that the consumer wishes to transfer resources towards sick states and that relaxation of the no-behavioral-response constraint is more valuable in sicker rather than healthier states. Ultimately, both the sign and magnitude of these effects in the real world are empirical questions.

Our goals in the remainder of the paper are twofold. First, we want to evaluate the quantitative importance of each of these forces in standard parameterizations of consumer preferences that have been estimated and used in the health insurance literature. Our second goal is to evaluate the extent to which these estimates imply a social benefit of moral hazard utilization that exceeds its social cost.

III Empirical Evidence

We implement our welfare decomposition in the context of parameterizations and estimates of consumer preferences in the health economics literature. This provides, to our knowledge, the first empirical decomposition of welfare into (i) the value that insurance can provide when changes in healthcare utilization are prohibited, and (ii) the incremental value it can provide when consumers can change their utilization in response to insurance.

III.A Parameterization of Utility

We present the decomposition based on parameter estimates from three papers: Einav et al. (2013), Marone and Sabety (2022), and Ho and Lee (2023). In each of these papers, utility takes the form $-\exp(-\psi(y + b(m;l)))$, where $b(\cdot;l)$ is quadratic in m and represents the consumer's value of healthcare in health state l, implying linear demand for healthcare (consistent with Pauly's original analysis). From Einav et al. (2013), we present results under each of the two parameterizations of utility used in the paper—what the authors call the "additive" versus "multiplicative" specifications. Table 1 presents the four specifications. Appendix B provides a detailed description of our procedure for constructing a population of consumers and applying the parameter estimates reported in each paper.

Table 1. Specifications of Utility

Spec.	Source of estimates	Parametrization of $b(m;l)$	Implied demand for healthcare
(1)	Marone and Sabety (2022)	$(m-l) - \frac{1}{2\omega}(m-l)^2$	$l + \omega(1-\bar{c})$
(2)	Einav et al. (2013) – Additive	$(m-l) - \frac{1}{2\omega}(m-l)^2$	$l + \omega(1-\bar{c})$
(3)	Einav et al. (2013) – Multiplicative	$(m-l) - \frac{1}{2\tilde{\omega}l}(m-l)^2 - \frac{\tilde{\omega}l}{2}$	$l + \tilde{\omega}l(1-\bar{c})$
(4)	Ho and Lee (2023)	$(m-l) - \frac{1}{2\tilde{\omega}l}(m-l)^2 - \frac{\tilde{\omega}l}{2}$	$l + \tilde{\omega}l(1-\bar{c})$

Notes: The table shows the four specifications we study. Specifications (1) and (2) share an "additive" parametric form for $b(m;l)$, in which $\omega \in \mathbb{R}_{++}$ is the dollar amount of over-utilization induced by full insurance. Specifications (3) and (4) share a "multiplicative" form, in which $\tilde{\omega} \in \mathbb{R}_{++}$ is the proportional amount of over-utilization. In both forms, the first-best level of healthcare utilization equals the health state: $m^{FB}(l) = l$. The final column reports implied demand $m(\bar{c};l)$ as a function of a linear out-of-pocket price $\bar{c} \in [0,1]$. The implied price elasticity of demand is $\varepsilon = -\bar{c}\omega/(l + (1-\bar{c})\omega)$ in the additive form and $\varepsilon = -\bar{c}\tilde{\omega}/(1 + (1-\bar{c})\tilde{\omega})$ in the multiplicative form. Einav et al. (2013) provides parameter estimates under both forms; Marone and Sabety (2022) estimates the additive form; Ho and Lee (2023) estimates the multiplicative form.

Under this class of parameterizations, a number of terms in our decomposition can be generically signed. First, all of the risk protection terms—$\Psi^N(x)$, $\Delta\Psi^{under}$, $\Delta\Psi^{over}$—will be (weakly) positive. Appendix A.3 provides additional discussion and the short proofs. Second, $SCMH^N$ and $\Delta SCMH^{under}$ will when summed equal zero, leaving $SCMH^{over}$ as the only relevant component of the social cost of moral hazard. Because the consumer's preferences are quasilinear in non-health consumption, over-utilization corresponds to the substitution effect of the price distortion, and is thus the only contributor to deadweight loss.[38] Instead of six terms in the decomposition of welfare, there are thus only four that are relevant here.

[38]Under these specifications, the consumer's interior optimal level of healthcare utilization is invariant to income, and thus as long as the consumer can afford her interior optimum in a given state, that amount will coincide with the first-best level of healthcare utilization. The income effect enters in the decomposition

III.B Results

Table 2 presents our main empirical results. The first three columns decompose healthcare utilization into its three components: non-moral-hazard utilization $m^N(l)$, under-utilization $\mathbb{E}_l[m^{FB}(l) - m^N(l)]$, and over-utilization $\mathbb{E}_l[m^*(l,x) - m^{FB}(l)]$. Each component is expressed as a share of total healthcare spending $\mathbb{E}_l[m^*(l,x)]$, such that the three columns sum to one within each row. The subsequent four columns decompose welfare $V(x)$ into the four relevant terms: risk protection value absent moral hazard utilization (Ψ^N), incremental risk protection value from under-utilization and from over-utilization ($\Delta\Psi^{under}$ and $\Delta\Psi^{over}$), and the social cost of over-utilization ($\Delta SCMH^{over} = SCMH$), presented as negative for clarity. Each column under "Decomposition of welfare" reports the proportion of total welfare represented by that term, on average across consumers. For example, an entry under Ψ^N represents the average value of $\frac{\Psi^N}{V}$ across consumers. These four columns thus also sum to one within each row.

Panel A reports the decomposition at the population level for each of the four specifications. Panel B reports the decomposition for subgroups of the Marone and Sabety (2022) population, based on consumer characteristics most relevant to the decomposition. Appendix Table A.1 presents subgroup decompositions for the other specifications. The first grouping is based on the extent to which consumers under-utilize healthcare absent insurance, as measured by their expected dollar amount of under-utilization. This stratification primarily captures differences in consumer income relative to their potential health shocks. Consumers with a high level of under-utilization are exposed, absent insurance, to substantial risk of ending up in a health state in which they would use less healthcare than they would have wanted to do in a first-best world. The second grouping is based on the extent to which consumers over-utilize healthcare with insurance, as measured by their expected amount of over-utilization under full insurance. This stratification captures differences in the over-utilization parameter ω.

In all four specifications, the average proportion of healthcare utilization attributable to non-moral hazard spending is between 80 and 85 percent. The remainder is almost entirely over-utilization, with the under-utilization gap close to zero on average in every population. The corresponding picture for welfare is similarly skewed: the average proportion of welfare achieved by the no-behavioral-response benchmark lies between 93 and 119 percent. For the average consumer in these populations, the risk protection value of insurance *absent* any ability to change utilization behavior thus accounts for the overwhelming majority of the welfare generated by insurance.

thus solely through the budget corner: in states where the consumer cannot afford her efficient healthcare, holding her behavior fixed leaves an income-mismatch loss $SCMH^N$, but allowing her to re-optimize removes it exactly—which is why $SCMH^N$ and $\Delta SCMH^{under}$ sum to zero.

Table 2. Decomposition of Welfare

	Decomposition of healthcare utilization			Decomposition of welfare				Relative change in welfare		
	Non-MH utiliz.	Under-utiliz.	Over-utiliz.	Ψ^N	$\Delta\Psi^{und.}$	$\Delta\Psi^{ovr.}$	$-SCMH$	Efficient coverage	Prevent over-util.	Prevent MH-util.
Panel A.										
(1) Marone and Sabety	0.81	<0.01	0.18	0.94	0.10	–	−0.04	<0.01	0.04	−0.06
(2) Einav et al. (Addv.)	0.81	<0.01	0.18	0.93	0.09	–	−0.02	<0.01	0.02	−0.07
(3) Einav et al. (Mult.)	0.80	<0.01	0.19	1.03	0.03	0.05	−0.11	0.02	0.06	0.02
(4) Ho and Lee	0.85	<0.01	0.14	1.19	0.02	<0.01	−0.22	0.05	0.21	0.19
Panel B. [Marone and Sabety]										
Tertile of under-utiliz.										
Lowest	0.82	–	0.18	1.04	–	–	−0.04	0.01	0.04	0.04
Middle	0.77	<0.01	0.23	0.84	0.18	–	−0.03	<0.01	0.03	−0.16
Highest	0.81	0.05	0.14	0.40	0.62	–	−0.01	<0.01	0.01	−0.62
Tertile of over-utiliz.										
Lowest	0.85	<0.01	0.15	0.90	0.12	–	−0.02	<0.01	0.02	−0.11
Middle	0.80	<0.01	0.20	0.93	0.11	–	−0.03	<0.01	0.03	−0.08
Highest	0.79	<0.01	0.21	0.99	0.06	–	−0.06	0.02	0.06	−0.01

Notes: This table shows the decomposition of healthcare utilization and welfare generated by full insurance. The first three columns decompose total expected healthcare spending into its three components, each expressed as a share of the total. The next four columns report the four relevant components of the welfare decomposition in Equation 10: Ψ^N, $\Delta\Psi^{under}$, $\Delta\Psi^{over}$, and $\Delta SCMH^{over} = SCMH$. Each entry represents the proportion of total welfare V represented by that term, on average across consumers. The final three columns report the relative changes in welfare that would occur in three scenarios: (a) if consumers were enrolled in their efficient level of coverage, rather than full insurance; (b) if over-utilization were prevented; and (c) if all moral hazard utilization were prevented. A relative change of 0.02 means a 2% increase in welfare. Panel A reports the population-average decomposition for each of the four specifications in Table 1. Panel B reports the decomposition for subgroups of the Marone and Sabety (2022) population.

The contribution of moral hazard utilization is captured by the remaining three welfare terms. The risk protection value of moral hazard utilization, $\Delta\Psi^{under} + \Delta\Psi^{over}$, varies across specifications. In the "additive" specifications (1) and (2), under-utilization is the only source of the risk protection from moral hazard utilization, and it makes up a sizable amount despite being a very small fraction of utilization itself.[39] In the latter two specifications, both the under- and over-utilization channels play a small role. The negative contribution to welfare from the social cost of moral hazard is reported under $-SCMH$, and ranges from between 2 to 22 percent of welfare. These amounts represent the social cost of over-utilization induced by full insurance. Heterogeneity in these magnitudes across specifications reflects both differences in the parameterization themselves (between (1)/(2) versus (3)/(4)) as well as between the underlying demographics and parameter estimates. Netting this social cost against

[39] The additive specifications imply a risk protection value from over-utilization that is zero because the consumer surplus generated by over-utilization, $\frac{1}{2}\omega(1-\bar{c})^2$, is the same in every health state. Under the multiplicative form, in contrast, this surplus is proportional to the health state l, so over-utilization is worth more in sicker states, and can thus provide risk protection.

$\Delta\Psi^{under} + \Delta\Psi^{over}$, moral hazard utilization is on net welfare-improving in the additive specifications, raising welfare by roughly 6 and 7 percent respectively, and welfare-reducing in the multiplicative specifications, lowering it by 2 and 19 percent, where the larger social cost dominates. This reversal is most pronounced under the Ho and Lee (2023) estimates, for which the population's overall value of insurance is smallest.[40]

Panel B shows that these averages conceal substantial heterogeneity across consumers. The risk protection value of moral hazard utilization is concentrated among the consumers who under-utilize the most absent insurance. For consumers with zero under-utilization (the *Lowest* tertile), Ψ^N accounts for the entirety of the risk protection value provided by insurance. These consumers have a high income relative to their potential health outcomes, and thus are predicted to never experience a health outcome in which, absent insurance, they would want to use a greater amount of healthcare than they could afford. Among consumers in the *Middle* tertile of under-utilization, the absolute level of under-utilization remains small—on average less than 1 percent of healthcare utilization—but nonetheless still generates a sizable amount of risk protection value. The ability to avoid low-probability tail events in which insufficient healthcare is accessible thus appears highly valuable. In the *Highest* tertile of under-utilization, consumers still experience a modest absolute level of under-utilization (5 percent of total healthcare spending), but the cost of this under-utilization is now sufficient to imply an enormous risk protection value of avoiding it. More than half of the risk protection value from insurance is now directly generated by the consumer's ability to increase her healthcare utilization in response to insurance. For these consumers, the value of being able to increase healthcare use up to the first-best level, rather than only to transfer non-health consumption across states, accounts for the majority of the welfare that insurance provides.

Sorting the same consumers by their propensity to over-utilize healthcare (measured by the parameter ω) tells a complementary story. Consumers with a higher ω have more elastic demand and over-utilize more care under full insurance, and the social cost of that over-utilization rises accordingly. Across tertiles, the social cost of moral hazard climbs from 2 percent of welfare for the consumers with the lowest ω, to 6 percent for those with the highest. In this additive specification, over-utilization carries no offsetting risk protection value, and is therefore pure deadweight loss.[41]

[40]This reversal reflects low risk aversion rather than unusually wasteful over-utilization: over-utilization is in fact a *smaller* share of healthcare spending under specification (4) than in any other. Ho and Lee (2023) estimate a low coefficient of absolute risk aversion in their population, smaller than in the Marone and Sabety (2022) and Einav et al. (2013) populations by factors of roughly 7 and 13, respectively. Because the risk protection value of moral hazard utilization vanishes as $\psi \to 0$, while its social cost is invariant to ψ, low risk aversion shrinks the former toward zero and leaves the latter unchanged.

[41]Under the multiplicative specifications (3) and (4), in contrast, over-utilization is more valuable in sicker

The final three columns present the relative change in welfare that would arise under three scenarios: (a) if consumers were enrolled in their efficient coverage level (second-best contract), rather than in full insurance;[42] (b) if over-utilization were prevented; and (c) if all moral hazard utilization were prevented. In the first scenario (*Efficient coverage*), enrolling consumers in their efficient contract rather than full insurance will, of course, weakly increase welfare. Our estimates suggest that it will not, however, have a large impact quantitatively, echoing findings reported in Marone and Sabety (2022) and Ho and Lee (2023). The (small) magnitude of welfare effects in this first scenario provides a benchmark for thinking about the relative importance of the others.

As is usual, consumers who have higher efficient coverage levels tend to face more uncertainty about their health status and have higher risk aversion parameters (ψ). They do not, however, necessarily have a smaller increase in healthcare utilization in response to insurance, in particular among consumers with lower incomes. Appendix Figure A.3 shows that among consumers in the bottom third of the income distribution, the relationship between efficient coverage level and expected moral hazard utilization is U-shaped. At relatively low levels of moral hazard utilization, the relationship with efficient coverage level is negative (as expected). But for consumers with high levels of moral hazard utilization, the relationship becomes positive. A primary indication for why moral hazard utilization may be quite high is if the consumer were substantially under-utilizing healthcare absent insurance. In such case, moral hazard utilization will tend to be worth its cost, increasing the efficient coverage level.[43]

The second scenario (*Prevent over-util.*) prevents over-utilization of healthcare, while still providing full insurance (corresponding to the second hypothetical discussed in Section II.B [p. 20] above). This counterfactual removes both the risk protection value of over-utilization, $\Delta\Psi^{over}$, and its social cost, $SCMH$. Under the additive specifications, (1) and (2), over-utilization provides no risk protection, so preventing it can only raise welfare—by 4 and 2 percent on average, with no consumer made worse off. Under multiplicative specification (3), eliminating over-utilization is not unequivocally good, because over-utilization there transfers resources toward sicker states and so provides some risk protection. Welfare still rises on average, by 6 percent, but for 21 percent of consumers it falls: the resource transfer that over-

states and can there provide some risk protection, so $\Delta\Psi^{over}$ is positive and rises with ω.

[42] A consumer's efficient level of coverage is defined as the welfare-maximizing contract from among a pre-specified set (Zeckhauser, 1970; Marone and Sabety, 2022). The set of contracts we use is shown in Appendix Figure B.1.

[43] This pattern has been absent in prior papers because the positivity constraint on non-health consumption has not been enforced. In effect, this allows the consumer to circumvent her budget constraint and achieve her first-best level of healthcare utilization regardless of whether she can afford it, leaving no possibility for under-utilization (see, e.g., Einav et al. (2013), Equation 2, page 184; Azevedo and Gottlieb (2017), page 73; Marone and Sabety (2022), page 309).

utilization accomplishes is valuable enough for these consumers that preventing it makes them worse off. Specification (4) is also multiplicative, but its estimated value of risk protection from over-utilization is negligible, so preventing over-utilization raises welfare for every consumer (by 21 percent on average).

Finally, the third scenario (*Prevent MH-util.*) eliminates moral hazard utilization entirely, both under-utilization and over-utilization (corresponding to the first hypothetical discussed in Section II.B [p. 19] above). Full insurance is still provided, but healthcare utilization is held fixed at the same level that prevailed absent insurance. Here, substantial risk protection value is lost. Under the additive specifications, (1) and (2), the risk protection value of moral hazard utilization on average exceeds its social cost. Preventing it thus reduces welfare, by 6 and 7 percent on average. Under the multiplicative specifications, (3) and (4), the social cost of moral hazard utilization on average exceeds its risk protection value. Preventing it thus raises welfare, by 2 and 19 percent on average. Panel B shows that preventing consumers from changing their behavior in response to insurance is especially costly for those with a substantial degree of under-utilization absent insurance. In specification (1), welfare for the average consumer would be higher by 7 percent under their efficient contract than under a full insurance contract that prevented them from changing their behavior. For consumers in the highest tertile of under-utilization, welfare would instead be a striking 62 percent higher. At the same time as the efficient contract delivers higher welfare, the cost to the insurer of providing it is also lower, on average by 26 percent. A *lower* fair premium (on a lower-coverage contract that allows moral hazard utilization) thus yields *higher* risk protection (than a full insurance contract that prevents it), implying that behavioral responses to insurance can be an efficient means of producing risk protection.

We draw two main conclusions from this analysis. First, it is clear that the most important driver of welfare from health insurance in these studies is the value of risk protection that arises even absent moral hazard utilization (Ψ^N). The ability to transfer non-health consumption across states, holding behavior fixed, is almost all that matters for most consumers in these populations. The welfare effects of moral hazard utilization—both positive and negative—are an order of magnitude smaller than the welfare effect of gaining insurance at all (though less so under the Ho and Lee (2023) estimates). Second, these welfare effects are nonetheless economically significant when measured relative to other counterfactual scenarios of interest. In specification (1), eliminating ex-ante asymmetric information (enabling the assignment of all consumers to their efficient coverage level) would increase welfare by less than a quarter as much as eliminating ex-post asymmetric information (enabling all over-utilization to be prevented). Moreover, eliminating *both* ex ante and ex post asymmetric information (enabling the first-best) would produce a welfare benefit about equivalent in magnitude to the welfare *harm* produced by a policy that eliminated moral hazard utilization entirely. The welfare effects of limiting moral hazard utilization, either in part or in whole, are thus very much on par with those of other policy-relevant objects.

An important caveat to both of these conclusions is that there is substantial heterogeneity across consumers. Consumers with lower incomes, for example, derive a substantially greater value from the ability to use a greater amount of healthcare with insurance, rather than just to increase their non-health consumption in sick states. Among those in the lowest third of the income distribution, moral hazard utilization accounts for between 7 and 29 percent of the value of risk protection provided by full insurance. In contrast, among those in the highest third of income, it accounts for essentially none of that value—at most 6 percent across the four specifications.

These quantitative results are of course specific to the parameterizations of utility and the estimates of structural primitives we have used. Our empirical analysis can thus be viewed as highlighting that existing parameterizations and estimates of the primitives governing demand for healthcare and health insurance fall within the space of theoretical possibilities where moral hazard utilization plays an important role in driving the risk protection value of insurance. We view further investigation of these primitives to be an important direction for future work.

IV Discussion

Our analysis offers a new perspective on a number of healthcare policy issues. First, it challenges the conventional view that in designing public health insurance, greater cost-sharing should be imposed on services for which demand is more responsive to coverage. While this logic is rooted in the goal of discouraging over-utilization, it overlooks the fact that part of the increase in utilization caused by insurance is efficient. If there are strong income effects in the demand for healthcare, or if individuals are resource constrained in some health states absent insurance, efforts to contain over-utilization may unwittingly exacerbate under-utilization, ultimately doing more harm than good. Our analysis suggests that healthcare services which are expensive and rarely needed are more likely to be under-utilized than over-utilized, suggesting caution in placing any substantial cost-sharing on their use.

A related implication is the importance of the consumer's income level in relation to their efficient levels of healthcare utilization. As discussed above, if the consumer is likely to not be able to afford her first-best level of utilization in many health states, insurance sufficiently generous to allow her to do so will avoid welfare losses from under-utilization. As long as bad health outcomes are unlikely, lower-income individuals are therefore all-else-equal more likely to have higher efficient levels of insurance coverage. Using income as a tag on which to differentiate the terms of public insurance coverage, as is done in many high-income countries, is consistent with this idea. Providing higher-coverage public insurance to lower-income individuals can thus be grounded directly in an efficiency argument, and need not be predicated on distributional considerations alone.

Our model also provides a way to interpret some features of market-based health insurance, for example prior authorization, case management, referral requirements, utilization review, and vertical integration. All of these tools can be considered to some extent efforts by the insurer to verify the realized state of the world, and thus to enact the first-best level of healthcare utilization while preventing over-utilization.[44] Managed care plans (HMOs) can be viewed as contracts under which consumers willingly grant the insurer more tools with which to monitor and verify the realized health state, thus further reducing the extent of over-utilization (and the premium alongside it). Different consumers will perceive the costs of this health-state verification exercise differently. One consumer may prefer to avoid it by paying the premium increase necessary to fund her own over-utilization (for example, under a PPO plan). Another may happily submit to hassle costs in exchange for a lower premium.

Finally, our model makes plain that unlike in the case of adverse selection, the market failure that leads to moral hazard in health insurance cannot be remedied by any special powers held by a benevolent regulator. In the face of ex-ante asymmetric information, the government can combat adverse selection by using the power of taxation, to compel insurance take-up by consumers otherwise unwilling to pay above-actuarially fair rates. But this power is not of use in the face of ex-post asymmetric information. Though the government can imitate tools used in private market, we see no strong reason to think the government would have a technological advantage over the market in designing contracts that elicit private information from consumers about realized health status. This observation might be placed in the "pro" column of an evaluation of privately-administered (versus publicly-administered) tax-funded health insurance.[45]

Beyond these policy issues, the logic of our analysis, though developed for ex-post moral hazard, extends also to the case of ex-ante moral hazard. In the canonical principal–agent model (Holmström, 1979), a risk-averse agent chooses unobservable effort that, together with an exogenous state of the world (e.g., the strength of market demand), determines her output, and a risk-neutral principal offers a wage contract that can condition on output but not effort. If effort were contractible, the first-best contract would specify its efficient level and insure the

[44] On this point, see also Albert Ma and Riordan (2002).

[45] See (Einav, Finkelstein and Polyakova, 2018, Section IV) for related discussion.

agent fully against market risk.[46] Instead, the optimal feasible contract must leave the agent exposed to some output risk in order to preserve her incentive to work, and her effort typically falls below its first-best level ("shirking"). But now consider the comparison between the second-best contract and no wage contract at all, where the agent is "self-employed." Bearing the full market risk alone, she exerts effort not only for its expected return, but also to limit her exposure to bad outcomes, potentially working herself to the bone to ensure she can feed her family no matter what market conditions arise. Her effort may thus exceed even its first-best level.[47] An observer who compares effort under the wage contract with effort under self-employment—the analog of comparing insured with uninsured healthcare utilization—might attribute the entire difference to shirking, but only the gap between the first-best and wage-contract levels warrants that label. The remainder is the efficient unwinding of the agent's precautionary over-exertion, and is risk protection delivered through her behavior, just as in our model. Porting our welfare decomposition to this environment would be an interesting extension of our framework.

V Conclusion

This paper provides a novel treatment of a standard health insurance model with ex-post moral hazard. The model is substantially general, and likely applies to a number of other insurance (or redistribution) settings. Given the presence of multiple consumption goods (a necessary condition for ex-post moral hazard), we recast the consumer's fundamental risk to be over her marginal indirect utility of income, rather than over consumption directly. This approach allows a number of novel transparencies, including (i) a formal characterization of the first-best contract and the socially efficient level of healthcare utilization, (ii) a decomposition of moral hazard utilization into under-utilization versus over-utilization, theoretically grounding a distinction between what one might call "high value" versus "low value" care, and (iii) a decomposition of the welfare generated by insurance into terms that isolate the consumer's

[46]Note that the first best can equivalently be achieved by contracting on the state rather than on effort. The agent could then hedge the market risk directly (say, in an Arrow–Debreu securities market) while choosing her effort freely. This formulation of the first best corresponds to the one used in our analysis, where transfers condition on the health state and healthcare utilization is chosen freely.

[47]Note that the agent's excess effort is a form of costly self-insurance in the sense of Ehrlich and Becker (1972), an action that reduces the severity of adverse states, and for which market insurance is a substitute. The value of reducing "excess effort" does not operate through a liquidity channel; the agent is missing not a capital market, with which to smooth resources over time, but a securities market, with which to transfer them across states. The example is thus distinct from, though complementary to, the decomposition of Chetty (2008), in which part of the behavioral response to unemployment insurance reflects a relaxed liquidity constraint rather than an incentive distortion.

ability to change her behavior.

We find that a consumer's ability to change her behavior in response to insurance can play a central role in the ability of insurance to protect her from risk. Part of the increase in utilization caused by insurance reflects choices the consumer would also have made in a first-best world, but not absent insurance. This reduction in "under-utilization" is made possible by the real income that insurance transfers towards sick states. Though this care does not appear to be valued at cost from an ex post perspective (i.e., under initial income w_0), it *is* valued at cost from an ex ante perspective (i.e., under first-best income w^{FB}). Quantitatively, we find that standard parameterizations and estimates in the literature imply an economically significant role for this channel. Though the risk protection achievable without any change in behavior accounts for the large majority of the value of insurance for the average consumer, the risk protection value of moral hazard utilization accounts for as much as 10 percent of welfare under full insurance, and among the consumers most likely to under-utilize healthcare absent insurance, more than half. Preventing consumers from changing their behavior in response to insurance would lower healthcare spending, but would also forfeit this risk protection, on net reducing welfare in two of our simulated populations.

We conclude from our analysis that care should be taken not to conflate moral hazard utilization with the social cost of moral hazard. The two concepts differ in units—the former being measured in units of healthcare utilization, the latter in units of initial income (resources). Both are positive in expectation, but only the latter is a measure of welfare loss. Expected moral hazard utilization carries no welfare interpretation of its own; its contribution to welfare is the net of the risk protection it generates and the deadweight loss it creates, and can take either sign (Equation 10). The deadweight loss is moreover not even proportional to the amount of expected moral hazard utilization. Incremental units of healthcare the consumer would have purchased at undistorted prices—given the income she would have allocated to that state in a first-best world—create no deadweight loss at all. The same observed increase in utilization can be mostly efficient for one consumer and mostly wasteful for another. Policies aimed at constraining moral hazard utilization in the name of efficiency may thus be eliminating more value than waste.

There are a number of important issues with which this paper does not engage. First, while our theoretical decomposition places few restrictions on the form of consumer utility, our empirical analysis inherits the parameterizations of utility under which the available estimates were produced. In all of these specifications, the consumer has constant absolute risk aversion and preferences that are quasilinear in non-health consumption, meaning that income affects the demand for healthcare only through the budget constraint. We view richer estimation of these primitives as an important direction for future work. Second, our model is static. It is best interpreted as a model of a one-shot lifetime, in which the consumer's income constitutes her lifetime resources and the health state reflects a single lifetime draw (for example, whether she was unlucky enough to be born with a genetic disorder). There is no modeling of decisions within the lifetime. A natural extension would introduce these dynamics, incorporating saving and borrowing, the timing and recurrence of health shocks, and the distinction between healthcare a consumer could never afford and healthcare she cannot finance in the short run. Lastly, healthcare utilization decisions in our model are assumed to be a single, unidimensional choice over a continuous support. Reality is lumpier. These decisions are also rarely made by patients alone, but rather arise from a joint decision-making process between patients and physicians, shaped by norms, incentives, and asymmetric information. Our analysis abstracts from these issues. As such, it is only an imperfect approximation of a fundamentally discrete and institutionally mediated process. While these simplifications facilitate tractability, they also highlight the need for further research.

References

Albert Ma, C.-T. and M. H. Riordan (2002). Health insurance, moral hazard, and managed care. *Journal of Economics & Management Strategy 11*(1), 81–107.

Arrow, K. J. (1973). Welfare analysis of changes in health coinsurance rates. R-1281-oeo, Rand Corporation, Santa Monica, CA.

Azevedo, E. M. and D. Gottlieb (2017). Perfect Competition in Markets With Adverse Selection. *Econometrica 85*(1), 67–105.

Baicker, K., S. Mullainathan, and J. Schwartzstein (2015). Behavioral hazard in health insurance. *The Quarterly Journal of Economics 130*(4), 1623–1667.

Baily, M. N. (1978). Some aspects of optimal unemployment insurance. *Journal of Public Economics 10*, 379–402.

Besley, T. J. (1988). Optimal reimbursement health insurance and the theory of ramsey taxation. *Journal of Health Economics 7*(4), 321–336.

Cardon, J. H. and I. Hendel (2001). Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *The RAND Journal of Economics 32*(3), 408–427.

Cawley, J. and C. J. Ruhm (2011). The economics of risky health behaviors. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 2, Chapter 3, pp. 95–199. Elsevier.

Chade, H., V. Marone, A. Starc, and J. Swinkels (2023, December). Multidimensional screening and menu design in health insurance markets. Working Paper.

Chandra, A., E. Flack, and Z. Obermeyer (2024, November). The health costs of cost sharing. *The Quarterly Journal of Economics 139*(4), 2037–2082.

Chetty, R. (2006). A general formula for the optimal level of social insurance. *Journal of Public Economics 90*(10–11), 1879–1901.

Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy 116*(2), 173–234.

Chetty, R. and A. Finkelstein (2013). Social insurance: Connecting theory to data. In A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Volume 5 of *Handbook of Public Economics*, pp. 111–193. Elsevier.

Chipman, J. S. and J. C. Moore (1980). Compensating variation, consumer's surplus, and welfare. *The American Economic Review 70*(5), 933–949.

Crew, M. (1969). Coinsurance and the welfare economics of medical care. *American Economic Review 59*(5), 906–908.

Cutler, D. M. and R. J. Zeckhauser (2000, January). The Anatomy of Health Insurance*. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1 of *Handbook of Health Economics*, pp. 563–643. Elsevier.

Dávila, E. and A. Schaab (2025). Welfare assessments with heterogeneous individuals. *Journal of Political Economy 133*(9).

de Meza, D. (1983). Health insurance and the demand for medical care. *Journal of Health Economics 2*(1), 47–54.

Deshpande, M. and L. M. Lockwood (2022, July). Beyond health: Nonhealth risk and the value of disability insurance. *Econometrica 90*(4), 1781–1810.

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. *Journal of Political Economy 80*(4), 623–648.

Einav, L. and A. Finkelstein (2018, 05). Moral Hazard in Health Insurance: What We Know and How We Know It. *Journal of the European Economic Association 16*(4), 957–982.

Einav, L., A. Finkelstein, and M. Polyakova (2018). Private provision of social insurance: Drug-specific price elasticities and cost sharing in medicare part d. *American Economic Journal: Economic Policy 10*(3), 122–153.

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.

Ericson, K. M., J. G. Jaspersen, and J. R. Sydnor (2025, April). Moral hazard under liquidity constraints. Working Paper 33648, National Bureau of Economic Research.

Feldstein, M. (1973). The welfare loss of excess health insurance. *Journal of Political Economy 81*(2), 251–80.

Finkelstein, A., N. Hendren, and E. F. P. Luttmer (2019). The value of medicaid: Interpreting results from the oregon health insurance experiment. *Journal of Political Economy 127*(6), 2836–2874.

Frick, K. D. and M. E. Chernew (2009). Beneficial moral hazard and the theory of the second best. *Inquiry: A Journal of Medical Care Organization, Provision and Financing 46*(2), 229–240.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Harberger, A. C. (1971). Three basic postulates for applied welfare economics: An interpretive essay. *Journal of Economic Literature 9*(3), 785–797.

Hausman, J. A. (1981). Exact consumer's surplus and deadweight loss. *The American Economic Review 71*(4), 662–676.

Hicks, J. R. (1943). The four consumer's surpluses. *The Review of Economic Studies 11*(1), 31–41.

Ho, K. and R. S. Lee (2023). Health insurance menu design for large employers. *RAND Journal of Economics 54*(4), 598–637.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics 10*(1), 74–91.

Kang, Z. Y. and S. Vasserman (2025). Robustness measures for welfare analysis. *American Economic Review 115*(8), 2449–2487.

Kesternich, I., A. Romahn, J. Van Biesebroeck, and M. Van Damme (2025, May). Cash or care? insights from the german long-term care system. CESifo Working Paper 11875, CESifo, Munich. Available at SSRN: https://ssrn.com/abstract=5254204.

King, M. A. (1983). Welfare analysis of tax reforms using household data. *Journal of Public Economics 21*(2), 183–214.

Lieber, E. M. J. and L. M. Lockwood (2019). Targeting with in-kind transfers: Evidence from medicaid home care. *American Economic Review 109*(4), 1461–1485.

Manning, W. G. and M. Marquis (1996). Health insurance: The tradeoff between risk pooling and moral hazard. *Journal of Health Economics 15*(5), 609–639.

Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, A. Leibowitz, and M. S. Marquis (1987). Health insurance and the demand for medical care: Evidence from a randomized experiment. *The American Economic Review 77*(3), 251–277.

Marone, V. R. and A. Sabety (2022, January). When should there be vertical choice in health insurance markets? *American Economic Review 112*(1), 304–42.

Marshall, J. M. (1976). Moral hazard. *American Economic Review 66*(5), 880–890.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. New York: Oxford University Press.

McKenzie, L. W. (1957). Demand theory without a utility index. *The Review of Economic Studies 24*(3), 185–189.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(2), 175–208.

Mommaerts, C. (2025). Long-term care insurance and the family. *Journal of Political Economy 133*(1), 1–52.

Newhouse, J. P. (1978). *The Economics of Medical Care*. Reading, MA: Addison-Wesley.

Newhouse, J. P. (2006). Reconsidering the moral hazard–risk avoidance tradeoff. *Journal of Health Economics 25*(6), 1005–1014.

Nyman, J. A. (1999a, December). The economics of moral hazard revisited. *Journal of Health Economics 18*(6), 811–824.

Nyman, J. A. (1999b). The value of health insurance: the access motive. *Journal of Health Economics 18*(2), 141–152.

Nyman, J. A. (2003). *The Theory of Demand for Health Insurance*. Stanford, CA: Stanford University Press.

Nyman, J. A. (2007). American health policy: Cracks in the foundation. *Journal of Health Politics, Policy and Law 32*(5), 759–783.

Nyman, J. A. (2008, November). Health insurance theory: The case of the missing welfare gain. *The European Journal of Health Economics 9*(4), 369–380.

Nyman, J. A., C. Koc, B. E. Dowd, E. McCreedy, and H. M. Trenz (2018). Decomposition of moral hazard. *Journal of Health Economics 57*, 168–178.

Oregon Public Employees Retirement System (2024, May). PERS by the Numbers. https://www.oregon.gov/pers/Documents/General-Information/PERS-by-the-Numbers.pdf. Accessed: 2025-07-17.

Pauly, M. V. (1968). The economics of moral hazard: Comment. *American Economic Review 58*(3), 531–537.

Pauly, M. V. (1983). More on moral hazard. *Journal of Health Economics 2*(1), 81–85.

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.

Samuelson, P. A. (1974). Complementarity: An essay on the 40th anniversary of the Hicks-Allen revolution in demand theory. *Journal of Economic Literature 12*(4), 1255–1289.

Schwartz, A. L., B. E. Landon, A. G. Elshaug, M. E. Chernew, and J. M. McWilliams (2014). Measuring low-value care in medicare. *JAMA Internal Medicine 174*(7), 1067–1076.

Slutsky, E. E. (1915). On the theory of the budget of the consumer, translated by o. ragusa. *Giornale degli Economisti e Annali di Economia 71(Anno 125)*(2/3), 173–200.

U.S. Census Bureau (2023). American community survey 1-year estimates. https://www.census.gov/programs-surveys/acs.

Zeckhauser, R. (1970). Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory 2*(1), 10–26.

Appendix A Theoretical Appendix

The model in the main text is deliberately general. Signing certain objects definitively naturally requires additional structure. Our aim in this appendix is to isolate clearly the set of conditions that imply each result. Two such conditions recur repeatedly, and we state them together here at the outset. As in the main text, we use increasing and decreasing in the weak sense of non-decreasing and non-increasing, with "strictly" added when needed. Similarly, we use positive and negative in the weak sense of non-negative and non-positive, with "strictly" added when necessary. All expectations taken over the health state are assumed to be finite.

Condition 1 (**Marginal Propensity to Spend on Healthcare is Higher when Sicker or Richer**). *The marginal propensity to spend on healthcare out of income at undistorted prices, $\frac{\partial m^*(l,x_0,w)}{\partial w}$, is increasing in both the health state l and the income level w.*

Substantively, Condition 1 orders the marginal propensity to spend on healthcare along two margins: a marginal dollar of income is devoted (weakly) more to healthcare when the consumer is sicker, and also when her income is higher. In terms of the income expansion path at undistorted prices drawn in Figure 1, the first margin says that the path rotates clockwise—becoming flatter—as the health state worsens, while the second says that any single path is concave in the (y, m) plane.

Condition 2 (**First-Best Income is Higher when Sicker**). *$w^{FB}(l)$ is increasing in l.*

Condition 2 says that the consumer's optimal redistribution of income, in a first-best world, moves income from healthier to sicker states. This condition is the canonical motivation for health insurance. When it fails, the consumer simply does not have more valuable uses of income in sicker states, and thus has no need for health insurance.

A.1 Properties on Decomposition of Moral Hazard Utilization

Lemma 1 (**Positive Expected Moral Hazard Utilization**). *Under Condition 1, expected moral hazard utilization is positive: $\mathbb{E}_l[m^*(l,x)] \geq \mathbb{E}_l[m^N(l)]$ for every contract $x \in \mathcal{X}$.*

Proof. Consider a non-null contract x with actuarially fair premium $\bar{p}(x) = \mathbb{E}_l[m^*(l,x) - c(m^*(l,x);x)] > 0$. Let $r(l,x) \equiv y^*(l,x) + m^*(l,x)$ denote the total resources consumed under contract x in state l, and write $\Delta(l,x) \equiv r(l,x) - w_0$ to denote the resources transferred towards a given state. Actuarial fairness gives $\mathbb{E}_l[\Delta(l,x)] = 0$.

Moral hazard utilization can be decomposed into an income effect (inclusive of the fair premium compensation) and a substitution effect:[48]

$$m^*(l,x) - m^N(l) \;=\; \underbrace{\big[m^*(l,x_0,r(l,x)) - m^*(l,x_0,w_0)\big]}_{\text{income effect}} \;+\; \underbrace{\big[m^*(l,x) - m^*(l,x_0,r(l,x))\big]}_{\text{substitution effect}}.$$

The substitution effect is positive in every state: At $m^*(l,x)$, the consumer's marginal rate of substitution between m and y satisfies $u_m/u_y = c_m(m^*(l,x);x) \leq 1$. At $m^*(l,x_0,r(l,x))$, $u_m/u_y = 1$. Both bundles lie on the budget line $y + m = r(l,x)$; since the marginal rate of substitution is decreasing in m along it, $m^*(l,x) \geq m^*(l,x_0,r(l,x))$.[49]

The income effect is positive in expectation: By the fundamental theorem of calculus,

$$m^*(l,x_0,r(l,x)) - m^*(l,x_0,w_0) \;=\; \bar{m}_w(l,x)\cdot\Delta(l,x),$$

and

$$\bar{m}_w(l,x) \;\equiv\; \int_0^1 \frac{\partial m^*}{\partial w}\big(l,\, x_0,\, w_0 + t\,\Delta(l,x)\big)\,dt.$$

The resource transfer $\Delta(l,x)$ is increasing in l since the insurer's realized cost is increasing in m and $m^*(l,x)$ is increasing in l (a monotone comparative static of $\partial_l(u_m/u_y) > 0$). The path-average $\bar{m}_w(l,x)$ is thus also increasing in l: By Condition 1, at each t, the integrand $\frac{\partial m^*}{\partial w}(l,x_0,w_0+t\Delta(l,x))$ is increasing in l both directly and through the increasing income path $w_0 + t\Delta(l,x)$. Then, since $\mathbb{E}_l[\Delta(l,x)] = 0$,

$$\mathbb{E}_l\big[m^*(l,x_0,r(l,x)) - m^*(l,x_0,w_0)\big] \;=\; \mathrm{Cov}_l\big(\bar{m}_w(l,x),\, \Delta(l,x)\big) \;\geq\; 0,$$

by Chebyshev's sum inequality. Combining, $\mathbb{E}_l[m^*(l,x) - m^N(l)] \geq 0$. □

Lemma 2 (**Positive Expected Under-Utilization**). *Under Condition 1 and Condition 2, under-utilization is positive on average:* $\mathbb{E}_l[m^{FB}(l)] \geq \mathbb{E}_l[m^N(l)]$.

Proof. Both $m^{FB}(l)$ and $m^N(l)$ lie on the income expansion path at undistorted prices and differ only in the income at which they are evaluated: $m^N(l) = m^*(l,x_0,w_0)$ and $m^{FB}(l) =$

[48] This decomposition parallels the utility-compensated decomposition described in footnote 21 and Appendix Figure A.1, with the intermediate bundle taken at equal total resources rather than at equal utility.

[49] At a corner solution $m^*(l,x_0,r(l,x)) = 0$, the inequality $m^*(l,x) \geq 0$ holds trivially. At a corner solution where the insured bundle exhausts income ($y^*(l,x) = 0$), the marginal rate of substitution at the insured bundle may strictly exceed $c_m(m^*(l,x);x)$; but there $m^*(l,x) = r(l,x)$ is the maximal healthcare level on the budget line, so $m^*(l,x) \geq m^*(l,x_0,r(l,x))$ directly.

$m^*(l, x_0, w^{FB}(l))$. Write $\Delta w^{FB}(l) \equiv w^{FB}(l) - w_0$. Actuarial fairness of the first-best allocation gives $\mathbb{E}_l[\Delta w^{FB}(l)] = 0$.

By the fundamental theorem of calculus,

$$m^{FB}(l) - m^N(l) \ = \ \bar{m}_w(l) \cdot \Delta w^{FB}(l),$$

and

$$\bar{m}_w(l) \ \equiv \ \int_0^1 \frac{\partial m^*}{\partial w}\big(l,\, x_0,\, w_0 + t\,\Delta w^{FB}(l)\big)\, dt.$$

The transfer $\Delta w^{FB}(l)$ is increasing in l by Condition 2. The path-average $\bar{m}_w(l)$ is thus also increasing in l: By Condition 1, at each t, the integrand $\frac{\partial m^*}{\partial w}(l, x_0, w_0 + t\Delta w^{FB}(l))$ is increasing in l both directly and through the increasing income path $w_0 + t\Delta w^{FB}(l)$. Then, since $\mathbb{E}_l[\Delta w^{FB}(l)] = 0$,

$$\mathbb{E}_l\big[m^{FB}(l) - m^N(l)\big] \ = \ \mathrm{Cov}_l\big(\bar{m}_w(l),\, \Delta w^{FB}(l)\big) \ \geq \ 0,$$

by Chebyshev's sum inequality. □

Lemma 3 (**Positive Expected Over-Utilization for Sufficiently Generous Insurance**). *Suppose preferences satisfy bliss-invariance, such that the state-dependent bliss point $\bar{m}(l)$ defined by $u_m(y, \bar{m}(l); l) = 0$ is independent of y and is bounded by some global upper bound M. Suppose also that initial income is strictly greater than expected bliss utilization, $w_0 > \mathbb{E}_l[\bar{m}(l)]$. Then, there exists a coverage threshold $\bar{c} > 0$ such that every contract $x \in \mathcal{X}$ whose maximum out-of-pocket exposure satisfies $\sup_m c(m; x) \leq \bar{c}$ generates positive expected over-utilization: $\mathbb{E}_l[m^*(l, x)] \geq \mathbb{E}_l[m^{FB}(l)]$.*

Proof. It is clear that the consumer will never choose healthcare utilization above her bliss point in a given state, and thus $m^*(l, x) \leq \bar{m}(l) \leq M$ for any $x \in \mathcal{X}$. If $\bar{m}(l) = 0$ almost surely, then $m^*(l, x)$ and $m^{FB}(l)$ equal zero almost surely and the lemma holds trivially. Assuming hereafter that interior first-best care occurs with positive probability, at such an allocation, $u_m/u_y = 1$ at $m^{FB}(l)$, so $u_m > 0$ and $m^{FB}(l) < \bar{m}(l)$.[50] Taking expectations, $2\eta \equiv \mathbb{E}_l[\bar{m}(l)] - \mathbb{E}_l[m^{FB}(l)] > 0$.

It remains to show that $\mathbb{E}_l[m^*(l, x)]$ can be brought within η of $\mathbb{E}_l[\bar{m}(l)]$ once out-of-pocket exposure is small enough. As $\sup_m c(m; x) \to 0$, the consumer's problem in each state converges to the full-insurance problem $\max_m u(w_0 - \bar{p}, m; l)$, where $\bar{p} \leq \mathbb{E}_l[\bar{m}(l)] < w_0$. Its solution is

[50]And at a corner, $m^{FB}(l) = 0 < \bar{m}(l)$.

the bliss quantity, and by bliss-invariance the solution is $\bar{m}(l)$ whatever the premium $\bar{p}$. Equilibrium utilization is continuous in the contract,[51] so $m^*(l,x) \to \bar{m}(l)$ in every state; and since $m^*(l,x) \leq M$ uniformly, the convergence carries to the expectation, $\mathbb{E}_l[m^*(l,x)] \to \mathbb{E}_l[\bar{m}(l)]$. As $\mathbb{E}_l[\bar{m}(l)] = \mathbb{E}_l[m^{FB}(l)] + 2\eta$, there is a threshold $\bar{c} > 0$ such that $\sup_m c(m;x) \leq \bar{c}$ implies $\mathbb{E}_l[m^*(l,x)] \geq \mathbb{E}_l[m^{FB}(l)] + \eta > \mathbb{E}_l[m^{FB}(l)]$. □

A.2 Properties on Decomposition of Welfare

Proposition 1. **(Health Insurance Transfers Real Income Towards Sicker States)** *For any $x \in \mathcal{X}$, equivalent income $\tilde{w}(l,x)$ is increasing in l.*

Proof. Recall that equivalent income is defined by identity $v(\tilde{w}(l,x);l) = u(y^*(l,x), m^*(l,x);l)$. Let $\tilde{m}(l,x) \equiv m^*(l,x_0,\tilde{w}(l,x))$ and $\tilde{y}(l,x) \equiv \tilde{w}(l,x) - \tilde{m}(l,x)$ denote the consumer's optimal bundle at undistorted prices given income $\tilde{w}(l,x)$, so that $v(\tilde{w}(l,x);l) = u(\tilde{y},\tilde{m};l)$. Differentiating the defining identity with respect to l and applying the envelope theorem to both sides yields

$$v_w(\tilde{w}(l,x);l) \cdot \tilde{w}_l(l,x) + u_l(\tilde{y},\tilde{m};l) = u_l(y^*, m^*;l),$$

and rearranging,

$$\tilde{w}_l(l,x) = \frac{u_l(y^*,m^*;l) - u_l(\tilde{y},\tilde{m};l)}{v_w(\tilde{w}(l,x);l)}.$$

The denominator is $v_w(\tilde{w}(l,x);l) = u_y(\tilde{y},\tilde{m};l) > 0$. To sign the numerator, observe that the insured bundle (y^*, m^*) and the undistorted-price bundle $(\tilde{y},\tilde{m})$ lie on the same indifference curve in state l, since by construction they deliver the same utility. At the insured bundle the consumer's marginal rate of substitution satisfies $u_m/u_y = c_m(m^*;x) \leq 1$, while at the undistorted-price bundle $u_m/u_y = 1$. By strict joint concavity of u, the marginal rate of substitution is strictly decreasing in m along any indifference curve, so the weakly lower marginal rate of substitution at the insured bundle places it at weakly more healthcare: $m^* \geq \tilde{m}$.[52]

[51] A standard regularity property: the consumer's maximizer varies continuously with the contract, here by the maximum theorem applied to the strictly concave objective with its unique limiting peak. Because the fair premium lies in the compact set $[0, \mathbb{E}_l[\bar{m}(l)]]$, and objective functions converge uniformly bounded by $\sup_m c(m;x) \cdot \sup u_y$, bliss-invariance ensures that the limit maximizer $\bar{m}(l)$ is uniform across all feasible premiums. This guarantees the existence of a uniform threshold $\bar{c}$.

[52] If the consumer is at a corner solution $\tilde{m} = 0$ at undistorted prices, the marginal rate of substitution satisfies $u_m/u_y \leq 1$ rather than equality; the conclusion $m^* \geq 0 = \tilde{m}$ holds trivially. If instead the insured bundle exhausts income ($y^* = 0$), its marginal rate of substitution may strictly exceed one. Every point on the shared indifference curve with less healthcare then has a marginal rate of substitution greater than one, so total expenditure $y + m$ falls as m rises along the curve, and the undistorted-price bundle coincides with the insured bundle, $\tilde{m} = m^*$, making the numerator zero.

Because u_m/u_y is increasing in l, the marginal value of the health state u_l is increasing in m along any indifference curve.[53] Since (y^*, m^*) and $(\tilde{y}, \tilde{m})$ share an indifference curve in state l with $m^* \geq \tilde{m}$, it follows that $u_l(y^*, m^*; l) \geq u_l(\tilde{y}, \tilde{m}; l)$, and therefore $\tilde{w}_l(l, x) \geq 0$. $\square$

Proposition 2. **(Value of Risk Protection is Positive)** *For any $x \in \mathcal{X}$, if $v_w(\tilde{w}(l,x); l)$ is increasing in l, then $\Psi(x) \geq 0$.*

Proof. By Equations 5 and 8, $\Psi(x)$ is defined implicitly by

$$\mathbb{E}_l \left[v(w_0 + \Psi(x) - SCMH(x); l) \right] = \mathbb{E}_l \left[v(\tilde{w}(l,x); l) \right].$$

Since $v(\cdot; l)$ is increasing, $\Psi(x) \geq 0$ if and only if $\mathbb{E}_l \left[v(w_0 - SCMH(x); l) \right] \leq \mathbb{E}_l \left[v(\tilde{w}(l,x); l) \right]$. It is therefore enough to show that $\Delta\tilde{v} \equiv \mathbb{E}_l \left[v\left(\tilde{w}(l,x); l\right) \right] - \mathbb{E}_l \left[v\left(w_0 - SCMH(x); l\right) \right] \geq 0$.

By concavity of $v(\cdot; l)$:

$$v\left(\tilde{w}(l,x); l\right) - v\left(w_0 - SCMH(x); l\right) \geq v_w\left(\tilde{w}(l,x); l\right) \left[\tilde{w}(l,x) - w_0 + SCMH(x) \right] \quad \forall\ l.$$

Taking expectations and using the fact that $\mathbb{E}_l \left[\tilde{w}(l,x) - w_0 + SCMH(x) \right] = 0$ (Equation 7) gives

$$\Delta\tilde{v} \geq \mathbb{E}_l \left[v_w\left(\tilde{w}(l,x); l\right) \left[\tilde{w}(l,x) - w_0 + SCMH(x) \right] \right] = \mathrm{Cov}_l\left(\tilde{w}(l,x), v_w\left(\tilde{w}(l,x); l\right) \right).$$

By Proposition 1, $\tilde{w}(l,x)$ is increasing in l. By assumption, $v_w(\tilde{w}(l,x); l)$ is also increasing in l. Therefore, the covariance is positive, which implies that $\Delta\tilde{v} \geq 0$, and hence, $\Psi(x) \geq 0$. $\square$

The next step is to sign the components of risk protection attributable to the consumer's ability to change her behavior in response to insurance. To that end, it is useful to first introduce additional notation, along with two ancillary lemmas. Recall from footnote 37 that the risk protection generated by reducing under-utilization is $\Delta\Psi^{under}(x) = \Psi^{FB}(x) - \Psi^{N}(x)$, and that generated by over-utilization is $\Delta\Psi^{over}(x) = \Psi(x) - \Psi^{FB}(x)$. Similarly, the corresponding changes in the social cost of moral hazard are $\Delta SCMH^{under}(x) \equiv SCMH^{FB}(x) - SCMH^{N}(x)$ and $\Delta SCMH^{over}(x) \equiv SCMH(x) - SCMH^{FB}(x)$. The definitions of these objects rely on two intermediate points in the space of state-contingent equivalent incomes for a given contract x, based on the two hypothetical utilization schedules at the no-insurance and first-best levels: $\tilde{\mathbf{w}}^{N}(x)$ and $\tilde{\mathbf{w}}^{FB}(x)$. The additional notation thus concerns these intermediate objects used in the welfare decomposition.

[53] Parameterizing the curve by m as $y = Y(m)$, the derivative of u_l along it is $u_{lm} - (u_m/u_y)\, u_{ly} = u_y \cdot (u_m/u_y)_l$, which is positive since $u_y > 0$ and $(u_m/u_y)_l > 0$.

For a contract $x \in \mathcal{X}$, define the change in equivalent income due to the reduction in under-utilization and to over-utilization:

$$\Delta\tilde{w}^{under}(l,x) \equiv \tilde{w}^{FB}(l,x) - \tilde{w}^{N}(l,x) \qquad \Delta\tilde{w}^{over}(l,x) \equiv \tilde{w}(l,x) - \tilde{w}^{FB}(l,x).$$

It will also be useful to recenter these vectors so as to isolate the redistribution each produces, stripping out the change in mean equivalent income. To that end, define the re-centered equivalent income vectors:

$$\tilde{z}^{under}(l,x) \equiv \tilde{w}^{FB}(l,x) - \mathbb{E}_l[\Delta\tilde{w}^{under}(l,x)] \qquad \tilde{z}^{over}(l,x) \equiv \tilde{w}(l,x) - \mathbb{E}_l[\Delta\tilde{w}^{over}(l,x)].$$

Just as the movement from E to $\tilde{w}(l,x)$ in Figure 2 decomposes into a movement along the certainty line (the social cost of moral hazard) and a mean-preserving redistribution along the iso-mean line at the intermediate point $w_0 - SCMH(x)$, the re-centered vectors decompose the movements from $\tilde{\mathbf{w}}^N(x)$ to $\tilde{\mathbf{w}}^{FB}(x)$ and from $\tilde{\mathbf{w}}^{FB}(x)$ to $\tilde{\mathbf{w}}(x)$ into a uniform level shift in equivalent income and a mean-preserving redistribution. Unlike the intermediate point, however, the re-centered vectors are not certain incomes: each lies on the iso-mean line rather than on the certainty line, differing from that benchmark only by a mean-preserving redistribution. That redistribution raises risk protection when it favors the states in which the marginal indirect utility of income is higher. By construction $\mathbb{E}_l[\tilde{z}^{under}(l,x)] = \mathbb{E}_l[\tilde{w}^N(l,x)]$ and $\mathbb{E}_l[\tilde{z}^{over}(l,x)] = \mathbb{E}_l[\tilde{w}^{FB}(l,x)]$.

The two ancillary lemmas, stated next, provide the inputs the propositions draw on. Lemma 4 shows that, under one additional condition, equivalent income under the hypothetical utilization schedule $m^{FB}(l)$ is also increasing in l. Lemma 5 gives primitive sufficient conditions under which upward level shifts do not reduce the value of risk protection.

Condition 3 (**Marginal Preferences at First-Best Utilization Hypothetical Under Insurance**). *Under the first-best utilization hypothetical, the consumer weakly prefers more healthcare at the insured marginal price, but not at the undistorted price:* $\frac{u_m}{u_y}$ *evaluated at* $(y^{FB}(l,x), m^{FB}(l))$ *lies within* $[c_m(m^{FB}(l);x), 1]$.

Lemma 4 ($\tilde{\mathbf{w}}^{FB}(\mathbf{l},\mathbf{x})$ **Increasing in** l). *For any* $x \in \mathcal{X}$*, under Condition 2 and Condition 3,* $\tilde{w}^{FB}(l,x)$ *is increasing in* l.

Proof. Equivalent income is defined by $v(\tilde{w}^{FB}(l,x);l) = u(y^{FB}(l,x), m^{FB}(l);l)$, where $y^{FB}(l,x) = w_0 - \bar{p}^{FB}(x) - c(m^{FB}(l);x)$. Let $\hat{m}^{FB}(l,x) \equiv m^*(l,x_0,\tilde{w}^{FB}(l,x))$ and $\hat{y}^{FB}(l,x) \equiv \tilde{w}^{FB}(l,x) -$

$\hat{m}^{FB}(l,x)$. Differentiating the defining identity and applying the envelope theorem gives

$$v_w(\tilde{w}^{FB};l)\tilde{w}_l^{FB} = \big[u_m - u_y c_m(m^{FB};x)\big]_{(y^{FB}(l,x),m^{FB}(l))} m_l^{FB} + \big[u_l(y^{FB}(l,x),m^{FB};l) - u_l(\hat{y}^{FB},\hat{m}^{FB};l)\big].$$

Since $v_w(\tilde{w}^{FB};l)$ is positive, the sign of $\tilde{w}_l^{FB}$ is determined by the right hand side of the expression. The first term on the right is nonnegative by the first inequality in Condition 3, together with the fact that $m^{FB}(l) = m^*(l, x_0, w^{FB}(l))$ is increasing in l by Condition 2, normality, and single-crossing.

For the second term, the bundles $(y^{FB}(l,x), m^{FB}(l))$ and $(\hat{y}^{FB}, \hat{m}^{FB})$ lie on the same indifference curve, along which the marginal rate of substitution is strictly decreasing in m by strict concavity, and u_l is increasing in m by single-crossing. If $\hat{m}^{FB} = 0$, then $\hat{m}^{FB} \leq m^{FB}(l)$ trivially. Otherwise, the marginal rate of substitution at $(\hat{y}^{FB}, \hat{m}^{FB})$ is at least one—equal to one at an interior optimum, and greater than one where the budget binds ($\hat{y}^{FB} = 0$)—while by the second inequality in Condition 3 it is at most one at $(y^{FB}(l,x), m^{FB}(l))$. The weakly lower marginal rate of substitution at the latter bundle places it at weakly more healthcare along the curve: $\hat{m}^{FB}(l,x) \leq m^{FB}(l)$. Hence the second term is also nonnegative. Therefore $\tilde{w}_l^{FB}(l,x) \geq 0$. □

Condition 4 (**Curvature and Sorting for Upward Level Shifts**). *Marginal indirect utility is convex in income, $v_{www}(w;l) \geq 0$, and $v_{ww}(w;l)$ is increasing in l for all w.*

Lemma 5 (**$\Psi(x)$ Increasing in Level Shifts of Equivalent Income**). *Under Condition 4, any equivalent-income vector $\tilde{w}(l,x)$ that is increasing in l and delivers positive risk protection will deliver greater risk protection under uniform upward shift in equivalent income, $\tilde{w}(l,x)+\epsilon$ for $\epsilon \geq 0$.*

Proof. Shift equivalent income uniformly by $\epsilon \geq 0$, from $\tilde{w}(l,x)$ to $\tilde{w}(l,x)+\epsilon$, and let Ψ' denote the scalar which solves

$$\mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}(l,x)] + \epsilon + \Psi';\, l)] = \mathbb{E}_l[v(\tilde{w}(l,x) + \epsilon;\, l)]\,,$$

so that, at $\epsilon = 0$, Ψ' is the value of risk protection associated with $\tilde{w}(l,x)$. Differentiating with respect to ϵ, the value Ψ' depending on ϵ,

$$\mathbb{E}_l[v_w(\mathbb{E}_l[\tilde{w}(l,x)] + \epsilon + \Psi';\, l)]\left(1 + \frac{d\Psi'}{d\epsilon}\right) = \mathbb{E}_l[v_w(\tilde{w}(l,x) + \epsilon;\, l)]\,,$$

so, the denominator below being positive because $v(\cdot;l)$ is increasing in income,

$$\frac{d\Psi'}{d\epsilon} = \frac{\mathbb{E}_l[v_w(\tilde{w}(l,x)+\epsilon;\, l)] - \mathbb{E}_l[v_w(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l)]}{\mathbb{E}_l[v_w(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l)]}.$$

The sign of $d\Psi'/d\epsilon$ is therefore the sign of the numerator, and we show it is positive at every $\epsilon \geq 0$. By Condition 4, $v_w(\cdot;l)$ is convex in income, so it lies weakly above its tangent at $\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi'$: for each l, $v_w(\tilde{w}(l,x)+\epsilon;\, l) - v_w(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l) \geq v_{ww}(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l)\left(\tilde{w}(l,x)-\mathbb{E}_l[\tilde{w}(l,x)]-\Psi'\right)$. Taking $\mathbb{E}_l$ and noting that the income gap $\tilde{w}(l,x)-\mathbb{E}_l[\tilde{w}(l,x)]-\Psi'$ has mean $-\Psi'$, the numerator is bounded below by

$$\begin{aligned}
&\mathbb{E}_l[v_w(\tilde{w}(l,x)+\epsilon;\, l)] - \mathbb{E}_l[v_w(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l)] \\
&\qquad \geq \underbrace{-\Psi'\,\mathbb{E}_l[v_{ww}(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l)]}_{\text{(I)}} + \underbrace{\operatorname{Cov}_l(\tilde{w}(l,x),\, v_{ww}(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';\, l))}_{\text{(II)}}.
\end{aligned}$$

Term (II) is positive: by the cross-state curvature property in Condition 4, $v_{ww}(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi';l)$ is increasing in l. $\tilde{w}(l,x)$ is also increasing in l by hypothesis, so these comonotone functions of l have a positive covariance. Therefore the numerator is bounded below by term (I) alone. Since the denominator in $d\Psi'/d\epsilon$ is positive,

$$\frac{d\Psi'}{d\epsilon} \geq \kappa(\epsilon)\Psi'(\epsilon), \qquad \kappa(\epsilon) \equiv \frac{-\mathbb{E}_l[v_{ww}(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi'(\epsilon);l)]}{\mathbb{E}_l[v_w(\mathbb{E}_l[\tilde{w}(l,x)]+\epsilon+\Psi'(\epsilon);l)]} \geq 0.$$

It follows that

$$\frac{d}{d\epsilon}\left[\Psi'(\epsilon)\exp\left(-\int_0^\epsilon \kappa(t)\,dt\right)\right] \geq 0.$$

At $\epsilon = 0$, $\Psi'(0) \geq 0$ by hypothesis. Hence $\Psi'(\epsilon) \geq 0$ for every $\epsilon \geq 0$. Returning to the bound above, term (I) is then positive, and term (II) is positive, so the numerator is positive throughout. Thus $d\Psi'/d\epsilon \geq 0$, and the value of risk protection is increasing in the uniform level shift. □

We now turn to signing the components of risk protection attributable to the consumer's ability to change her behavior in response to insurance: $\Delta\Psi^{under}(x)$ and $\Delta\Psi^{over}(x)$.

Proposition 3 ($\Delta\Psi^{under}(x) \geq 0$). *Suppose that Conditions 2, 3, and 4 hold, and that:*

(i) reducing under-utilization redistributes equivalent income toward sicker states: $\Delta\tilde{w}^{under}(l,x)$ is increasing in l;

(ii) the marginal indirect utility of income is higher in sicker states along the re-centered vector: $v_w(\tilde{z}^{under}(l,x);l)$ is increasing in l;

(iii) reducing under-utilization weakly lowers expected deadweight loss relative to the no-insurance utilization hypothetical: $\Delta SCMH^{under}(x) \leq 0$*, equivalently* $\mathbb{E}_l[\Delta\tilde{w}^{under}(l,x)] \geq 0$*;*

Then $\Delta\Psi^{under}(x) \geq 0$*.*

Proof. By Conditions 2 and 3 and Lemma 4, $\tilde{w}^{FB}(l,x)$ is increasing in l; since $\tilde{z}^{under}(l,x)$ differs from it by the constant $\mathbb{E}_l[\Delta\tilde{w}^{under}(l,x)]$, it is increasing in l as well. With $\mathbb{E}_l[\tilde{z}^{under}(l,x)] = \mathbb{E}_l[\tilde{w}^N(l,x)]$, concavity of $v(\cdot;l)$ gives

$$\mathbb{E}_l[v(\tilde{z}^{under}(l,x);l)] - \mathbb{E}_l[v(\tilde{w}^N(l,x);l)] \geq \mathrm{Cov}_l\big(\Delta\tilde{w}^{under}(l,x), v_w(\tilde{z}^{under}(l,x);l)\big) \geq 0,$$

the final inequality by conditions (i), (ii), and Chebyshev's sum inequality. Because the two schedules share the mean $\mathbb{E}_l[\tilde{w}^N(l,x)]$ and $\mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^N(l,x)]+s;l)]$ is increasing in s, the value of risk protection associated with $\tilde{z}^{under}(l,x)$ is at least $\Psi^N(x)$; and since $\tilde{z}^{under}(l,x)$ and $v_w(\tilde{z}^{under}(l,x);l)$ are both increasing in l, $\mathrm{Cov}_l(\tilde{z}^{under}, v_w(\tilde{z}^{under};l)) \geq 0$, so this value of risk protection is also positive.

Finally, $\tilde{w}^{FB}(l,x) = \tilde{z}^{under}(l,x)+\mathbb{E}_l[\Delta\tilde{w}^{under}(l,x)]$ is a uniform upward shift by $\mathbb{E}_l[\Delta\tilde{w}^{under}(l,x)] = -\Delta SCMH^{under}(x) \geq 0$ (condition (iii)); since $\tilde{z}^{under}(l,x)$ is increasing in l with positive value of risk protection, Lemma 5 (under Condition 4) implies the shift does not reduce it. Hence $\Psi^{FB}(x)$ is at least the value of risk protection associated with $\tilde{z}^{under}(l,x)$, which is at least $\Psi^N(x)$; that is, $\Delta\Psi^{under}(x) \geq 0$. □

Proposition 4 ($\Delta\Psi^{over}(x) \geq 0$). *Suppose that:*

(i) over-utilization redistributes equivalent income toward sicker states: $\Delta\tilde{w}^{over}(l,x)$ *is increasing in* l*;*

(ii) the marginal indirect utility of income is higher in sicker states along the re-centered vector: $v_w(\tilde{z}^{over}(l,x);l)$ *is increasing in* l*;*

(iii) over-utilization weakly raises expected deadweight loss relative to the efficient-utilization hypothetical: $\Delta SCMH^{over}(x) = SCMH(x) - SCMH^{FB}(x) \geq 0$*;*[54]

[54]Because $SCMH(x) \geq 0$ always—by revealed preference $\tilde{w}(l,x) \leq y^*(l,x) + m^*(l,x)$, so the equilibrium bundle's equivalent income never exceeds the resources it consumes—the condition $\Delta SCMH^{over}(x) = SCMH(x) - SCMH^{FB}(x) \geq 0$ always holds if the efficient-utilization hypothetical carries no deadweight loss under contract x, $SCMH^{FB}(x) = 0$. This occurs automatically under income-independent healthcare demand (as in the quasilinear/CARA case of Appendix A.3), or if there is no gap in non-health consumption at the efficient-utilization hypothetical, $y^{FB}(l,x) = y^{FB}(l)$; in either case the efficient-utilization bundle is the undistorted optimum at its own resources.

(iv) the expected-utility gain from over-utilization's redistribution toward sicker states is at least the expected-utility loss from the fall in mean equivalent income it produces: $\mathrm{Cov}_l\big(\Delta\tilde{w}^{over}(l,x), v_w(\tilde{z}^{over}(l,x);l)\big) \geq \mathbb{E}_l[v(\tilde{z}^{over}(l,x);l)] - \mathbb{E}_l[v(\tilde{w}(l,x);l)]$.

Then $\Delta\Psi^{over}(x) \geq 0$.

Proof. Recall $\tilde{z}^{over}(l,x) = \tilde{w}(l,x) - \mathbb{E}_l[\Delta\tilde{w}^{over}(l,x)]$, so $\mathbb{E}_l[\tilde{z}^{over}(l,x)] = \mathbb{E}_l[\tilde{w}^{FB}(l,x)]$ and $\tilde{z}^{over}(l,x) = \tilde{w}(l,x) + \Delta SCMH^{over}(x)$, a uniform upward shift of $\tilde{w}$ by $\Delta SCMH^{over}(x) = -\mathbb{E}_l[\Delta\tilde{w}^{over}(l,x)] \geq 0$ (condition (iii)). We first show $\mathbb{E}_l[v(\tilde{w};l)] \geq \mathbb{E}_l[v(\tilde{w}^{FB};l)]$, then translate it into $\Psi(x) \geq \Psi^{FB}(x)$.

The move from $\tilde{w}^{FB}$ to $\tilde{w}$ passes through $\tilde{z}^{over}$, splitting into a mean-preserving redistribution and a uniform downward shift. *Redistribution* ($\tilde{w}^{FB} \to \tilde{z}^{over}$): the two share the mean $\mathbb{E}_l[\tilde{w}^{FB}]$ and differ by $\tilde{z}^{over} - \tilde{w}^{FB} = \Delta\tilde{w}^{over} - \mathbb{E}_l[\Delta\tilde{w}^{over}]$, so concavity of $v(\cdot;l)$ gives $\mathbb{E}_l[v(\tilde{z}^{over};l)] - \mathbb{E}_l[v(\tilde{w}^{FB};l)] \geq \mathrm{Cov}_l(\Delta\tilde{w}^{over}, v_w(\tilde{z}^{over};l))$, a lower bound on the gain, positive by (i), (ii), and Chebyshev's sum inequality. *Level shift* ($\tilde{z}^{over} \to \tilde{w}$): a uniform downward shift of size $\Delta SCMH^{over}(x)$, with expected-utility loss $\mathbb{E}_l[v(\tilde{z}^{over};l)] - \mathbb{E}_l[v(\tilde{w};l)]$. Subtracting,

$$\mathbb{E}_l[v(\tilde{w};l)] - \mathbb{E}_l[v(\tilde{w}^{FB};l)] = \big(\mathbb{E}_l[v(\tilde{z}^{over};l)] - \mathbb{E}_l[v(\tilde{w}^{FB};l)]\big) - \big(\mathbb{E}_l[v(\tilde{z}^{over};l)] - \mathbb{E}_l[v(\tilde{w};l)]\big) \;\geq\; 0,$$

since the first bracket is at least $\mathrm{Cov}_l(\Delta\tilde{w}^{over}, v_w(\tilde{z}^{over};l))$ and this covariance is at least the second bracket by condition (iv).

To translate, $\Psi(x)$ and $\Psi^{FB}(x)$ solve $\mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}] + \Psi(x);l)] = \mathbb{E}_l[v(\tilde{w};l)]$ and $\mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^{FB}] + \Psi^{FB}(x);l)] = \mathbb{E}_l[v(\tilde{w}^{FB};l)]$. Since $\mathbb{E}_l[\tilde{w}^{FB}] \geq \mathbb{E}_l[\tilde{w}]$ (condition (iii)) and $v(\cdot;l)$ is increasing,

$$\begin{aligned}
\mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^{FB}] + \Psi(x);l)] &\geq \mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}] + \Psi(x);l)] \\
&= \mathbb{E}_l[v(\tilde{w};l)] \\
&\geq \mathbb{E}_l[v(\tilde{w}^{FB};l)] \\
&= \mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^{FB}] + \Psi^{FB}(x);l)].
\end{aligned}$$

As $s \mapsto \mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^{FB}] + s;l)]$ is increasing, $\Psi(x) \geq \Psi^{FB}(x)$; that is, $\Delta\Psi^{over}(x) \geq 0$. □

It is worth emphasizing that the hypotheses of Propositions 3 and 4 are sufficient but not necessary. The CARA/quasilinear case below illustrates two distinct departures from these sufficient conditions. For under-utilization, the proof of Proposition 3 can be adapted by replacing Condition 4 and Lemma 5 with the direct level-shift invariance implied by CARA utility. For over-utilization, we use a separate argument rather than Proposition 4: condition (iv) is not invoked, and the sign is instead obtained by comparing the re-centered and first-best-utilization vectors directly and then applying level-shift invariance.

A.3 Properties Under CARA/Quasilinear Utility

The specifications collected in Table 1 all take the form $u(y, m; l) = -\exp\big(-\psi\,[\,y + b(m; l)\,]\big)$, with $\psi > 0$, $l \geq 0$, and $b(\cdot; l)$ quadratic. Utility is thus a strictly increasing transformation of the composite index $z \equiv y + b(m; l)$, and the consumer's preferences are quasilinear in non-health consumption: her marginal rate of substitution $u_m/u_y = b_m(m; l)$ does not depend on her income. Three implications organize the entire analysis of this section—demand at undistorted prices is independent of income, the marginal indirect utility of income collapses onto a single shadow price along each indifference curve, and the value of risk protection under CARA is invariant to level shifts in the consumer's income—and we collect the primitives they generate before signing the terms of the decomposition. Throughout, $B(l) \equiv b(l; l) - l$ denotes the intercept of indirect utility, and a contract is *undistorted* when its out-of-pocket price equals one (x_0). We maintain that initial income is high enough that the first best is interior in every state, so that $m^{FB}(l) = l$; the explicit requirement is recorded in Equation 11 below.

Lemma 6 (**Primitives of the CARA Specifications**). *For each specification in Table 1:*

(a) Demand at undistorted prices is income-independent up to the budget constraint, $m^*(l, x_0, w) = \min(l, w)$, with first-best utilization $m^{FB}(l) = l$ and no-behavioral-response utilization $m^N(l) = \min(w_0, l)$.

(b) Indirect utility at undistorted prices is

$$v(w; l) = \begin{cases} -\exp\big(-\psi(w + B(l))\big) & \textit{if } w \geq l, \\ -\exp\big(-\psi\, b(w; l)\big) & \textit{if } w < l, \end{cases} \qquad B(l) = \begin{cases} -l & \textit{(additive)}, \\ -(1 + \frac{\tilde{\omega}}{2})\, l & \textit{(multiplicative)}. \end{cases}$$

(c) Deadweight loss is $D \equiv R - \tilde{w}$, with $R \equiv y + m$; it is generally positive, and zero when $m = l$. Wherever the bundle's own equivalent income is interior ($\tilde{w} \geq l$), it depends only on the utilization gap,

$$D = \tfrac{1}{2\omega}(m - l)^2 \ \textit{(additive)}, \qquad D = \tfrac{1}{2\tilde{\omega} l}(m - l)^2 \ \textit{(multiplicative)}.$$

(d) On the undistorted region $w \geq l$, the curvature ratio $\rho(w; l) \equiv v_{wl}(w; l)/|v_{ww}(w; l)|$ is income-independent and equals $-B'(l)$:

$$\rho(w; l) = 1 \ \textit{(additive)}, \qquad \rho(w; l) = 1 + \tfrac{\tilde{\omega}}{2} \ \textit{(multiplicative)}.$$

(e) *The value of risk protection depends only on the spread of equivalent income across states, not on its level, wherever the incomes involved satisfy $w \geq l$. For a state-contingent income vector $\big(w(l)\big)$, the scalar Ψ solving $\mathbb{E}_l[v(\mathbb{E}_l[w(l)] + \Psi;\, l)] = \mathbb{E}_l[v(w(l);\, l)]$ is equivalent to the scalar Ψ' solving*

$$\mathbb{E}_l[v(\mathbb{E}_l[w(l)] + \epsilon + \Psi';\, l)] = \mathbb{E}_l[v(w(l) + \epsilon;\, l)]$$

for every level shift ϵ.

Proof. Since u is a strictly increasing transformation of $z = y + b(m;l)$, the consumer's problem at undistorted prices is to maximize $z = (w - m) + b(m;l)$ over $m \in [0, w]$. The first-order condition is $b_m(m;l) = 1$, whose solution is $m = l$ by the quadratic form of b and is independent of w. Since $b_m(m;l) > 1$ for every $m < l$, the objective is strictly increasing on $[0, w]$ whenever $w < l$, so the constraint binds and $m = w$; otherwise the interior solution $m = l$ is feasible. In any case $m^*(l, x_0, w) = \min(l, w)$, which gives $m^{FB}(l) = m^*(l, x_0, w^{FB}(l)) = l$ under the maintained first-best interiority and $m^N(l) = m^*(l, x_0, w_0) = \min(w_0, l)$, establishing (a).

For (b), evaluating z at the optimum yields $v(w;l) = -\exp(-\psi z^*)$. When $w \geq l$ the optimum is interior with $z^* = (w - l) + b(l;l) = w + B(l)$, $B(l) \equiv b(l;l) - l$; substituting $b(l;l) = 0$ (additive) and $b(l;l) = -\tilde{\omega} l/2$ (multiplicative) gives the stated intercepts. When $w < l$ the optimum is the corner $m = w$, $y = 0$, so $z^* = b(w;l)$. This establishes (b).

For (c), non-negativity $D \geq 0$ is the revealed-preference property of Equation 6. Equivalent income satisfies $v(\tilde{w};l) = u(y, m; l)$; when $\tilde{w} \geq l$ this reads $\tilde{w} + B(l) = y + b(m;l)$ on the interior branch of (b), so that $\tilde{w} = y + b(m;l) - b(l;l) + l$ and $D = R - \tilde{w} = (m - l) - \big[b(m;l) - b(l;l)\big]$. Plugging in the two forms of b and simplifying yields $D = \frac{1}{2\omega}(m - l)^2$ and $D = \frac{1}{2\tilde{\omega} l}(m - l)^2$ respectively, each zero exactly when $m = l$. This establishes (c).

For (d), differentiating the interior branch gives $v_w = \psi e^{-\psi(w + B(l))}$, $v_{ww} = -\psi^2 e^{-\psi(w+B(l))}$, and $v_{wl} = -\psi^2 B'(l)\, e^{-\psi(w+B(l))}$. Therefore $\rho(w;l) = v_{wl}/|v_{ww}| = -B'(l)$, which is independent of w; substituting $B'(l) = -1$ (additive) and $B'(l) = -(1 + \tilde{\omega}/2)$ (multiplicative) gives the stated values.

For (e), fix a vector $\big(w(l)\big)$ with $w(l) \geq l$ in every state and a shift $\epsilon \geq 0$, and let Ψ and Ψ' denote the solutions of the unshifted and shifted identities in (e), respectively, so that $CE_0 \equiv \mathbb{E}_l[w(l)] + \Psi$ is the certainty-equivalent income associated with $\big(w(l)\big)$. Write $v(w;l) = -e^{-\psi Z(w;l)}$ as in (b). Since $b_m(m;l) \geq 1$ for $m \leq l$, $Z_w(w;l) \geq 1$ everywhere, with equality where $w \geq l$. Hence, for any income CE and $\delta \geq 0$,

$$v(CE + \delta; l) \;\geq\; e^{-\psi\delta}\, v(CE; l),$$

with equality whenever $CE \geq l$. Because $w(l) \geq l$ in every state, the vector side scales exactly, and the unshifted identity gives

$$\mathbb{E}_l[v(w(l)+\epsilon; l)] = e^{-\psi\epsilon}\, \mathbb{E}_l[v(CE_0; l)] \;\leq\; \mathbb{E}_l[v(CE_0+\epsilon; l)].$$

Since the left-hand side of the shifted identity is strictly increasing in Ψ', it follows that $\Psi' \leq \Psi$. If $CE_0 \geq l$ in every state, the inequality in the display is an equality, so $\Psi' = \Psi$; and a downward shift is the same comparison read from its endpoint, with the hypotheses placed on the lower vector. The solution is thus the same for every level shift under the bound in (e). $\square$

Lemma 6 verifies the standing hypotheses of the earlier appendix at once. Single-crossing $\partial_l(u_m/u_y) = \partial_l b_m > 0$ holds in both forms,[55] so Proposition 1 applies and equivalent income $\tilde{w}(l,x)$ is increasing in l. Moreover, since $B'(l) < 0$, the first-best income schedule $w^{FB}(l) = w_0 + \mathbb{E}_l[B] - B(l)$ is increasing in l, so Condition 2 holds; it equals $w_0 + (l - \bar{l})$ in the additive form and $w_0 + (1+\tilde{\omega}/2)(l-\bar{l})$ in the multiplicative form, where $\bar{l} \equiv \mathbb{E}_l[l]$. First-best interiority, $w^{FB}(l) \geq l$ in every state, is then the requirement

$$w_0 \;\geq\; \bar{l} \quad \text{(additive)}, \qquad w_0 \;\geq\; \bar{l} + \tfrac{\tilde{\omega}}{2}\,(\bar{l} - \underline{l}) \quad \text{(multiplicative)}, \tag{11}$$

where $\underline{l}$ is the lowest health state; the multiplicative bound binds first in the healthiest state, where the desired first-best transfer is smallest.

Two remaining standing conditions in the general model do *not* hold under these specifications. Condition 4 is violated because, since $B'(l) < 0$, the income curvature $v_{ww}(w;l) = -\psi^2 e^{-\psi(w+B(l))}$ is *decreasing* in l. Nonetheless, by Lemma 6(e) the CARA parametrization satisfies the property that Condition 4 implies.

Condition 1 is also violated because, by part (a) the marginal propensity to spend on healthcare out of income, $\partial m^*(l, x_0, w)/\partial w = \mathbf{1}\{w < l\}$, is increasing in l but *decreasing* in w. The orderings of utilization that Condition 1 delivers in general nonetheless hold here, by the stronger state-by-state argument of Lemma 7 below.

Lemma 7 (**Utilization Ordering under Full Insurance**). *Suppose the consumer's income is high enough so that she can afford the full-insurance premium ($w_0 \geq \bar{l} + \omega$ in the additive form, $w_0 \geq (1+\tilde{\omega})\bar{l}$ in the multiplicative form). Then, under full insurance,*

$$m^N(l) \;\leq\; m^{FB}(l) \;\leq\; m^*(l,x),$$

[55] $\partial_l b_m = 1/\omega > 0$ in the additive form and $\partial_l b_m = m/(\tilde{\omega} l^2) > 0$ in the multiplicative form.

in every health state, so non-moral hazard utilization, under-utilization, and over-utilization are each positive state-by-state. The under-utilization gap is $m^{FB}(l) - m^N(l) = (l - w_0)^+$, and the over-utilization gap $m^*(l, x) - m^{FB}(l)$ equals the constant ω in the additive form and $\tilde{\omega}\, l$ in the multiplicative form.

Proof. By Lemma 6(a), $m^N(l) = \min(w_0, l) \leq l = m^{FB}(l)$, with equality when $l \leq w_0$ and strict inequality when $l > w_0$, so the under-utilization gap is $(l - w_0)^+$. Under full insurance the out-of-pocket price is zero, and by the affordability hypothesis $y^*(l, x) = w_0 - \bar{p} \geq 0$, so the consumer chooses her bliss quantity $\bar{m}(l)$, defined by $b_m(\bar{m}(l); l) = 0$; the quadratic form gives $\bar{m}(l) = l + \omega$ (additive) and $\bar{m}(l) = l(1 + \tilde{\omega})$ (multiplicative). Since $\bar{m}(l) \geq l = m^{FB}(l)$ in either form, $m^*(l, x) = \bar{m}(l) \geq m^{FB}(l)$, and the over-utilization gap equals ω and $\tilde{\omega} l$ respectively. Therefore each of the three components is positive, and since the state considered was arbitrary, the orderings hold state-by-state. □

The over-utilization gap is the first point at which the two forms part company, and it does so in a way that will drive the sign of the over-utilization risk-protection term: the additive form over-consumes by the same dollar amount ω in every state, while the multiplicative form over-consumes in proportion to the health state. We return to this distinction when signing the over-utilization risk-protection term below.

Proposition 5 (**Risk Protection Absent Behavioral Response is Positive**). *Under either specification and the full-insurance contract, $\Psi^N(x) \geq 0$.*

This sign requires neither first-best interiority (Equation 11) nor any restriction beyond $\psi > 0$: the no-behavioral-response hypothetical pools risk whether or not the consumer can afford her efficient healthcare.

Proof. Since $s \mapsto \mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^N(l, x)] + s; l)]$ is strictly increasing, $\Psi^N(x) \geq 0$ if and only if $\mathbb{E}_l[v(\tilde{w}^N(l, x); l)] \geq \mathbb{E}_l[v(\mathbb{E}_l[\tilde{w}^N(l, x)]; l)]$, where $\mathbb{E}_l[\tilde{w}^N(l, x)] = w_0 - SCMH^N(x)$. Inserting the certain income w_0 splits this difference into two parts,

$$\mathbb{E}_l\big[v(\tilde{w}^N; l)\big] - \mathbb{E}_l\big[v(w_0 - SCMH^N; l)\big] = \underbrace{\mathbb{E}_l\big[v(w_0; l)\big] - \mathbb{E}_l\big[v(w_0 - SCMH^N; l)\big]}_{\text{(I)}} + \underbrace{\mathbb{E}_l\big[v(\tilde{w}^N; l)\big] - \mathbb{E}_l\big[v(w_0; l)\big]}_{\text{(II)}},$$

and we show each is positive.

For (I), recall $SCMH^N(x) = \mathbb{E}_l[D^N(l,x)]$ with $D^N = [y^N + m^N] - \tilde{w}^N$. The no-behavioral-response bundle $(y^N, m^N(l))$ is feasible at undistorted prices given resources $y^N + m^N(l)$, so $v(y^N + m^N(l); l) \geq u(y^N, m^N(l); l) = v(\tilde{w}^N(l); l)$; since v is increasing in income, $y^N + m^N(l) \geq \tilde{w}^N(l)$, that is $D^N(l,x) \geq 0$ in every state, including any in which $\tilde{w}^N(l) < l$. Therefore $SCMH^N \geq 0$, and since $v(\cdot; l)$ is increasing in income, term (I)$= \mathbb{E}_l[v(w_0; l)] - \mathbb{E}_l[v(w_0 - SCMH^N; l)]$ is positive.

For (II), write $v(w;l) = -\exp(-\psi Z(w;l))$, where by Lemma 6(b) the achieved index is $Z(w;l) = w + B(l)$ for $w \geq l$ and $Z(w;l) = b(w;l)$ for $w < l$. By construction $v(\tilde{w}^N(l); l) = u(y^N, m^N(l); l) = -\exp(-\psi z^N(l))$ with $z^N(l) \equiv y^N + b(m^N(l); l)$, and $v(w_0; l) = -\exp(-\psi z^0(l))$ with $z^0(l) \equiv Z(w_0; l)$. Both the no-behavioral-response and the no-insurance allocations consume the uninsured quantity $m^N(l) = \min(w_0, l)$, so they share the common health term $Q(l) \equiv b(m^N(l); l)$ and differ only in non-health consumption: under full insurance it is the constant $y^N = w_0 - \mu$, where $\mu \equiv \mathbb{E}_l[\min(w_0, l)]$, while without insurance it is $y^0(l) \equiv w_0 - m^N(l) = \max(w_0 - l, 0)$. Thus $z^N(l) = Q(l) + (w_0 - \mu)$ and $z^0(l) = Q(l) + y^0(l)$, and since $\mathbb{E}_l[y^0(l)] = w_0 - \mu$, the two index lotteries have the same mean.

Notice that $Q(l)$ and $y^0(l)$ are each decreasing in l: $y^0 = \max(w_0 - l, 0)$ is decreasing; and $Q(l) = b(\min(w_0, l); l)$ equals $b(l; l) = l + B(l)$ for $l \leq w_0$—constant under the additive form, equal to $-\tilde{\omega} l/2$ under the multiplicative form—and equals $b(w_0; l)$ for $l > w_0$, which decreases in l as the benefit of a fixed quantity falls with worsening health. Consequently $e^{-\psi Q(l)}$ and $e^{-\psi y^0(l)}$ are both increasing in l. Term (II) is positive if and only if $\mathbb{E}_l[e^{-\psi z^0}] \geq \mathbb{E}_l[e^{-\psi z^N}]$, that is

$$\mathbb{E}_l\left[e^{-\psi Q}\, e^{-\psi y^0}\right] \;\geq\; e^{-\psi(w_0-\mu)}\, \mathbb{E}_l\left[e^{-\psi Q}\right].$$

By Chebyshev's sum inequality, the comonotone functions $e^{-\psi Q}$ and $e^{-\psi y^0}$ satisfy $\mathbb{E}_l[e^{-\psi Q} e^{-\psi y^0}] \geq \mathbb{E}_l[e^{-\psi Q}]\, \mathbb{E}_l[e^{-\psi y^0}]$; and by Jensen's inequality applied to the convex map $t \mapsto e^{-\psi t}$, $\mathbb{E}_l[e^{-\psi y^0}] \geq e^{-\psi \mathbb{E}_l[y^0]} = e^{-\psi(w_0-\mu)}$. Since $\mathbb{E}_l[e^{-\psi Q}] > 0$, combining the two yields the display, so term (II) is positive.

Therefore $\mathbb{E}_l[v(\tilde{w}^N; l)] \geq \mathbb{E}_l[v(w_0 - SCMH^N; l)]$, and $\Psi^N(x) \geq 0$. $\square$

Incremental Risk-protection. We now turn to signing the under- and over-utilization risk-protection terms. We establish $\Delta\Psi^{under}(x) \geq 0$ by verifying conditions (i)–(iii) of Proposition 3, noting that Condition 3 holds under full insurance because $b_m(l; l) = 1$ and $c_m(l; x) = 0$, and that the proposition's appeal to Lemma 5 is unnecessary in light of the level-shift invariance property described in Lemma 6(e). We establish $\Delta\Psi^{over}(x) \geq 0$ by comparing the value of risk protection associated with $\tilde{w}^{FB}(l,x)$ and $\tilde{z}^{over}(l,x)$ directly—by equality in the additive specification and by Jensen's inequality in the multiplicative specification—and then use the level-shift invariance property from Lemma 6(e) to pass from $\tilde{z}^{over}(l,x)$ to $\tilde{w}(l,x)$. Conditions (i)–(iii) of Proposition 4 are checked below only to characterize the redistribution and the accompanying level shift; they are not used as a sufficient application of that proposition.

Throughout, we write $\bar{l} \equiv \mathbb{E}_l[l]$ and $\mu \equiv \mathbb{E}_l[\min(w_0, l)]$, and use that the curvature ratio of Lemma 6(d) is the constant $\rho = 1$ (additive) and $\rho = 1 + \tilde{\omega}/2$ (multiplicative) on the undistorted region. Accordingly, in addition to first-best interiority 11 and the affordability requirement below, we maintain that the re-centered vector $\tilde{z}^{under}(l,x)$ of Proposition 3 and the certainty-equivalent incomes along its level shift lie on the interior region in every state.[56] For the over-utilization comparison, the affordability requirement places the relevant equivalent-income vectors on the interior region, so the exact level-shift invariance of Lemma 6(e) applies.

Under-utilization. The no-behavioral-response equivalent income is $\tilde{w}^N(l,x)$, and $\tilde{w}^{FB}(l,x) = w_0 - \bar{l} + l$ since the first-best utilization bundle is undistorted. For condition (i), the under-utilization gap is $\tilde{w}^{FB}(l,x) - \tilde{w}^N(l,x)$. When $l \le w_0$ the no-behavioral-response bundle is itself undistorted, $\tilde{w}^N(l,x) = w_0 - \mu + l$, and the gap equals the constant $\mu - \bar{l}$; when $l > w_0$ the no-insurance utilization hypothetical $m^N(l) = w_0$ lies strictly below the efficient level and $\tilde{w}^N(l,x)$ is decreasing in l.[57] Hence $\partial_l\big(\tilde{w}^{FB} - \tilde{w}^N\big) = 1 - \partial_l \tilde{w}^N \ge 0$ everywhere, and Condition (i) holds.

For condition (ii), the re-centered vector $\tilde{z}^{under}(l,x) = \tilde{w}^{FB}(l,x) - \mathbb{E}_l[\Delta\tilde{w}^{under}(l,x)]$ differs from $\tilde{w}^{FB}(l,x)$ by a constant, so it has unit slope and is interior under the maintained interiority above. Wherever a schedule is interior, v_w along it is increasing in l exactly when its slope is at most the curvature ratio; here $1 \le \rho$, with equality in the additive form and strict inequality in the multiplicative form, so condition (ii) holds. Condition (iii) is immediate: $\Delta SCMH^{under}(x) = -SCMH^N(x) \le 0$, since $SCMH^{FB}(x) = 0$ (Lemma 6(a),(c)). With $\tilde{w}^{FB}(l,x)$ increasing directly from its explicit form, the redistribution argument of Proposition 3 applies, and its level-shift step holds by the invariance; hence $\Delta\Psi^{under}(x) \ge 0$ under both specifications. The additive specification sits at the boundary of condition (ii): the slope

[56] The binding requirements are $w_0 \ge \bar{l} + SCMH^N(x)$, which places the re-centered vector on the interior region, and $\mathbb{E}_l[\tilde{w}^N(l,x)] + \Psi^N(x) \ge l$ in every state, which places the certainty-equivalent incomes there. In economic terms, the first requires that initial income cover expected undistorted care together with the no-behavioral-response social cost; the second requires that the certain income the consumer regards as equivalent to her insured no-behavioral-response position cover undistorted care even in the sickest state. Each is a joint restriction on initial income and the support of the health state, and each may fail for consumers with low income facing sufficiently severe health states. Where the second fails, an upward level shift can reduce the value of risk protection.

[57] Differentiating the identity $v(\tilde{w}^N; l) = u(y^N, m^N(l); l)$ in l for $l > w_0$, where $m^N(l) = w_0$ is fixed, gives $v_w(\tilde{w}^N; l)\,\partial_l \tilde{w}^N = u_l(y^N, w_0; l) - v_l(\tilde{w}^N; l)$. The right-hand side is non-positive: along the consumer's indifference curve the marginal value of the health state is larger at higher healthcare, and $m^N(l) = w_0$ is weakly below the quantity the consumer would choose at equivalent income $\tilde{w}^N$. Since $v_w > 0$, $\partial_l \tilde{w}^N \le 0$.

of the re-centered vector equals the curvature ratio, so the marginal indirect utility of income is constant across states and the covariance bound in the proof is zero, though the sign is unaffected.

Over-utilization. The relevant points are the first-best utilization point $\tilde{w}^{FB}(l,x)$ and the equilibrium point $\tilde{w}(l,x)$. We take the full-insurance allocation to be affordable, $y^*(l,x) \geq 0$—equivalently $w_0 \geq \bar{l} + \omega$ in the additive form and $w_0 \geq (1+\tilde{\omega})\bar{l}$ in the multiplicative form, as the income floor of Appendix B ensures for the parameter estimates we use—under which the equilibrium equivalent income is itself interior, $\tilde{w}(l,x) \geq l$. Under full insurance the consumer consumes her bliss quantity, so by Lemma 7 the over-utilization quantity $m^*(l,x) - m^{FB}(l)$ equals the constant ω (additive) and $\tilde{\omega}\, l$ (multiplicative), and by Lemma 6(c) the equilibrium deadweight loss is $D(l,x) = \omega/2$ (additive) and $\tilde{\omega}\, l/2$ (multiplicative). Because the insurer's cost moves one-for-one with utilization, total resources at the equilibrium and first-best utilization bundles coincide in the additive case and differ by $\tilde{\omega}(l-\bar{l})$ in the multiplicative case, so the equivalent-income over-gap is

$$\tilde{w}(l,x) - \tilde{w}^{FB}(l,x) = \begin{cases} -\,\omega/2 & \text{(additive)}, \\ \frac{\tilde{\omega}}{2}\,(l-2\bar{l}) & \text{(multiplicative)}. \end{cases}$$

The gap between the re-centered vector of Proposition 4 and the first-best utilization point, $h_{over}(l) \equiv \tilde{z}^{over}(l,x) - \tilde{w}^{FB}(l,x) = \left[\tilde{w} - \tilde{w}^{FB}\right] + \left[SCMH - SCMH^{FB}\right]$, then equals 0 (additive) and $\frac{\tilde{\omega}}{2}(l-\bar{l})$ (multiplicative), using $SCMH = \omega/2$ and $\tilde{\omega}\bar{l}/2$ respectively and $SCMH^{FB} = 0$ (Lemma 6(c) and (a)).

Under the additive specification, $h_{over} \equiv 0$: the re-centered and first-best-utilization vectors coincide, and the movement from $\tilde{w}^{FB}(l,x)$ to $\tilde{w}(l,x)$ is a pure uniform downward shift. By Lemma 6(e), this level shift leaves the value of risk protection unchanged, so $\Delta\Psi^{over}(x) = 0$. Under the multiplicative specification, $h_{over}(l) = \frac{\tilde{\omega}}{2}(l-\bar{l})$ is increasing in l, while $\tilde{z}^{over}(l,x) = \tilde{w}^{FB}(l,x) + h_{over}(l)$ has slope $1+\tilde{\omega}/2 = \rho$. Hence its utility index $\tilde{z}^{over}(l,x) + B(l)$ is constant across states. Since $\tilde{z}^{over}(l,x)$ and $\tilde{w}^{FB}(l,x)$ have the same mean, whereas the utility index associated with $\tilde{w}^{FB}(l,x)$ is non-constant, Jensen's inequality implies

$$\mathbb{E}_l[v(\tilde{z}^{over}(l,x);l)] > \mathbb{E}_l[v(\tilde{w}^{FB}(l,x);l)].$$

The value of risk protection associated with $\tilde{z}^{over}(l,x)$ is therefore strictly greater than $\Psi^{FB}(x)$. Finally, $\tilde{w}(l,x) = \tilde{z}^{over}(l,x) - \Delta SCMH^{over}(x)$ is a uniform level shift, which leaves the value of risk protection unchanged under the maintained interiority. Thus $\Psi(x) > \Psi^{FB}(x)$ and $\Delta\Psi^{over}(x) > 0$.

The Social Cost of Moral Hazard. By Lemma 6(a) the first-best schedule is undistorted, $m^{FB}(l) = l$, so the first-best utilization hypothetical carries no deadweight loss, $D^{FB}(l,x) = 0$ in every state and $SCMH^{FB}(x) = 0$. Since $\Delta SCMH^{under}(x) = SCMH^{FB}(x) - SCMH^{N}(x)$ by definition, the no-response social cost and its under-utilization increment cancel,

$$SCMH^{N}(x) + \Delta SCMH^{under}(x) = SCMH^{FB}(x) = 0,$$

and the social cost of moral hazard reduces to its over-utilization component, $SCMH(x) = \Delta SCMH^{over}(x)$. Because preferences are quasilinear in non-health consumption, demand at undistorted prices is independent of income, so the only deadweight loss is the substitution effect of the price distortion on healthcare—over-utilization—while the income transfers that insurance achieves are frictionless except where the budget binds.[58] Combining with the sign results above, the six-term welfare decomposition collapses to the four relevant terms Ψ^{N}, $\Delta\Psi^{under}$, $\Delta\Psi^{over}$, and $\Delta SCMH^{over}$, the first three weakly positive and the additive specification further setting $\Delta\Psi^{over} = 0$.

A.4 The Marginal Approach to Insurance Valuation

The welfare measure studied in this paper, $V(x)$, evaluates a discrete change in insurance coverage: contract x versus status quo contract x_0.[59] One could alternatively evaluate the welfare generated by a *marginal* change in coverage. This approach differentiates the consumer's well-being with respect to coverage, comparing a status quo contract x to a marginally higher-coverage contract "$x + \epsilon$", with the derivative then commonly expressed in terms of estimable sufficient statistics (Baily, 1978; Chetty, 2006). Applied to health insurance, this approach delivers the familiar formula in which the value of marginally higher coverage is a covariance between the marginal utility of consumption and the consumption transfer that incremental coverage delivers, net of the cost of the behavioral response (Finkelstein, Hendren and Luttmer, 2019). The welfare effect of any discrete change in coverage can also be evaluated using this approach, by integrating the derivative over the relevant range of coverage.

The central focus of this paper is a decomposition of the welfare generated by insurance into that which would be generated absent any behavioral response, $V^{N}(x)$, and the social value of the behavioral response, $\Delta V^{MH}(x)$ (Equation 9). Taking x to be full insurance, our quantitative analysis demonstrates that the social value of the behavioral response to

[58] The no-response deadweight loss $SCMH^{N}$ is exactly the income-mismatch cost incurred at budget-corner states, where the consumer cannot afford her efficient healthcare; it is undone, dollar for dollar, by the under-utilization step, leaving over-utilization as the sole net contributor to the social cost.

[59] The status quo contract need not be the null contract, but we will use that here for simplicity.

the discrete step in coverage from x_0 to x can be positive and economically large. Viewed at any infinitesimal margin of coverage, however, the behavioral response is a pure cost. It raises the consumer's fair premium but, by the envelope theorem, has no further first-order effect on her utility in any state, and thus leaves the relevant covariance unchanged. The behavioral response to marginal coverage will therefore always have a negative social value. Nonetheless, these two results are not in tension. The risk protection value of the behavioral response to insurance—of moral hazard utilization—is an intrinsically non-marginal object. While it is zero at any marginal increase in coverage, it accumulates over a discrete change because the behavioral response raises the utilization *levels* at which each subsequent marginal change is evaluated. In this section, we formalize this observation by presenting our welfare decomposition in its derivative form.[60]

Throughout this section, we consider a one-parameter family of contracts in $\mathcal{X}$, taking the out-of-pocket cost function $c(m;x)$ to be twice continuously differentiable in (m,x). Contracts are ordered so that increases in x denote more generous coverage. Formally, $c_x(m;x) \leq 0$ and $c_{mx}(m;x) \leq 0$, so that an increase in x makes healthcare weakly cheaper at every utilization level, and also marginally cheaper at higher utilization levels. We also assume that the consumer's income is sufficiently high that she is always at her interior optimum, that her optimal utilization is differentiable in x, and that her second-order condition holds strictly, so that her optimal utilization is continuously differentiable in the health state and in income.

Marginal Value of Coverage. Recall the consumer's expected utility at the fair premium is given by $U(x, w_0 - \bar{p}(x)) = \mathbb{E}_l\big[u\big(w_0 - \bar{p}(x) - c(m^*(l,x);x)\,,\, m^*(l,x); l\big)\big]$, and her fair premium by $\bar{p}(x) = \mathbb{E}_l\big[m^*(l,x) - c(m^*(l,x);x)\big]$. The effect of a marginal increase in coverage on expected utility is then:

$$\frac{d}{dx}\,U(x, w_0 - \bar{p}(x)) = \mathbb{E}_l\Big[\,\underbrace{(u_m - u_y\,c_m)}_{\substack{=0 \text{ by consumer}\\ \text{optimization}}}\,\frac{\partial m^*}{\partial x} - u_y\,c_x\Big] - \mathbb{E}_l[u_y]\cdot\underbrace{\mathbb{E}_l\Big[(1-c_m)\,\frac{\partial m^*}{\partial x} - c_x\Big]}_{\text{fair premium adjustment}},$$

where the derivatives of u are evaluated at $(y^*(l,x), m^*(l,x); l)$, those of c at $(m^*(l,x);x)$, and those of m^* at (l,x).[61] The term $(u_m - u_y\,c_m)$ equals zero by the consumer's first-order condition for $m^*(l,x)$, leaving the behavioral response $\frac{\partial m^*}{\partial x}$ present only in the premium adjustment. Collecting the c_x terms into a covariance, the expression becomes the familiar

[60]We thank Nathan Hendren for correspondence that prompted this section.

[61]In the main text we measure welfare in dollars of initial income, $V(x)$. Here, we work directly with expected utility, which keeps the expressions closer to the sufficient-statistics literature. The two are related by $V'(x) = \frac{d}{dx}U(x, w_0 - \bar{p}(x))/\mathbb{E}_l[v_w(w_0 + V(x); l)]$.

optimal insurance formula:

$$\frac{d}{dx}U(x,w_0-\bar{p}(x)) = \underbrace{\mathrm{Cov}_l\big(u_y,\,-c_x\big)}_{\substack{\text{risk protection value of}\\ \text{incremental coverage}}} - \underbrace{\mathbb{E}_l\big[u_y\big]\cdot\mathbb{E}_l\Big[(1-c_m)\,\frac{\partial m^*}{\partial x}\Big]}_{\substack{\text{cost of incremental}\\ \text{moral hazard utilization}}}. \tag{12}$$

Incremental coverage generates risk protection value only insofar as out-of-pocket cost reductions are concentrated in states where the marginal utility of non-health consumption is high. The cost is the expected marginal deadweight loss from moral hazard utilization, which the consumer pre-pays through the fair premium.

Marginal Value of Coverage Absent Behavioral Response. Now consider the value of the same marginal increase in coverage under a no-behavioral-response hypothetical: in every state, utilization remains at the status quo schedule $m^*(l,x)$, and the coverage increase is priced fairly given that behavior. Differentiating as before, with $\partial m^*/\partial x$ replaced by zero, the fair premium adjustment reduces to $-\mathbb{E}_l[c_x]$, leaving:

$$\frac{d}{dx}U(x,w_0-\bar{p}(x))\bigg|_{\frac{\partial m^*}{\partial x}=0} = \mathrm{Cov}_l\big(u_y,\,-c_x\big), \tag{13}$$

with every term evaluated at the status quo allocation, exactly as in Equation 12. Preventing the behavioral response to a marginal increase in coverage thus removes the cost term and nothing else.

Value of a Discrete Increase in Coverage. The value of a discrete increase in coverage can now be calculated by integrating the expression in Equation 12 over the relevant range. Write $\mathrm{Cov}^*(t)$ for the covariance in Equation 12 evaluated at contract t, and similarly for the terms of the cost integrand. The value of the discrete increase in coverage from x_0 to x can then be written as:

$$U(x,w_0-\bar{p}(x))-U(x_0,w_0) = \int_{x_0}^{x}\mathrm{Cov}^*(t)\,dt - \int_{x_0}^{x}\mathbb{E}_l\big[u_y\big]\,\mathbb{E}_l\Big[(1-c_m)\,\frac{\partial m^*}{\partial t}\Big]\,dt. \tag{14}$$

Value of a Discrete Increase in Coverage Absent Behavioral Response. Recall the no-behavioral-response hypothetical of the main text, under which the consumer is provided contract x but constrained to utilize $m^N(l) = m^*(l,x_0)$ in every state, paying fair premium $\bar{p}^N(x) = \mathbb{E}_l\big[m^N(l) - c(m^N(l);x)\big]$ and receiving non-health consumption $y^N(l,x) = w_0 - \bar{p}^N(x) - c(m^N(l);x)$. Expected utility under the hypothetical is $U^N(x,w_0-\bar{p}^N(x)) \equiv$

$\mathbb{E}_l\big[u\big(y^N(l,x),\, m^N(l);l\big)\big]$, and differentiating with respect to coverage level gives:

$$\frac{d}{dx}\,U^N(x, w_0 - \bar{p}^N(x)) \;=\; \mathrm{Cov}_l\big(u_y^N,\, -c_x\big), \tag{15}$$

where $u_y^N(l,x) \equiv u_y\big(y^N(l,x),\, m^N(l);l\big)$ and where c_x is now evaluated at $(m^N(l);x)$. There is again no cost term, but note that the covariance is no longer evaluated at the equilibrium allocation, and thus differs from the one in Equations 12 and 13 as long as $x \neq x_0$.

Writing $\mathrm{Cov}^N(t)$ for the no-behavioral-response covariance in Equation 15 at contract t, and integrating over the range of coverage from x_0 to x gives the value of the discrete increase in coverage when utilization is held fixed at the status quo level, or in other words, when the behavioral response is prevented:

$$U^N(x, w_0 - \bar{p}^N(x)) - U(x_0, w_0) \;=\; \int_{x_0}^{x} \mathrm{Cov}^N(t)\,dt. \tag{16}$$

Value of Moral Hazard Utilization. Subtracting Equation 16 from Equation 14, the consumer's value of the ability to change her behavior in response to the discrete increase in coverage from x_0 to x, expressed in utils and using the marginal approach, is:

$$\begin{aligned} U(x, w_0 - \bar{p}(x)) - U^N(x, w_0 - \bar{p}^N(x)) \;=\; & \underbrace{\textstyle\int_{x_0}^{x} \big[\mathrm{Cov}^*(t) - \mathrm{Cov}^N(t)\big]\,dt}_{\substack{\text{risk protection value of moral}\\ \text{hazard utilization}}} \\ & - \underbrace{\textstyle\int_{x_0}^{x} \mathbb{E}_l\big[u_y\big]\,\mathbb{E}_l\Big[(1-c_m)\,\frac{\partial m^*}{\partial t}\Big]\,dt.}_{\substack{\text{cost of moral hazard}\\ \text{utilization}}} \end{aligned} \tag{17}$$

The cost term accumulates the per-margin cost of the behavioral response from Equation 12. The covariance gap is zero at x_0 and grows as $m^*(l,t)$ rises above $m^N(l)$. Incremental coverage pays the consumer more when her utilization is $m^*(l,t)$ than it would at $m^N(l)$, and the gap is biggest in the sick states. The behavioral response at each margin thus enters its own margin only as a cost, but raises the relevant covariance at every subsequent marginal increase in coverage.

Appendix B Computational Appendix

B.1 Population Construction and Application of Estimates

This appendix describes our procedure for constructing a population of consumers and applying the parameter estimates used in each of the three papers (and four specifications) we study. The four specifications are outlined in Table 1. Across all specifications, we model consumers as households, where a household is a group of individuals. Each individual is characterized by an age, a gender, and a health risk score. We construct a population of households to match characteristics of the non-elderly US population reported in the American Community Survey (ACS) for 2023 (U.S. Census Bureau, 2023). The construction of each household begins with a "head of household," who is female with 50 percent probability and has a uniform distribution of age between 22 and 65. We assume that households have a spouse present with a 90 percent chance, and that the spouse is of the opposite gender to the head of household when present. The age of the spouse is drawn from a normal distribution with mean equal to the age of the head of household and a standard deviation of 4, subject to bounds between 22 and 65. Each household can have between 0 and 4 children, where each child independently exists with 15 percent probability. Conditional on existing, a child is female with 50 percent probability and her age is drawn from a uniform distribution between 0 and 18.

Each paper we study uses risk scores as a measure of individual health risk. To model this, all individuals in our population draw a risk score from a log-normal distribution with mean positively related to age, such that for individual i: $\log(riskscore_i) \sim \mathcal{N}(\frac{age_i}{20}, 1)$. The right tail of the risk score distribution is censored such that no individual has a risk score greater than five standard deviations above the uncensored mean. Our baseline population contains 1,000 households; increasing the number of households does not change our results. The household-size and age draws are held fixed across the four specifications in Table 1; summary statistics on these characteristics are reported in the Demographics top section of Table B.1. What varies across specifications is the source of parameter estimates for the structural primitives that describe each consumer: initial income w_0, the health state distribution F, the risk aversion parameter ψ, and the "moral hazard" parameter ω (or $\tilde{\omega}$).

We express monetary amounts in thousands of dollars and rescale each paper's coefficient of absolute risk aversion accordingly. Two adjustments are applied uniformly across populations. First, the risk aversion parameter is capped at $\psi = 5$ per \$000.[62] Second, the income draws are floored at $1.3 \times \mathbb{E}_l\big[(m^{FB}(l) + m^*(l, x_{\text{bliss}}))/2\big]$ to ensure each household can in expectation

[62] Among the three populations we consider, the cap binds only for a small upper tail in Marone and Sabety (2022); the risk aversion estimates from Einav et al. (2013) and Ho and Lee (2023) lie well below the cap.

afford its bliss-point level of healthcare utilization. This floor binds for a small minority of households in each population and provides numerical headroom for the optimization that recovers $m^*(l, x)$ in high-coverage contracts.

Marone and Sabety (2022). We use the parameter estimates reported in Column 3 of Table 3 and Appendix Table A.8 of Marone and Sabety (2022). Each household draws $(\psi, \omega, \mu, \sigma, \kappa)$ from the estimated joint distribution, where μ, σ, and κ are the parameters of a shifted log-normal health state distribution, $l = \exp(\mu + \sigma z) - \kappa$ with $z \sim \mathcal{N}(0, 1)$. The data underlying this paper do not include household income; we calibrate the income distribution to a log-normal with mean \$55,000 and coefficient of variation 0.4.[63]

Summary statistics on the resulting household types are reported in Panel A of Table B.1. Facing an equal-odds gamble between \$0 and \$100, the average household would have a certainty equivalent of \$48.83, reflecting risk aversion. For the average household, there is a 1 percent probability of realizing a health state in which their first-best level of healthcare utilization exceeds their initial income, implying there will necessarily be under-utilization absent insurance. This measure is highly skewed: even the 75th-percentile household never experiences being constrained in this way. The average household would have total healthcare spending equal to \$9,768 under a full insurance contract, but only \$8,333 under a null contract, reflecting moral hazard utilization. Under first-best contracts, the average household would have total healthcare spending of \$8,419.

Einav et al. (2013). We use the parameter estimates reported in Table 7A of Einav et al. (2013) for the additive specification (2), and the estimates from Online Appendix Table A6 for the multiplicative specification (3). The estimated parameters are functions of household-level covariates (family composition, demographics, treatment group, income, job tenure, and risk-score quartile). We assign every household to the "Switch 2006" treatment group (the largest in the EFRSC sample) and set job tenure to its group mean of 5.7 years. Since Einav et al. (2013) report only the mean *individual* income (Table 1), we construct household income by scaling this number by the mean number of adults per household and drawing from a log-normal with the same coefficient of variation as in specification (1).[64]

Summary statistics on the resulting household types are reported in Panel B (additive)

[63] The mean is calibrated to be consistent with the ACS mean income for employees in Oregon in 2010, since the Marone and Sabety (2022) estimates of the health state distribution come from a sample of Oregon public school employees between 2008 and 2013. The mean household income for Oregon employees in 2010 was \$61,095 (ACS Table S1902); we adjust the mean down to \$55,000 to capture the fact that public school employees have below-median wages (Oregon Public Employees Retirement System, 2024).

[64] The mean individual income reported in Einav et al. (2013) is \$31,292; with a mean of 1.9 adults per household in our population, this gives a mean household income of \$59,454.

Table B.1. Population Demographics

Sample demographic	Mean	Percentile 10	25	Median	75	90
Demographics						
Number of adults	1.9	2.0	2.0	2.0	2.0	2.0
Number of children	0.6	–	–	–	1.0	2.0
Average age of household adults	43.4	26.5	32.5	43.4	54.1	61.1
Panel A. Marone and Sabety (2022)						
Initial income, w_0 ($000)	55.4	31.9	40.2	51.5	66.7	86.9
Risk aversion parameter, ψ	0.9	0.2	0.4	0.7	1.2	2.0
Moral hazard parameter (additive), ω ($000)	1.3	0.8	1.0	1.3	1.7	2.0
CE of equal-odds gamble between $0 and $100 ($)	48.8	47.6	48.5	49.2	49.5	49.7
Prob. of realizing health state in which $w_0 < m^{FB}(l)$	0.01	–	–	–	–	0.01
Utilization under null contract, $\mathbb{E}[m^N(l)]$ ($000)	8.3	2.6	3.9	6.8	11.4	16.4
full insurance, $\mathbb{E}[m^*(l,x)]$ ($000)	9.7	3.7	5.2	8.2	12.9	18.2
Panel B. Einav et al. (2013) (Additive spec)						
Initial income, w_0 ($000)	60.1	34.6	43.6	56.0	72.4	94.3
Risk aversion parameter, ψ	1.8	0.7	1.0	1.5	2.3	3.4
Moral hazard parameter (additive), ω ($000)	1.1	0.2	0.2	0.3	1.0	2.9
CE of equal-odds gamble between $0 and $100 ($)	47.7	45.8	47.1	48.1	48.7	49.1
Prob. of realizing health state in which $w_0 < m^{FB}(l)$	<0.01	–	–	–	–	<0.01
Utilization under null contract, $\mathbb{E}[m^N(l)]$ ($000)	4.5	0.7	1.7	3.3	6.0	10.0
full insurance, $\mathbb{E}[m^*(l,x)]$ ($000)	5.7	0.9	2.1	3.9	7.6	12.6
Panel C. Einav et al. (2013) (Multiplicative spec)						
Initial income, w_0 ($000)	60.1	34.6	43.6	56.0	72.4	94.3
Risk aversion parameter, ψ	0.6	0.0	0.0	0.1	0.5	1.8
Moral hazard parameter (multiplicative), $\tilde{\omega}$	0.3	0.0	0.1	0.1	0.3	0.7
CE of equal-odds gamble between $0 and $100 ($)	49.3	47.8	49.4	49.8	49.9	50.0
Prob. of realizing health state in which $w_0 < m^{FB}(l)$	<0.01	–	–	–	–	<0.01
Utilization under null contract, $\mathbb{E}[m^N(l)]$ ($000)	3.0	0.4	0.7	1.6	3.9	7.9
full insurance, $\mathbb{E}[m^*(l,x)]$ ($000)	3.8	0.5	0.9	2.0	4.9	9.5
Panel D. Ho and Lee (2023)						
Initial income, w_0 ($000)	78.9	34.5	45.8	68.5	100.9	136.6
Risk aversion parameter, ψ	0.135	0.135	0.135	0.135	0.135	0.135
Moral hazard parameter (multiplicative), $\tilde{\omega}$	0.18	0.04	0.04	0.26	0.26	0.26
CE of equal-odds gamble between $0 and $100 ($)	49.8	49.8	49.8	49.8	49.8	49.8
Prob. of realizing health state in which $w_0 < m^{FB}(l)$	<0.01	–	–	–	–	<0.01
Utilization under null contract, $\mathbb{E}[m^N(l)]$ ($000)	5.9	1.0	1.8	3.9	8.3	13.9
full insurance, $\mathbb{E}[m^*(l,x)]$ ($000)	7.0	1.1	2.2	4.6	10.0	16.1

Notes: This table shows the distribution of consumer demographics and types used in our numerical analysis. Each population consists of 1,000 consumers (households). The top block reports characteristics fixed by the shared ACS demographic scaffold; each of Panels A–D reports the population-specific characteristics for one of the four specifications. "CE" stands for certainty equivalent. "FB" stands for first best. Specifications (1)–(4) are in reference to the parameterizations of utility presented in Table 1 in the main text.

and Panel C (multiplicative) of Table B.1. The average household certainty equivalent for an equal-odds gamble between \$0 and \$100 is \$47.70 (additive) and \$49.30 (multiplicative). The average household would have total healthcare spending of \$5,710 (additive) and \$3,803 (multiplicative) under a full insurance contract, and \$4,520 (additive) and \$3,009 (multiplicative) under a null contract, reflecting moral hazard utilization.

Ho and Lee (2023). We use the parameter estimates reported in Online Appendix Table B4 of Ho and Lee (2023). Because the risk aversion estimate in this table is so low as to imply no valuation of insurance for most households, we use the risk aversion estimate from Online Appendix Table B6 ($\psi = e^{-2} \approx 0.135$ per \$000) rather than the Table B4 estimate.[65] The estimated parameters are functions of household-level categorical variables (union status, family type, salary tier, age, and risk-score quartile). The moral hazard parameter is a two-type mixture, with $\tilde{\omega} \approx 0.04$ for the low type (36 percent of households) and $\tilde{\omega} \approx 0.26$ for the high type (64 percent). Household income is drawn from a three-tier salary structure: we assign each household to a tier with probabilities that depend on union status, and draw within-tier income from a log-normal with coefficient of variation 0.4.[66] Finally, Ho and Lee (2023) model the consumer's choice on both the extensive and intensive margins of healthcare utilization; we incorporate the extensive margin by setting health states $l < 2\tilde{\omega}\zeta$ to zero in each household's distribution, where ζ is the estimated household-specific hassle cost.[67]

Summary statistics on the resulting household types are reported in Panel D of Table B.1. Because ψ is constant across households, the certainty equivalent for an equal-odds gamble between \$0 and \$100 is fixed at \$49.83 for every household. The average household would have total healthcare spending of \$6,985 under a full insurance contract and \$5,883 under a null contract, reflecting moral hazard utilization.

[65] The Table B4 main estimate ($\psi \approx 3.07 \times 10^{-3}$ per \$000) is near risk-neutrality and produces a negative value of full insurance for most households, making the welfare decomposition uninformative. The Table B6 value is the one the authors use when imposing more conservative bounds on the consumer's choice problem; it also brings the risk aversion estimate closer to those of Marone and Sabety (2022) and Einav et al. (2013).

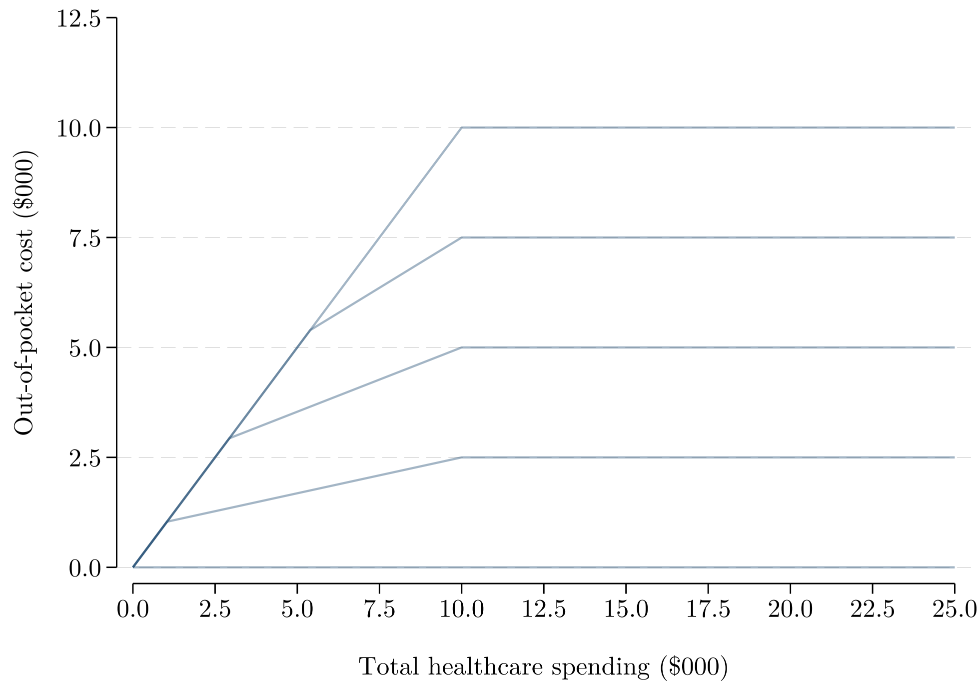
[66] Tier boundaries follow Harvard employee salary brackets; Ho and Lee (2023) does not report a within-tier income distribution.

[67] Below the participation threshold the consumer would optimally choose zero healthcare utilization. At $l = 0$, any positive care yields negative infinite utility, so the consumer strictly chooses $m = 0$. Because $b(0; 0) = 0$ in this limit of our parametrization, replacing the realized health state with zero in this region is observationally equivalent to the consumer's optimal extensive-margin choice.

B.2 Insurance Contracts

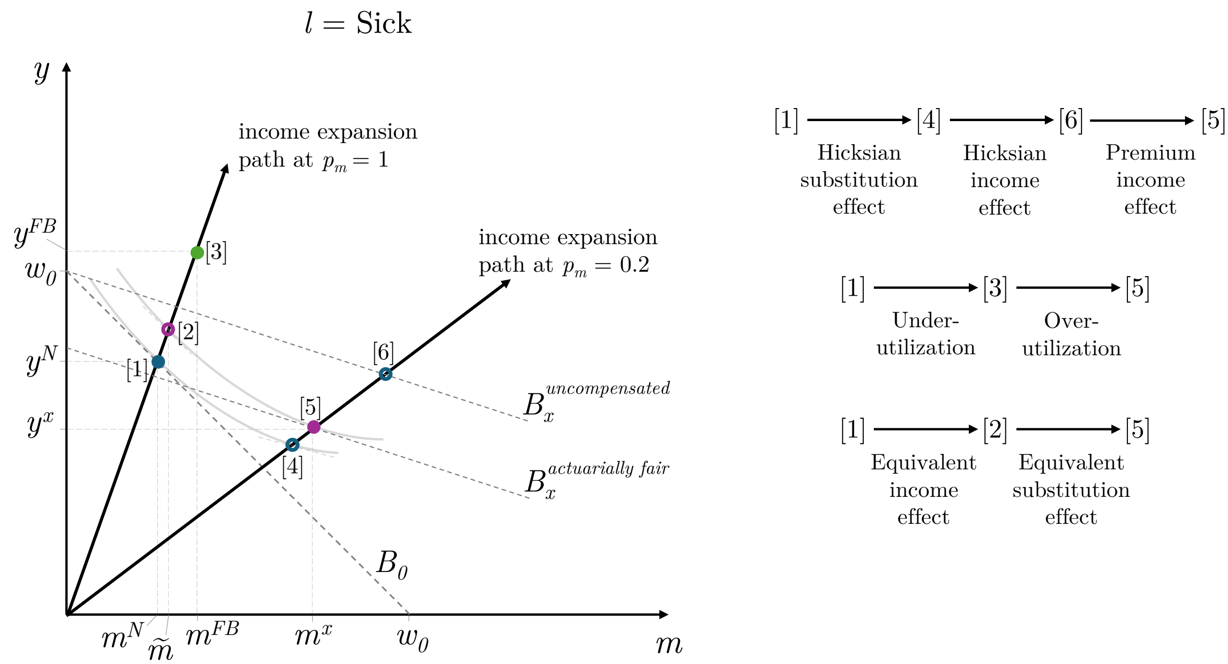
We consider a set of five contracts that are piecewise linear, with a deductible, coinsurance region, and out-of-pocket maximum design. The lowest level of coverage considered is a "Catastrophic" contract with a deductible and out-of-pocket maximum of $10,000. The highest level of coverage is full insurance. The out-of-pocket cost functions for the five contracts are depicted in Figure B.1. Analysis in Chade et al. (2023) suggests that adding additional contracts would not have material effects on achievable second-best welfare. The contracts' deductibles, coinsurance rates, and out-of-pocket maximums are: $5,846, 40%, $7,500 for Bronze; $3,182, 27%, $5,000 for Silver; and $1,125, 15%, $2,500 for Gold. In our population of consumers, the actuarial value of the five contracts are: 0.40, 0.49, 0.61, 0.79, and 1.00.

Figure B.1. Potential Contracts



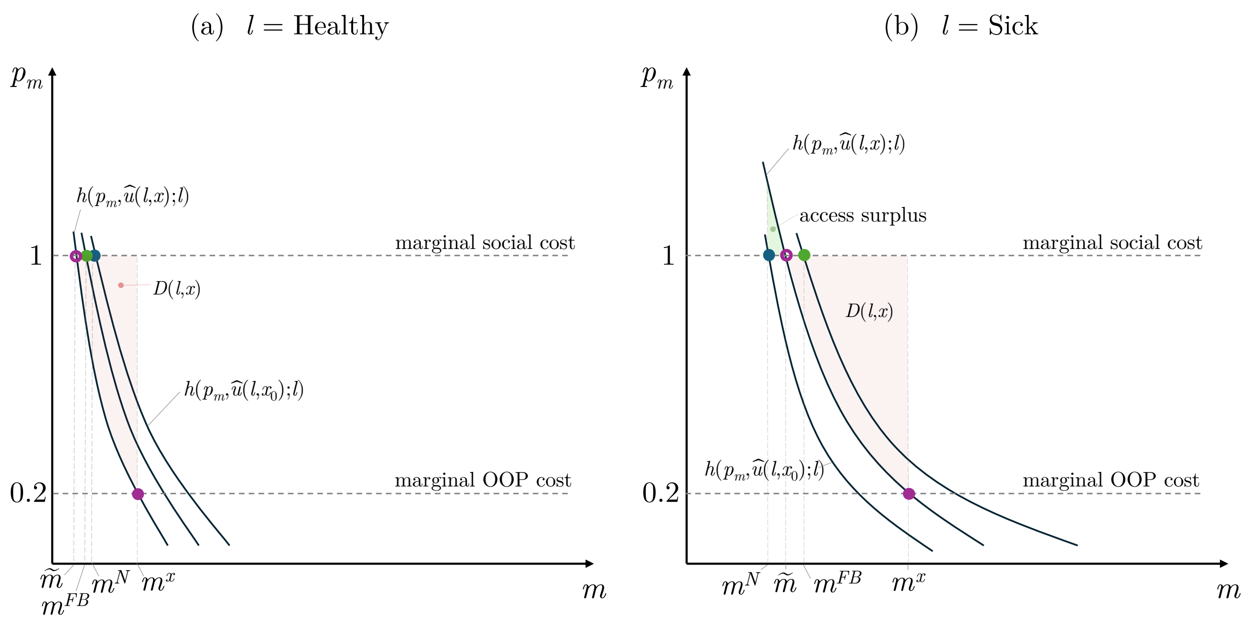
Notes: The figure shows the set of potential contracts from which the consumers' efficient level of coverage is determined.

Figure A.1. Income vs. Substitution Effects



Notes: This figure depicts three alternative decompositions of the movement from the no-insurance bundle [1] to the insured bundle [5] in the sick state under contract x (20 percent coinsurance), corresponding to the two-state example in Figure 1. The six numbered bundles are: [1] the no-insurance optimal bundle; [2] the equivalent-income optimal bundle (on the income expansion path at $p_m = 1$ at equivalent income $\tilde{w}(l, x)$); [3] the first-best bundle (on the income expansion path at $p_m = 1$ at first-best income $w^{FB}(l)$); [4] the Hicksian compensated bundle at the no-insurance utility level and insured price (on the no-insurance indifference curve at the $p_m = 0.2$ tangency); [5] the equilibrium bundle under contract x (on the income expansion path at $p_m = 0.2$ on actuarially fair budget line B_x); [6] the optimal bundle at insured prices and original income (on the income expansion path at $p_m = 0.2$, on budget line $B_x^{uncompensated}$). Three decompositions of the total change in demand [1]→[5] are shown. The first is the canonical Hicksian decomposition—with substitution and income effects compensated at the no-insurance utility level (the compensating variation)—augmented by a separate premium income effect. The second is the decomposition between under- and over-utilization. The third is also a Hicksian-type decomposition, but with substitution and income effects compensated at the equilibrium utility level (the equivalent variation), and with the two income effects collapsed together. This figure is referenced in Section II.A.

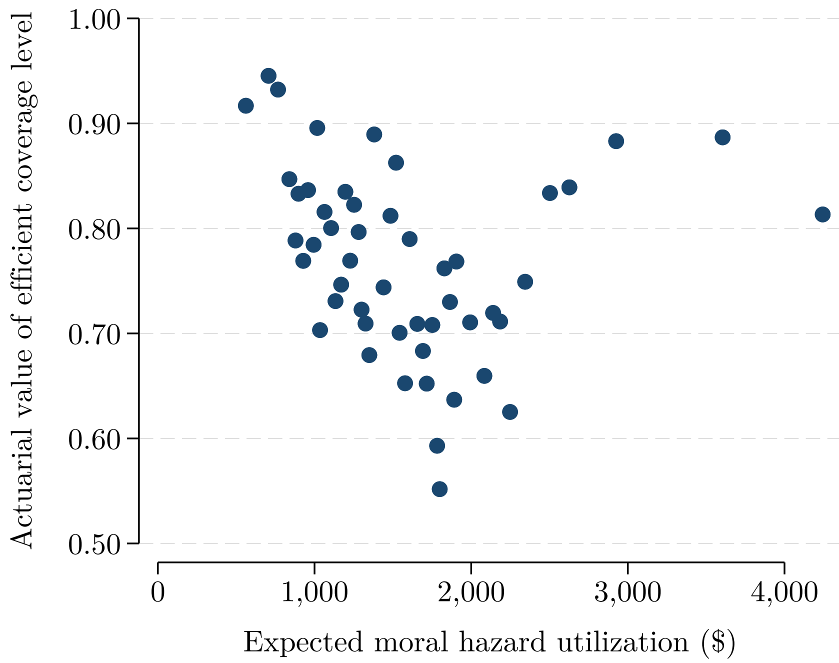
Figure A.2. Compensated Demand Curves for Healthcare



Notes: This figure presents the demand-based analog to Figure 1, using price–quantity axes (and compensated demand curves) instead of commodity-commodity axes (and indifference curves). It is thus the the analog to Figure 1 in Pauly (1968), once income effects are incorporated. The vertical axis is the consumer price of healthcare p_m, and the horizontal axis is healthcare utilization m. The social cost of care is fixed at $p_m = 1$. Each panel shows the compensated demand curves for healthcare $h(p, \hat{u}; l)$ that correspond to each of the indifference levels indicated on Figure 1, representing the marginal rate of substitution u_m/u_y along an indifference curve (and thus the derivative of that indifference curve). $\hat{u}(l, x_0)$ is the no-insurance utility level; $h(p, \hat{u}(l, x_0); l)$ thus passes through the no-insurance utilization level m^N at the undistorted price $p_m = 1$. $\hat{u}(l, x)$ is the equilibrium utility level under contract x; $h(p, \hat{u}(l, x); l)$ thus passes through the equilibrium utilization level m^x at the insured price $p_m = 0.2$. The under-/over-utilization decomposition splits the change in utilization $m^N \to m^x$ at m^{FB}, while the income/substitution decomposition splits it at $\tilde{m}$ (corresponding to bundle [2] in Figure A.1).

The red shaded region $D(l, x)$ is the deadweight loss in each state, equal to the area between the equilibrium demand curve and marginal social cost from $\tilde{m}$ to m^x (the substitution effect on m at the equilibrium utility level). In the sick state, insurance raises the consumer's equivalent income from w_0 to $\tilde{w}(l, x)$, and with it her optimal utilization at undistorted prices from m^N to $\tilde{m}$. We call the green shaded region above the marginal-social-cost line between these two quantities the *access surplus*: the value, in excess of resource cost, of this additional utilization, measured along the equilibrium compensated demand curve. Its sign tracks the direction of the income transfer insurance provides. In the sick state, insurance raises equivalent income, and the access surplus is thus positive (given healthcare is a normal good). In the healthy state, insurance instead lowers equivalent income, so $\tilde{m}$ falls below m^N and the access surplus is negative. To avoid clutter, this region is not drawn in panel (a). This figure is referenced in Section II.A.

Figure A.3. Efficient Coverage Level versus Expected Moral Hazard Utilization



Notes: The figure shows the relationship between a consumer's efficient coverage level (measured in actuarial value) and their expected increase in healthcare utilization when moving from no insurance to full insurance (expected moral hazard utilization). It is a binned scatter plot using 50 points, across the 334 households in the bottom third of the income distribution using specification (1). This figure is referenced in Section III.B.

Table A.1. Decomposition of Welfare: Population Subgroups

	Decomposition of healthcare utilization			Decomposition of welfare				Relative change in welfare		
	Non-MH utiliz.	Under-utiliz.	Over-utiliz.	Ψ^N	$\Delta\Psi^{und.}$	$\Delta\Psi^{ovr.}$	$-SCMH$	Efficient coverage	Prevent over-util.	Prevent MH-util.
Panel B.1 [Einav et al. (Addv.)]										
All	0.81	<0.01	0.18	0.93	0.09	–	−0.02	<0.01	0.02	−0.07
Tertile of under-utiliz.										
Lowest	0.82	–	0.18	1.02	–	–	−0.02	<0.01	0.02	0.02
Middle	0.79	<0.01	0.21	0.55	0.46	–	−0.01	<0.01	<0.01	−0.45
Highest	0.80	0.01	0.18	0.47	0.55	–	−0.02	<0.01	0.02	−0.53
Tertile of over-utiliz.										
Lowest	0.87	<0.01	0.13	0.88	0.12	–	−0.01	<0.01	<0.01	−0.12
Middle	0.86	<0.01	0.14	0.93	0.07	–	−0.01	<0.01	<0.01	−0.07
Highest	0.71	<0.01	0.29	1.00	0.06	–	−0.06	0.01	0.06	−0.01
Panel B.2 [Einav et al. (Mult.)]										
All	0.80	<0.01	0.19	1.03	0.03	0.05	−0.11	0.02	0.06	0.02
Tertile of under-utiliz.										
Lowest	0.81	–	0.19	1.06	–	0.05	−0.11	0.02	0.06	0.06
Middle	0.76	<0.01	0.23	0.93	0.09	0.06	−0.07	0.03	0.01	−0.07
Highest	0.82	0.04	0.14	0.89	0.20	0.03	−0.12	0.02	0.09	−0.17
Tertile of over-utiliz.										
Lowest	0.93	<0.01	0.07	0.98	0.05	<0.01	−0.03	<0.01	0.03	−0.03
Middle	0.85	<0.01	0.15	1.05	0.03	0.02	−0.09	0.03	0.08	0.05
Highest	0.69	<0.01	0.30	1.04	0.02	0.11	−0.17	0.03	0.06	0.03
Panel B.3 [Ho and Lee]										
All	0.85	<0.01	0.14	1.19	0.02	<0.01	−0.22	0.05	0.21	0.19
Tertile of under-utiliz.										
Lowest	0.86	–	0.14	1.22	–	<0.01	−0.23	0.05	0.22	0.22
Middle	0.88	<0.01	0.12	1.20	0.04	<0.01	−0.25	0.07	0.25	0.20
Highest	0.83	0.02	0.15	0.90	0.18	<0.01	−0.08	<0.01	0.08	−0.11
Tertile of over-utiliz.										
Lowest	0.95	<0.01	0.05	1.09	0.03	<0.01	−0.11	0.02	0.11	0.09
Middle	0.85	<0.01	0.15	1.30	0.02	<0.01	−0.33	0.09	0.32	0.30
Highest	0.79	<0.01	0.21	1.16	0.02	<0.01	−0.19	0.02	0.18	0.16

Notes: This table extends the subgroup analysis of Panel B of Table 2 to the three specifications other than Marone and Sabety (2022). Each panel reports the decomposition of healthcare utilization and welfare generated by full insurance for one specification, following the same construction as Table 2: the first three columns express each component of expected healthcare utilization as a share of total expected utilization; the next four columns express each component of the welfare decomposition in Equation 10 as a share of total welfare V, on average across consumers; and the final three columns report relative changes in welfare in three scenarios (enrollment in the efficient coverage level, prevention of over-utilization, and prevention of all moral hazard utilization). Within each panel, the *All* row reports the population average—matching the corresponding row of Panel A of Table 2—and the subgroup rows stratify consumers into tertiles by their expected dollar amount of under-utilization absent insurance and, separately, by their propensity to over-utilize under full insurance (the parameter ω). This figure is referenced in Section III.B.