Efficient Health Insurance with Adverse Selection and Ex-Post Moral Hazard

Alex Boche*

February 15, 2026

(Job Market Paper)

Updated regularly. Click here for the latest version.

Abstract

I study consumer-optimal health insurance under adverse selection and ex-post moral hazard while viewing risk from an ex-ante perspective. I show that, under a simplification of incentive constraints known as the first-order approach (FOA), optimal linear coverage will be set higher than would be efficient without adverse selection. This distortion relaxes upward incentive constraints for lower-risk consumers, thereby allowing for greater cross-subsidization from lower-risk to higher-risk consumers, which is important for insuring dynamic risk. I then characterize the optimal menu of nonlinear coinsurance schedules under the FOA. Results inform the design of both private-sector dynamic health insurance contracts and social health insurance programs with a participation mandate administered by governments or employers.

1 Introduction

Healthcare is a major part of all developed economies. Since healthcare costs can be volatile, health insurance is a crucial instrument for managing consumer risk. However, the efficiency

*Northwestern Kellogg, Strategy/MEDS. I gratefully acknowledge financial support from the Kellogg School of Management at Northwestern University. There are no conflicts of interest to report. I am grateful to my committee members Dan Barron (most especially), David Dranove, Alessandro Pavan, and Robert Porter for invaluable feedback and mentorship on this paper. Mike Powell also provided excellent feedback. For valuable comments on the project, I also thank Eduardo Azevedo, Meghan Busse, Shengmao Cao, Piotr Dworczak, George Georgiadis, Mikhail Golosov, Ben Golub, Igal Hendel, Sean Higgins, Victoria Marone, Luis Rayo, Amanda Starc, Bruno Strulovici, Jeroen Swinkels, and seminar participants at Northwestern.

of health insurance markets is limited by three key frictions: ex-post moral hazard, adverse selection, and dynamic risk. Ex-post moral hazard occurs when consumer health states are not contractible and insured consumers overuse care since they don't internalize its full cost [Zeckhauser 1970].[1] This motivates some degree of partial coverage, leaving the consumer responsible for a share known as coinsurance. Adverse selection occurs when consumer *risk* over future health states is not contractible and high-risk consumers differentially select into high-coverage insurance plans; this can lead to high premiums for those plans and insufficient coverage for other plans [Rothschild and Stiglitz 1976]. Dynamic risk is the risk consumers face over what their future premium — and expected out-of-pocket liability — will be if they turn out to be a high-risk individual. Short-term contracts may fail to insure against dynamic risk either due to explicit rating or due to adverse selection [Handel et al. 2015].

This paper studies the design of efficient health insurance when all three frictions are present. I first show that without adverse selection, the combination of moral hazard and dynamic risk can be handled by addressing each separately. More interestingly, I then show that when adverse selection is present, a tradeoff arises between the other two frictions: distorting the solution to the moral hazard problem can help better insure dynamic risk.

I study this interplay through a model in which a risk-averse consumer signs a dynamic contract with a risk-neutral insurer before learning their health risk. The dynamic contract is a commitment to choose from a menu of continuation plans, each of which consists of a premium and coinsurance schedule. After contracting, the consumer privately observes a signal — which I call their "risk-type" — about their future health state, at which point they must choose a plan from the menu and pay the associated premium. They then privately learn their actual health state, after which they choose their treatment amount and pay their corresponding coinsurance. Risk-types are ordered so that higher types are more likely to have more severe health states. Treatment and monetary consumption occur only in the final period. In the model, moral hazard and adverse selection arise due to the privacy of health state and risk-type, while dynamic risk arises through the evolution of risk information.

Although the model describes a private-sector dynamic contract with a single consumer, it can be reinterpreted as a social health insurance program in which a population of consumers with pre-existing risk differences is required to participate. Under this reinterpretation, the model captures part of the motivation for many government and employer-based health insurance programs.

Regardless of the interpretation, the optimal menu of contracts aims to do two things.

[1] This contrasts so-called ex-ante moral hazard which refers to unobservable actions taken by consumers after contracting such as exercise or reckless driving. Throughout, I use "moral hazard" to mean ex-post moral hazard.

First, it addresses moral hazard by making coverage partial in each continuation plan. Second, it mitigates dynamic risk by committing the consumer to a cross-subsidization scheme between their potential low-risk and high-risk future selves. However, since the risk-type is private, such cross-subsidization is limited by a screening problem in that subsidized plans intended for high-risk types may attract the low-risk types meant to be subsidizing them. Further, since coverage will be partial due to moral hazard, attempting to solve this screening problem by simply pooling types into a common plan will not fully address dynamic risk: higher-risk types will still be exposed to higher expected out-of-pocket liability despite a common premium. Instead, the paper's key insight is that we can ease the screening problem by distorting coverage upward for higher-risk plans. Doing so relaxes upward incentive constraints for lower-risk types and thus enables greater cross-subsidization to high-risk types.

The paper develops this insight in three steps, building intuition by adding frictions sequentially. First, I review the standard static moral hazard problem of choosing, for a given risk-type, the consumer-optimal coinsurance schedule (and premium) subject to an actuarial fairness constraint [Blomqvist 1997]. The optimal schedule balances risk-protection and incentives in the most efficient way given the distribution of health states associated with the risk-type. Second, I add dynamic risk but maintain observable risk-types. In this case, the optimal contract exhibits a property I call interim efficiency: each risk-type receives a plan with exactly the coinsurance schedule that would be optimal for them in the static moral hazard problem. Moreover, cross-subsidization fully insures dynamic risk, equalizing expected marginal utilities across risk-types (Proposition 1).

Third, I introduce adverse selection by making risk-types private information, which changes the optimal contract structure by creating a tradeoff between interim efficiency and insuring dynamic risk. The main result (Theorem 1) — which restricts attention to linear coinsurance schedules and a simplified version of incentive constraints known as the first-order approach (FOA) — shows how we can improve cross-subsidization (and thus dynamic risk) at only second-order cost to interim efficiency. Specifically, it states that each risk-type will receive coverage strictly above the interim-efficient level (except the lowest type). This upward distortion serves a specific purpose: when coverage is increased with premiums also raised to offset the budgetary impact, lower-risk types find this compound change less attractive than do higher-risk types, since they value additional coverage less, expecting to use it less. This relaxes upward incentive constraints for lower-risk types and thus enables greater cross-subsidization toward higher-risk types. This is beneficial since higher-risk types have higher expected marginal utility given incomplete coverage due to moral hazard. Crucially, the welfare gain from increased cross-subsidization exceeds the

efficiency loss from distorted coverage levels, which is only second-order at the margin.

I extend this analysis to nonlinear coinsurance schedules again under the FOA (Proposition 3), showing that the same economic force operates but can be targeted more precisely. The distortion relative to the interim optimum now varies across spending levels: larger distortions occur at dollar amounts where a given risk-type's probability of exceeding those amounts is differentially high relative to the risk-types just below them. Since each risk-type's valuation of marginal coverage at a given spending amount depends on their probability of exceeding that amount, such targeting allows for greater relaxation of incentive constraints per unit of distortion and hence for more efficient cross-subsidization.

Since the key results for the private-risk case apply only under the first-order approach (FOA) — which I define as the relaxed contracting problem that considers only upward-adjacent incentive constraints — I provide some intuition for when the FOA is valid in the sense of yielding the same solution as the unrelaxed problem. Since higher-risk types value coverage more than lower-risk types, an allocation of coverage levels can only be implemented under private risk if it gives weakly higher coverage to higher-risk types (Lemma 6). Thus, if interim-efficient coverage levels are increasing in risk-type, interim efficiency accords with implementability. In this case, intuition suggests the FOA is likely to be valid, and I confirm this formally in the two risk-type case with linear coinsurance (Proposition F.2). However, if interim-efficient coverage is sufficiently *anti-monotonic*, it may be possible for the FOA to fail — a possibility that is most intuitive in the linear case. Specifically, if the FOA fails when restricting to linear coinsurance, pooling of risk-types will occur (Proposition F.1), which is the closest one can get to implementing anti-monotonic coverage.[2] That being said, monotonicity of interim-efficient coverage is not *necessary* for validity of the FOA since the cross-subsidization force — which benefits from strictly monotonic coverage via relaxation of upward incentive constraints — may override the interim-efficiency force.

Whether the model is interpreted as a private-sector dynamic contract or a static social insurance program, the same lessons apply. Increasing coverage for plans intended for higher-risk consumers can allow for more cross-subsidization without attracting lower-risk consumers into those better-subsidized plans. In particular, coverage for those plans should be set above the level that would be optimal in a static moral hazard problem for the enrollees therein. Further, such upward distortion can be done at lowest cost to interim efficiency by raising marginal coverage at dollar amounts that higher-risk consumers have high differential likelihood of exceeding relative to lower-risk ones.

An additional lesson is that, to the extent that non-contractibility of risk information

[2]Pooling will never occur when nonlinear coinsurance is allowed even if the FOA fails, but I don't further analyze the structure of the non-FOA solution in the nonlinear case.

stems from community rating laws (i.e., bans) rather than intrinsically private information, social planners — or private insurers offering dynamic contracts — should consider using that information to directly *assign* consumers to plans appropriate to their risk profiles. In principle, such assignment can fully achieve the twin objectives of cross-subsidization and risk-tailored coinsurance schedules. Allowing choice (rather than assignment) of plan is, instead, more appropriate to the extent that risk profiles are unobservable (or at least unverifiable).

Outline of Paper

Section 2 discusses the relation to the literature. Section 3 presents the model. Section 4 formulates and analyzes the static and observable-risk benchmarks used to study the private-risk dynamic problem, which is then formulated in Section 5. Section 6 gives the main distortion result for linear coinsurance (Thm. 1). Section 7 studies the optimal menu of nonlinear contracts. Section 8 develops the aforementioned dual interpretations of the model and discusses connections to national health insurance programs like the Affordable Care Act. Section 9 concludes and offers several directions for future research. All proofs are in the appendix, which also contains extra results and discussion.

2 Literature

My analysis contributes to two different literatures: health insurance and dynamic income taxation. I discuss each in turn. The former papers are closer in terms of the application, while the latter are closer in methodology.

2.1 Health Insurance Literature

The most closely related paper is [Zwart 2025], who studies optimal risk adjustment in a regulated-competition program under moral hazard and adverse selection while (in an extension) allowing a redistributive objective. His Proposition 6 is related in spirit to my main result: in both papers, the contract intended for high-risk types is more generous than in the corresponding observable-risk benchmark in order to provide the screening incentives needed to support cross-subsidization toward high-risk types. My paper and my result differ from his in two main ways. First, in my model, the desire for cross-subsidization from low- to high-risk types arises endogenously under a utilitarian objective with concave utility over monetary consumption: high-risk types have higher expected marginal utility and so transfers to them are of higher social value in expectation. By contrast, [Zwart 2025] builds

in redistribution more directly by specifying higher welfare weights on high-risk consumers. Relatedly, whereas the desire for insurance in my model comes from concave utility over monetary consumption, Zwart's consumer has linear utility over monetary consumption but with an extra disutility weight on copayments made ex-post relative to premiums paid ex-ante.[3] Second, whereas I study a single-payer system where contract features are fully determined by the social planner, he studies the design of risk-adjustment policies set by the planner in a regulated marketplace of competing insurers.[4]

Three other closely related papers are [Marone and Sabety 2022] — along with its follow-up [Chade et al. 2024] — and [Ho and Lee 2023], which all study ex-post moral hazard and adverse selection together, but abstract from dynamic risk. The key difference is that their social planner's objective (which is similar in both papers) uses each consumer's ex-ante consumer surplus from insurance coverage (i.e., willingness-to-pay minus actual premium) to measure their welfare gain, and then aggregates this measure across consumers. My paper builds on these papers by instead using a classical utilitarian objective based on ex-post interpersonal comparisons of utility. This objective — which also appears in [Hendren 2021] who is concerned with welfare measurement, not contract design — leads to the dynamic redistributive force that drives my main result (Thm. 1).

[Boone 2015] and [Boone 2018] also study the combination of moral hazard and adverse selection but, again, without dynamic risk.[5]

[Besley 1988], in an extension, studies moral hazard and (effectively) dynamic risk — with the latter reinterpreted as a static redistributive motive — but without adverse selection (i.e., with observable risk profiles). Although in his formal analysis, his planner is restricted to imposing a common premium across individuals (with individualized coinsurance), he informally notes that if the premiums were allowed to vary by individual, allocative efficiency and redistribution could be addressed separately (p. 332).[6] His comment thus informally captures the spirit of my observable-risk benchmark, Proposition 1.

Several papers study ways to address dynamic risk but without moral hazard. [Luz 2022]

[3] Zwart does have concave utility over health status as do I.

[4] The contrast between my paper and [Zwart 2025] is echoed — albeit considerably more distantly — by that between my paper and [Bijlsma et al. 2017]. The latter is concerned with reinsurance (i.e., insurance for insurers) and resulting insurer-side moral hazard — i.e., reduced incentives for the insurer to cut costs, which they model in a reduced-form way.

[5] [Boone 2015] examines the relative roles of government insurance and supplementary private insurance, while capturing consumer risk aversion through mean-variance utility rather than concave utility (as in my model). [Boone 2018] considers risk-neutral consumers who are subject to a liquidity constraint so that, as in [Nyman 1999], the role of insurance is to provide *access* to care for consumers who could not otherwise afford it rather than consumption smoothing (as in my model).

[6] He writes, "If both discriminatory premia and insurance rates were permitted, then we would be back to the rule given in (4.2) in which considerations of distribution do not matter" (p. 332).

studies dynamic contracts, but with adverse selection at the ex-ante stage (and without moral hazard); my model only has adverse selection at the interim stage. Considering a series of one-year contracts, [Handel et al. 2015] empirically studies the tradeoff between dynamic risk due to explicit risk-rating and dynamic risk due to market unraveling from adverse selection when risk-rating is banned. Several others study dynamic contracts with *limited commitment* on the part of consumers: [Cochrane 1995], [Hendel and Lizzeri 2003], [Pauly et al. 1995], [Ghili et al. 2023], and [Atal et al. 2025].[7] The role of commitment in dynamic contracts like the ones I study is to prevent consumers from retroactively opting out of the cross-subsidization scheme should they turn out to be low risk. Essentially, these limited-commitment papers show how to disincentivize such opting out by front-loading premiums for earlier periods. Such front-loaded premiums fund later cross-subsidization while ensuring continued voluntary participation even by those who turn out to be low-risk. My model assumes full commitment or, under the social insurance interpretation, the ability of the planner to mandate consumer participation.

Finally, a group of papers studies optimal patient cost-sharing (i.e., coinsurance) under ex-post moral hazard mostly isolated from other frictions.[8] [Blomqvist 1997] and [Martinon et al. 2018] study nonlinear coinsurance. [Chernew et al. 2000] studies a binary treatment for a disease whose presence is observable but whose severity is not. Though their framing is quite different, their model can be viewed as a very special case of the observable-risk benchmark I consider in section 4.2 — see Appendix A for related discussion.

2.2 Dynamic Income Taxation Literature

As observed by [Zeckhauser and Spence 1971], [Blomqvist 1997], and [Kapička 2013], the ex-post moral hazard problem in health insurance is mathematically similar to the optimal income tax problem of [Mirrlees 1971]. The latter thinks of income differences as endogenously resulting from exogenous differences in a privately-known productivity parameter that captures ability to earn money per unit of effort. In these models, labor choice is analogous to treatment choice, and private productivity is analogous to a private pre-treatment health-state. A recent set of dynamic taxation papers allow for the productivity parameter to evolve over time either exogenously [Farhi and Werning 2013; Golosov et al. 2016; Kapička

[7] Whereas these papers assume observable risk profiles, [Pauly et al. 2011] considers the case where risk profiles are observed only by the incumbent insurer, not by others to which the consumer may defect.

[8] The ex-post moral hazard literature began with [Pauly 1968], [Zeckhauser 1970], and [Zeckhauser and Spence 1971]. Later papers include [Ma and Riordan 2002], [Ellis and Manning 2007], [Goldman and Philipson 2007], [Baicker et al. 2015], [Ellis et al. 2015], [Zhang et al. 2022], [Boche 2025].

2013] or endogenously [Makris and Pavan 2020, 2021].[9]

I adapt the techniques of and insights from the dynamic taxation literature to the health insurance context. In doing so, I go beyond the tax literature in several ways. First and most importantly, I show upward distortion of linear coverage under the FOA in my main result (Thm. 1), as discussed in the introduction. [Kapička 2013] informally suggests a similar result, but does not formally state or prove it.[10]

Second, I use a somewhat different information structure. My model features dynamic evolution of information but only one round of treatment choice. By contrast, the taxation papers feature multiple rounds of labor choice and assume that a worker's productivity in one period is positively correlated with their future productivity and indeed a sufficient statistic for their information thereabout.[11] My information structure leads to a new implementability restriction under linear contracts. Specifically, an assignment of linear coinsurance schedules can be implemented (by premiums) only if it gives weakly higher coverage to higher-risk types (Lem. 6). The analogous restriction does not hold in these taxation papers since they use prior-period labor choice as an additional screening instrument beyond just tax-plan choice.[12]

Third, my information structure and its associated linear implementability restriction allow me to provide a novel intuition for what drives validity of the FOA in that structure, namely increasingness of *statically* efficient coverage (Prop. F.2). Although [Kapička 2013] identified a sufficient condition for FOA-validity — namely, increasingness of marginal rents with respect to type in the *dynamic* problem (his Lemma 12) — he leaves unaddressed what causes such increasingness. I identify increasingness of statically efficient coverage as such a cause in my context.

[9] [Stantcheva 2020] surveys the literature.

[10] He writes "... the social planner is forced to provide insurance that is, from the ex-post perspective, too high." (p. 1028). See also his more technical discussion at the end of the last paragraph of his section 7.1 (p. 1043). There, he informally suggests a result about distortion of the marginal rent given to marginally higher types. Such marginal rent can be thought of as a proxy for (lack of) expected marginal coverage since the benefit from being lower-risk is higher when coverage is lower.

[11] Thus, the relation between my information structure and theirs is analogous to that between [Courty and Li 2000] and [Pavan et al. 2014] in the quasi-linear case.

[12] To understand why not, it is useful to adapt their information structure to the health setting so that we can see how the restriction fails when prior-period treatment choice is used as a screening instrument. To that end, suppose we extended my model to allow two rounds of treatment with the new round — which would be the first of the two — occurring concurrently with plan choice (which occurs only once). Suppose also that the severity of pre-treatment health state in the first round fully determines risk over the second-round health state, with severe first-round state implying high risk for the second round. In this case, it would be possible to implement an allocation of coverage levels that is *decreasing* in risk-type (contra my Lem. 6) using a mechanism in which the only way to access a high-coverage plan would be to choose low treatment in the first round. Specifically, since high-risk consumers are those with severe first-round health state, they may be willing to accept a low-coverage plan (for the second round) in order to get high treatment in the first round.

3 Model

There are two players: a consumer and an insurer.

Timing

In terms of information, there are three periods: 0, 1, and 2.

Period 0.A: The consumer proposes a dynamic contract to the insurer under symmetric information. The dynamic contract consists of a menu of continuation contracts $M = (c_\tau(\cdot), \pi_\tau)_{\tau=1}^{T}$, each consisting of a coinsurance schedule $c_\tau(\cdot)$ and premium π_τ.

Period 0.B: The insurer decides whether to accept or reject the proposed contract. If it accepts, we proceed with menu M. If it rejects, we proceed as if the following trivial menu had been proposed and accepted: a single continuation contract with $c(m) = m$ for all m and $\pi = 0$.

Period 1.A: The consumer privately realizes a signal $t \in \mathcal{T} \equiv \{1, ..., T\}$, called their *risk-type* with probability f_t.

Period 1.B: They then must choose a continuation-contract $(c_\tau(\cdot), \pi_\tau)$ from the menu M, and pay the premium π_τ to the insurer.

Period 2.A: They then privately realize their pre-treatment health *severity* $s \in \mathbb{R}$ according to their distribution $G_t(\cdot)$.

Period 2.B: They then choose their treatment level m (in dollars) and pay the out-of-pocket amount $c_\tau(m)$ to the (unmodeled) healthcare provider. The insurer pays the remainder: $m - c_\tau(m)$.

Information

At Period 0 (whether A or B), the consumer and the insurer both know the distributions $(G_t(\cdot))_t$ over s and the probabilities f_t. To repeat, the types t and s will be private to the consumer once they are learned.

I assume G_t has a density g_t supported on a positive ray $[\underline{s}, \bar{s}]$ with $\bar{s} = \infty$ and $\underline{s} \in \mathbb{R}$. Unless otherwise stated, I assume throughout that the distributions $(G_t)_t$ are strictly ordered by the monotone likelihood ratio property (MLRP): for any two risk-types $t > t'$, the ratio $\frac{g_t(s)}{g_{t'}(s)}$ is strictly increasing in s. I denote this relation as $G_t \succ_{\text{MLRP}} G_{t'}$ or $g_t \succ_{\text{MLRP}} g_{t'}$.

Preferences

Insurer Preferences: The insurer is risk-neutral over profit, which is given by premiums minus liabilities: $\pi_\tau - [m - c_\tau(m)]$ for continuation contract τ and treatment spending m.

Consumer Preferences: The consumer maximizes the expectation of utility $U(y,h) : \mathbb{R}_2 \to \mathbb{R}$, where y is monetary consumption and h is post-treatment health state. $h(m,s)$ is a function of dollars-worth-of-treatment m and pre-treatment health severity $s \in \mathbb{R}$. I assume consumer baseline wealth is constant across risk-types t and normalize it to zero unless otherwise stated.

Throughout the paper, I assume the following, where subscripts denote partial derivatives.

Standing Assumptions on Utility:

A1. $U(y,h)$ is C^2

A2. $U_1 > 0$ and $U_2 > 0$

A3. $U_{11} < 0$ and $U_{22} \leq 0$

A4. $\frac{U_1(y,h)}{U_2(y,h)}$ is weakly decreasing in y and weakly increasing in h. [Diminishing MRS]

Although these assumptions are always in place unless otherwise stated, I sometimes use the phrase "for general utility" or "let utility be general" to emphasize that no other assumptions on U are being made.[13]

The cleanest results will be given under the further assumption that utility has no wealth effects, which I refer to as "money-loss CARA" utility. First, by "money-loss" utility (not necessarily CARA), I mean: $U(y,h) = u(y+h)$ with $u' > 0$ and $u'' < 0$.[14] It means preferences over deterministic outcomes are quasi-linear, though wealth effects may still arise in risk-preferences. I say utility is money-loss *CARA* if, further, u exhibits constant absolute risk aversion (CARA): $u(x) = -\frac{1}{\psi}\exp(-\psi x)$ for some $\psi > 0$.

Turning to the health function h, I assume the following throughout the paper.

Standing Assumptions on $h(m,s)$:

B1. $h(m,s)$ is C^2 on $\mathbb{R}^2$.

B2. $h_2(m,s) < 0$ for all m, s. [Higher Severity $\longrightarrow$ Worse Health given Same Treatment]

B3. $h_{11}(m,s) < 0$ for all m, s and $\lim_{m\to\infty} h_1(m,s) < 0$ for all s [Diminishing Returns and Eventual Harm from Treatment]

B4. $h_{12} > 0$ for all m, s [Returns-to-Treatment Increasing in Severity]

[13] The diminishing MRS assumption (A4) ensures that health h — which is purchased indirectly through treatment — is a weakly normal good. It can be seen to be equivalent to quasi-concavity.

[14] This form is characterized by $\frac{U_1(y,h)}{U_2(y,h)} = 1$ everywhere, so that health is only a weakly (not strictly) normal good.

B5. If utility is money-loss, there exists some $s^* < \bar{s}$ such that $h_1(0, s^*) > 1$. For general utility, for every $\pi \in \mathbb{R}$, there exists $s^*_\pi < \bar{s}$ such that
$-U_1\left(-\pi, h(s^*_\pi, 0)\right) \; + \; U_2\left(-\pi, h(s^*_\pi, 0)\right) \cdot h_1(s^*_\pi, 0) \; > \; 0$. [Non-Triviality]

B2 says that higher severity s means strictly worse (post-treatment) health conditional on treatment choice.[15] B3 says that there are strictly diminishing returns to treatment and that, for any s, treatment eventually becomes harmful. B4 says that the marginal benefit of treatment strictly increases with pre-treatment health severity. B5 ensures that, for all t, treatment demand will be non-zero for a positive measure of s, even without insurance.

Comments on Model

Although in practice, insurers are the ones proposing the contracts, I have the consumer play the role of the principal so that the agent may be viewed as representing a competitive market of insurers in a reduced-form way.[16] The model abstracts from the role of the clinician, particularly as an *auditor* of the patient's pre-treatment severity s; I sketch how the model can be reinterpreted to address auditing in Appendix A. [17]

4 Benchmarks

I develop the solution to the main dynamic-contracting problem by first studying two benchmarks, thereby adding frictions sequentially. I first review the static moral hazard problem and then study the observable-risk dynamic problem.

4.1 Static Moral Hazard Problem

Treatment Demand A static contract (c, π) consists of a coinsurance schedule c and a premium π, so that the consumer's total liability when purchasing m dollars of care under (c, π) is $\pi + c(m)$. I will consider coinsurance schedules c in some admissible set $\mathcal{A}$ of

[15] Assumption B2 is used to prove Lemma 2 which underlies Theorem 1. Its strict inequality $h_2(m, s) < 0$ rules out a situation where a treatment can fully cure both mild and severe cases of the same health problem for the same cost. Weakenings of this assumption will be explored in a future draft.

[16] Reversing these roles would not have a major effect on results, and would have virtually no effect when utility has no wealth effects. This is because there is no ex-ante asymmetric information either way.

[17] Although I don't formalize it, concern about limited commitment in a private-sector dynamic contract could be mitigated by having the ex-ante consumer pay the continuation menu's lowest premium immediately upon signing the dynamic contract. That way, if they turn out to be low risk — the risk-type for which a participation constraint (if one was added) would be most likely to bind — they have already paid their premium and thus already have "free" access to discounted health care (via coinsurance), thus giving them no reason to drop out of the dynamic contract. Such premium front-loading would be in the spirit of the limited-commitment papers described in the literature review above, e.g., [Pauly et al. 1995].

functions with domain $\mathbb{R}_+$. I leave $\mathcal{A}$ abstract for now, but in specific results, it will be either essentially unrestricted or the set of linear schedules.

I next define the consumer's ex-post demand function for treatment. Given a coinsurance schedule c and premium $\pi \in \mathbb{R}$, a consumer with health severity s demands $m_{c,\pi}(s)$ dollars worth of treatment, where

$$m_{c,\pi}(s) \equiv \arg\max_{m \geq 0} U\left(-\pi - c(m),\, h\left(m, s\right)\right). \tag{4.1}$$

$m_{c,\pi}(s)$ balances better health $h\left(m, s\right)$ with greater out-of-pocket expense $c(m)$.[18]

Static Problem Formulation I now define risk-type t's indirect expected utility $V_t(c, \pi)$ from static contract (c, π) and the insurer's expected profit $P_t(c, \pi)$ from providing (c, π) to t. Let

$$V_t\left(c, \pi\right) \equiv \int U\left(-\pi - c\left(m_{c,\pi}(s)\right),\, h\left(m_{c,\pi}(s), s\right)\right) dG_t(s); \qquad \text{and} \tag{4.2}$$

$$P_t\left(c, \pi\right) \equiv \pi - k_t\left(c, \pi\right), \quad \text{where} \quad k_t\left(c, \pi\right) \equiv \int \left[m_{c,\pi}(s) - c\left(m_{c,\pi}(s)\right)\right] dG_t(s). \tag{4.3}$$

In words, the insurer's expected profit $P_t(c, \pi)$ is the premium minus their expected liability from insurance claims.

I say the coinsurance schedule c is *linear* when $c(m) = (1 - x)m$ for some fixed coverage rate $x \in \mathbb{R}$ (so that higher x means more coverage). When restricting to linear coinsurance, I abuse notation by writing $V_t(x, \pi)$, $P_t(x, \pi)$, and $k_t(x, \pi)$, with the notation x being used a stand-in for the function $c(m) = (1 - x)m$.

The static moral hazard problem for risk-type t is then

$$\max_{(c,\pi) \in \mathcal{A} \times \mathbb{R}} V_t(c, \pi) \tag{4.4a}$$

$$\text{s.t.} \quad \text{(IR/Act-Fair)} \quad P_t(c, \pi) \geq 0. \tag{4.4b}$$

Here, the consumer chooses a contract (c_t, π_t) to maximize their expected utility subject to the insurer's participation constraint 4.4b. Since the insurer is risk-neutral, that constraint takes the form of an actuarial fairness constraint with respect to the risk distribution G_t. It will also be useful to consider what I'll call the *S-subsidized* static problem which replaces

[18] I assume existence and uniqueness of $m_{c,\pi}(s)$ throughout, but some sufficient conditions for these properties are mentioned in Footnote 19 in the next subsection.

the actuarial fairness constraint 4.4b with the "S-subsidized" version thereof:

$$(S\text{ - Act-Fair}) \quad P_t(c,\pi) \geq -S \tag{4.5}$$

This constraint says that the insurer's expected losses from the contract must not exceed S, which I call the subsidy. The S-subsidized static problem will be relevant to the dynamic contracting problem because the latter will involve cross-subsidization between risk-types.

Definition 1. Let $(\hat{c}_{t,S}, \hat{\pi}_{t,S})$ denote the optimal solution to t's S-subsidized static problem. In the case of linear contracts, $(\hat{x}_{t,S}, \hat{\pi}_{t,S})$ denotes the same. When $S = 0$, I drop the subscript S so that $\hat{c}_t \equiv \hat{c}_{t0}$, $\hat{x}_t \equiv \hat{x}_{t0}$, and $\hat{\pi}_t \equiv \hat{\pi}_{t,0}$.

I refer to $\hat{c}_{t,S}$ (or $\hat{x}_{t,S}$) as t's *statically efficient* coinsurance schedule (or coverage level) given subsidy S. The next lemma shows why money-loss CARA utility is simplest to work with.

Lemma 1. *Let utility be money-loss CARA. $\hat{c}_{tS}$ is constant in S. [So we just write $\hat{c}_t$ or $\hat{x}_t$].*

Given the distribution G_t, the optimal coinsurance schedule $\hat{c}_t$ balances a tradeoff between risk-protection and incentives to temper treatment, with higher coverage promoting the former but hurting the latter. For nonlinear coinsurance, optimally balancing these objectives given G_t involves imposing high marginal coinsurance rates c' (ceteris paribus) at levels of marginal spending m where the consumer has a high chance of landing at the end of the coverage period [Blomqvist 1997], [Kleef et al. 2009].

4.2 Observable-Risk Problem

Having introduced the static moral hazard problem, the next step is to analyze the dynamic problem with observable risk. I will show that the latter's optimal solution is *interim efficient* in the sense that the coinsurance schedule it assigns to each risk-type will be that type's statically efficient schedule as defined in the previous section. It will follow that any coinsurance distortions that arise in the private-risk problem (discussed next) are caused not from dynamics alone, but rather from the combination of dynamics and interim private information [Fudenberg et al. 1990].

Recall that $\mathcal{A}$ denotes the set of admissible coinsurance schedules and that f_t denotes the consumer's probability of realizing risk-type t. The observable-risk problem is

$$\max_{(c_t,\pi_t)_{t=1}^T \in (\mathcal{A}\times\mathbb{R})^T} \sum_{t=1}^T f_t\, V_t\left(c_t, \pi_t\right) \tag{4.6a}$$

$$\text{s.t.} \quad \text{(Ex-Ante IR)} \quad \sum_{t=1}^T f_t\, P_t\left(c_t, \pi_t\right) \geq 0. \tag{4.6b}$$

This problem chooses a menu of static contracts $(c_t, \pi_t)_{t=1}^T$ to maximize the consumer's expected utility from the point of view of Period 0, before risk-types t are realized; recall the timing from Section 3. The participation constraint (4.6b) requires that the insurer make weakly positive expected profit from the dynamic contract, though not necessarily on each continuation plan (c_t, π_t). Since risk-types t are observable, each t can be directly assigned to a plan (c_t, π_t).

Definition 2. Let $(\tilde{c}_t, \tilde{\pi}_t)_{t=1}^T$ denote the optimal solution to the observable-risk problem. Let $\tilde{S}_t \equiv -P_t(\tilde{c}_t, \tilde{\pi}_t)$, the endogenous subsidy to type t under the solution $(\tilde{c}_t, \tilde{\pi}_t)_{t=1}^T$.

I also define t's expected *marginal* utility of monetary consumption at plan (c, π):

$$\lambda_t^{c,\pi} \equiv \int U_1\left(-\pi - c\left(m_{c,\pi}(s)\right), h\left(m_{c,\pi}(s), s\right)\right) dG_t(s). \tag{4.7}$$

By the envelope theorem, it's easy to see that $\frac{\partial}{\partial \pi} V_t(c, \pi) = -\lambda_t^{c,\pi}$. For brevity, let $\tilde{\lambda}_t \equiv \lambda_t^{\tilde{c}_t, \tilde{\pi}_t}$, which is t's expected marginal utility at the optimal solution to the observable-risk problem.

Proposition 1 (Observable-Risk Characterization)**.** *Let $\mathcal{A}$ be the set of measurable functions $c : \mathbb{R}_+ \to \mathbb{R}$ for which treatment-demand $m_{c,\pi}(s)$ exists and is unique for all π, s.*[19] *Suppose that $(\tilde{c}_t, \tilde{\pi}_t)_{t=1}^T$ is a solution to the observable-risk problem (4.6) with the $\mathcal{A}$ just specified. Then the following hold.*

i) *(Interim Efficiency): If utility is money-loss CARA, then $\tilde{c}_t = \hat{c}_t$ for all t. For general utility, $\tilde{c}_t = \hat{c}_{t\tilde{S}_t}$ for all t.*[20]

[19] Existence of $m_{c,\pi}(s)$ is guaranteed when c is weakly increasing and upper semi-continuous (Fact G.1 part i). As for uniqueness of $m_{c,\pi}(s)$, a sufficient condition involving convexity of c is given in Fact G.2.

[20] Technically, I don't guarantee existence or uniqueness of the observable-risk solution $(\tilde{c}_t, \tilde{\pi}_t)_{t=1}^T$, but this result is phrased accordingly. In case of multiplicity, the notations $\tilde{S}_t$ and $\tilde{\lambda}_t$ should understood to correspond to the specific solution $(\tilde{c}_t, \tilde{\pi}_t)_{t=1}^T$ whose existence is "supposed" near the start of the proposition. As for the optimal *static* solution $\hat{c}_{t\tilde{S}_t}$, in case there are multiple of them, the result should be understood as saying $\tilde{c}_t = \hat{c}_{t\tilde{S}_t}$ *for some* statically optimal $\hat{c}_{t\tilde{S}_t}$. Finally, it is not necessary to assume that $\hat{c}_{t\tilde{S}_t}$ exists, since if it didn't, $\tilde{c}_t$ wouldn't exist either and the latter's existence is "supposed" near the start of the result statement.

ii) *(Full Cross-Subsidization): If utility is money-loss, then expected marginal utility is equalized across risk-types:* $\tilde{\lambda}_t = \tilde{\lambda}_\tau$ *for all risk-types* $t, \tau \in \mathcal{T}$*. For general utility, if* $\tilde{c}_t$ *is* C^1 *for all* $t \in \mathcal{T}$*, then* $\tilde{\lambda}_t \cdot \left[\frac{\partial}{\partial \pi} P_t(\tilde{c}_t, \tilde{\pi}_t)\right]^{-1} = \tilde{\lambda}_\tau \cdot \left[\frac{\partial}{\partial \pi} P_\tau(\tilde{c}_\tau, \tilde{\pi}_\tau)\right]^{-1}$ *for all* $t, \tau \in \mathcal{T}$*.*

Part i says that coinsurance is statically efficient given the endogenous actuarial subsidy $\tilde{S}_t$; in this sense, the optimal dynamic contract is interim efficient. To see why, suppose $\tilde{c}_t \neq \hat{c}_{t\tilde{S}_t}$, so that $(\tilde{c}_t, \tilde{\pi}_t)$ is a suboptimal solution to the $\tilde{S}_t$-subsidized static problem. By instead setting $(c_t, \pi_t) = \left(\hat{c}_{t\tilde{S}_t}, \hat{\pi}_{t\tilde{S}_t}\right)$, we could improve t's expected utility while maintaining insurer profits from t at $-\tilde{S}_t$.[21]

Part ii says that for money-loss utility (without assuming CARA), expected marginal utilities $\tilde{\lambda}_t$ are equalized across types; for general utility, a similar equality holds which corrects for the effects of premium changes on treatment choice and ultimately the budget. If the equalities in part ii didn't hold, we could redistribute money between risk-types through the premiums in a way that increases ex-ante expected utility while preserving the budget. As I show in Proposition D.1 in Appendix D.1, when nonlinear coinsurance is allowed, more complex monetary transfers between risk-types are possible — beyond just premium transfers — and so the condition $\tilde{\lambda}_t = \tilde{\lambda}_\tau$ has a closer wealth-effects analog than *the last* condition in part ii.[22]

To emphasize the extent of the cross-subsidization being prescribed in part ii, note that the optimal contract here may well assign a strictly lower premium to higher-risk types than to lower-risk types in order to compensate the former for their higher expected out-of-pocket liability for any given c. More precisely (though still informally), when L is a lower risk-type than H, we will have $\tilde{\pi}_H < \tilde{\pi}_L$ unless $\tilde{c}_L$ is sufficiently lower-coverage than $\tilde{c}_H$. It is thus unsurprising that when we study the private-risk problem, upward IC will bind. That being said, the present result (Prop. 1 ii) illustrates how conditioning contracts on observed consumer risk profiles can — if done in service of ex-ante consumer-welfare rather than interim profit maximization — actually protect consumers from dynamic risk rather than expose them to it.

5 Private-Risk Problem

This section introduces the private-risk problem and shows that upward IC binds therein. The latter will motivate the coverage distortions explored in the next two sections for linear

[21] To gain a deeper understanding of the observable-risk solution under wealth effects, we need to understand how $\hat{c}_{tS_t}$ depends on S_t. This is explored in Appendix C.2, for linear contracts.

[22] That analogous condition involves treatment-choice-*preserving* transfers. Such transfers specialize to premium-transfers only in the money-loss case.

and nonlinear contracts.

The private-risk problem, which is the focus of the paper, is the following:

$$\max_{(c_t,\pi_t)_{t=1}^T \in (\mathcal{A}\times\mathbb{R})^T} \sum_{t=1}^T f_t\, V_t\left(c_t, \pi_t\right) \tag{5.1a}$$

$$\text{s.t.} \quad \text{(Interim IC)} \quad (c_t, \pi_t) \in \arg\max_{\hat{t}} V_t\left(c_{\hat{t}}, \pi_{\hat{t}}\right) \quad \forall t \tag{5.1b}$$

$$\text{(IR)} \quad \sum_{t=1}^T f_t\, P_t\left(c_t, \pi_t\right) \geq 0. \tag{5.1c}$$

The difference from the observable-risk problem (4.6) is the addition of the interim incentive compatibility (IC) constraints (5.1b). Those constraints require that consumers be incentivized to choose the plan intended for their risk-type.[23]

I will focus on the simplified version of the above problem that has only upward-adjacent IC constraints:

$$\text{(FOA-Relaxed IC)} \quad V_t\left(c_t, \pi_t\right) \geq V_t\left(c_{t+1}, \pi_{t+1}\right) \quad \forall t < T \tag{5.2}$$

I refer to this relaxed version of (5.1) with (5.2) substituted for (5.1b) as the *FOA-relaxed* problem (or just the "relaxed" problem) and to the unrelaxed version of (5.1) as the "full" (or "unrelaxed") problem. The *first-order approach* (FOA) is the approach of solving the full problem (5.1) by solving the FOA-relaxed problem [Battaglini and Lamba 2019].[24] For our purposes, the most important property of the FOA is that it ignores downward ICs; however, to be clear, it also ignores non-adjacent upward ICs. As codified in the next definition, I use $*$ and $**$ to indicate the solutions to the relaxed and full problems respectively.

Definition 3. Let $(c_t^*, \pi_t^*)_{t=1}^T$ denote the optimal solution to the *FOA-relaxed* problem. When $\mathcal{A}$ is restricted to linear schedules, let $(x_t^*, \pi_t^*)_{t=1}^T$ denote the corresponding solution with $c_t^*(m) = (1 - x_t^*)m$. Let $(c_t^{**}, \pi_t^{**})_{t=1}^T$ denote the optimal solution to the *full* problem, and define $(x_t^{**}, \pi_t^{**})_{t=1}^T$ analogously.

Definition 4. For each risk-type t, let $S_t^* \equiv -P_t(c_t^*, \pi_t^*)$ and $\lambda_t^* \equiv \lambda_t^{c_t^*, \pi_t^*}$. Similarly, let $S_t^{**} \equiv -P_t(c_t^{**}, \pi_t^{**})$ and $\lambda_t^{**} \equiv \lambda_t^{c_t^{**}, \pi_t^{**}}$.

[23] Since t is not directly payoff-relevant, there is no concern about double deviations whereby a consumer of risk-type t misreports as some $\hat{t} \neq t$ in anticipation of choosing some m different from $m_{c_{\hat{t}}, \pi_{\hat{t}}}(s)$ for certain potentially realized s's.

[24] This definition follows that paper's Definition 1, except that — in a merely semantic reversal — they instead consider only *downward*-adjacent ICs.

In words, S_t^* is the actuarial subsidy given to risk-type t and λ_t^* is t's expected marginal utility — both under the relaxed solution (c_t^*, π_t^*). S_t^{**} and λ_t^{**} are defined analogously.[25]

We saw that with observable risk, optimal dynamic contracts use cross-subsidization to fully insure dynamic risk in the sense that expected marginal utility $\tilde{\lambda}_t$ is equalized across risk-types (Prop. 1 ii).

However, when risk-types are private, incentives for plan choice will constrain the amount of cross-subsidization that can take place. Specifically, *upward* IC will bind[26] as increasing cross-subsidization will incentivize lower-risk types to deviate to the better-subsidized plans intended for higher-risk types. Given limited cross-subsidization, it's natural to expect that higher-risk types will have higher expected marginal utility in the optimal contract, both FOA-relaxed and full. The next two lemmas confirm these properties in the no wealth-effects case.

Lemma 2 (Expected Marginal Utility Increasing in t)**.** *Let utility be money-loss CARA. Suppose $(c_t, \pi_t)_{t=1}^T$ satisfies the upward-adjacent IC constraints (5.2) of the private-risk problem. Then, expected marginal utility $\lambda_t^{c_t,\pi_t}$ is strictly increasing in t. In particular, λ_t^* and λ_t^{**} are both strictly increasing in t.*[27]

Thus, in the optimal menu for the FOA-relaxed problem, upward-adjacent IC constraints must bind since otherwise, we could redistribute money upward through the premiums, thereby increasing ex-ante expected utility. The next lemma states this more precisely and, in the linear case, extends to the full problem.

Lemma 3 (Upward IC Binds)**.** *Let utility be money-loss CARA and suppose $(c_t^*, \pi_t^*)_{t=1}^T$ is an optimal solution to the FOA-relaxed problem. All upward-adjacent ICs bind: $V_{t-1}\left(c_{t-1}^*, \pi_{t-1}^*\right) = V_{t-1}\left(c_t^*, \pi_t^*\right)$ for all $t > 1$. If we restrict to linear coinsurance, the same holds not only for the FOA-relaxed solution $(x_t^*, \pi_t^*)_{t=1}^T$ but for the full-problem solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$ as well.*[28]

In Lemma 3′ (in Appendix D.2), I extend Lemma 3 to account for wealth effects under extra hypotheses, though only for the linear-coinsurance case.[29]

[25] In case of multiplicity of $(x_t^*, \pi_t^*)_{t=1}^T$, the notations S_t^* and λ_t^* should be understood to correspond to the contract $(c_t^*, \pi_t^*)_{t=1}^T$ mentioned just prior to their use. Similarly, for the full-problem versions.

[26] To be precise, when I say a constraint "binds", I mean it holds with equality at the optimal solution, or at some other solution if one has been specified.

[27] This lemma is proved in Appendix G.5.1. It is the key channel through which the money-loss CARA assumption will be used in the proof of Theorem 1. It is also the channel through which the standing assumption B2 that $b_2(m, s) < 0$ will be used in the proof of the main result Theorem 1.

[28] For nonlinear coinsurance, I have not extended this result to the full problem since with nonlinear, we lack the single-crossing property established for linear in the next section (Lem. 5). However, the proof shows it does extend to the intermediate problem that only ignores non-adjacent upward ICs while still considering all downward ICs; for details, see Footnote 98 in the proof of Lem 3′ which (as mentioned in the next footnote), subsumes Lem. 3.

[29] I prove only Lemma 3′ which subsumes Lemma 3. The proof of Lemma 3′ is in Appendix G.5.1.

The fact that upward IC binds here contrasts the well-known [Rothschild and Stiglitz 1976] model where, instead, *downward* IC binds (and upward does not). The difference is that in their model, there is no cross-subsidization so that their contracts feature actuarially-fair premiums for each type. Although that paper ignores moral hazard, the same idea applies in its presence, as I show for the two-type case in Fact E.1 (in Appendix E.1).

6 Linear Coinsurance

I now study the private-risk problem (5.1) (relaxed and full) under the restriction of $\mathcal{A}$ to linear coinsurance schedules. Linear schedules are totally-ordered from less-coverage to more-coverage, which will facilitate a clean comparison between the solutions to the static and dynamic problems in the main result.

The next step is to show that higher-risk types place greater value on extra coverage than do lower-risk ones. Recall that the set of risk-types $\mathcal{T} \equiv \{1, ..., T\}$ is such that if $L < H \in \mathcal{T}$, then $G_L \prec_{\text{MLRP}} G_H$.

6.1 Single Crossing

Recalling from (4.7) that $\lambda_t^{c,\pi}$ denotes expected t's marginal utility of monetary consumption at plan (c, π), I again abuse notation by writing $\lambda_t^{x,\pi}$ in the linear case. As observed before, the envelope theorem implies $\frac{\partial}{\partial \pi} V_t(x, \pi) = -\lambda_t^{x,\pi}$. Note that $\frac{\partial}{\partial \pi} V_t(x, \pi) < 0$ and $\frac{\partial}{\partial x} V_t(x, \pi) > 0$ are the effects of a unit increase in premium and coverage respectively on t's expected utility starting at (x, π). Thus, by the implicit function theorem, $\frac{\frac{\partial}{\partial x} V_t(x,\pi)}{\lambda_t^{x,\pi}}$ is the amount of extra premium that t would be willing to pay for an extra unit of coverage starting at (x, π). As shown in the next lemma, this amount is strictly greater for higher-risk types. Essentially, this is because greater coverage will save them relatively more money since it will apply to a (likely) greater spending amount than with lower-risk types. This captures the idea of *adverse selection* into higher-coverage plans.

Lemma 4 (Willingness-to-Pay)**.** *Let $L < H \in \mathcal{T}$, and let (x, π) be a linear contract with $x \in [0, 1]$.*[30,31] *Then*

$$\frac{\frac{\partial}{\partial x} V_L(x, \pi)}{\lambda_L^{x,\pi}} < \frac{\frac{\partial}{\partial x} V_H(x, \pi)}{\lambda_H^{x,\pi}}.$$

[30] This lemma generalizes a result of [Chade et al. 2024] to allow general utility rather than just money-loss CARA; see their Proposition 3, part ii. My proof technique is based on theirs.

[31] This lemma is the channel through which the standing assumption B4 that $b_{12}(m, s) > 0$ will be used in the proof of the main result Theorem 1.

This implies a *single-crossing* property: if lower-risk types weakly prefer a higher-coverage contract to a lower-coverage one, then higher-risk types must *strictly* share that preference.

Lemma 5 (Single Crossing). *Let $L < H \in \mathcal{T}$ and $x', x \in [0,1]$. If $x' < x$ and $V_L(x', \pi') \leq V_L(x, \pi)$ for some $\pi', \pi \in \mathbb{R}$, then $V_H(x', \pi') < V_H(x, \pi)$.*

This implies that allocations satisfying the *full* set of ICs must be weakly increasing in coverage (and premium).

Lemma 6 (Implementability). *Suppose $(x_t, \pi_t)_{t=1}^T$ satisfies all IC constraints (5.1b) of the full problem (5.1) with $x_t \in [0,1]$ for all t. Then $(x_t)_{t=1}^T$ is weakly increasing in t, and so is $(\pi_t)_{t=1}^T$. Further, $x_t < x_{t+1}$ if and only if $\pi_t < \pi_{t+1}$.*[32]

6.2 Main Result

We saw that with observable risk, optimal dynamic contracts are interim efficient (Prop. 1 i). We now show the failure of that result under private-risk and specify the nature of the interim inefficiency. Since the results are cleanest for money-loss CARA utility, I restrict to that case here, leaving wealth-effects extensions to Appendix D.2.

Now for the main result of the paper, which essentially shows the upward distortion of coverage $x_t^* > \hat{x}_t$ in the *relaxed* problem. A follow-up proposition will address the *full* problem.

Theorem 1 (Main Result: Overinsurance). *Let utility be money-loss CARA, and suppose the static problem is strictly concave for every t.* [33] *Suppose $(x_t^*, \pi_t^*)_{t=1}^T$ is an optimal solution to the FOA-relaxed problem with $x_t^* \geq 0$ for all t. Then $x_t^* > \hat{x}_t$ for all $t > 1$.*[34]

Essentially, Theorem 1 states that, under the FOA and without wealth effects, optimal coverage will be distorted strictly above its statically efficient level $\hat{x}_t$. I now describe the intuition behind it. Recall S_t^* denotes the endogenous subsidy to t in the relaxed-optimal solution. We can reformulate the (S_t^*-subsidized) static problem (4.4) to optimize only over coverage x, letting the premium automatically adjust to keep the insurer's profit from t's contract at $-S_t^*$.[35] Starting from t's statically efficient coverage $\hat{x}_t$, a marginal increase in

[32] In the dynamic income-tax papers mentioned in the literature review (Section 2.2), the analog of this lemma does not hold, specifically the part about $(x_t)_{t=1}^T$ being increasing. As I mentioned there, the reason is that they have an additional screening instrument in that tax-plan choice is bundled with labor choice from the previous period.

[33] For a particular form of the health function $h(m, s)$, Fact G.3 shows strict concavity of the static problem under money-loss utility.

[34] This assertion still holds if $(x_t^*, \pi_t^*)_{t=1}^T$ is replaced by an optimal solution to the slightly less relaxed problem that considers all upward ICs (adjacent or otherwise) but still no downward ones.

[35] This is shown in Appendix G.1.2.

coverage x accompanied by such a budget-preserving increase in premium keeps t indifferent to first-order. But by Lemma 4, this compound deviation relaxes $(t-1)$'s upward-adjacent IC constraint, thus enabling an additional premium transfer from $t-1$ to t (and potentially to higher types too), which is desirable since $\lambda^*_{t-1} < \lambda^*_t$ (Lem. 2).

Extension to Full Problem For this argument to work in the full problem, we need to make sure $(t+1)$'s downward-adjacent IC doesn't bind so that it's not violated by the initial premium-compensated increase in t's coverage. For that, it suffices that $x^{**}_t < x^{**}_{t+1}$ (as opposed to $x^{**}_t = x^{**}_{t+1}$), given single crossing (Lem. 5) and the fact that upward IC binds (Lem. 3). Thus, we have the next proposition.

Proposition 2 (Overinsurance in Full Problem under Separation)**.** *Let utility be money-loss CARA, and suppose the static problem is strictly concave for every t. Suppose $(x^{**}_t, \pi^{**}_t)_{t=1}^T$ is an optimal solution to the full private-risk problem with all IC constraints (5.1b) with $x^{**}_t \geq 0$. If $t > 1$ and $x^{**}_t < x^{**}_{t+1}$, then $x^{**}_t > \hat{x}_t$.*

The following corollary is immediate.

Corollary 1. *Assume the hypotheses of Proposition 2. If $(x^{**}_t)_{t=1}^T$ is strictly increasing in t throughout, then $x^{**}_t > \hat{x}_t$ for all $t > 1$.*

Outside money-loss CARA utility, the analogs of the distortion results $x^*_t > \hat{x}_t$ and $x^{**}_t > \hat{x}_t$ are $x^*_t > \hat{x}_{tS^*_t}$ and $x^{**}_t > \hat{x}_{tS^{**}_t}$ respectively. In Theorem 1′ (in Appendix D.2), I give sufficient conditions for these analogs under general-utility (i.e., with wealth effects).[36] Theorem 1 and Proposition 2 are both subsumed in and proved as part of that more general result. The proof is in Appendix G.5.2.[37]

Validity of the FOA We say the FOA is *valid* if optimality in the FOA-relaxed problem implies optimality in the full problem.

In Appendix F (Propositions F.1 and F.2), I go into further depth on conditions for validity of the FOA under linear contracts. There, I show — as foreshadowed by Corollary 1 above — that the FOA will be valid if x^*_t is increasing in t (Prop. F.1 ii), a similar condition

[36] The key condition is that $\lambda^*_t \cdot \left[\frac{\partial}{\partial \pi} P_t(x^*_t, \pi^*_t)\right]^{-1}$ be strictly increasing. This guarantees that *budget-neutral* premium transfers from lower- to higher-risk types increase ex-ante utility. Since $\frac{\partial}{\partial \pi} P_t(x^*_t, \pi^*_t) = 1$ for money-loss utility, this condition is guaranteed under money-loss *CARA* by Lem. 2.

[37] The fact that upward IC binds (Lem. 3) — though conceptually important for understanding Thm. 1 (the relaxed-problem result) — is technically only needed to prove Prop. 2 (the full-problem result). Instead, to prove Thm. 1, we only need $\lambda^*_{t-1} < \lambda^*_t$, which Lem. 2 established. Lem. 3's role for proving Prop. 2 is only to show — together with single crossing (Lem. 5) — that $(t+1)$'s downward-adjacent IC doesn't bind when $x^{**}_t < x^{**}_{t+1}$.

to that found in [Kapička 2013] (Lemma 12). If, instead, we have $x_t^* > x_{t+1}^*$ for some t, then Proposition F.1 iv shows the full solution will feature *pooling* — likely between t and $t+1$,[38] i.e., $x_t^{**} = x_{t+1}^{**}$ and $\pi_t^{**} = \pi_{t+1}^{**}$ — due to the monotonicity restriction imposed by Lemma 6.

Although, as just alluded to, the fact that increasingness of the FOA-relaxed allocation x_t^* is sufficient for FOA-validity was — albeit in a different framing — essentially shown already in [Kapička 2013], the question of what drives such increasingness has been left open by the literature to my knowledge. I point to increasingness of *statically* efficient coverage $\hat{x}_t$ — or $\hat{x}_{tS_t^*}$ under wealth effects — as a key factor in my context. Proposition F.2 confirms this formally in the two risk-type case. However, increasingness of $\hat{x}_t$ is not *necessary* for FOA-validity since the desire to enable cross-subsidization while respecting upward IC may push — due to the same differential willingness-to-pay for coverage that drives Thm. 1 — towards increasingness of x_t^* despite mild decreasingness of $\hat{x}_t$.

7 Nonlinear Coinsurance

This section studies the private-risk problem with nonlinear coinsurance. Its key result, Proposition 3, uses first-order necessary conditions to characterize the optimal menu under the FOA with no wealth effects. I extend it to allow for wealth effects in Appendix I and for a continuum of risk-types in Supplemental Appendix S1.

Proposition 3 gives several conceptual insights. First, it shows that the dynamic redistributive force behind Theorem 1 extends to nonlinear contracts. As in the linear case, an increase in marginal coverage — together with a compensating increase in premium — relaxes upward interim ICs, thus allowing for more cross-subsidization from lower- to higher-risk types. Second, it shows how nonlinear contracts sharpen that force: the amount of IC-relaxation achieved between adjacent types t and $t-1$ (and hence redistribution towards the former) is related to their relative *likelihood of exceeding* the spending level at which marginal coverage is increased for t. This is because consumers only benefit from extra marginal coverage at m if they exceed m dollars in spending. This leads to higher coverage distortions (ceteris paribus) at dollar amounts where that relative likelihood is highest. Third, it shows how coverage in each type's plan is partially *tailored* to their risk-distribution in order to balance the intra-type risk-incentive tradeoff, notwithstanding distortions from the dynamic force.

In order to state Proposition 3 precisely in section (7.2), some technical setup will be needed regarding direct mechanisms and the first-order approach.

[38] Technically, Prop. F.1 iv only guarantees such pooling for *some* pair of consecutive types $(\tau, \tau+1)$, not necessarily $\tau = t$, but the latter seems very likely.

7.1 Technical Setup

Direct vs. Indirect Formulations As formulated in (5.1), the private-risk problem optimized over menus $(c_t, \pi_t)_{t=1}^T$ with coinsurance schedules c_t in some admissible class $\mathcal{A}$ (and premiums $\pi_t \in \mathbb{R}$) and then let these determine treatment choices through the demand functions $m_{c_t,\pi_t}(s)$. Here, I instead adopt a direct-mechanism approach that optimizes directly over treatment levels $m_t(s)$ and corresponding reimbursement levels $r_t(s)$, subject to an explicit ex-post IC constraint.[39] I defer a full presentation of the direct formulation and its relationship to the indirect formulation to Appendix H.1, but they are essentially equivalent and I informally sketch the correspondence here. Given functions $m_t(s)$ and $r_t(s)$ on the support of G_t (which is t-independent), we can define (c_t, π_t) as follows. First let $\pi_t = -r_t(\underline{s}) + m_t(\underline{s})$ where $\underline{s}$ is the lowest severity in the support of G_t; for intuition, note that when $m_t(\underline{s}) = 0$ (as would be natural), this means $\pi_t = -r_t(\underline{s})$. Then let $c_t(m) = \left[m - r_t\left(s_t^{-1}(m)\right)\right] - \pi_t$, where $s_t^{-1}(\cdot)$ is the inverse function of $m_t(\cdot)$. I refer to the menu $(c_t, \pi_t)_{t=1}^T$ obtained in this way as the indirect menu *induced* by the direct menu $(m_t(\cdot), r_t(\cdot))_{t=1}^T$.

The First-Order Approach I again study the FOA-relaxed problem, thus simplifying the interim IC constraints (5.1b) to (5.2). However, since we study the direct mechanism here, we also apply the FOA to the ex-post stage where IC constraints governing treatment choice now arise. Since I always assume a continuum of ex-post types s, the FOA applied to ex-post ICs replaces them with a *local* optimality condition, formally captured by an envelope condition [Milgrom and Segal 2002].[40] Thus, the FOA-relaxed version of the direct formulation becomes the following:

$$\max_{m_t, r_t, W_t} \sum_{t=1}^T f_t \int U\left(-m_t(s) + r_t(s),\, h\left(m_t(s), s\right)\right)\, dG_t(s)$$

$$\text{s.t.} \quad \sum_{t=1}^T f_t \int r_t(s)\, dG_t(s) = 0, \tag{7.1a}$$

$$W_t(s) = U\left(-m_t(s) + r_t(s),\, h\left(m_t(s), s\right)\right) \quad \forall t, \tag{7.1b}$$

$$\int W_t(s)\, dG_t(s) \geq \int W_{t+1}(s)\, dG_t(s) \quad \forall t \leq T-1, \tag{7.1c}$$

$$W_t'(s) = U_2\left(-m_t(s) + r_t(s),\, h\left(m_t(s), s\right)\right) \cdot h_2\left(m_t(s), s\right) \quad \forall t. \tag{7.1d}$$

I refer to (7.1) as the FOA-relaxed *direct* problem. The first constraint (7.1a) is the ex-

[39] For more on the *indirect*-mechanism approach with nonlinear contracts, see Supplemental Appendix S3.

[40] At the ex-post stage, failure of the FOA will coincide with bunching of multiple severities s into a single treatment level m. [Martinon et al. 2018] address bunching in a static model of health insurance.

ante budget constraint, familiar from the indirect formulation (5.1). The second constraint (7.1b) is merely to introduce the extra variable $W_t(s)$ that will set up the last two constraints. The third constraint (7.1c) is the FOA-relaxed interim-IC constraint, also familiar from the indirect formulation. The fourth constraint (7.1d) is the envelope condition for ex-post IC.[41]

7.2 No Wealth-Effects Case

I now restrict to *money-loss CARA* utility. For wealth effects, see Appendix I.

Notation Fixing a particular direct menu $(m_t(\cdot), r_t(\cdot))_{t=1}^T$, let $u'_{ts} \equiv u'\left(-m_t(s) + r_t(s) + h\left(m_t(s), s\right)\right)$ be the marginal utility of money for type s under t's plan. Then, for risk-types $t, \tau \in \mathcal{T}$, define the following:

$$\lambda_t^\tau \equiv \int u'_{\tau s}\, g_t(s)\, ds, \quad \lambda_t \equiv \lambda_t^t, \tag{7.2a}$$

$$\psi_t^\tau(s) \equiv \int_s^{\bar{s}} u'_{\tau s}\, \frac{g_t(\hat{s})}{1 - G_\tau(s)}\, d\hat{s} \text{ for } s < \bar{s}, \quad \psi_t(s) \equiv \psi_t^t(s)\ \forall s. \tag{7.2b}$$

Although the dependence is not made explicit by the notation, λ_t^τ and ψ_t^τ should always be understood to correspond to the specific menu $(m_t(\cdot), r_t(\cdot))_{t=1}^T$ fixed just prior to their use; the dependence operates through $u'_{\tau s}$ within which $m_\tau(\cdot)$ and $r_\tau(\cdot)$ appear (see just above). λ_t^τ is the increase in t's expected marginal utility associated with a one-dollar decrease in τ's premium when purchasing τ's plan. I'll call such a decrease a *premium-transfer* to τ. $\psi_t^\tau(s)$ is, roughly, t's expected marginal utility conditional on severity being at least s, when choosing τ's plan. However, the normalization quotient $1 - G_\tau(s)$ in ψ_t^τ is τ's probability (not t's), so $[1 - G_\tau(s)]\, \psi_t^\tau(s)$ is the change in t's expected marginal utility from a one-dollar increase in reimbursement $r_\tau(\hat{s})$ for all $\hat{s} > s$ when purchasing τ's plan. I'll call such an increase in reimbursement a *$\psi(s)$-transfer* to τ — and a "premium-compensated" one when it's paired with a premium increase that keeps τ indifferent overall.

Valuing Nonlinear Coverage Note that $[1 - G_\tau(s)]\, \frac{\psi_t^\tau(s)}{\lambda_t^\tau}$ is t's willingness-to-pay (WTP) for a $\psi(s)$-transfer to τ when choosing τ's plan. For linear contracts, Lemma 4 showed that higher-risk types have higher WTP for higher coverage, and this was key to Theorem 1. The analogous result for nonlinear contracts says that they also have higher WTP for an increase in *marginal* coverage at $m_t(s)$ or, equivalently, for a $\psi(s)$-transfer. The

[41] I discuss conditions for validity of the FOA under nonlinear contracts in Appendix H.3, specifically in Proposition H.1 part i which adapts a result of [Kapička 2013].

intuition is that since $G_{t-1} \prec_{\text{MLRP}} G_t$, t is strictly more likely than $t-1$ to *exceed* s (or, equivalently, $m_t(s)$) and thus benefit from the $\psi(s)$-transfer to t upon choosing t's plan.

Lemma 7. *Fix a direct menu $(m_t(\cdot), r_t(\cdot))_{t=1}^T$. Let $t' < t \in \mathcal{T}$ so that $G_{t'} \prec_{MLRP} G_t$. Then for any $\tau \in \mathcal{T}$ and $s \in (\underline{s}, \bar{s})$, we have $\frac{\psi_{t'}^{\tau}(s)}{\lambda_{t'}^{\tau}} < \frac{\psi_t^{\tau}(s)}{\lambda_t^{\tau}}$. In particular, setting $t' = t-1$ and $\tau = t$, we have $\frac{\psi_{t-1}^{t}(s)}{\lambda_{t-1}^{t}} < \frac{\psi_t(s)}{\lambda_t}$.*[42,43]

The Optimal Menu The next proposition characterizes the optimal menu under the FOA. It is interpreted at length through the rest of this section.[44] Let F be the CDF on risk-types $\mathcal{T} \equiv \{1, ..., T\}$ so that $f_t = F(t) - F(t-1)$ for $t > 1$, and let $F_t \equiv F(t)$.

Proposition 3 (Optimal Menu)**.** *Suppose $(c_t, \pi_t)_{t=1}^T$ is the indirect menu induced by an optimal solution $(m_t(\cdot), r_t(\cdot))_{t=1}^T$ to the FOA-relaxed direct problem (7.1) under money-loss CARA utility. For each $t \in \mathcal{T}$ and $m \in \mathbb{R}_+$, if there exists s such that $m_t(s) = m$ and the derivative $c_t'(m)$ exists, then the following condition is satisfied:*

$$1 - c_t'\left(m_t(s)\right) = \frac{1 - G_t(s)}{g_t(s)} \frac{h_{12}\left(m_t(s), s\right)}{\lambda} \cdot \Bigg\{ \underbrace{\lambda \underbrace{\left[-1 + \frac{\psi_t(s)}{\lambda_t}\right]}_{\textit{static term}}}_{} + \underbrace{\frac{1 - F_{t-1}}{f_t} \underbrace{\left[\frac{\psi_t(s)}{\lambda_t} - \frac{\psi_{t-1}^t(s)}{\lambda_{t-1}^t}\right]}_{\textit{slackening}} \underbrace{\left[-\lambda + \sum_{i=t}^{T} \frac{f_i}{1 - F_{t-1}} \lambda_i\right]}_{\textit{redistr. force}}}_{\textit{dynamic term}} \Bigg\}, \tag{7.3}$$

where $\lambda \equiv \sum_{t=1}^T f_t \lambda_t$ and, by convention, $F_0 \equiv 1$.[45,46]

Equation (7.3) can be decomposed into two parts, corresponding to the static and dynamic forces. As explained below, the dynamic term captures the same upward distortionary force we studied in the linear case.

[42] This lemma is proved in Appendix J.2 as part of a more general result, Lemma J.1.

[43] The spirit of this idea is found in [Crocker and Snow 2011], but in their framework, which is a multidimensional analog of [Rothschild and Stiglitz 1976], their utility function depends purely on monetary consumption; they do not model the consumer's health state or treatment decision.

[44] Proposition 3 is proved as a special case of its wealth-effects generalization Proposition I.1, which is stated in Appendix I.1. The proof is in Appendix J.3.1-J.3.3.

[45] The convention $F_0 \equiv 1$ works here because $\sum_{k=1}^T f_k \left[\lambda_k - \lambda\right] = 0$ as can be seen easily from the formula $\lambda \equiv \sum_{t=1}^T f_t \lambda_t$ at the end of the Proposition. This convention implies that the dynamic term disappears for $t = 1$ since $F_{1-1} = F_0 \equiv 1$ (so that $1 - F_{1-1} = 0$), thus manifesting the principle of no distortion at the "top".

[46] Note (7.3) is an *implicit* characterization in that c_t (as induced by m_t and r_t) appears inside the utility function in several objects on the right-hand-side (recall 7.2).

The static force is represented by the left-hand side (LHS) together with the term outside the curly braces on the RHS and the first term inside them (which is labeled "static term"). Overall, the static force captures the usual risk-incentives tradeoff from static (nonlinear) moral hazard problems [Blomqvist 1997]. The LHS, which is the marginal coverage rate $1 - c_t'(m_t(s))$, captures the rate at which the budget would be burdened by an increase in (t, s)'s treatment choice. The first term in curly braces represents the utility value of the money saved by a premium-compensated $\psi(s)$-transfer to t. This is strictly positive since such a compound transfer redistributes money to higher marginal utility consumers; mathematically, $\psi_t(s) > \lambda_t$ for $s > \underline{s}$. All transfers are viewed as being made from the ex-ante budget — rather than from any particular type — and thus incurring a (per-dollar) utility-cost λ, the multiplier on the budget constraint (7.1a). The appearance of $\frac{1-G_t(s)}{g_t(s)}$ indicates how t's contract is partially *tailored* to G_t, notwithstanding distortion resulting from the dynamic term.

The dynamic force is represented by the second term in curly braces, excluding the inverse hazard $\frac{1-F_{t-1}}{f_t}$. Paralleling the linear case, a premium-compensated $\psi_t(s)$-transfer slackens upward IC (Lem. 7), thereby allowing for greater cross-subsidization to higher risk-types. Specifically, the slackening allows for $\frac{\psi_t(s)}{\lambda_t} - \frac{\psi^t_{t-1}(s)}{\lambda^t_{t-1}} > 0$ dollars to be premium-transferred to t (from the budget) while keeping $t-1$ indifferent between t's contract and her own. In turn, such a transfer to t slackens t's upward-adjacent IC constraint, thus allowing for a similar transfer to $t+1$ and so on. This explains the last factor on the right in (7.3): $\left[-\lambda + \sum_{i=t}^{T} \frac{f_i}{1-F_{t-1}} \lambda_i\right]$. Crucially, the presence of the slackening factor $\frac{\psi_t(s)}{\lambda_t} - \frac{\psi^t_{t-1}(s)}{\lambda^t_{t-1}}$ in the dynamic term suggests that the coverage rate $1-c_t'(m_t(s))$ will be increasing (ceteris paribus) in the difference between the exceedance probabilities $1 - G_t(m_t(s))$ and $1 - G_{t-1}(m_t(s))$.

The inverse hazard rate $\frac{1-F_{t-1}}{f_t}$ determines the relative importance of the dynamic and static forces, holding fixed the size of each. Ceteris paribus, optimal marginal coverage $1 - c_t'(m_t(s))$ will be increasing in this inverse hazard rate for the following reason. The cost of a premium-compensated increase in t's marginal coverage — which comprises the static force — is proportional to f_t while the amount of interim redistribution it enables is proportional to $1 - F_{t-1}$, the fraction of interim individuals towards whom it will be carried out.[47]

Without the dynamic term, (7.3) reduces to the analogous formula for t's *static* moral

[47] I say "ceteris paribus" because as t gets lower and $1-F_{t-1}$ thus gets higher, the wedge $-\lambda + \sum_{\ell=t}^{T} \frac{f_\ell}{1-F_{t-1}} \lambda_\ell$ shrinks, thus lessening the value of redistribution conditional on the number of beneficiaries thereof .

hazard problem — or, more precisely, the (microeconomic) dual thereof:[48]

$$1 - c_t'\left(m_t(s)\right) \;= \frac{1 - G_t(s)}{g_t(s)} h_{12} \underbrace{\left[-1 + \frac{\psi_t(s)}{\lambda_t}\right]}_{\text{static term}}. \tag{7.4}$$

Given that the dynamic term is strictly positive — by Lemma 7 (for the slackening term) and Lemma 2 (for the redistribution term) — this suggests a nonlinear analog to the linear distortion result $x_t > \hat{x}_t$ (Thm. 1), such as $1 - c_t'\left(m_t(s)\right) > 1 - \hat{c}_t'\left(m_t(s)\right)$. I discuss this question in Appendix H.4.

8 Reinterpretation as Social Health Insurance

The preceding sections analyzed the private-risk optimization problem (5.1), which is the consumer's problem in the model from Section 3. In this section, I reinterpret that analysis as pertaining to a social health insurance program with a consumer-participation mandate rather than a (perfectly competitive) private-sector dynamic insurance market. The discussion here is informal, but I more formally describe a model of the social insurance interpretation in Appendix B. It will be easy to see that the two models are equivalent in the sense that they lead to the same constrained optimization problem. Thus all the same insights from elsewhere in the paper apply to the social insurance model as well.

The social insurance model thinks of a utilitarian social planner designing a menu of static contracts for a population of consumers subject to a consumer-participation mandate. The population begins with heterogeneous privately-known risk profiles but homogeneous wealth. Relative to the original model from Section 3, the social insurance model dispenses with the ex-ante stage Period 0 where the dynamic contract was proposed and accepted. It instead begins immediately at Period 1A where consumers privately realize their risk-types. A new stage is then inserted where the social planner — who replaces the insurer in the new model — designs a menu of plans to *impose* on the population. Periods 1B, 2A, and 2B then proceed identically as in the original model, with the insurer replaced by the planner. At the aforementioned design stage, the planner designs the menu to maximize expected consumer welfare subject to an expected zero-budget constraint. This constraint guarantees that the program can be fully funded by premiums, assuming there is no aggregate risk.

The social insurance model captures part of the motivation for the Affordable Care Act

[48] I mean "dual" in the sense of minimizing (expected) insurer losses subject to achieving a fixed expected utility for the consumer, not in the sense of "Lagrange duality".

(ACA) program in the U.S. and similar institutions in other countries,[49] though an important qualification about regulated-competition programs is discussed at the end of Section 8.1. The social insurance model also applies to employer-based health insurance programs which can be viewed as providing intra-firm social insurance.

The equivalence of the two models shows that the key insights of the preceding sections have implications for the design of social insurance programs as well. In particular, Theorem 1 suggests that, if risk is non-contractible, coverage levels in social insurance programs should be set above what would be statically efficient, in order to enable greater cross-subsidization between consumers with pre-existing risk differences.

8.1 The Private and Social Models in Practice (incl. ACA)

The original model envisions a private dynamic contract with symmetric information at the ex-ante stage and private information at the interim (and ex-post) stage. This informational structure might occur if, for example, consumers sign the dynamic contract at age 30, when most people are assumed to be fairly healthy and of similar risk. They then would choose continuation plans at 31 after potentially learning new risk information, and finally choose treatment at 32 given their current needs. The private model abstracts from treatment that may occur between 30 and 32.

Turning to the social insurance model, although it captures much of the motivation behind the Affordable Care Act (ACA), there is a key difference.[50] In the ACA and other regulated-competition programs internationally, private insurers set premiums competitively as part of a regulated market, rather than — as in my social model — the government directly setting premiums (and coinsurance schedules) as the *single payer* in the system. Despite the participation mandate, when multiple plans are offered, premiums set directly by (competing) profit-maximizing insurers will reflect the risk pool of each plan and thus won't cross-subsidize between plans. Partly to address this problem, the ACA and many peer programs use so-called *risk adjustment* schemes which force monetary transfers between plans that attract unequal risk pools [Layton et al. 2018].[51]

Applying the insights from my analysis to a regulated-competition program with risk adjustment is a potentially interesting direction for future research, though (as discussed in Section 2.1) the topic is explored already by [Zwart 2025]. To repeat, in his model, the

[49] The ACA initially had a participation mandate, though it was subsequently repealed.

[50] A further difference is that, unlike the ACA, my social insurance model is fully funded by consumer premiums, whereas the ACA uses external government funding and also redistributes money to low-income consumers rather than just those with high health risk.

[51] Risk adjustment is also intended to disincentivize subtle forms of cream-skimming techniques by insurers that cannot be directly banned.

regulator first sets lump-sum risk-adjustment transfers between insurers depending on the risk-type of their enrollees, and then competing insurers choose premiums and plan features. In particular, when the regulator increases the degree of risk-adjustment transfers for high-risk enrollees (beyond a certain point), an upward distortion of the high-risk contract's generosity will be carried out by the insurers in order to attract the high-risk consumers and thereby benefit from their risk-adjustment subsidies. Nevertheless, because [Zwart 2025] works with parsimonious type spaces and (more importantly) a relatively reduced-form preference specification — where aversion to copayments and concern for reclassification risk are captured through differential weights rather than a micro-founded ex-ante expected-utility framework — there likely remains room for future work on socially optimal risk adjustment in a richer environment like the one developed in this paper.

9 Conclusion

This paper studies a model of dynamic health insurance under evolving risk with moral hazard and adverse selection. The analysis can be thought of in terms of either a *commercial* dynamic contract or a static *social* health insurance program with a consumer-participation mandate.

Before studying the private-risk case (i.e., adverse selection), I analyze the observable-risk case and show that its solution involves interim-efficient coinsurance schedules for all risk-types and full insurance of dynamic risk in a certain sense. To the extent that real consumer risk profiles are observable and verifiable, this analysis can inform policy. While regulators sometimes ban insurers from conditioning contracts on consumers' risk information in order to protect them from dynamic risk, my analysis shows how, in principle, such information can — given the right objective — be used to actually improve protection from dynamic-risk while also tailoring coinsurance schedules to address moral hazard efficiently.

Turning to the private-risk solution, I show that, unlike the observable-risk solution, it faces a tradeoff between interim efficiency and protection from dynamic risk. My main result — which applies to linear coinsurance — states that coverage will be higher than the interim-efficient level for each risk-type except the lowest. This upward distortion of coverage, when paired with a budget-balancing increase in premium, relaxes upward IC and thus enables more cross-subsidization from low- to high-risk types. With nonlinear coinsurance, the same dynamic force that led to the upward distortion in the linear case is still present. However now, the amount of IC-relaxation — and hence the amount of enabled redistribution — depends on the relative likelihood that the type in question (versus the type below them) will *exceed* the level of spending at which marginal coverage is increased. This suggests that

greater upward coverage distortions are optimal when that relative likelihood is higher.

Future Directions

My analysis suggests several directions for future research. First, as discussed in Section 8.1, one could embed risk adjustment into a regulated-competition environment à la [Zwart 2025], but using the richer framework developed here. Second, one could restrict to a fixed number of plans while having a continuum of risk-types so that some types are infra-marginal while others are marginal; this issue is studied by [Chade et al. 2024], albeit under a different objective. Third, whereas I assume the consumer can fully commit to staying in the dynamic contract should they turn out to be low risk, one could study a similar model with limited consumer commitment. This would involve extending papers such as [Ghili et al. 2023] or [Cochrane 1995] to account for moral hazard.[52] Fourth, one could study a model with multiple rounds of treatment rather than just one. Coverage levels would then depend not only on "plan choice" but also on prior treatment choice since treatment choice may reveal private information about risk of future needs in addition to information about current needs.[53,54] Relatedly, a model with multiple treatment rounds could also allow for treatment to endogenously affect risk of future health needs, as with preventative care [Ellis and Manning 2007], [Li 2023].[55]

Fifth, while this paper considers only a single dimension of health and treatment, real consumers face many potential health problems, each with their own set of treatment options. Thus, insurers can condition coinsurance on what kind of treatment is being purchased rather than just on the total spending amount. In a companion paper [Boche 2025], I study the implications of such multi-dimensionality for optimal coinsurance under moral hazard; prior work on the topic includes [Besley 1988], [Goldman and Philipson 2007] and [Ellis et al. 2015].

[52] See also [Golosov and Iovino 2021].

[53] [Luz 2022] also allows for such dependence on treatment history (and also considers its prohibition) but with no moral hazard – hence "treatment history" not "treatment choice".

[54] Technically, standard year-long contracts with nonlinear coinsurance already exhibit dependence on within-year treatment history (e.g., through deductibles) [Brot-Goldberg et al. 2017], but it's not clear this dependence is designed to achieve dynamic screening. Designers may instead have a static, single-treatment model in mind.

[55] Mathematically, this would parallel the "learning by doing" phenomenon in optimal taxation [Makris and Pavan 2021]. See also [Luz and Parra 2023].

References

Atal, Juan Pablo, Hanming Fang, Martin Karlsson, and Nicolas R. Ziebarth (2025). "German Long-Term Health Insurance: Theory Meets Evidence". In: *J. Polit. Econ.*

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein (2015). "Behavioral Hazard in Health Insurance". In: *Q. J. Econ.* 130.4, pp. 1623–1667.

Battaglini, Marco and Rohit Lamba (2019). "Optimal dynamic contracting: The first-order approach and beyond". In: *Theor. Econ.* 14.4.

Besley, Timothy J (1988). "Optimal reimbursement health insurance and the theory of Ramsey taxation". In: *J. Health Econ.* 7.4, pp. 321–336.

Bijlsma, Michiel, Jan Boone, and Gijsbert Zwart (2017). "The complementarity between risk adjustment and community rating: Distorting market outcomes to facilitate redistribution". In: *J. Public Econ.* 155, pp. 21–37.

Blomqvist, Åke (1997). "Optimal non-linear health insurance". In: *J. Health Econ.* 16.3, pp. 303–321.

Boche, Alex (2025). "Efficiency of Bundled Copays in Multidimensional Health Insurance".

Boone, Jan (2015). "Basic versus supplementary health insurance: Moral hazard and adverse selection". In: *J. Public Econ.* 128, pp. 50–58.

— (2018). "Basic versus supplementary health insurance: Access to care and the role of cost effectiveness". In: *J. Health Econ.* 60, pp. 53–74.

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad (2017). "What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics". In: *Q. J. Econ.* 132.3, pp. 1261–1318.

Chade, Hector, Victoria Marone, Amanda Starc, and Jeroen Swinkels (2024). "Multidimensional Screening and Menu Design in Health Insurance Markets".

Chernew, Michael E., William E. Encinosa, and Richard A. Hirth (2000). "Optimal health insurance: The case of observable, severe illness". In: *J. Health Econ.* 19.5.

Clarke, Francis (2013). *Functional Analysis, Calculus of Variations and Optimal Control.* London: Springer.

Cochrane, John H. (1995). "Time-Consistent Health Insurance". In: *J. Polit. Econ.* 103.3, pp. 445–473.

Courty, Pascal and Hao Li (2000). "Sequential Screening". In: *Rev. Econ. Stud.* 67.4, pp. 697–717.

Crocker, Keith J and Arthur Snow (2011). "Multidimensional Screening in Insurance Markets with Adverse Selection". In: *J. Risk Insur.* 78.2, pp. 287–307.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J. Notowidigdo (2018). "The Economic Consequences of Hospital Admissions". In: *Am. Econ. Rev.* 108.2, pp. 308–352.

Edlin, Aaron S and Chris Shannon (1998). "Strict Monotonicity in Comparative Statics". In: *J. Econ. Theory* 81.1, pp. 201–219.

Ellis, Randall P, Shenyi Jiang, and Willard G Manning (2015). "Optimal health insurance for multiple goods and time periods". In: *J. Health Econ.* 41, pp. 89–106.

Ellis, Randall P and Willard G Manning (2007). "Optimal health insurance for prevention and treatment". In: *J. Health Econ.* 26.6, pp. 1128–1150.

Farhi, Emmanuel and Iván Werning (2013). "Insurance and Taxation over the Life Cycle". In: *Rev. Econ. Stud.* 80.2 (283), pp. 596–635.

Fudenberg, Drew, Bengt Holmstrom, and Paul Milgrom (1990). "Short-term contracts and long-term agency relationships". In: *J. Econ. Theory* 51.1.

Ghili, Soheil, Ben Handel, Igal Hendel, and Michael D Whinston (2023). "Optimal Long-Term Health Insurance Contracts: Characterization, Computation, and Welfare Effects". In: *Rev. Econ. Stud.*

Goldman, Dana and Tomas J Philipson (2007). "Integrated Insurance Design in the Presence of Multiple Medical Technologies". In: *Am. Econ. Rev.* 97.2, pp. 427–432.

Golosov, Mikhail and Luigi Iovino (2021). "Social insurance, information revelation, and lack of commitment". In: *J. Polit. Econ.* 129.9.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski (2016). "Redistribution and social insurance". In: *Am. Econ. Rev.* 106.2, pp. 359–386.

Handel, Ben, Igal Hendel, and Michael D. Whinston (2015). "Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk". In: *Econometrica* 83.4.

Hendel, Igal and Alessandro Lizzeri (2003). "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance". In: *Q. J. Econ.* 118.1, pp. 299–328.

Hendren, Nathaniel (2021). "Measuring Ex Ante Welfare in Insurance Markets". In: *Rev. Econ. Stud.* 88.3, pp. 1193–1223.

Ho, Kate and Robin S Lee (2023). "Health insurance menu design for large employers". In: *RAND J. Econ.* 54.4, pp. 598–637.

Jahn, Johannes (2020). *Introduction to the Theory of Nonlinear Optimization, Fourth edition.*

Kapička, Marek (2013). "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach". In: *Rev. Econ. Stud.* 80.3, pp. 1027–1054.

Kleef, R. C. van, W.P.M.M. van de Ven, and R.C.J.A. van Vliet (2009). "Shifted deductibles for high risks: More effective in reducing moral hazard than traditional deductibles". In: *J. Health Econ.* 28.1, pp. 198–209.

Layton, Timothy J, Ellen Montz, and Mark Shepard (2018). "Health Plan Payment in US Marketplaces: Regulated Competition With a Weak Mandate". In: *Risk Adjust. Risk Shar. Prem. Regul. Heal. Insur. Mark. Theory Pract.* Ed. by Thomas G. McGuire and Richard C. van Kleef. Academic Press. Chap. 17, pp. 491–522.

Li, Anran (2023). "Commitment, Competition, and Preventive Care Provision".

Lollivier, Stefan and Jean-Charles Rochet (1983). "Bunching and second-order conditions: A note on optimal tax theory". In: *J. Econ. Theory* 31.2, pp. 392–400.

Luz, Vitor Farinha (2022). "Optimal Dynamic Insurance Contracts". In: *Work. Pap.*, pp. 1–48.

Luz, Vitor Farinha and Alvaro Parra (2023). "Health Insurance Markets with Endogenous Risks".

Ma, Ching To Albert and Michael H. Riordan (2002). "Health insurance, moral hazard, and managed care". In: *J. Econ. Manag. Strateg.* 11.1.

Makris, Miltiadis and Alessandro Pavan (2020). "Wedge Dynamics with Evolving Private Information".

— (2021). "Taxation under learning by doing". In: *J. Polit. Econ.* 129.6.

Marone, Victoria R. and Adrienne Sabety (2022). "When Should There Be Vertical Choice in Health Insurance Markets?" In: *Am. Econ. Rev.* 112.1.

Martinon, Pierre, Pierre Picard, and Anasuya Raj (2018). "On the design of optimal health insurance contracts under ex post moral hazard". In: *Geneva Risk Insur. Rev.* 43.2, pp. 137–185.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995). *Microeconomic Theory*. Oxford University Press.

Milgrom, Paul and Ilya Segal (2002). "Envelope Theorems for Arbitrary Choice Sets". In: *Econometrica* 70.2, pp. 583–601.

Mirrlees, J A (1971). "An Exploration in the Theory of Optimum Income Taxation". In: *Rev. Econ. Stud.* 38.2, pp. 175–208.

Nyman, John A. (1999). "The value of health insurance: the access motive". In: *J. Health Econ.* 18.2, pp. 141–152.

Pauly, Mark V. (1968). "The Economics of Moral Hazard: Comment". In: *Am. Econ. Rev.* 58.3, pp. 531–537.

Pauly, Mark V., Howard Kunreuther, and Richard A. Hirth (1995). "Guaranteed Renewability in Insurance". In: *J. Risk Uncertain.* 10.2, pp. 143–156.

Pauly, Mark V., Kai Menzel, Howard Kunreuther, and Richard A. Hirth (2011). "Guaranteed renewability uniquely prevents adverse selection in individual health insurance". In: *J. Risk Uncertain.* 43.2, pp. 127–139.

Pavan, Alessandro, Ilya Segal, and Juuso Toikka (2014). "Dynamic Mechanism Design: A Myersonian Approach". In: *Econometrica* 82.2, pp. 601–653.

Rothschild, Michael and Joseph Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information". In: *Q. J. Econ.* 90.4, pp. 629–649.

Stantcheva, Stefanie (2020). "Dynamic taxation". In: *Annu. Rev. Econom.* 12.

Werquin, Nicolas, Aleh Tsyvinski, and Mikhail Golosov (2015). "A Variational Approach to the Analysis of Tax Systems". In: NBER Working Paper 244.

Zeckhauser, Richard (1970). "Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives". In: *J. Econ. Theory* 2.1.

Zeckhauser, Richard and Michael Spence (1971). "Insurance, Information, and Individual Action". In: *Am. Econ. Rev.* 61.2, pp. 380–387.

Zhang, Yan, Yonghong Wu, and Haixiang Yao (2022). "Optimal health insurance with constraints under utility of health, wealth and income". In: *J. Ind. Manag. Optim.* 18.3, pp. 1519–1540.

Zwart, Gijsbert (2025). "Moral hazard and risk adjustment". In: *J. Health Econ.* 99.

A Prior Authorization and Gatekeeping

I now return to the observable-risk model from Section 4.2 and informally reinterpret it as a model of non-binary prior-authorization. Rather than thinking of the signal t as a signal of consumer "risk" of health needs s that might materialize in the *future*, we can view it as a noisy audit of the patient's *current* needs s immediately prior to the treatment decision. For example, a surgeon might assess the necessity of surgery for a patient on, say, a 5-point scale like the following: 1) inappropriate, to 3) beneficial but not necessary, to 5) essential. If we think of t as the surgeon's assessment score, the patient's copay for the surgery could then be conditioned on this assessment as in my observable-risk benchmark.

Proposition 1, which characterized the observable-risk solution, suggests how such conditioning might work. First, recall part i of that proposition. Essentially, it said that patient cost-sharing conditional on the signal t should efficiently balance risk-protection and incentives given the *residual* uncertainty about the patient's degree of need. This suggests that the copay for the surgery should be decreasing in the degree of necessity perceived by the surgeon since higher perceived necessity means that incentives are less likely to be needed. Second, recall part ii of that proposition, which prescribes cross-subsidization across signal realizations t until expected marginal utilities are equalized. This suggests that it may be beneficial to offer a rebate to patients who — despite a verifiable signal suggesting at least some benefits from the procedure — decline the costly surgery. The reason is that patients for whom the signal suggests the procedure may be helpful face risk of having to pay the copay which makes their expected marginal utility higher than that of patients for whom the signal shows the procedure is completely unnecessary. This last point about rebates is shown formally by [Chernew et al. 2000].[56] In practice, copays/coinsurance would, of course, also depend on plan choice (or observed *a priori* risk-signals) in addition to these ex-post signals of severity. I leave such extensions for future work, but I note that analysis of these issues must also account for multi-dimensionality of health conditions and treatments — a topic explored in [Besley 1988], [Goldman and Philipson 2007], [Ellis and Manning 2007] [Ellis et al. 2015], and [Boche 2025].

B Formal Model for Social Insurance Interpretation

This section more formally describes the social insurance model which was introduced informally in Section 8 and which underlies the social insurance interpretation of my analysis from

[56] Their model features binary treatment and a binary signal that indicates whether the patient has the condition at all but not its severity. Reinterpreted in the way suggested in this section, my observable-risk benchmark can roughly be viewed as a generalization of their model.

the rest of the paper. It is easy to see that the planner's constrained optimization problem in this new model coincides with that of the ex-ante consumer in the original private-dynamic model from Section 3. Thus, the solutions coincide as well.

Players: A social planner and a *population* of consumers.

Timing:

Period 1.A: Each consumer privately realizes their risk-type $t \in \mathcal{T} \equiv \{1, ..., T\}$ with probability f_t.

Design Stage: The social planner chooses a menu of contracts $(c_\tau(\cdot), \pi_\tau)_{\tau \in \mathcal{T}}$ and imposes it on the population of consumers.[57]

Period 1.B: Each consumer then must choose a plan $(c_\tau(\cdot), \pi_\tau)$ from the menu $(c_\tau(\cdot), \pi_\tau)_{\tau \in \mathcal{T}}$ and pay the associated premium π_τ to the planner.

Period 2.A: Each consumer then privately realizes their pre-treatment health *severity* $s \in \mathbb{R}$ according to their distribution $G_t(\cdot)$.

Period 2.B: Each consumer then chooses their treatment level m (in dollars) and pays the out-of-pocket amount $c_\tau(m)$ to the (unmodeled) healthcare provider. The planner pays the remainder: $m - c_\tau(m)$.

Preferences:

Consumer Preferences: Each consumer maximizes the expectation of their utility $U(y, h)$, where (recall) y and h denote monetary consumption and final health state respectively. Note that all consumers start with the same baseline wealth $w = 0$.

Planner Preferences: At the Design Stage, the planner chooses the menu to maximize aggregate consumer expected utility in the population — that is, $\sum_{t=1}^{T} f_t \int U(y, h) dG_t(s)$ — subject to an expected zero-budget constraint, where y and h denote monetary consumption and final health state respectively.

C Extra Properties of Static Solution (Linear) — Bounds and Comparative Static in S

C.1 Bounds on $\hat{x}_t$

Intuitively, it is optimal to have at least some coverage because when there is zero coverage, the marginal moral hazard cost of extra coverage is only second-order but the insurance

[57] The order of Period 1A and the Design Stage can also be reversed without affecting the analysis.

motive is first-order. The next result shows this is true under money-loss and that under general utility, it is profitable to marginally increase coverage at $x = 0$, thus showing $\hat{x}_{tS} \neq 0$ and giving strong evidence that $\hat{x}_{tS} > 0$. I prove (and restate) it later in this subsection.

Proposition C.1 ($\hat{x} > 0$)**.** *i) Let utility be money-loss. For any S and G_t, if $\hat{x}_{tS}$ exists, $\hat{x}_{tS} > 0$.*[58] *ii) Let utility be general with $U_{12} \leq 0$, and suppose $h\left(m_{c,\pi}(s), s\right)$ is weakly decreasing in s. For any S and G_t, $\alpha'_{tS}(0) > 0$ (where, recall, α takes x as argument).*[59]

Is it also guaranteed that $\hat{x}_{tS} < 1$? Not in general. The following inessential remark explains.

Remark 1. $\hat{x}_{tS} > 1$ can occur if when $U_{12}(y, h) < 0$ (e.g. with money-loss where $U_{12} = u'' < 0$) *and* treatment does not fully cure health problems so that $h\left(m_{x,\pi}(s), s\right)$ is strictly decreasing in s.[60] This is because, in such a case, even with full coverage of health spending, the consumer's poor post-treatment health h leaves them with a high marginal utility of monetary consumption U_1. And as a result, it may be optimal to give them a cash indemnity as an increasing function of their fully-covered treatment spending, which is exactly what $x > 1$ does.[61,62]

C.2 Wealth Effects in the Static Problem (Linear Contracts)

In this section, I study how baseline wealth (or, equivalently, subsidy) affects statically efficient linear coverage. In doing so, I consider two channels: the effect of baseline wealth on risk-attitude (henceforth, "ex-ante wealth effects") and the effect of wealth on treatment-sensitivity to coverage and premiums (henceforth, "ex-post wealth effects"). Considering that statically efficient coverage balances risk exposure with treatment incentives, it should be clear that both kinds of wealth effects are relevant. Regarding ex-ante wealth effects, I show

[58] In case there are multiple optimal static solutions $\hat{x}_{tS}$, this result should be understood as saying they're all > 0.

[59] Presumably, some version of this "strictly positive coinsurance" result is already in the literature, but I'm not sure where to attribute it.

[60] Such lack of "full cure" may occur either because the technology for full cure does not exist, or because full cure would not be efficient given its cost. At least with money-loss utility, lack of full-cure may also be interpreted as a reduced form for a loss of income that is caused by health problems (see, e.g., [Dobkin et al. 2018]). Several papers including [Marone and Sabety 2022] use the form $h(m, s) = -\frac{1}{2\omega}(m - s)^2 + (m - s)$ for $\omega > 0$, which leads to "full cure" in the sense that endogenous post-treatment health state $h(m, s)$ is constant in s.

[61] If the idea of $\hat{x}_{tS} > 1$ seems odd, recall that I'm abstracting from much of the complexity of healthcare including the role of the physician as an "auditor" and gatekeeper to costly care provision. Again, see A for more on the latter issues.

[62] In any case, Fact G.1 gives conditions to ensure that even if coverage is greater than full, $m_{c,\pi}(s)$ is still finite.

that under money-loss (which eliminates ex-post wealth effects), $\hat{x}_{tS}$ is strictly decreasing in S if utility satisfies of decreasing absolute risk aversion (DARA), a plausible condition. On the other hand, the ex-post wealth-effects channel appears to be ambiguous, thus also leaving the overall effect of S on $\hat{x}_{tS}$ ambiguous for general utility.

This analysis informs the study of the dynamic problem (both private- and observable-risk) since when cross-subsidization occurs, it affects statically efficient coverage $\hat{x}_{tS_t^*}$ (or $\hat{x}_{t\tilde{S}_t}$), which can then affect the tradeoff between static efficiency and cross-subsidization — thus creating a feedback loop. The endogenous subsidy S_t^* (or $\tilde{S}_t$) is very likely to be strictly increasing in t — see Prop. E.1 for a formal result for the two-type private-risk case — and obviously crosses zero between some intermediate risk-types given the budget constraint 5.1c. Thus, if $\hat{x}_{tS}$ is decreasing in S, as will be the case for money-loss DARA, then we'll have $\hat{x}_{tS_t^*} > \hat{x}_t$ for lower t and $\hat{x}_{tS_t^*} < \hat{x}_t$ for higher t, and similarly for $\tilde{S}_t$. If instead, $\hat{x}_{tS}$ is increasing in S, then the opposite will hold.

C.2.1 Ex-Ante Wealth Effects

As shown in Lemma G.15 (Appendix G.2.3), statically efficient coverage $\hat{x}_t$ is increasing in risk-aversion under money-loss utility. Putting this together with the definition of DARA shows that statically efficient coverage is decreasing in baseline wealth. This is the next proposition, which is proved in Appendix G.2.3. To set up its notation, for any $w \in \mathbb{R}$, let $\hat{x}_{tS}^w$ denote t's optimal linear static coverage with baseline wealth w and actuarial subsidy S.[63] w and S play interchangeable roles in the static problem when the other is held fixed.

Proposition C.2 (S Comparative Static, Money-Loss)**.** *Let utility be money-loss and assume the static problem has a unique solution for all S and baseline wealth levels w. If u satisfies strict DARA, IARA, or CARA, $\hat{x}_{tS}^w$ is (respectively) strictly decreasing, strictly increasing, or constant in both S and w.*[64]

C.2.2 Adding Ex-Post Wealth Effects

Without money-loss utility, health is a strictly normal good (by assumption A4 in Section 3), so we must also take into account the effect of baseline wealth (or, equivalently, subsidy) on treatment-sensitivity to coverage and premium $\frac{\partial}{\partial x} m_{x,\pi}(s)$ and $\frac{\partial}{\partial \pi} m_{x,\pi}(s)$.[65] This is important for optimal static coverage since, say, higher treatment sensitivity to coverage makes moral hazard more problematic thus pushing towards lower optimal coverage. [The effect of baseline

[63] Equation G.27 in the Appendix G.2.3 gives a more formal definition.

[64] The CARA part extends the prior result Lemma 1, which dealt only with subsidies not baseline wealth levels.

[65] The effect of baseline wealth on these sensitivities are $-\frac{\partial^2}{\partial x \partial \pi} m_{x,\pi}(s)$ and $-\frac{\partial^2}{\partial \pi^2} m_{x,\pi}(s)$ respectively.

wealth on $\frac{\partial}{\partial \pi} m_{x,\pi}(s)$ comes into play because changing coverage will lead to surplus or deficit in the budget that must be dispersed through the premium which, in turn, affects treatment choice.]

Informally, the effect of baseline wealth on statically efficient coverage will operate both through this treatment-sensitivity channel and through the effective risk-aversion channel discussed in the last subsection. The former channel may push either way, so the overall comparative static of baseline wealth on optimal static coverage is ambiguous, even if the latter can be signed with a DARA assumption (e.g. in the separable utility case) as suggested above.

Ambiguity in the sensitivity comparative static $-\frac{\partial^2}{\partial x \partial \pi} m_{x,\pi}(s)$ stems in part from ambiguity of third-order derivatives such as U_{111} and U_{112}.[66] However, rather than giving a detailed analysis of the sources of ambiguity, I instead focus on pointing out one force that may push towards lower treatment for higher baseline-wealth consumers.

With high enough baseline wealth, for most s's, the consumer will be buying close to the "best" available treatment. More formally, $h_1\left(m_{\pi,x}(s), s\right)$ will be close to zero and there isn't much more treatment they can buy with positive value — recall the assumptions $h_{11} < 0$ and $\lim_{m\to\infty} h_1(m) < 0$, so that h_1 will eventually cross zero at which point the "best" treatment is obtained. In that case the sensitivity $\frac{\partial}{\partial x} m_{x,\pi}(s)$ (which is positive) will likely be very small for most s's given the very low benefit and still-present out-of-pocket cost of increased treatment. If so, moral hazard will be relatively less of a concern for very wealthy people, so their statically efficient coverage will be higher, all else equal.

If, as just suggested (at least for high enough initial wealth), marginally higher wealth decreases treatment sensitivity to coverage, then the ex-ante and ex-post wealth effects will push in opposite directions leading to an ambiguous effect on statically efficient coverage.

D Wealth Effects in Dynamic Problem (Obs.- and Private-Risk)

D.1 Observable Risk: Wealth Effects

Recall Proposition 1 ii, which characterized the interim-redistributive aspect of the observable risk-solution. There is an important sense in which the condition $\tilde{\lambda}_t \cdot \left[\frac{\partial}{\partial \pi} P_t(\tilde{c}_t, \tilde{\pi}_t)\right]^{-1} = \tilde{\lambda}_\tau \cdot \left[\frac{\partial}{\partial \pi} P_\tau(\tilde{c}_\tau, \tilde{\pi}_\tau)\right]^{-1}$ from that result should not be viewed as the general-utility analog of the

[66] I conjecture that $U_{111} < 0$ will imply $\frac{\partial}{\partial x} m_{x,\pi}(s)$ is decreasing in baseline wealth (at for separable utility with no risk-aversion on health), thus reinforcing the treatment satiation force mentioned next. However, however this is precisely the negation of the standard prudence assumption $U_{111} > 0$.

money-loss condition $\tilde{\lambda}_t = \tilde{\lambda}_\tau$ (also given there). Both conditions rule out profitable redistribution through premium-transfers, but it is only under money-loss that premium-transfers coincide with *treatment-choice-preserving* transfers. The latter are of interest because they isolate the interim-redistributive aspect of the problem from the static efficiency aspect, and because they appear in the multiplier on the budget constraint 4.6b.[67]

Accordingly, in this section, I explicitly construct treatment-preserving transfers under general utility, which I refer to as ν-transfers, and establish the necessary optimality condition $\tilde{\nu}_t = \tilde{\nu}_\tau$ — which generalizes $\tilde{\lambda}_t = \tilde{\lambda}_\tau$ — where $\tilde{\nu}_t$ is the expected utility of a dollar transferred to t via a ν-transfer (at the observable-risk solution). It should be noted that outside money-loss, performing ν-transfers transfers will require the flexibility of nonlinear coinsurance, so establishing the condition $\tilde{\nu}_t = \tilde{\nu}_\tau$ will as well. I also address the special case of separable utility.

Let $D^{c,\pi}(m) \equiv \frac{-U_{11}c' + U_{21}h_1}{U_1}$, where the derivatives of U are evaluated at $\left(-\pi - c(m), h\left(m, s_{c,\pi}^{-1}(m)\right)\right)$, where where $s_{c,\pi}^{-1}(\cdot)$ is the inverse function of $m_{c,\pi}(\cdot)$ if it exists. This can be equivalently defined as $D^{c,\pi}(m) \equiv -\frac{\mathbb{I}_{\text{NG}}^{c,\pi}(m)}{\mathbb{E}_{\text{NG}}^{(c),c,\pi}(m)}$, where and $\mathbb{E}_{\text{NG}}^{(c),c,\pi}(m) < 0$ is the derivative of the compensated treatment-demand with respect to out-of-pocket price and $\mathbb{I}_{\text{NG}}(m)$ is the derivative of treatment-demand with respect to income; for completeness, I define these objects as follows:

$$\mathbb{E}_{\text{NG}}^{(c),c,\pi}(m) \equiv -\frac{-U_1}{Y} \quad \mathbb{I}_{\text{NG}}^{c,\pi}(m) \equiv -\frac{-U_{11}c' + U_{21}h_1}{Y},$$

where $Y \equiv U_{11}\left(c'\right)^2 - 2U_{12}c'h_1 + U_{22}h_1^2 + U_2 h_{11}$ is simply shortcut notation, and c and π are hidden within the suppressed arguments of the constituent expressions. Note that for money-loss, $D^{c,\pi}(m) = 0$ because $D^{c,\pi}(m) = -u''\left[c' + h_1\right] = 0$ by the FOC for treatment choice; for separable, $D^{c,\pi}(m)$ reduces to $D^{c,\pi}(m) = \frac{-U_{11}c'}{U_1}$. Also note that $\mathbb{I}_{\text{NG}}^{c,\pi}(m) > 0$ outside money-loss by diminishing MRS (assumption A4 in the model section).

For all $t \in \mathcal{T}$ and all (c, π), let

$$L_{c,\pi,s}(\hat{s}) \equiv \exp\left(-\int_{m_{c,\pi}(s)}^{m_{c,\pi}(\hat{s})} D^{c,\pi}(\hat{m})d\hat{m}\right), \tag{D.1}$$

$$\nu_t^{c,\pi} \equiv \left[\int_{\underline{s}}^{\bar{s}} L_{c,\pi,\underline{s}}(\hat{s})dG_t(\hat{s})\right]^{-1} \cdot \int_{\underline{s}}^{\bar{s}} W_1^{c,\pi}(\hat{s})L_{c,\pi,\underline{s}}(\hat{s})dG_t(\hat{s}). \tag{D.2}$$

Before interpreting, I note that for money-loss, $L_{c,\pi,s}\left(\hat{s}\right) = 1$ for all $s < \hat{s}$ since $D^{c,\pi}(m) = 0$. Turning to interpretation, $\nu_t^{c,\pi}$ is the increase in t's expected *utility* from (starting at c, π)

[67] I solve for the multiplier explicitly in the private-risk case in Sections 7 (end of Prop. 3) and I (Eqn. (I.5)).

increasing consumption at each s in a way that keeps treatment choices constant at the margin and costs one *dollar* in expectation (w.r.t. G_t). This is accomplished by decreasing c at each supported $m(s)$ by the following dollar amount

$$\beta_t^{c,\pi}\left(m(s)\right) \equiv \underline{\beta}_t^{c,\pi} L_{c,\pi,\underline{s}}\left(s_{c,\pi}^{-1}(m)\right), \tag{D.3}$$

where $\underline{\beta}_t^{c,\pi}$ is just a normalization constant so that the change costs one dollar in expectation.[68] I refer to such a change as a *ν-transfer* to t at (c,π). I say a direction of perturbation $d(\cdot)$ "keeps treatment choices constant at the margin (at a given (c,π))" if the Gateaux derivative of $m_{c,\pi}(s)$ in direction $d(\cdot)$ is zero — that is, if $0 = \frac{\partial}{\partial \mu}\left[m_{(c+\mu d),\pi}(s)\right]_{|\mu=0}$ — for all s. The next lemma confirms formally that ν-transfers do so, and computes their effect on interim expected utility.

Lemma D.1. *i) For all (c,π), if $c \in C^2$, then $\frac{\partial}{\partial \mu}\left[m_{c+\mu\beta_t^{c,\pi},\pi}(s)\right]_{|\mu=0} = 0$ for all s.*[69] *Further, $\int \beta_t^{c,\pi}\left(m(s)\right) dG_t(s) = 1$.*

ii) $\frac{\partial}{\partial \mu}\left[V_t\left(c+\mu\beta_t^{c,\pi},\pi\right)\right]_{\mu=0} = -\nu_t^{c,\pi}$.

In the special case of separable utility, it will turn out that $\nu_t^{c,\pi}$ equals $\theta_t^{c,\pi} \equiv \left[\int 1/\left[W_1^{c,\pi}(s)\right] dG_t(s)\right]^{-1}$, the reciprocal of t's expected *inverse* marginal utility. And in the special case of money-loss utility, $\nu_t^{c,\pi} = \lambda_t^{c,\pi}$, in which case I refer to ν-transfers as "premium transfers" for obvious reasons. In this way, the $\tilde{\lambda}_t = \tilde{\lambda}_\tau$ result from Proposition 1 part ii is a special case of the next proposition, which is the main result of this subsection.

Let $\tilde{\nu}_t \equiv \tilde{\nu}_t^{\tilde{c}_t,\tilde{\pi}_t}$ and $\tilde{\beta}_t \equiv \beta_t^{\tilde{c}_t,\tilde{\pi}_t}$ for all t.

Proposition D.1. *For all t, suppose $\tilde{c}_t$ exists and $\tilde{c}_t \in C^2$.*

Then the following hold:

i) For general utility, if arbitrary nonlinear coinsurance is allowed, $\tilde{\nu}_t = \tilde{\nu}_\tau$ for all $t,\tau \in \mathcal{T}$.

ii) Under separable utility $\nu_t^{c,\pi} = \theta_t^{c,\pi}$ for all t and (c,π).

iii) Under money-loss utility, $\nu_t^{c,\pi} = \lambda_t^{c,\pi}$ for all t and (c,π).

The intuition for part i is that if $\tilde{\nu}_t \neq \tilde{\nu}_\tau$ for some $t,\tau \in \mathcal{T}$, we could use ν-transfers between t and τ to increase ex-ante expected utility in a budget-neutral way.

D.2 Private Risk: Wealth Effects (Linear)

In this section, I extend the results from Sections (5) and (6.2) to allow for wealth effects. Specifically, I generalize Lemma 3 as well as the main result Theorem 1 — which concerns

[68] Formally, $\underline{\beta}_t^{c,\pi} \equiv \left[\int L_{c,\pi,\underline{s}}(s) dG_t(s)\right]^{-1}$ so that $\int \beta_t^{c,\pi}\left(m(s)\right) dG_t(s) = 1$.

[69] The proof of part i uses methods from [Werquin et al. 2015].

only the FOA-relaxed problem — and its sister result Proposition 2 which concerns the full problem. Those original results, which assumed money-loss CARA, are subsumed in these more general ones. In particular, Theorem 1 and Proposition 2 are both subsumed in Theorem 1′; that result also gives some additional generality to address the possibility of pooling. Proofs will be given only for these more general results (not the ones in the main text) and will be given in Appendix G.5.

The next lemma generalizes Lemma 3 from the main text, again giving conditions for upward-adjacent IC to bind. As discussed in Footnote 37 of the main text, it is not needed to prove Theorem 1 or any generalization showing upward distortion in the *FOA-relaxed* problem. Instead it is only needed to show distortion in the *full* problem as in Proposition 2.[70] Its role in the latter — and thus in Theorem 1′ below (which, again, subsumes both) — is only to show that $(t+1)$'s downward-adjacent IC doesn't bind when $x_t < x_{t+1}$. To show upward distortion in the FOA-relaxed problem (à la Theorem 1), we only need to show that budget-neutral premium transfers from lower to higher types will be strictly beneficial. For this, we need only strict increasingness of λ_t^* (under money-loss CARA), or of $\lambda_t^* \left[\frac{\partial}{\partial \pi} P_t\left(x_t^*, \pi_t^*\right)\right]^{-1}$ under general utility. We've shown the former condition under money-loss CARA (Lem. 2), and the latter condition will be taken as hypothesis, though in a generalized form that accounts for (possible) pooling.

Lemma 3′ (Generalized Upward IC Binds)**.** *Let utility be general. Suppose $\left(x_t^{**}, \pi_t^{**}\right)_{t=1}^T$ is an optimal solution to the private-risk problem with $x_t^{**} \geq 0$ for all t. Let $\mathcal{P}$ be the set of pools, i.e., the set of equivalence classes of types defined by the relation $t \sim t'$ iff $x_t^{**} = x_{t'}^{**}$.*

Suppose either utility is money-loss CARA, or $\left[\sum_{t \in p} f_t \lambda_t^{**}\right] \cdot \left[\sum_{t \in p} f_t \frac{\partial}{\partial \pi} P_t\left(x_t^{**}, \pi_t^{**}\right)\right]^{-1}$ is strictly increasing in p over the set of pools $\mathcal{P}$ and $x_t^{**} \leq 1$ for all t. Then all upward-adjacent ICs bind at $\left(x_t^*, \pi_t^*\right)_{t=1}^T$: $V_{t-1}\left(x_{t-1}^*, \pi_{t-1}^*\right) = V_{t-1}\left(x_t^*, \pi_t^*\right)$.

Further, the same result holds if $\left(x_t^{**}, \pi_t^{**}\right)_{t=1}^T$ is replaced by any optimal solution $\left(x_t^*, \pi_t^*\right)_{t=1}^T$ to the FOA-relaxed problem.[71]

Finally, in the nonlinear coinsurance case, upward-adjacent IC also binds at the FOA-relaxed solution under money-loss CARA, as asserted in Lemma 3.[72]

As foreshadowed before the statement just above, the main difference from the money-loss

[70] Since both of these (in particular the latter) are subsumed in Theorem 1′ below, it (i.e. Lem. 3′) qualifies as a lemma supporting that result.

[71] This replacement includes replacing λ_t^{**} with λ_t^*. It also includes adjusting $\mathcal{P}$ to be defined via the relation $t \sim t'$ iff $x_t^* = x_{t'}^*$ (rather than $x_t^{**} = x_{t'}^{**}$).

[72] For nonlinear coinsurance, I have not extended this result to the full problem since with nonlinear, we lack the single crossing property established for linear in Lem. 5. However, the proof shows it does extend to the intermediate problem that only ignores non-adjacent upward ICs while still considering all downward ICs.

CARA result (Lemma 3) is the additional hypothesis that $\left[\sum_{t\in p} f_t\lambda_t^*\right]\cdot\left[\sum_{t\in p} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^*,\pi_t^*\right)\right]^{-1}$ is strictly increasing in p over the set of pools $\mathcal{P}$.[73] That quantity is the amount of money by which we can decrease the premium in pool p for the cost of spending one extra dollar from the budget. Thus, the condition says that budget-neutral premium transfers from t to $t+1$ would be ex-ante welfare-enhancing (at the optimum). Under money-loss (CARA or not), this condition reduces to strict increasingness of λ_t^* (since then $\frac{\partial}{\partial\pi}P(x,\pi)=1$), which holds by Lemma 2. However, outside money-loss, premium transfers affect treatment choice and thus the budget, so the amount of the transfer must be adjusted by the $[\frac{\partial}{\partial\pi}P]^{-1}$-factors to ensure budget neutrality.

Now for the extended version of the main result; it has three parts. Part i generalizes Theorem 1 about the FOA-relaxed solution, while part ii generalizes Proposition 2 about the solution to the full problem. Part iii is novel relative to the main text. Like part ii, it addresses the fully optimal solution, but shows a distortion relative to the pooled distribution $G_p(s)\equiv\sum_{t\in p}\frac{f_t}{\sum_{\tau\in p}f_\tau}G_t(s)$.

Theorem 1′ (Generalized Main Result)**.** *Let utility be general. Suppose that the static problem is strictly concave for every t and S.*

Suppose $(x_t^*,\pi_t^*)_{t=1}^T$ is an optimal solution to either the FOA-relaxed problem or the problem with only upward ICs (including non-adjacent ones) with $x_t\geq 0$.

i) Suppose that either utility is money-loss CARA, or $\lambda_t^{x_t^*,\pi_t^*}\left[\frac{\partial}{\partial\pi}P_t\left(x_t^*,\pi_t^*\right)\right]^{-1}$ is strictly increasing in t and $x_t^*\leq 1$ for all t.[74] *Then $x_t^*>\hat{x}_{tS_t^*}$ for all $t>1$.*

For parts ii and iii, suppose $(x_t^{**},\pi_t^{**})_{t=1}^T$ is an optimal solution to the full private-risk problem with $x_t^{**}\leq 1$ for all t. Let $\mathcal{P}$ be the set of pools — i.e., the set of equivalence classes of types defined by the relation $t\sim t'$ iff $x_t^{**}=x_{t'}^{**}$. Since all pools consist of an interval of risk-types, let $\mathcal{P}$ inherit the order from the set of risk-types $\mathcal{T}$. For all pools $p\in\mathcal{P}$, let t_p^+ denote the highest risk-type in p.

ii) Suppose that either utility is money-loss CARA or the following hypotheses hold: $\left[\sum_{t\in p} f_t\lambda_t^{**}\right]\cdot\left[\sum_{t\in p} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^{**},\pi_t^{**}\right)\right]^{-1}$ is strictly increasing in p over the set of pools $\mathcal{P}$, $x_t^{**}\leq 1$ for all t,[75] *and for all $p\in\mathcal{P}$, $\left[\sum_{t\in p\setminus\{t_p^+\}} f_t\lambda_t^{**}\right]\cdot\left[\sum_{t\in p\setminus\{t_p^+\}} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^{**},\pi_t^{**}\right)\right]^{-1}<$*

[73] The other added hypothesis is $x_t^*\leq 1$. Outside money-loss utility, this condition is needed to invoke Lemmas 5 and 6, which is the only place it will be used. Technically, it is not needed for non-CARA money-loss, though the result still assumes it there.

[74] Under money-loss CARA, the increasingness condition holds by Lem. 2 (and by $\frac{\partial}{\partial\pi}P_t=1$) at the optimal solutions to both the FOA-relaxed and full problems. The latter condition ($x_t^{**}\leq 1$) is not needed under money-loss in order to apply Lemmas 5 and 6, which are the only places it is otherwise used.

[75] As mentioned above, $\left[\sum_{t\in p} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^*,\pi_t^*\right)\right]^{-1}$ is the amount of money by which we can decrease the premium in pool p for the cost of spending one extra dollar from the budget.

$\lambda^{**}_{t_p^+}\left[\frac{\partial}{\partial\pi}P_{t_p^+}\left(x^{**}_{t_p^+},\pi^{**}_{t_p^+}\right)\right]^{-1}$. *[Note that a sufficient condition for the latter is that* $|\frac{\partial}{\partial\pi}m_{x_t^{**},\pi_t^{**}}(s)|$ *is weakly decreasing in* s*.] Then* $x^{**}_{t_p^+} > \hat{x}_{t_p^+ S^{**}_{t_p^+}}$ *for all* $p \in \mathcal{P}$ *such that* $t_p^+ > 1$.

iii) Suppose that either utility is money-loss CARA or the following hypotheses hold: $\left[\sum_{t\in p} f_t\lambda_t^{**}\right]\cdot\left[\sum_{t\in p} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^{**},\pi_t^{**}\right)\right]^{-1}$ *is strictly increasing in* p *over the set of pools* $\mathcal{P}$ *and* $x_t^{**} \leq 1$ *for all* t*. Then* $x_p^{**} > \hat{x}_{pS_p^{**}}$ *for all* $p \in \mathcal{P}$ *satisfying* $p \cap \{1\} = \emptyset$*, where* $(x_p^{**},\pi_p^{**}) \equiv (x_t^{**},\pi_t^{**})$ *for all* $t \in p$*,* $S_p^{**} \equiv -\sum_{t\in p}\frac{f_t}{\sum_{\tau\in p} f_\tau}P_t(x_t^{**},\pi_t^{**})$*, and* $\hat{x}_{pS_p^{**}}$ *is defined as follows.* $\hat{x}_{pS_p^{**}}$ *denotes the unique solution to the static problem (4.4) with distribution* $G_p(s) \equiv \sum_{t\in p}\frac{f_t}{\sum_{\tau\in p} f_\tau}G_t(s)$ *and subsidy* S_p^{**}*.*[76]

Relative to Theorem 1, the key difference with part i of the above result is the increasingness condition on $\lambda_t^{x_t,\pi_t}\left[\frac{\partial}{\partial\pi}P_t\left(x_t,\pi_t\right)\right]^{-1}$ which under money-loss reduces to λ_t^{x,π_t} and holds automatically under money-loss CARA by Lemma 2. This condition guarantees that upward, budget-neutral premium transfers will increase ex-ante expected utility. Similarly, relative to Proposition 2, part ii of the present result adds hypotheses that play a similar role but account for the possibility of pooling. Specifically, they ensure welfare-increasingness of a transfer where the transfer may be from one entire pool of types p to the next pool $p+1$, or (in part ii only) from one pool excluding its highest member type (i.e., $p\backslash\{t_p^+\}$) to that highest member type t_p^+.

E Cross-subsidization

The two results in this section are proved in Appendix G.6. The first result supports the second, but may be of independent interest.

E.1 IC Without Cross-subsidization

If there is no cross-subsidization so that each risk-type gets an actuarially-fair contract, lower-risk types will get more attractive contracts, all else equal. To develop this point further, suppose L and H are risk-types with H being higher risk. Since $(\hat{c}_L,\hat{\pi}_L)$ is also more statically efficient for L than is $(\hat{c}_H,\hat{\pi}_H)$ by definition, the following result is not surprising. [Note it doesn't assume $T=2$.]

Fact E.1. *Let* $L, < H \in \mathcal{T}$*.* $V_L(\hat{c}_L,\hat{\pi}_L) > V_L(\hat{c}_H,\hat{\pi}_H)$.

Effectively, this result says that, absent cross-subsidization, *upward* IC will not bind. On the other hand, *downward* IC typically will bind without cross-subsidization, unless $\hat{c}_L$ is

[76] To see why $\hat{x}_{pS_p^{**}}$ is unique, note that since the static problem is assumed concave for all t, it must also be so for G_p since its object and constraints are convex combinations of those from each t's static problem.

sufficiently lower-coverage than $\hat{c}_H$ relative to the difference in risk levels, L and H.[77] I leave the latter as an informal claim. As mentioned in the main text, the idea that downward (not upward) IC binds under actuarially-fair contracts is familiar from the static adverse selection literature [Rothschild and Stiglitz 1976]. My analysis here is novel only in that it allows for moral hazard.

E.2 Cross-Subsidization Occurs: $S_H^* > 0 > S_L^*$

Given the dynamic insurance motive, it is natural to expect that subsidy S_t^* will be strictly increasing in risk-type t, moving from strictly negative to strictly positive. When there are just two risk-types, the next result shows this is true for money-loss CARA utility and gives a sufficient condition under general utility. Naturally, the latter is the same condition used in Lem. 3′ to ensure upward IC binds under wealth effects. This result does assume only two risk-types, though I expect a T-type extension can be shown.[78]

Proposition E.1. *Suppose $T = 2$ so that the set of risk-types is $\{1, 2\}$, and relabel these as $L \equiv 1$ and $H \equiv 2$ so that $L < H$. Suppose $(x_t^*, \pi_t^*)_{t=L,H}$ is an optimal private-risk solution with $x_t^* \geq 0$ for all t. Suppose either utility is CARA money-loss, or $\lambda_H^* \left[\frac{\partial}{\partial \pi} P_H\left(x_H^*, \pi_H^*\right)\right]^{-1} > \lambda_L^* \left[\frac{\partial}{\partial \pi} P_L\left(x_L^*, \pi_L^*\right)\right]^{-1}$ and (outside money-loss) $x_L^*, x_H^* \leq 1$. Suppose also that there exist a unique solution $\hat{x}_{LS_t^*}$ to L's S_L^*-subsidized static problem (4.4). Then $S_H^* > 0 > S_L^*$.*

F FOA-Validity and Separation-vs-Pooling (Linear Contracts)

In this section, I discuss when the FOA is valid for linear contracts and connect this to the issue of separation-versus-pooling of risk-types. I first show that, roughly, FOA-validity will hold when the solution to the FOA-relaxed *dynamic* problem features increasing coverage in risk-type (Prop. F.1 ii). Although this echoes a similar result in [Kapička 2013], both results leave open what determines such increasingness. I identify increasingness of *statically* efficient coverage as such a determinant, thereby providing a novel intuition for what drives FOA-validity, which is slightly closer to primitives.

Although as we'll see, with linear contracts, FOA-validity essentially coincides with optimality of complete separation of risk-types (i.e. no pooling), these phenomena come apart

[77] To see why this latter condition may be needed, consider a case where statically efficient coverage level for L is very low but is high for H. [Note: it is unclear to me whether this is possible given $L \prec_{\text{MLRP}} H$, but I haven't ruled it out.] In that case, the lower premium associated with L's contract $(\hat{c}_L, \hat{\pi}_L)$ will not be worth it for H since she strongly values high coverage.

[78] Namely, that S_t^* is strictly increasing in t and spans both sides of zero. Similarly, for $\tilde{S}_t$.

when linearity is not imposed. With nonlinear contracts, whereas I strongly conjecture that pooling of risk-types will never be optimal (or at the very least, this will certainly be true for mathematically generic parameter values), the FOA may have scope to fail. Thus, FOA-validity is a conceptually more important question than separation-versus-pooling.

In what follows, I use the terms "(FOA-)relaxed-optimal" and "fully optimal" (or "full-optimal") to refer to optimal solutions to the FOA-relaxed and full problems respectively. Analogous terms apply to feasibility.

The next proposition describes the relation between FOA-validity, monotonicity (i.e. weak increasingness) in the relaxed-optimal solution, and separation-versus-pooling in the full-optimal solution. Roughly speaking, it shows these are all equivalent up to non-generic exceptions. Its proof is based on single crossing (Lem. 5) and the monotonicity restriction that imposes on the full problem (Lem. 6). The proposition is much more precise.

Proposition F.1 (FOA-Validity)**.** *Let utility be general. Suppose $(x_t, \pi_t)_{t=1}^T$ is feasible in the FOA-relaxed problem (henceforth just "relaxed problem"). If utility is not money-loss CARA, suppose further that all upward-adjacent ICs bind*[79] *and that $x_t \leq 1$ for all t.*

i) If $(x_t)_t$ is weakly increasing (in t) and $\pi_t = \pi_{t+1}$ whenever $x_t = x_{t+1}$, then $(x_t, \pi_t)_{t=1}^T$ is feasible in the full problem. If further, $x_t < x_{t+1}$ for some t, then all downward ICs pointing to t hold with slack, i.e., $V_{t+i}(x_{t+i}, \pi_{t+i}) > V_{t+i}(x_t, \pi_t)$ for all $i \geq 1$. Additionally, if $(x_t)_t$ is strictly increasing (in t) throughout, then all extra ICs not imposed by the relaxed problem hold with slack at $(x_t, \pi_t)_{t=1}^T$.

ii) Suppose $(x_t, \pi_t)_{t=1}^T$ is also optimal in the relaxed problem. The FOA is valid if and only if both $(x_t)_t$ is weakly increasing and $\pi_t = \pi_{t+1}$ whenever $x_t = x_{t+1}$. Thus, strict increasingness of $(x_t)_t$ implies strict increasingness of $(x_t)_t$ is optimal in the full problem.

iii) Suppose the relaxed problem is strictly concave (by which I include the assumption that its feasible set is convex). Then the FOA will be valid if the following holds: there exists an optimal solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$ to the full problem with $x_t^{**} \leq 1$ (for all t) such that $(x_t^{**})_t$ is strictly increasing and all upward-adjacent ICs bind there.

iv) Again suppose the relaxed problem is strictly concave (by which I include the assumption that its feasible set is convex). If the optimal solution $(x_t^*, \pi_t^*)_{t=1}^T$ to the relaxed problem has $x_t^* > x_{t+1}^*$ for some t, then $x_\tau^{**} = x_{\tau+1}^{**}$ (and $\pi_\tau^{**} = \pi_{\tau+1}^{**}$) for some $\tau \in \mathcal{T}$ in any optimal solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$ to the full problem.

Part i shows that (strict) monotonicity of $(x_t)_t$ in an FOA-feasible solution implies all additional ICs hold (with slack). Part ii then easily follows by containment of feasible sets,

[79] This is guaranteed for money-loss CARA by Lem. 2.

asserting essentially that monotonicity of $(x_t)_t$ in the relaxed solution is equivalent to FOA-validity.[80] As a corollary, if the monotonicity in the relaxed solution is strict rather than weak — as will generically be the case — separation is fully optimal. Part iii says (by contrapositive): if the FOA fails, pooling will occur, at least provided the relaxed problem is strictly concave. Its converse essentially holds (but for a pathological exception): if some pooling is full-optimal, the only way the FOA could not fail would be the *non-generic* event that pooling happened to be optimal in the relaxed problem — by a pure coincidence — despite strictly decreasing coverage ($x_{t+1} < x_t$) being implementable unlike in the full problem (Lem. 6).[81] Part iv, which follows from part iii, says that if $x_t^* > x_{t+1}^*$ in the relaxed problem for some t, then we will have pooling in the full solution between some pair of consecutive types. It seems very likely that that pair would be t and $t+1$, but technically I haven't established that.

As previewed above, although Proposition F.1 ii gives a sufficient condition for FOA-validity, namely monotonicity of coverage in the relaxed problem, it leaves open what drives that monotonicity. The answer is two factors. First, there is the (monotonicity of the) *statically* efficient coverage $(\hat{x}_t)_t$ — or $(\hat{x}_{tS_t^*})_t$ with wealth effects[82] — towards which coverage will be partially tailored in the private-risk problem (relaxed or otherwise) and fully tailored under observable-risk.[83] Second, there's the *dynamic* force of Theorem 1 which distorts the relaxed-optimal coverage above $\hat{x}_t$ (or above $\hat{x}_{tS_t^*}$) for $t > 1$.

Intuitively, the first factor means that monotonicity of $(\hat{x}_t)_t$ (or $(\hat{x}_{tS_t^*})_t$) pushes towards monotonicity of x_t in the relaxed problem and thus (by Prop. F.1 ii) towards FOA-validity, and pushes harder the greater the difference $\hat{x}_t - \hat{x}_{t-1}$ (or $\hat{x}_{t,S_t^*} - \hat{x}_{t-1,S_{t-1}^*}$). In my view, this intuition is the main novel insight of this section, and it suggests the following conjecture: under money-loss CARA utility and strict concavity of all static problems, monotonicity of $\hat{x}_t$ implies monotonicity of x_t in the relaxed-optimal solution and thus (by Lem. F.1 ii) FOA-validity. The difficulty in establishing this conjecture is that, with $t > 2$, once the

[80] The direction of part ii (of Prop. F.1) that gives a sufficient condition for FOA-validity is closely related to [Kapička 2013]'s Lemma 12, which I adapt to my setting in Proposition H.1 i (Appendix H.3). The sufficient condition identified by the latter results is monotonicity of *marginal rents* in the relaxed solution — i.e. $V_t(c_\tau, \pi_\tau) - V_{t-1}(c_\tau, \pi_\tau)$ being increasing in τ for $t, \tau \in \mathcal{T}$ — whereas Prop. F.1 ii shows the same for monotonicity of linear *coverage*. As I show in Proposition H.1 ii, the latter condition essentially implies the former: $V_t(x_\tau, \pi_\tau) - V_{t-1}(x_\tau, \pi_\tau)$ will be increasing in τ if $(x_\tau)_\tau$ (and (π_τ)) are. Note also that the "marginal rent" approach is more general in that it allows for nonlinear contracts.

[81] A reasonable alternative definition of FOA-validity would be that at the full optimum, the extra IC constraints all hold with slack. The first assertion of Prop. F.1 iii can be thought of as essentially saying that this alternative definition implies the original, as long as the relaxed problem is strictly concave. The converse will hold generically, for the reasons just noted before this footnote.

[82] Technically, I should write $\hat{x}_{tS_t}$ (rather than $\hat{x}_{tS_t^*}$) where S_t is the endogenous subsidy to t at the relaxed-optimal solution (rather than full-optimal solution).

[83] In the latter, it's $\hat{x}_{t\tilde{S}_t}$ not $\hat{x}_{tS_t^*}$.

dynamic force has distorted x_t marginally above $\hat{x}_t$ and x_{t-1} marginally above $\hat{x}_{t-1}$, it's not inconceivable that the dynamic force could eventually push x_{t-1} up harder than it does x_t, thus resulting in $x_t < x_{t-1}$ in the relaxed solution. However, with just two risk-types (i.e. $T = 2$), the dynamic force distorts *only* x_2 upwards, thus reinforcing static monotonicity. As a result, when $T = 2$, we can show that monotonicity of statically efficient coverage guarantees FOA-validity and separation, as the next result shows for CARA.

Proposition F.2 (Static Monotonicity drives FOA-Validity, two types)**.** *Let $T = 2$. Suppose $(x_t, \pi_t)_{t=1,2}$ is an optimal solution to the FOA-relaxed problem. Suppose also that the static problem is strictly concave and $\hat{x}_{tS}$ exists for all t, S.*

i) Let utility be money-loss CARA. If $\hat{x}_1 \leq \hat{x}_2$, then $x_1 < x_2$ and the FOA is valid, so $(x_t, \pi_t)_{t=1}^T$ is optimal in the full problem.

ii) Now let utility be general and suppose $\lambda_t^{x_t,\pi_t} \left[\frac{\partial}{\partial \pi} P_t\left(x_t, \pi_t\right)\right]^{-1}$ is strictly increasing in t. Let $S_t \equiv -P_t(x_t, \pi_t)$ for all t. If $\hat{x}_{1S_1} \leq \hat{x}_{2S_2}$, then $x_1 < x_2$ and the FOA is valid so $(x_t, \pi_t)_{t=1}^T$ is optimal in the full problem.

But what, in turn, determines the *statically* efficient coverage $\hat{x}_t$ (or $\hat{x}_{tS_t^*}$)? Again, two things: the distribution G_t and, under wealth effects, the subsidy S_t^* (or S_t for the subsidy in the relaxed solution). I have not been able to fully understand the effect of G_t on $\hat{x}_t$.[84]

As for the effect of S_t^* on $\hat{x}_{tS_t^*}$ under wealth effects, this is discussed in Appendix C.2. How that channel impacts separation versus pooling (and hence FOA-validity), is also discussed briefly discussed there but here I note only the following. S_t^* clearly must take both positive and negative signs as t varies (given actuarial fairness (5.1c)), and it's natural to expect S_t^* will be strictly increasing in t. The latter was formally confirmed in the two-type case by Prop. E.1 in Appendix E.2, under the same hypotheses as other wealth-effects results.

[84] One factor is how much uncertainty G_t contains. Intuitively, the statically efficient level of coverage will be low for a risk-type who can predict their future severity with high precision (whether it be mild or severe), since there is little variance to insure but the moral hazard cost of coverage remains. However, this is of little help since when risk-types are ordered by MLRP, they will not be ordered in terms of how much uncertainty they contain, at least not in the sense of mean-preserving spreads.

G First set of Proofs

G.1 Supporting Lemmas and Definitions for Propositions 1-4 & Theorem 1

G.1.1 Care-Demand Lemmas

Fact G.1. *The following two conditions are each sufficient for existence of the maximizer $m_{c,\pi}(s)$ in (4.1) (for every s and π). i) c is weakly increasing and upper semi-continuous. ii) c is Lipschitz continuous and the health function $h(\cdot,\cdot)$ is such that $h_1(m,s) \to -\infty$ as $m \to \infty$.*

Proof. Fix s and π.

i) Since $\lim_{m\to\infty} h_1(m,s) < 0$ (by standing assumption B3 in the model Section 3), we can take $\hat{m}$ sufficiently large that for all $m \geq \hat{m}$, $h_1(m,s) < 0$. Since c, h, and U are upper semi-continuous, the following maximum exists $D \equiv \max_{m\in[0,\hat{m}]} U(-\pi - c(m), h(m,s))$ by a standard result. It then suffices to show that no $m > \hat{m}$ leads to higher utility than D. For all $m > \hat{m}$,

$$U(-\pi - c(m), h(m,s)) < U(-\pi - c(m), h(\hat{m},s)) \leq U(-\pi - c(\hat{m}), h(\hat{m},s)) \leq D,$$

where first equality follows by $U_2 > 0$ and $h_1(m,s) < 0$ for all $m \geq \hat{m}$; the second inequality follows by the assumption that $c(m)$ is weakly increasing; the third inequality by definition of D. This completes the proof.

ii) Suppose for contradiction that $\max_{m\geq 0} U(-\pi - c(m), h(m,s))$ does not exist. Then there is a sequence $\left(m^{(i)}\right)_{i=1}^{\infty}$ such that $\left(U\left(-\pi - c\left(m^{(i)}\right), h\left(m^{(i)},s\right)\right)\right)_{i=1}^{\infty}$ is strictly increasing and converges to ∞. Since $U(-\pi - c(m), h(m,s))$ is continuous in m (since c is Lipschitz), $U(-\pi - c(m), h(m,s))$ cannot be unbounded on a compact interval, so we must have $\left(m^{(i)}\right)_{i=1}^{\infty} \to \infty$. Taking a subsequence if necessary, it is without loss to assume $\left(m^{(i)}\right)_{i=1}^{\infty}$ is strictly increasing.

Recall the standing assumption that, for all m, s, $h_{11} < 0$ and $h_1(m,s) \to -\infty$ as $m \to \infty$. Accordingly, let $\hat{m}$ be such that for all $m \geq \hat{m}$, $h_1(m,s) < -\varepsilon$ for some ε. Passing to a subsequence, it is again without loss to assume each of $\left(m^{(i)}\right)_{i=1}^{\infty}$ are greater than $\hat{m}$. Since $U(y,h)$ is strictly increasing in both arguments, after taking subsequences it is without loss to assume $\left(c(m^{(i)})\right)_{i=1}^{\infty}$ is strictly decreasing since otherwise $\left(U\left(-\pi - c\left(m^{(i)}\right), h\left(m^{(i)},s\right)\right)\right)_{i=1}^{\infty}$ would be decreasing given that $\left(m^{(i)}\right)_{i=1}^{\infty}$ are greater than $\hat{m}$. We have now shown that in the sequence $\left(U\left(-\pi - c\left(m^{(i)}\right), h\left(m^{(i)},s\right)\right)\right)_{i=1}^{\infty}$, the first argument $-\pi - c\left(m^{(i)}\right)$ is strictly increasing and the second argument $h\left(m^{(i)},s\right)$ is strictly decreasing.

By our standing assumption that MRS is weakly decreasing in y and weakly increasing in h, this implies $\text{MRS}(m) \equiv \frac{U_1(-\pi - c(m), h(m,s))}{U_2(-\pi - c(m), h(m,s))}$ is weakly decreasing. Let $M > 0$ be the Lipschitz constant for c. Let $N \equiv \frac{U_1\left(-\pi - c\left(m^{(1)}\right), h\left(m^{(1)}, s\right)\right)}{U_2\left(-\pi - c\left(m^{(1)}\right), h\left(m^{(1)}, s\right)\right)}$; then $\text{MRS}(m) \leq N$ for all $m \geq m^{(1)}$, and recall $m^{(i)} \geq m^{(1)}$ for all i. Since $h_1(m,s) \to -\infty$ as $m \to \infty$ (by assumption), choose $\tilde{m} \geq \hat{m}$ such that $h_1(m,s) < -NM$ for all $m > \tilde{m}$. Passing to a subsequence again, assume without loss that $\left(m^{(i)}\right)_{i=1}^{\infty} > \tilde{m}$, where N remains fixed even as we take subsequences that change what we call $m^{(1)}$. Note that for all $m \geq \tilde{m}$

Then, for each i,

$$U\left(-\pi - c\left(m^{(i+1)}\right), h\left(m^{(i+1)}, s\right)\right) - U\left(-\pi - c\left(m^{(i)}\right), h\left(m^{(i)}, s\right)\right) \tag{G.1}$$

$$= \int_{m^{(i)}}^{m^{(i+1)}} \left[U_1\left(-\pi - c(m), h(m,s)\right)\left(-c'(m)\right) + U_2\left(-\pi - c(m), h(m,s)\right) h_1(m)\right] dm \tag{G.2}$$

$$\leq \int_{m^{(i)}}^{m^{(i+1)}} \left[U_1\left(-\pi - c(m), h(m,s)\right) M + U_2\left(-\pi - c(m), h(m,s)\right)(-NM)\right] dm \tag{G.3}$$

$$= \int_{m^{(i)}}^{m^{(i+1)}} \left\{U_2\left(-\pi - c(m), h(m,s)\right)\left[\text{MRS}(m)M + (-NM)\right]\right\} dm \tag{G.4}$$

$$\leq \int_{m^{(i)}}^{m^{(i+1)}} \left\{U_2\left(-\pi - c(m), h(m,s)\right)\left[NM + (-NM)\right]\right\} dm \tag{G.5}$$

$$= 0. \tag{G.6}$$

where we use the fact that c is Lipschitz and thus differentiable almost everywhere so the fundamental theorem is valid. The first inequality is by the fundamental theorem of calculus. The second inequality is by the Lipschitz assumption on c and the definition of $\tilde{m}$. The last inequality is from $\text{MRS}(m) \leq N$ for all $m \geq m^{(1)}$ and the fact that $m^{(i)} \geq m^{(1)}$ for all i. Putting together the above chain of relations from (G.1) to (G.6) contradicts that $\left(U\left(-\pi - c\left(m^{(i)}\right), h\left(m^{(i)}, s\right)\right)\right)_{i=1}^{\infty}$ is increasing. This completes the proof. □

Definition 5. We say U satisfies *strictly* diminishing MRS if $\frac{U_1(y,h)}{U_2(y,h)}$ strictly decreasing in y and strictly increasing in h.[85]

Lemma G.1. *i) Let $c' \geq 0$ everywhere. Then for all π, s, when $m = m_{c,\pi}(s)$ (for any of $m_{c,\pi}(s)$ if there are multiple), $U_{11}c'(m) - U_{12}h_1(m,s) \leq 0$ and $-U_{21}c' + U_{22}h_1(m,s) \leq 0$, where the U terms are evaluated at $(-\pi - c(m), h(m,s))$, and the inequalities are strict when $c' > 0$ everywhere and MRS is strictly diminishing. [For applications, note $U_{12} = U_{21}$ since U is C^2.]*

[85] This can be seen to be equivalent to strict quasi-concavity of U.

ii) Let $c' \leq 0$ everywhere. Then for all π, s, when $m = m_{c,\pi}(s)$ (for any of $m_{c,\pi}(s)$ if there are multiple), $U_{11}c'(m) - U_{12}h_1(m,s) \geq 0$ and $-U_{21}c' + U_{22}h_1(m,s) \geq 0$, where the U terms are evaluated at $(-\pi - c(m), h(m,s))$, and the inequalities are strict when $c' > 0$ everywhere and MRS is strictly diminishing.

Proof. i) Since $c' \geq 0$, $h_1(m_{c,\pi}(s), s) \geq 0$ by the FOC for $m_{c,\pi}$, so evaluating $U_{11}c'(m) - U_{12}h_1(m,s)$ at $m = m_{c,\pi}(s)$ we see that it is directly proportional to $\left[U_{11}\frac{c'}{h_1} - U_{12}\right]$, unless $h_1(m_{c,\pi}(s)) = 0$ in which case $c'(m_{c,\pi}(s)) = 0$ in which case we're done. Then using the FOC again, the latter is equal to:

$$\begin{aligned} &\left[U_{11}\frac{U_2}{U_1} - U_{21}\right] \\ &= U_2\left[\frac{U_{11}}{U_1} - \frac{U_{21}}{U_2}\right] \\ \propto \left[\frac{U_{11}}{U_1} - \frac{U_{21}}{U_2}\right] &= \frac{\partial}{\partial y}\log\left[\frac{U_1(y,h)}{U_2(y,h)}\right] \leq (<)0. \end{aligned} \tag{G.7}$$

The proportionality step ($\propto$) is by $U_2 > 0$. The inequality is by the standing diminishing-MRS assumption and the fact that log is a strictly increasing transformation, where $\frac{\partial}{\partial y}\log\left[\frac{U_1(y,h)}{U_2(y,h)}\right]$ is evaluated at $(y,h) = (-\pi - c(m), h(m,s))$ with $m = m_{c,\pi}(s)$. Of course the inequality becomes strict if we strengthen the decreasing MRS assumption as indicated in the last part of the result statement. The argument for $-U_{21}c' + U_{22}h_1$ evaluated at $m_{c,\pi}(s)$ is similar: since $h_1 > 0$, it is directly proportional to $-U_{21}\frac{c'}{h_1} + U_{22}$. By the FOC, the latter equals

$$\begin{aligned} &-U_{21}\frac{U_2}{U_1} + U_{22} \\ = U_1\left[-\frac{U_{21}}{U_1} + \frac{U_{22}}{U_2}\right] \quad &\propto \left[-\frac{U_{21}}{U_1} + \frac{U_{22}}{U_2}\right] \\ = \frac{\partial}{\partial h}\log\left[\frac{U_2(y,h)}{U_1(y,h)}\right] &\leq (<)0, \end{aligned}$$

again by the standing decreasing-MRS assumption, where, again, the last expression is evaluated at $(y,h) = (-\pi - c(m), h(m,s))$ with $m = m_{c,\pi}(s)$. This completes the proof.

ii) Since $c' \leq 0$, $h_1(m_{c,\pi}(s), s) \leq 0$ by the FOC for $m_{c,\pi}$, so evaluating $[U_{11}c'(m) - U_{12}h_1(m,s)]$ at $m = m_{c,\pi}(s)$ we see that it is directly proportional to $-\left[U_{11}\frac{c'}{h_1} - U_{12}\right]$, unless $h_1(m_{c,\pi}(s)) = 0$ in which case $c'(m_{c,\pi}(s)) = 0$ in which case we're done. But $-\left[U_{11}\frac{c'}{h_1} - U_{12}\right]$ is weakly (strictly, under strictly diminishing MRS) positive by the argument from part i (see just before (G.7)). The argument that $-U_{21}c' + U_{22}h_1(m,s) \geq (>)0$ is similar and thus omitted. □

Fact G.2. *Let utility be general. If c is twice differentiable with $c' \geq 0$ and $c'' \geq 0$, then $m_{c,\pi}(s)$ is unique for all s, π.*

Proof. First note the FOC for $m_{c,\pi}(s)$ implies $h_1 > 0$ since $c' > 0$. To complete the proof, we show strict concavity by checking the second derivative:

$$\begin{aligned}\frac{\partial^2}{\partial m^2} U\left(-\pi - c(m), h\left(m, s\right)\right) &= \frac{\partial}{\partial m}\left\{-U_1\left(-\pi - c(m), h\left(m, s\right)\right) c' + U_2\left(-\pi - c(m), h\left(m, s\right)\right) h_1\right\} \\ &= U_{11}\left(c'\right)^2 - U_{12} c' h_1 - U_1 c'' - U_{21} h_1 + U_{22} h_1 + U_2 h_{11} \\ &= c'\left[U_{11} c' - U_{12} h_1\right] + h_1\left[-U_{21} c' + U_{22} h_1\right] - U_1 c'' + U_2 h_{11} \\ &< c'\left[U_{11} c' - U_{12} h_1\right] + h_1\left[-U_{21} c' + U_{22} h_1\right] \leq 0.\end{aligned} \tag{G.8}$$

The last inequality follows by Lemma G.1. □

Recall from Section 3 that for each $\pi \in \mathbb{R}$, Assumption B5 posits an $s_\pi^* < \bar{s}$ at which point the first unit of treatment is cost-effective even if the consumer fully internalizes the cost. Further, since $h_{12} > 0$ (Assumption B4), this is true for all $s \geq s_\pi^*$. The next result shows that this implies $m_{c,\pi}(s) > 0$ for $s \geq s_\pi^*$ as long as $c' \leq 1$. Note that for money-loss utility, s_π^* is independent of π and is thus just called s^*.

Lemma G.2. *Suppose $\pi \in \mathbb{R}$ and $c' \leq 1$ throughout. Let s_π^* be as defined in Assumption B5 in Section 3. If $s \geq s_\pi^*$, then $m_{c,\pi}(s) > 0$, so that $m_{c,\pi}(s)$ is an interior solution to the demand problem (4.1).*

Proof. First note that is without loss of generality to assume $c(0) = 0$ since non-zero $c(0)$ can simply be included as part of the premium. Further, by Assumption B5, $-U_1\left(-\pi, h(0, s_\pi^*)\right) + U_2\left(-\pi, h(0, s_\pi^*)\right) h_1(0, s_\pi^*) > 0$. From the demand objective (4.1), the derivative of utility in m at $m = 0$ is $-U_1\left(-\pi - c(0), h(0, s)\right) c'(0) + U_2\left(-\pi, h(0, s)\right) h_1(0, s)$, and this will be strictly positive at all $s \geq s_\pi^*$ for the following reason. First, it will be strictly positive at s_π^* since $c(0) = 0$ and $c' \leq 1$, and then $h_{12} > 0$ (Assumption B4) implies that the same is true a fortiori for $s > s_\pi^*$. Thus the marginal return to treatment under $s = 0$ is strictly positive at $m = 0$. This strictly positive derivative at $m = 0$ implies $m_{c,\pi}(s) > 0$ since the model always requires $m_{c,\pi}(s) \geq 0$. □

Remark 2. When restricting to linear coinsurance, I abuse notation as usual by writing $m_{x,\pi}(s)$ instead of $m_{c,\pi}(s)$ where the notation x functions as a stand-in for the function $c(m) = (1 - x)m$.

Lemma G.3. *Let (c,π) be a static contract. i) For money-loss utility, $\frac{\partial}{\partial\pi}m_{\pi,c}(s)=0$. Now let utility be general and $c\in C^1$. ii) $m_{\pi,c}(s)$ is C^1 in π. iii) If $c'\geq(\leq)0$, $\frac{\partial}{\partial\pi}m_{\pi,c}(s)\leq(\geq)0$. Further, if the inequalities $\frac{\partial}{\partial\pi}m_{\pi,c}(s)\leq(\geq)0$ can be made strict for $s\geq s^*_\pi$ when MRS is strictly diminishing (recall Def. 5), $c'>(<)0$, and $c'\leq 1$. iv) For linear contracts, $m_{x,\pi}(s)$ is C^1 in x. Further $\frac{\partial}{\partial x}m_{\pi,x}(s)>0$ for $s\geq s^*$ under money-loss for any $x\geq 0$, and $\frac{\partial}{\partial x}m_{\pi,x}(s)>(\geq)0$ for $s\geq s^*_\pi$ under general utility when $0\leq x<(\leq)1$.*

Proof. i) For money-loss, $\frac{\partial}{\partial\pi}m_{\pi,c}(s)=0$ is obvious since under money-loss, 4.1 reduces to $\arg\max_{m\geq 0}\left[-\pi-c\left(m\right)+h\left(m,s\right)\right]$.

Now consider the separable model. Taking the (m,x) cross-derivative of $u\left(-\pi-(1-x)m\right)+v\left(h(m,s)\right)$ is $u'-m(1-x)u''+v'h_{12}+v''h_1h_2>0$ for $x\leq 1$ given assumptions on u, v, and h. The (m,π) cross-derivative of $u\left(-\pi-c(m)\right)+v\left(h(m,s)\right)$ is $c'(m)u''\leq(<)0$ for $c'(m)\geq(>)0$ and $c'(m)u''\geq(>)0$ for $c'(m)\leq(<)0$.

The comparative statics for money-loss and separable then follow from standard monotone comparative statics results.

General Utility (π-comparative static):

ii) Since u and h are C^1 (indeed C^2, by a standing assumption), it follows from the implicit function theorem (IFT) that m is C^1 with respect to both π (even with nonlinear coinsurance when $c\in C^1$) and x, where the IFT is applied to the FOC for $m_{c,\pi}(s)$.

iii) We will show $\frac{\partial^2}{\partial\pi\partial m}U\left(-\pi-c\left(m\right),h\left(m,s\right)\right)<0$ when evaluated at $m=m_{c,\pi}(s)$. That $m_{c,\pi}(s)$ is weakly decreasing in π will then follow from standard monotone comparative statics results [Edlin and Shannon 1998] since $c\in C^1$. The strict decreasingness will follow as long as $m_{c,\pi}(s)$ is interior, i.e. as long as $m_{c,\pi}(s)>0$.

$$\begin{aligned}&\frac{\partial^2}{\partial\pi\partial m}U\left(-\pi-c(m),h\left(m,s\right)\right)\\&=\frac{\partial}{\partial\pi}\left\{-U_1\left(-\pi-c(m),h\left(m,s\right)\right)c'+U_2\left(-\pi-c(m),h\left(m,s\right)\right)h_1\right\}\\&=U_{11}c'-U_{21}h_1,\end{aligned}\tag{G.9}$$

suppressing the argument $(-\pi-c(m),h(m,s))$. When evaluated at $m=m_{c,\pi}(s)$, Lemma G.1 implies the last expression is weakly negative (positive) when $c'\geq 0$ ($c'\leq 0$) and strictly so when $c'>0$ ($c'<0$) and MRS is strictly diminishing. As previewed above, to establish the strict comparative static, we need to make sure — per [Edlin and Shannon 1998] Corollary 1 (of their Theorem 1 "Strict Monotonicity") — $m_{c,\pi}(s)>0$ so that $m_{c,\pi}(s)$ is an interior solution to the demand problem (4.1). For this it suffices by Lemma G.2 that $s\geq s^*_\pi$ and $c'\leq 1$, both of which are hypotheses.

iv) To show $m_{x,\pi}(s)$ is C^1 in x, see part ii.

x-Comparative Static: The (m, x) cross-derivative of $-\pi - (1-x)m + h(m,s)$ is $1 > 0$. The result then follows from [Edlin and Shannon 1998] Corollary 1

Now for general utility: Now we show $\frac{\partial^2}{\partial x \partial m} U\left(-\pi - (1-x)m, h\left(m,s\right)\right) > 0$ at $m = m_{c,\pi}(s)$ for $0 \leq x < 1$. That $m_{c,\pi}(s)$ is strictly increasing in π will then follow from the same strict monotonicity result just mentioned since $c \in C^1$.

$$\begin{aligned}
&\frac{\partial^2}{\partial x \partial m} U\left(-\pi - (1-x)m, h\left(m,s\right)\right) \\
&= \frac{\partial}{\partial x}\left\{-U_1\left(-\pi - (1-x)m, h\left(m,s\right)\right)(1-x) + U_2\left(-\pi - (1-x)(m), h\left(m,s\right)\right)h_1\right\} \\
&= -U_{11}m(1-x) + U_1 + U_{21}h_1 \cdot m \\
&> m\left[-U_{11}\underbrace{(1-x)}_{c'} + U_{21}h_1\right].
\end{aligned}$$

When evaluated at $m = m_{x,\pi}(s)$, the bracketed expression in the last line is, by Lemma G.1, weakly positive when $x \leq 1$(i.e. $c' \geq 0$) and strictly positive when $x < 1$ (i.e. $c' > 0$). To conclude the proof, it remains only to finish establishing the strict comparative static by establishing interiority of $m_{x,\pi}(s)$. Similarly to above, for this it suffices by Lemma G.2 that $s \geq s^*_\pi$ and $x \geq 0$ (i.e. $c' \leq 1$), both of which are hypotheses. □

Lemma G.4. *Let (c, π) be a static contract. i) Let utility be money-loss. $m_{c,\pi}(s)$ is weakly increasing in s, and if $c \in C^1$ and $s \geq s^*$ and $c' \leq 1$, then $\frac{\partial}{\partial s} m_{c,\pi}(s) > 0$ (and $m_{c,\pi}(s)$ is C^1 in s). ii) Let utility be separable. If c is weakly increasing, then $m_{c,\pi}(s)$ is weakly increasing. iii) Let utility be general and $c' \geq 0$ everywhere. Then $m_{c,\pi}(s)$ is weakly increasing in s. iv) If in addition, c is C^1, then $m_{c,\pi}(s)$ is C^1 in s and for $s \geq s^*_\pi$ and $c' \leq 1$, we have $\frac{\partial}{\partial s} m_{c,\pi}(s) > 0$.*

Proof. i) For money-loss utility, $-\pi - c(m) + h(m,s)$ clearly satisfies increasing differences in m and s since $h_{12} > 0$. Thus by standard monotone comparative statics results, it follows $m_{c,\pi}(s)$ is weakly increasing in s. If $c \in C^1$, $m_{c,\pi}(s)$ is C^1 in s by the implicit function theorem (see also part iii), and we can take (m,s) cross-derivative of $-\pi - c(m) + h(m,s)$ to get $h_{12} > 0$. Then $\frac{\partial}{\partial s} m_{c,\pi}(s) > 0$ by [Edlin and Shannon 1998] Corollary 1 (of their Theorem 1 "Strict Monotonicity"). To apply this result, we need to check that $m_{c,\pi}(s)$ is interior to its domain, i.e., that $m_{c,\pi}(s) > 0$. It is so by Lemma G.2 since $s \geq s^*$ and $c' \leq 1$.

ii) Let utility be separable, and suppose c is weakly increasing. Then it follows from optimality of $m_{c,\pi}(s)$ that $h_1\left(m_{c,\pi}(s), s\right) \geq 0$. By standard monotone comparative statics results, to show $m_{\pi,c}(s)$ is weakly increasing in s, it suffices to show $u\left(-\pi - c(m)\right) + v\left(h(m,s)\right)$ satisfies weakly increasing differences in m and s. Since $u(-\pi - c(m))$ doesn't depend on

s, we just need to show $v\left(h(m,s)\right)$ satisfies weakly increasing differences. By a standard result, for this, it suffices to show the cross derivative is weakly positive. Computing this gives $\frac{\partial^2}{\partial m \partial s} v\left(h(m,s)\right) = v'h_{12} + v''h_1h_2 > 0$ since $h_1 \geq 0$, as desired.

iii) Let utility be general and $c' \geq 0$. We show $\frac{\partial}{\partial s}\frac{\partial}{\partial m} U\left(-\pi - c(m), h(m,s)\right)|_{m=m_{c,\pi}(s)} \geq 0$, which will imply $m_{c,\pi}(s)$ is weakly increasing (again) by standard monotone comparative statics results. Actually we will show the inequality is strict, but we won't need that until part iv. We now suppress the argument $(-\pi - c(m), h(m,s))$.

$$\frac{\partial^2}{\partial s \partial m} U\left(-\pi - c(m), h(m,s)\right) = \tag{G.10}$$

$$= -U_{12}c'h_2 + U_{22}h_1h_2 + U_2h_{12}$$

$$= h_2\left[-U_{12}c' + U_{22}h_1\right] + \underbrace{U_2h_{12}}_{>0} \tag{G.11}$$

Since $h_2 < 0$ and $U_2h_{12} > 0$, to show (G.10) is strictly positive, it suffices to show $\left[-U_{12}c' + U_{22}h_1\right] \leq 0$ when evaluated at $m = m_{c,\pi}(s)$. This follows from Lemma G.1 since $c' \geq 0$.

iv) That $\frac{\partial}{\partial s} m_{\pi,c}(s)$ exists and is continuous when c is C^1 follows from the implicit function theorem and the definition of $m_{c,\pi}(s)$ (4.1), given that U and h are C^1 (indeed C^2, by a standing assumption). Since the objective is C^1 and (G.10) is strictly positive, [Edlin and Shannon 1998] Corollary 1 (again) implies $\frac{\partial}{\partial s} m_{c,\pi}(s) > 0$ as long as $m_{c,\pi}(s)$ is an interior solution to the demand problem (4.1), i.e., that $m_{c,\pi}(s) > 0$. As in part ii, for that, it suffices by Lemma G.2 that $c' \leq 1$ and $s \geq s^*_\pi$, both of which are hypotheses. □

Definition 6. Let

$$W^{c,\pi}(s) \equiv U\left(-\pi - c\left(m_{c,\pi}(s)\right), h\left(m_{c,\pi}(s), s\right)\right), \tag{G.12}$$

the endogenous utility at contract (c,π) and severity s, and let $W_1^{c,\pi}(s)$ denote marginal utility of consumption at the same argument (U_1 rather than U). [This means we can write the notations V_t and λ_t more compactly: $V_t(c,\pi) = \int W^{c,\pi}(s) dG_t(s)$ and $\lambda_t^{c,\pi} = \int W_1^{c,\pi}(s) dG_t(s)$.]

Lemma G.5. *Fix a static contract (c,π). i) $W^{c,\pi}(s)$ is strictly decreasing in s. ii) If U is money-loss, $W_1^{c,\pi}(s)$ is strictly increasing in s. iii) Suppose $U_{12} \leq 0$, $c \in C^1$, and that both $c' > (\geq) 0$ and $h\left(m_{c,\pi}(s), s\right)$ is weakly (strictly) decreasing in s; the parentheticals indicate an alternate pair of hypotheses. Then $W_1^{c,\pi}(s) \equiv U_1\left(-\pi - c\left(m_{c,\pi}(s)\right), h\left(m_{c,\pi}(s), s\right)\right)$ is strictly increasing in s.*

Proof. i) Let $s' > s$. Then

$$\begin{aligned} W^{c,\pi}(s) &\equiv U\left(-\pi - c\left(m_{c,\pi}(s)\right), h\left(m_{c,\pi}(s), s\right)\right) \\ &\geq U\left(-\pi - c\left(m_{c,\pi}(s')\right), h\left(m_{c,\pi}(s'), s\right)\right) \\ &> U\left(-\pi - c\left(m_{c,\pi}(s')\right), h\left(m_{c,\pi}(s'), s'\right)\right) \equiv W^{c,\pi}(s'). \end{aligned}$$

The first and last equalities are by definition of $W^{c,\pi}$. The weak inequality is by definition of $m_{c,\pi}(s)$. The strict inequality is by the standing assumption $h_2(m, s) < 0$ (Assumption B3).

ii) Part ii then follows from the fact that under money-loss, for all $x, y \in \mathbb{R}$, $u(x) > u(y)$ iff $u'(x) < u'(y)$ since u is strictly increasing and strictly concave.

iii) Again, let $s' > s$. Then

$$\begin{aligned} W_1^{c,\pi}(s) &\equiv U_1\left(-\pi - c\left(m_{c,\pi}(s)\right), h\left(m_{c,\pi}(s), s\right)\right) \\ &\leq U_1\left(-\pi - c\left(m_{c,\pi}(s')\right), h\left(m_{c,\pi}(s'), s'\right)\right) \equiv W_1^{c,\pi}(s'), \end{aligned}$$

where the inequality holds because on the one hand, $U_{11} < 0$ and $U_{12} \leq 0$, and on the other, the first argument of U_1 has strictly (weakly) decreased (from s to s') while the second argument of U_1 has weakly (strictly) decreased as well. For the first argument, this is because because $c' > (\geq) 0$ and because $m_{c,\pi}(s)$ is strictly increasing in s by Lemma G.4 iv, which applies since $c \in C^1$ and (at least) $c' \geq 0$. For the second argument, this is by assumption, whether weak or strict. □

G.1.2 Reformulating the static problem over coinsurance only

The goal of this subsection is to reformulate the static problem (4.4) as an unconstrained maximization problem over c only (not π), thus dispensing with the budget constraint (4.5) and the decision variable π; the result will be the problem (G.14) below.

Under the money-loss model, we saw that $m_{c,\pi}(s) = m_c(s)$ so c fully determines the expected insurer liability: $k_t(c, \pi) = \int \left[m_c(s) - c\left(m_c(s)\right)\right] dG_t(s)$, which is then the actuarially-fair premium. However, for more general utility, $m_{c,\pi}(s)$ does depend on π — in particular it's decreasing in π (Lemma G.3) — so (for any subsidy S) we must *implicitly* define the premium $\pi_{t,c,S}$ which makes $(c, \pi_{t,c,S})$ actuarially fair for t given S.

Unpacking P_t, we see that $P_t(c, \pi) = -S$ can be viewed as a "fixed-point equation" in π when c is fixed. We use this equation to define $\pi_{t,c,S}$. [For results in this section, I fix an arbitrary risk-type t and subsidy S, thus avoiding the need to repeatedly quantify there-over.]

Definition 7. Let $\pi_{t,c,S}$ be the unique solution to the fixed point equation $P_t(c,\pi) = -S$ in π, and let $\pi_{tc} \equiv \pi_{tc0}$. In the linear case, we write $\pi_{t,x,S}$ when $c(m) = (1-x)m$.

[I sometimes omit commas in the subscript of $\pi_{t,c,S}$ or $\pi_{t,x,S}$.] The next lemma shows that $\pi_{t,c,S}$ is well-defined.

Lemma G.6. *Let c be C^1. If either utility is money-loss or $c' \geq 0$, there is a unique premium $\pi_{t,c,S}$ that satisfies $P_t(c,\pi_{t,c,S}) = -S$. The static contract $(c,\pi_{t,c,S})$ satisfies the budget constraint (4.5) of the S-subsidized static problem. Further, for money-loss utility, the above holds for arbitrary c for which demand is finite at every s.*

Proof. For the money-loss model this is obvious and already shown in the discussion right above the statement of the lemma in the main text.

We want to define $\pi_{t,c,S}$ as the unique solution π (for fixed c, t, S) to $P_t(c,\pi) = -S$. By Lemma G.9 below (without circularity), if $c \in C^1$, we have $\frac{\partial}{\partial \pi} P_t(c,\pi) > 1$ as long as $c' \geq 0$ or utility is money-loss, thus establishing uniqueness of any solution $\pi_{t,c,S}$ under the hypothesis of this result. To show existence, we use the intermediate value theorem (IVT). $\frac{\partial}{\partial \pi} P_t(c,\pi) > 1$ implies $P_t(c,\pi)+S$ clearly takes positive and negative signs for large and small enough values of π. Since $P_t(c,\cdot)$ is differentiable, the IVT applies, thus showing existence of the solution $\pi_{t,c,S}$. This completes the proof. □

Note $\pi_{t,c,S}$ is a function of $c(\cdot)$, as well as S and t.

We now use Lemma G.6 to reformulate the static problem (4.4) in terms of only one decision variable $c(\cdot)$, which then determines first π_{tcS} and finally $m_{c,\pi_{tcS}}(s)$ in turn. This allows us to dispense with the budget constraint, thus leading to an unconstrained optimization problem.

Definition 8. Let $\alpha_{tS} : \mathcal{A} \to \mathbb{R}$ be as follows:

$$\alpha_{tS}(c) = V_t(c, \pi_{tcS}), \tag{G.13}$$

where π_{tcS} is the function of c just defined in Lemma G.6. [As usual, I abuse notation by letting $\alpha_{tS}(x) \equiv V_t(x, \pi_{txS})$.] I refer to

$$\max_{c \in \mathcal{A}} \alpha_{tS}(c) \tag{G.14}$$

as t's *unconstrained* (S-subsidized) static problem.

By Lemma G.10 $\frac{\partial}{\partial \pi} V_t(c,\pi) < 0$. On the other hand, $\frac{\partial}{\partial \pi} P_t(c,\pi) > 1 > 0$ as we will later show (without circularity) in Lemma G.9 part iii (slightly below), when utility is money-loss or when we have the conjunction of $c \in C^1$, $c' \geq 0$, and $c' \leq 1$. It follows that the budget

constraint inequality $P_t(c,\pi) \geq -S$ is binding at the optimal solution to the static problem (4.4). Accordingly, it follows from Lemma G.6 (above) that when utility is money-loss or the admissible set $\mathcal{A}$ requires C^1, $c' \geq 0$, and $c' \leq 1$, the static problem (4.4) can be equivalently reformulated as the unconstrained static problem (G.14). I have now justified the following Lemma which follows from Lemma G.6. No further proof is needed.

Lemma G.7. *Suppose either utility is money-loss or $\mathcal{A}$ is contained in the set of C^1 functions c with $c' \geq 0$ and $c'(\cdot) \leq 1$. The following holds: $\hat{c}_{tS} \in \arg\max_{c\in\mathcal{A}} \alpha_{tS}(c)$, $\hat{\pi}_{tS} = \pi_{t,\hat{c}_{tS},S}$, and $\max_{c\in\mathcal{A}} \alpha_{tS}(c)$ equals the optimal value of the S-subsidized static problem 4.4.*

Proof. See paragraph immediately above the statement. □

Comparative Statics of π_{tcS} (and $P_t(c,\pi)$):

Lemma G.8. *i) For money-loss utility, π_{txS} is C^1 in x and $\frac{\partial \pi_{txS}}{\partial x} = \int m_{x,\pi}(s)dG_t(s) + x\int \left[\frac{\partial}{\partial x} m_{x,\pi}(s)\right] dG_t(s)$.*

ii) Let $x \leq 1$ unless utility is money-loss. π_{txS} is locally C^1 in x. It's derivative $\frac{\partial \pi_{txS}}{\partial x}$ equals $\frac{1}{\frac{\partial}{\partial \pi} P_t(x,\pi_{txS})}$ $\left[\underbrace{\int m_{x,\pi}(s)dG_t(s)}_{\text{Direct Effect}} + \underbrace{x\int \left[\frac{\partial}{\partial x} m_{x,\pi}(s)\right] dG_t(s)}_{\text{Substitution Effect}}\right]$ (where $\frac{\partial}{\partial \pi} P_t(x,\pi_{txS})$ is computed in G.17 below). Further, $\frac{\partial \pi_{txS}}{\partial x} \geq 0$, $\frac{\partial}{\partial x} k_t(x,\pi) \geq 0$, and $\frac{\partial}{\partial x} P_t(x,\pi) \leq 0$ with strict inequalities if $x < 1$ or under money-loss.

iii) For any $S \in \mathbb{R}$,

$$\frac{d}{dx} P_t(x,\pi_{txS}) = \frac{\partial}{\partial x} P_t(x,\pi_{txS}) + \frac{\partial}{\partial \pi} P_t(x,\pi_{txS}) \cdot \frac{\partial \pi_{txS}}{\partial x} = 0, \tag{G.15}$$

where $\frac{d}{dx}$ in $\frac{d}{dx} P_t(x,\pi_{txS})$ denotes the total derivative with respect to x.

Proof. i) For money-loss utility, $\frac{\partial}{\partial \pi} m_{x,\pi}(s) = 0$ by Lemma G.3, so the desired formula for $\frac{\partial \pi_{txS}}{\partial x}$ (and the C^1 property) is immediate from part ii.

ii) Recall $P_t(x,\pi) \equiv \pi - k_t(x,\pi)$, where $k_t(x,\pi) = \int x m_{x,\pi}(s)dG_t(s)$. Since, under the hypothesis of the result, $m_{\pi,x}(s)$ is C^1 in π and x by Lemma G.3, so is $k_t(x,\pi)$ and thus $P_t(x,\pi)$. Then applying the implicit function theorem (IFT) to $P_t(x,\pi) = -S$ establishes that π_{txS} is locally a C^1 function of x (since P_t is C^1, since [in turn] $m_{x,\pi}(s)$ is by Lemma G.3). The IFT also gives the formula $\frac{\partial \pi_{txS}}{\partial x} = -\frac{\frac{\partial}{\partial x} P_t(x,\pi_{txS})}{\frac{\partial}{\partial \pi} P_t(x,\pi_{txS})}$. Then we compute the formula with direct effect and substitution effect by unpacking the definitions of P_t and k_t above. The comparative statics are immediate from the definitions and the x-comparative statics in Lemma G.3 (either under money-loss or $x \leq (<)1$).

iii) By definition of π_{txS}, $P_t(x, \pi_{txS}) = -S$ for all x. Thus the total derivative $\frac{d}{dx}P_t(x, \pi_{txS}) = 0$. The formula with two terms in (G.15) follows immediately from chain rule. □

Lemma G.9. *i) For money-loss utility, $\frac{\partial}{\partial \pi}k_t(c, \pi) = 0$ and $\frac{\partial}{\partial \pi}P_t(c, \pi) = 1$ for any c; in particular these derivatives exist.*

ii) Now let $c \in C^1$ going forward. Then $\frac{\partial}{\partial \pi}k_t(c, \pi)$ and $\frac{\partial}{\partial \pi}P_t(c, \pi)$ exist and

$$\frac{\partial}{\partial \pi}k_t(c, \pi) = \int \frac{\partial}{\partial \pi}m_{c,\pi}(s)\left[1 - c'\left(m_{c,\pi}(s)\right)\right] dG_t(s), \tag{G.16}$$

$$\frac{\partial}{\partial \pi}P_t(c, \pi) = 1 - \frac{\partial}{\partial \pi}k_t(c, \pi). \tag{G.17}$$

iii) Assume further that $c' \geq 0$ (eqv. $x \leq 1$) and $c'(\cdot) \leq 1$ (eqv. $x \geq 0$) going forward. Then $\frac{\partial}{\partial \pi}k_t(c, \pi) \leq 0$ and $\frac{\partial}{\partial \pi}P_t(c, \pi) \geq 1$, with strict inequalities when $c' > 0$ and MRS is strictly diminishing (recall Def. 5).

iv) Further, $\frac{\partial \pi_{tcS}}{\partial S} = -\left[1 - \frac{\partial}{\partial \pi}k_t(c, \pi)\right]^{-1}$.

v) Further, if for two risk-types $H > L \in \mathcal{T}$ so that $G_H \succ_{MLRP} G_L$, then $\frac{\partial}{\partial \pi}P_L(x, \pi) \geq (\leq) \frac{\partial}{\partial \pi}P_H(x, \pi)$ if $\frac{\partial}{\partial \pi}m_{x,\pi}(s)$ is weakly increasing (decreasing) in s.[86] The last inequalities are strict when $\frac{\partial}{\partial \pi}m_{x,\pi}(s)$ is strictly increasing (decreasing) in s.[87]

Proof. Parts i-ii follow from the definitions of k_t and P_t (Eqn. 4.3) and Lemma G.3 about the comparative statics of $m_{c,\pi}(s)$. Part iii follows from part ii (spec. (G.16) and (G.17)) and the added assumption $c' \leq 0$ which helps sign (G.16).

iv): To calculate $\frac{\partial \pi_{tcS}}{\partial S}$, recall π_{tcS} is defined by the expression $P_t(c, \pi_{txS}) = -S$. Applying the implicit function theorem (IFT) gives $\frac{\partial \pi_{tcS}}{\partial S} = -\left[\frac{\partial}{\partial \pi}P_t(c, \pi)\right]^{-1}$, and then (G.17) gives the desired formula. [Note the IFT is valid since $\frac{\partial}{\partial \pi}P_t \geq 1 > 0$ by iii.]

v): Follows from G.16 and G.17, and the fact that G_H first-order stochastically dominates G_H (since $G_H \succ_{\text{MLRP}} G_L$). □

G.2 Proofs for Appendix C

Lemma 1 (Restated)**.** *Let utility be money-loss CARA. Then $\hat{c}_{tS}$ is constant in S.*

Proof. The proof is exactly the proof of Proposition C.2 part ii which is the same result but in a slightly more general framework that allows for variation in baseline wealth. □

Recall $V_t(c, \pi)$ and $\lambda_t^{c,\pi}$ from the main text ((4.2) and (4.7)), and recall ex-post utility $W^{c,\pi}(s)$ and marginal utility $W_1^{c,\pi}(s)$ from Definition 6 in Appendix G.1.1 above.

[86] Recall, $\frac{\partial}{\partial \pi}m_{x,\pi}(s)$ is negative when $x < 1$ (i.e. $c' > 0$) by Lemma G.3.

[87] Part v is used (with the weak inequality) in the proof of the main result Thm. 1′ ii.

Let

$$A_t^{x,\pi} \equiv \frac{\partial}{\partial x} V_t(x,\pi). \tag{G.18}$$

Lemma G.10. *For any (c,π) and t, $\frac{\partial}{\partial \pi} V_t(c,\pi)$ exists and*

$$\frac{\partial}{\partial \pi} V_t(c,\pi) = -\lambda_t^{c,\pi} < 0. \tag{G.19}$$

For any (x,π) and t, $\frac{\partial}{\partial \pi} V_t(c,\pi)$ exists and

$$A_t^{x,\pi} \equiv \frac{\partial}{\partial x} V_t(x,\pi) = \int m_{x,\pi}(s) W_1^{x,\pi}(s) dG_t(s) > 0. \tag{G.20}$$

Proof. By the envelope theorem, for all s, $\frac{\partial}{\partial \pi} W^{c,\pi}(s) = -W_1^{c,\pi}(s)$ and $\frac{\partial}{\partial x} W^{x,\pi}(s) = m_{c,\pi}(s) W_1^{c,\pi}(s)$ for all π and x. Then taking expectations over s, (G.19) and (G.20) follow from the definitions of $V_t(c,\pi)$, $\lambda_t^{c,\pi}$, and $A_t^{x,\pi}$. Those definitions also imply the signs $\lambda_t^{c,\pi} > 0$, $A_t^{x,\pi} > 0$ due to the following factors: $W_1^{c,\pi}(s) > 0$, $m_{x,\pi}(s) > 0$ for $s \geq s_\pi^*$ (by Lemma G.2), and $1 - G_t(s_\pi^*) > 0$ (where s_π^* was defined in Assumption B5) for all t by the fact that the G_t's have densities g_t with common support. □

Lemma G.11. *For any t, S, x, we have that α'_{tS} exists and $\alpha'_{tS}(x) = A_t^{x,\pi_{txS}} - \lambda_t^{x,\pi_{txS}} \frac{\partial \pi_{txS}}{\partial x}$; as a result, $\frac{\partial \pi_{txS}}{\partial x}|_{\hat{x}_{tS}} = A_t^{\hat{x}_{tS},\hat{\pi}_{tS}} / \lambda_t^{\hat{x}_{tS},\hat{\pi}_{tS}}$ for any optimal solution $(\hat{x}_{tS}, \hat{\pi}_{tS})$ to t's S-subsidized static problem. Further, if utility is money-loss,*

$$\alpha'(x) = A_t^{x,\pi_{txS}} - \lambda_t^{x,\pi_{txS}} \left[\int m_{x,\pi}(s) dG_t(s) + x \int \left[\frac{\partial}{\partial x} m_{x,\pi}(s) \right] dG_t(s) \right].$$

Proof. Recall $\alpha_{tS}(x) \equiv V_t(x, \pi_{txS})$. V_t is differentiable in both arguments by Lemma G.10 (just above) and the fact that the function π_{txS} is C^1 in x by Lemma G.8. From chain rule,

$$\alpha'_{tS}(x) = \frac{\partial}{\partial x} V_t(x, \pi_{txS}) + \frac{\partial}{\partial \pi} V_t(x, \pi_{txS}) \frac{\partial \pi_{txS}}{\partial x}. \tag{G.21}$$

The first assertion of the result then follows by substituting the derivatives of V_t from Lemma G.10. If $(\hat{x}_{tS}, \hat{\pi}_{tS})$ is statically optimal for t with subsidy S, then $\alpha'_{tS}(\hat{x}_{tS}) = 0$ by Lemma G.7. The result then follows from the first assertion by simple algebra. The money-loss part follows from (G.21) by Lemma G.8 part i which computes $\frac{\partial \pi_{txS}}{\partial x}$. □

G.2.1 Proof of Proposition C.1

Recall $W_1^{x,\pi}(s)$ from Definition 6 and $\lambda_t^{x,\pi}$ from (4.7).

Lemma G.12 (Helper for Prop. C.1)**.** *Consider any x, π, t. Assume either utility is money-loss or $x \leq 1$. If $W_1^{x,\pi_{txS}}(s)$ is strictly increasing (decreasing) in s, then $\int_{\underline{s}}^{\bar{s}} m_{x,\pi}(s) \left[W_1^{x,\pi_{txS}}(s) - \lambda_t^{x,\pi_{txS}}\right] dG_t(s) > (<) 0$.*

Proof. I show the result for the case where $W_1^{x,\pi_{txS}}(s)$ is weakly increasing, which will be used in the proof of Proposition C.1; the other case is similar. I henceforth suppress the superscripts on $W_1^{x,\pi_{txS}}(s)$ and $\lambda_t^{x,\pi_{txS}}$ and the subscripts on $m_{x,\pi}(s)$. We break up the integral

$$\int_{\underline{s}}^{\bar{s}} m(s) \left[W_1(s) - \lambda_t\right] dG_t(s) \tag{G.22}$$

into two pieces. Ideally there is a s^* such that $\lambda_t = W_1(s^*)$. If so, we break up the integral into a piece below and a piece above that point. Since $W_1(s)$ is strictly increasing in s by hypothesis, we will have $\lambda_t \geq W_1(s)$ for $s \in [\underline{s}, s^*]$ and $\lambda_t \leq W_1(s)$ for $s \in [s^*, \bar{s}]$ both by definition of λ_t, and strictly so on the interiors of these intervals. Since all relevant functions are continuous in s, such an s^* will exist by the intermediate value theorem. However, for extra generality, I address the case where G_t is discontinuous so that s^* doesn't exist. Let $s_{\text{lower}} = \max\{s | \lambda_t \geq W_1(s)\}$ and let $s_{\text{upper}} = \min\{s | \lambda_t \leq W_1(s)\}$. Then similarly, we will have $\lambda_t \geq W_1(s)$ for $s \in [\underline{s}, s_{\text{lower}}]$, and $\lambda_t \leq W_1(s)$ for $s \in [s_{\text{upper}}, \bar{s}]$, and $s_{\text{lower}} \leq s_{\text{upper}}$, strictly so on the interiors of these intervals. Then (G.22) equals

$$\begin{aligned}
&\int_{\underline{s}}^{s_{\text{lower}}} m(s) \left[W_1(s) - \lambda_t\right] dG_t(s) + \int_{s_{\text{upper}}}^{\bar{s}} m(s) \left[W_1(s) - \lambda_t\right] dG_t(s) \\
> m(s_{\text{lower}}) &\int_{\underline{s}}^{s_{\text{lower}}} \left[W_1(s) - \lambda_t\right] dG_t(s) + m(s_{\text{upper}}) \int_{s_{\text{upper}}}^{\bar{s}} \left[W_1(s) - \lambda_t\right] dG_t(s) \\
&\geq m(s_{\text{lower}}) \int_{\underline{s}}^{\bar{s}} \left[W_1(s) - \lambda_t\right] dG_t(s) = 0,
\end{aligned}$$

where the first inequality follows from strict increasingness of $m(s)$ in s (by Lemma G.4 i or iv corresponding to the disjunctive hypothesis "money-loss or $x \leq 1$") and the facts that $\lambda_t \geq W_1(s)$ for $s \in [\underline{s}, s_{\text{lower}}]$ and $\lambda_t \leq W_1(s)$ for $s \in [s_{\text{upper}}, \bar{s}]$, with strict inequality on the interiors. The second inequality follows from the latter property and from $m(s_{\text{lower}}) \leq m(s_{\text{upper}})$ which holds since $s_{\text{lower}} \leq s_{\text{upper}}$ and $m(s)$ is increasing in s. The last equality follows by definitions of λ_t. □

We are now ready to prove Proposition C.1 which I restate below.

Proposition C.1 ($\hat{x} > 0$: Restated)**.** *i) Let utility be money-loss. For any S and G_t, if $\hat{x}_{tS}$ exists, $\hat{x}_{tS} > 0$. ii) Let utility be general with $U_{12} \leq 0$, and suppose $h\left(m_{c,\pi}(s), s\right)$ is weakly decreasing in s. For any S and G_t, $\alpha'_{tS}(0) > 0$ (where, recall, α takes x as argument).*

Proof. I prove i and ii simultaneously. Suppose for a contradiction that $\hat{x}_{t,S} \leq 0$. Our goal is to show $\alpha'_{tS}(\hat{x}_{tS}) > 0$, which will contradict optimality of $\hat{x}_{tS}$. Let λ_t be short for $\lambda_t^{x,\pi_{txS}}$.

By Lemma G.11, for any t, S, x: $\alpha'_{tS}(x) = A_t^{x,\pi_{txS}} - \lambda_t^{x,\pi_{txS}} \frac{\partial \pi_{txS}}{\partial x}$. Plugging in the formula for $A_t^{x,\pi}$ from (G.20), and the formula for $\frac{\partial \pi_{txS}}{\partial x}$ in Lemma G.8 ii, $\alpha'_{tS}(x)$ equals

$$\int m_{x,\pi}(s) W_1^{x,\pi_{txS}}(s) dG_t(s) - \lambda_t \frac{1}{\frac{\partial}{\partial \pi} P_t\left(x, \pi_{txS}\right)} \left[\int m_{x,\pi}(s) dG_t(s) + x \int \left[\frac{\partial}{\partial x} m_{x,\pi}(s) \right] dG_t(s) \right].$$

For money-loss, $\frac{\partial}{\partial \pi} P_t\left(\hat{x}_{tS}, \pi_{t\hat{x}_{tS}S}\right) = 1$, and it's easy to see the same is true for general utility when $\hat{x}_{tS} = 0$. $\hat{x}_{tS} \leq 0$ by assumption and $\frac{\partial}{\partial x} m_{\hat{x}_{tS},\pi}(s) \geq 0$ by Lemma G.3, so $\hat{x}_{tS} \int \left[\frac{\partial}{\partial x} m_{\hat{x}_{tS},\pi}(s)\right] dG_t(s) \leq 0$. Thus by removing that term we get the inequality

$$\begin{aligned} \alpha'_{tS}(\hat{x}_{tS}) &\geq \int m_{x,\pi}(s) W_1^{x,\pi_{txS}}(s) dG_t(s) - \lambda_t \left[\int m_{x,\pi}(s) dG_t(s) \right] \\ &= \int m_{x,\pi}(s) \left[W_1^{x,\pi_{txS}}(s) - \lambda_t \right] dG_t(s). \end{aligned}$$

To prove ii as well as to establish the contradiction and thus proof i, it remains to show the last expression is strictly positive. This will follow from Lemma G.12 if we can show $W_1(s)$ is weakly increasing in s. And that holds by Appendix G.5 ii or iii under the hypotheses of i or ii (resp.) of the present result.[88] This concludes the proof. □

G.2.2 Interlude: Static Problem with Special Functional Form

Below, I refer to

$$h(m, s) = -\frac{1}{2\omega}(m - s)^2 + (m - s), \tag{G.23}$$

for some $\omega > 0$ as the special form.

Lemma G.13. *Let utility be money-loss model and h take the special form (G.23). Then i) $m_x(s) = s + x\omega$, and so $\frac{\partial}{\partial x} m_x(s) = \omega$. Further ii): $h\left(m_x(s), s\right)$ is constant in s.*

Proof. i) follows from taking the FOC with respect to $m_x(s)$ and solving for $m_x(s)$. ii) follows by plugging in $m_x(s) = s + x\omega$ into $h(m_x(s), s)$. □

[88] To verify the hypotheses of Appendix G.5 ii or iii, note linear coinsurance is C^1 and that we are considering the case $x \leq 0$ so we have $x < 1$ (i.e,. $c' > 0$) in particular.

Fact G.3. *Let utility be money-loss and assume h takes the special form (G.23). i) For any risk-type t, the unconstrained static problem $\max_x \alpha_{t,S}(x)$ is strictly concave in x for every subsidy S. ii) The optimum $(\hat{x}_{tS}, \hat{\pi}_{tS})$ (if it exists) is unique for all S.*

Proof. By Lemma G.14 ii, S simply functions to increase the baseline wealth by S, so it is without loss to prove the result for $S = 0$ since the baseline wealth w, which we (as usual) omit will be seen to obviously not affect the argument of the proof.

I suppress the t subscript in G_t. In the money-loss model

$$\alpha_{tS}(x) = \int u\left(-\underbrace{x\int m_x(s)dG_t(s)}_{\pi_{tx}} - (1-x)m_x(s) + h\left(m_{\pi_{tx},x}(s), s\right)\right) dG(s)$$

By the envelope theorem (note formula wraps),

$$\alpha'(x) = \int \left[-\int \left[m_x(\hat{s}) + x\frac{\partial}{\partial x} m_x(\hat{s})\right] dG(\hat{s}) + m_x(s)\right]$$
$$\cdot\, u'\left(-x\int m_x(s)dG_t(s) - (1-x)m_x(s) + h\left(m_{\pi_{tx},x}(s), s\right)\right) dG(s)$$

Differentiating again and suppressing the arguments of u', u'',

$$\alpha''_{tS}(x) =$$
$$\int \left[-\int \left[m_x(\hat{s}) + \frac{\partial}{\partial x} m_x(\hat{s})\right] dG(\hat{s}) + m_x(s)\right]^2 u'' dG(s) +$$
$$+ \int \left\{ \left[-\int \left[2\frac{\partial}{\partial x} m_x(\hat{s}) + \frac{\partial^2}{\partial x^2} m_x(\hat{s})\right] dG(\hat{s}) + \frac{\partial}{\partial x} m_x(s)\right] u' \right\} dG(s)$$

The first term is clearly strictly negative since $u'' < 0$ and since the squared term is strictly positive [since $\frac{\partial}{\partial x} m_x(\hat{s}) = \omega > 0$ for all $\hat{s}$ by the Lemma G.13]. Now consider the second term. We use the assumption $h(m, s) = -\frac{1}{2\omega}(m-s)^2 + (m-s)^2$ so that $m_x^*(s) = s + x\omega$, $\frac{\partial}{\partial x} m_x(\hat{s}) = \omega$ and $\frac{\partial^2}{\partial x^2} m_x(\hat{s}) = 0$, so the whole second term becomes $-2\omega + \omega = -\omega < 0$. Thus, for all x, $\alpha''_{tS}(x) < 0$ as desired.

By Lemma G.7, $\hat{x}_{tS}$ satisfies $\alpha'_{tS}(\hat{x}_{tS}) = 0$ so part i of this result implies $\hat{x}_{tS}$ is unique. Uniqueness of $\hat{\pi}_{tS}$ then follows from Lemma G.6 under the money-loss assumption and the fact that $V_t(x, \pi)$ is strictly decreasing in π (Lemma G.10). □

G.2.3 Proof of Comparative Statics in Subsidy and Baseline Wealth (Prop. C.2)

Definition 9. For any baseline wealth level $w \in \mathbb{R}$, we make the following definitions

$$V_t^w(x,\pi) \equiv \int U\left(w - \pi - (1-x)m_{x,-w+\pi}(s),\, h\left(m_{x,-w+\pi}(s), s\right)\right) dG_t(s) \tag{G.24}$$

$$P_{t,S}^w(x,\pi) \equiv S + \pi - (1-x)\int m_{x,-w+\pi}(s) dG_t(s) \tag{G.25}$$

$$R_t(w,S) \equiv \max_{(x,\pi):P_{t,S}^w(x,\pi)=0} V_t^w(x,\pi) \tag{G.26}$$

$$\left(\hat{x}_{t,S}^w, \hat{\pi}_{t,S}^w\right) \equiv \arg\max_{(x,\pi):P_{t,S}^w(x,\pi)=0} V_t^w(x,\pi) \tag{G.27}$$

Abusing notation, the notation in Definition 9 can be used with an arbitrary coinsurance schedule c instead of the linear coverage level x used above. I state the next helper-lemma in this extra generality before returning to the linear. I also allow general utility, whereas the other results in this section will be for money-loss only.

Lemma G.14 (Helper for Prop. C.2)**.** *Let utility be general and let $w \in \mathbb{R}$ denote a baseline wealth level. i) For any $w, w' \in \mathbb{R}$, $V_t^w(c,\pi) = V_t^{w'}(c, \pi + (w'-w))$. ii) For all w, w', S, $R_t(w,S) = R_t(w', S-(w'-w))$ and*

$$\left(\hat{c}_{t,S}^w, \hat{\pi}_{t,S}^w\right) = \left(\hat{c}_{t,S-(w'-w)}^{w'}, \hat{\pi}_{t,S-(w'-w)}^{w'} + (w'-w)\right) \tag{G.28}$$

Proof. Part i follows directly from (G.24). Intuitively, what matters is the "effective premium" which I define as $w - \pi$, and this is invariant to changes that simultaneously increase baseline wealth and premium by the same amount.

Part ii: Let $w, w' \in \mathbb{R}$. By (G.27), $P_{tS}^w\left(\hat{c}_{t,S}^w, \hat{\pi}_{t,S}^w\right) = 0$. Then $P_{t,S-(w'-w)}^{w'}\left(\hat{c}_{t,S}^w, \hat{\pi}_{t,S}^w + (w'-w)\right) = 0$ also since the $(w'-w)$ terms cancel in the first two terms of (G.25) and the second subscript on m, which is the effective premium, hasn't changed. This implies the inequality in

$$R_t\left(w', S-(w'-w)\right) \geq V_t^{w'}\left(\hat{c}_{t,S}^w, \hat{\pi}_{t,S}^w + (w'-w)\right) = V_t^w\left(\hat{c}_{t,S}^w, \hat{\pi}_{t,S}^w\right) = R_t(w,S) \tag{G.29}$$

where the first equality is by part i of the present result and the second is immediate from the definitions. A symmetric argument gives the chain:

$$R_t\left(w, S\right) \geq V_t^w\left(\hat{c}_{t,S-(w'-w)}^{w'}, \hat{\pi}_{t,S-(w'-w)}^{w'} + (w-w')\right) = \tag{G.30}$$

$$= V_t^{w'}\left(\hat{c}_{t,S-(w'-w)}^{w'}, \hat{\pi}_{t,S-(w'-w)}^{w'}\right) = R_t\left(w', S-(w'-w)\right) \tag{G.31}$$

Thus equalities hold throughout (G.29) and (G.30). This implies $R_t(w,S) = R_t\left(w', S-(w-w')\right)$ and (G.28). □

Lemma G.15. *Let utility be money-loss model but with two different utility functions u, v, where v is strictly more risk averse than u. Fix any t and assume the static problem is strictly concave with respect to both utility functions. Let $\hat{x}_t^u$ and $\hat{x}_t^v$ be the solutions to the unsubsidized static problem (4.4) under u and v respectively. Then $\hat{x}_t^v > \hat{x}_t^u$.*[89]

Proof. A standard characterization of "strictly more risk averse than" implies that for all x, $v(x) = \phi(u(x))$ for some strictly increasing concave function ϕ. Then $v'(x) = \phi'\left(u(x)\right)u'(x)$.

Recalling the function α_t from equation (G.13) which was defined for a fixed utility function, denote by α_t^u, α_t^v the functions corresponding to utility u and v respectively. From Lemma G.7, $\hat{x}_t^v \in \arg\max_x \alpha_t^v(x)$. Thus by expanding out $\alpha_t^{u\prime}(\hat{x}_t^u)$ using Lemma G.11, the FOC for the optimal x (under u) is the following expression evaluated at $x = \hat{x}_t^u$ (note the integral wraps):

$$\underbrace{\int \left[\underbrace{-\int \left[m_x(\hat{s}) + \frac{\partial}{\partial x} m_x(\hat{s}) \right] dG(\hat{s})}_{\frac{\partial}{\partial x}\pi_x \text{ constant in } s} + m_x(s) \right]}_{B(s)} \cdot u'\left(-\left[\underbrace{\int x m_x(\hat{s}) dG(\hat{s})}_{\pi_x} + (1-x) m_x(s) \right] + h\left(m_x(s), s\right) \right) dG_t(s) = 0 \tag{G.32}$$

when x is evaluated at $x = \hat{x}_t^u$. On the other hand, using $v'(x) = \phi'\left(u(x)\right)u'(x)$ we see that $\alpha_t^{v\prime}(\hat{x}_t^u)$ equals the following expression evaluated at $x = \hat{x}_t^u$ (note the integral wraps):

$$\int \left[-\int \left[m_x(\hat{s}) + \frac{\partial}{\partial x} m_x(\hat{s}) \right] dG(\hat{s}) + m_x(s) \right] u'\left(-\left[\underbrace{\int x m_x(\hat{s}) dG(\hat{s})}_{\pi_x} + (1-x) m_x(s) \right] + h\left(m_x(s), s\right) \right)$$
$$\cdot \phi'\left(u\left(-\left[\underbrace{\int x m_x(\hat{s}) dG(\hat{s})}_{\pi_x} + (1-x) m_x(s) \right] + h\left(m_x(s), s\right) \right) \right) dG_t(s).$$

The last expression is the same as (G.32) except for the presence of the factor ϕ'. Since ϕ

[89] The proof technique is from [Mas-Colell et al. 1995] Example 6.C.2 (p. 192).

is increasing and concave, ϕ' is positive and decreasing. The argument of ϕ' (which is just ex-post utility) is strictly decreasing in s by Lemma G.5 part i, so that the whole ϕ' term is positive and increasing in s. On the other hand, the first term in square brackets denoted by $B(s)$ is strictly increasing in s (since $m_x(s)$ is, Lemma G.3) and crosses zero once from below. The ϕ' term gives higher weight to the higher (i.e., more positive) parts of the $B(s)$ term (which occur for higher s) and lower weight to the lower (i.e., more negative) parts (which occur for lower s). Overall this raises the value of the outer integral, since the u' term is positive.

We've now shown that $\alpha_t^{v\prime}(\hat{x}_t^u) > \alpha_t^{w\prime}(\hat{x}_t^u) = 0$ which implies that $\alpha'_{S'}(\hat{x}_{tS''}) > 0$. Then $\hat{x}_t^v > \hat{x}_t^u$ follows from the concavity of the assumed strict concavity of the static programs. □

Proposition C.2 (S Comparative Static, Money-Loss: Restated)**.** *Let utility be money-loss and assume the static problem's objective $\alpha_{tS}(\cdot)$ is strictly concave for all S and baseline wealth levels w. If u satisfies strict DARA, IARA, or CARA, $\hat{x}_{tS}^w$ is (respectively) strictly decreasing, strictly increasing, or constant in both S and w.*[90]

Proof. Part i. For all $w \in \mathbb{R}$, $v_w(x) = u(w + x)$. If $w' > w''$, then by strict DARA, $v_{w''}$ is more risk-averse than $v_{w'}$. By Lemma G.15, $\hat{x}_{tS}^{w''} > \hat{x}_{tS}^{w'}$ for any S. Now fix w. Let $S' > S''$. Again, by strict DARA, $v_{S''}$ is more risk-averse than $v_{S'}$, so $\hat{x}_{t,0}^{w+S''} > \hat{x}_{t,0}^{w+S'}$ by Lemma G.15. But by Lemma G.14, $\hat{x}_{t,0}^{w+S'} = \hat{x}_{t,S'}^{w}$ and $\hat{x}_{t,0}^{w+S''} = \hat{x}_{t,S''}^{w}$, so $\hat{x}_{tS''}^{w} > \hat{x}_{tS'}^{w}$. Since w was arbitrary, we have shown, in the notation which drops w, $\hat{x}_{tS''} > \hat{x}_{tS'}$, as desired.

We now turn to part ii about CARA. We prove the result for nonlinear schedules $c(\cdot)$ with $c(m) = (1-x)m$ as a special case. Observe that for all c, S, t, we have $\pi_{t,c,S} = \pi_{t,c,0} - S$ since treatment choice $m_{c,\pi} = m_c$ is independent of π in the money-loss model (Lem. G.3).

Then, recalling from (G.13) the definition of the objective $\alpha_{tS}(c(\cdot))$ of the unconstrained static problem,

$$\begin{aligned}
\alpha_{tS}(c) &\equiv \int u\left(-\pi_{tcS} - (1-x)m_c(s) + h\left(m_c(s), s\right)\right) dG_t(s) \\
&= \int u\left(-\pi_{tc0} + S - (1-x)m_c(s) + h\left(m_c(s), s\right)\right) dG_t(s) \\
&= u(S) \cdot \int u\left(-\pi_{tc0} - (1-x)m_c(s), h\left(m_c(s), s\right)\right) dG_t(s) \\
&= u(S) \cdot \alpha_{t0}(c)
\end{aligned}$$

where the second-to-last equality uses the exponential form of u under CARA. Thus for all

[90] The CARA part extends the result of Lemma 1, which dealt only with subsidies not baseline wealth levels.

S, since $u(S)$ is a constant in $c(\cdot)$,

$$\arg\max_c \alpha_{tS}(c) = \arg\max_c u(S)\alpha_{t0}(c) = \arg\max_c \alpha_{t0}(c).$$

The argument for wealth differences is similar. This completes the proof of part ii. □

G.3 Proof: Observable-Risk Benchmark (Prop. 1 and Prop. D.1)

Proposition 1 (Observable-Risk Characterization: Restated)**.** *Let $\mathcal{A}$ be any subset of measurable functions, mapping $\mathbb{R}_+ \to \mathbb{R}$. If $(\tilde{c}_t, \tilde{\pi}_t)_{t=1}^T$ is a solution to the observable risk problem (4.6) (with that $\mathcal{A}$), the following hold.*

i) If utility is money-loss CARA, $\tilde{c}_t = \hat{c}_{t\tilde{S}_t}$ for all t. For general utility, $\tilde{c}_t = \hat{c}_t$ for all t.

ii) If utility is money-loss, expected marginal utility is equalized across types: $\lambda_L^{\tilde{c}_L,\tilde{\pi}_L} = \lambda_H^{\tilde{c}_H,\tilde{\pi}_H}$. For general utility, if $\tilde{c}_t$ is C^1 for all t, then for all $t, \tau \in \mathcal{T}$, $\tilde{\lambda}_t \cdot \left[\frac{\partial}{\partial\pi} P_t(\tilde{c}_t, \tilde{\pi}_t)\right]^{-1} = \tilde{\lambda}_\tau \cdot \left[\frac{\partial}{\partial\pi} P_\tau(\tilde{c}_\tau, \tilde{\pi}_\tau)\right]^{-1}$.

Proof. i) To show $\tilde{c}_t = \hat{c}_{t\tilde{S}_t}$ for all t, suppose not so that for some t, $\hat{c}_{t\tilde{S}_t} \neq \tilde{c}_t$. By definition of $\hat{c}_{t\tilde{S}_t}$ and $\tilde{S}_t$, replacing t's contract $(\tilde{c}_t, \tilde{\pi}_t)$ with $(\hat{c}_{t\tilde{S}_t}, \hat{\pi}_{t\tilde{S}_t})$ gives $V_t\left(\hat{c}_{t\tilde{S}_t}, \hat{\pi}_{t\tilde{S}_t}\right) \geq V_t\left(\tilde{c}_t, \tilde{\pi}_t\right)$ and keeps the ex-ante budget balanced unchanged since the expected actuarially loss on t's static contract remains at $\tilde{S}_t$. This contradicts the optimality of $(\tilde{c}_\tau, \tilde{\pi}_\tau)_{\tau=1}^T$.

ii) For money-loss, $\frac{\partial}{\partial\pi} P_t(c, \pi) = 1$ for money-loss utility for any c (Lemma G.9), so we just need to prove the general formula in ii. Note that if $\tilde{c}_t$ is C^1, then $\frac{\partial}{\partial\pi} P_t(\tilde{c}_t, \tilde{\pi}_t)$ exists by Lemma G.9. Recall f_t denotes the probability of risk-type t. Suppose for contradiction that for two arbitrary risk-types $L, H \in \mathcal{T}$, $\lambda_L^{\tilde{c}_L,\tilde{\pi}_L} \cdot \left[\frac{\partial}{\partial\pi} P_L(\tilde{c}_L, \tilde{\pi}_L)\right]^{-1} < \lambda_H^{\tilde{c}_H,\tilde{\pi}_H} \cdot \left[\frac{\partial}{\partial\pi} P_H(\tilde{c}_H, \tilde{\pi}_H)\right]^{-1}$; note we are not assuming $L < H$ so there is no loss of generality. In the space of coordinates (π_L, π_H), consider the direction-of-deviation

$$\left(\frac{f_H}{f_L}\left[\frac{\partial}{\partial\pi} P_L(\tilde{c}_L, \tilde{\pi}_L)\right]^{-1}, -\left[\frac{\partial}{\partial\pi} P_H(\tilde{c}_H, \tilde{\pi}_H)\right]^{-1}\right), \tag{G.33}$$

so that we are increasing π_L and decreasing π_H, keeping $\tilde{c}_L, \tilde{c}_H$ constant and leaving constant $(\tilde{c}_t, \tilde{\pi}_t)$ for all $t \neq L, H$.[91] I henceforth suppress the arguments of $\frac{\partial}{\partial\pi} P_t(\tilde{c}_t, \tilde{\pi}_t)$ for $t = L, H$. The marginal effect on ex-ante expected utility of moving in the above direction is (by chain

[91] In the case of the reverse strict inequality, we would instead use the deviation: $\left(-\left[\frac{\partial}{\partial\pi} P_L(\tilde{c}_L, \tilde{\pi}_L)\right]^{-1}, \frac{f_L}{f_H}\left[\frac{\partial}{\partial\pi} P_H(\tilde{c}_H, \tilde{\pi}_H)\right]^{-1}\right)$.

rule):

$$f_L\left(\frac{f_H}{f_L}\right)\left[\frac{\partial}{\partial\pi}P_L\right]^{-1}\frac{\partial}{\partial\pi}V_L\left(\tilde{c}_L,\tilde{\pi}_H\right)-f_H\left[\frac{\partial}{\partial\pi}P_H\right]^{-1}\frac{\partial}{\partial\pi}V_H\left(\tilde{c}_H,\tilde{\pi}_H\right)$$
$$=f_H\left[\frac{\partial}{\partial\pi}P_L\right]^{-1}\left(-\lambda_L^{\tilde{c}_L,\tilde{\pi}_L}\right)-f_H\left[\frac{\partial}{\partial\pi}P_H\right]^{-1}\left(-\lambda_H^{\tilde{c}_H,\tilde{\pi}_H}\right)$$
$$=f_H\left[-\lambda_L^{\tilde{x}_L,\tilde{\pi}_L}\left[\frac{\partial}{\partial\pi}P_L\right]^{-1}+\lambda_H^{\tilde{x}_H,\tilde{\pi}_H}\left[\frac{\partial}{\partial\pi}P_{H\tilde{S}_H}\right]^{-1}\right]>0,$$

where the first equality follows by G.10 and the last inequality is by our supposition toward contradiction.

The effect on the ex-ante budget is:

$$f_L\left(\frac{f_H}{f_L}\right)\left[\frac{\partial}{\partial\pi}P_L\right]^{-1}\frac{\partial}{\partial\pi}P_L-f_H\left[\frac{\partial}{\partial\pi}P_H\right]^{-1}\frac{\partial}{\partial\pi}P_H$$
$$=f_H-f_H=0.$$

Since the deviation is budget-feasible at the margin and increases expected utility, this contradicts optimality of the pair of contracts $(\tilde{x}_L,\tilde{\pi}_L)$, $(\tilde{x}_H,\tilde{\pi}_H)$. This concludes the proof of ii. □

Lemma D.1. *i) For all (c,π), if $c\in C^2$, then $\frac{\partial}{\partial\mu}\left[m_{c+\mu\beta_t^{c,\pi},\pi}(s)\right]_{|\mu=0}=0$ for all s. Further, $\int\beta_t^{c,\pi}\left(m(s)\right)dG_t(s)=1$.*

ii) $\frac{\partial}{\partial\mu}\left[V_t\left(c+\mu\beta_t^{c,\pi},\pi\right)\right]_{\mu=0}=-\nu_t^{c,\pi}$.

Proof. Fix c,π. Part i) We construct a direction of perturbation $d^{c,\pi}(\cdot)$ — which will turn out to be $\beta_t^{c,\pi}$ defined in (D.3), — that keeps treatment constant at the margin (starting at c,π).

In Supplemental Appendix S4.1 (Eqn. S4.3) I use the implicit function theorem to re-derive a result of [Werquin et al. 2015] (their proposition 1 equation 8) that when $c\in C^2$, for any C^1 direction of perturbation $d(\cdot)$, the following holds:

$$\frac{\partial}{\partial\mu}\left[m_{c+\mu d(\cdot),\pi}(s)\right]_{|\mu=0}=\left[1-\mathbb{E}_{\mathrm{NG}}^{(c)}(m)c''(m)\right]^{-1}\left[\mathbb{E}_{\mathrm{NG}}^{(c)}(m)d'(m)-\mathbb{I}_{\mathrm{NG}}(m)d(m)\right],\qquad\text{(G.34)}$$

where recall $\mathbb{E}_{\mathrm{NG}}^{(c)}(m)$ is the derivative of the compensated treatment-demand with respect to out-of-pocket price and $\mathbb{I}_{\mathrm{NG}}(m)$ is the derivative of treatment-demand with respect to income (defined in Supplemental Appendix S3.2). It follows that a C^1 direction of perturbation d preserves treatment at the margin (i.e. satisfies $\frac{\partial}{\partial\mu}\left[m_{c+\mu d,\pi}(s)\right]_{|\mu=0}=0$) if and only if

$d'(m) = -D^{c,\pi}(m)d(m)$ for all m, where (recall) $D^{c,\pi}(m) = -\frac{\mathbb{I}^{c,\pi}_{\text{NG}}(m)}{\mathbb{E}^{(c),c,\pi}_{\text{NG}}(m)}$. Solving this ODE with initial condition $d\left(m(\underline{s})\right) = \underline{d}$ (for arbitrary $\underline{d} \in \mathbb{R}$) gives $d(m) = \underline{d} L_{c,\pi,\underline{s}}\left(s^{-1}_{c,\pi}(m)\right)$ where $L_{c,\pi,s}(\hat{s})$ was defined in (D.1). Then, repeating the following definition (from (D.3))

$$\beta^{c,\pi}_t(m) \equiv \underline{\beta}^{c,\pi}_t L_{c,\pi,\underline{s}}\left(s^{-1}_{c,\pi}(m)\right), \quad \text{where} \quad \underline{\beta}^{c,\pi}_t \equiv \left[\int L_{c,\pi,\underline{s}}\left(s^{-1}_{c,\pi}(m)\right) dG_t(s)\right]^{-1}, \tag{G.35}$$

it follows that $\frac{\partial}{\partial \mu}\left[m_{c+\mu\beta^{c,\pi}_t(m),\pi}(s)\right]_{|\mu=0} = 0$ and $\int \beta^{c,\pi}_t\left(m(s)\right) dG_t(s) = 1$, as desired.[92]

Part ii) $\frac{\partial}{\partial \mu}\left[V_t\left(c + \mu\beta^{c,\pi}_t, \pi\right)\right]_{\mu=0} = -\nu^{c,\pi}_t$ by the definitions of $\beta^{c,\pi}_t(\cdot)$ and ν_t ((G.35) and (D.2)). □

Proposition D.1. *For all t, assume $\tilde{c}_t \in C^2$. Then the following hold: Then:*

i) For general utility, if arbitrary nonlinear coinsurance is allowed, $\tilde{\nu}_t = \tilde{\nu}_\tau$ for all $t, \tau \in \mathcal{T}$.

ii) Under separable utility $\nu^{c,\pi}_t = \theta^{c,\pi}_t$ for all t and (c, π).

iii) Under money-loss utility, $\nu^{c,\pi}_t = \lambda^{c,\pi}_t$ for all t and (c, π).

Proof. By Lemma D.1 and the definition of $\tilde{\beta}_t(\cdot) \equiv \beta^{\tilde{c}_t,\tilde{\pi}_t}_t(\cdot)$, the direction of deviation $\tilde{\beta}_t$ keeps treatment choices constant at the margin starting at $(\tilde{c}_t, \tilde{\pi}_t)$.

Part i) Intuitively, if $\tilde{\nu}_t \neq \tilde{\nu}_\tau$, we can use "$\nu$-transfers" to send money from the type with lower ν to the one with higher ν in a budget-neutral way, which will be a profitable deviation. We now make this for formal.

Without loss of generality, suppose (for contradiction) that for some arbitrary $L, H \in \mathcal{T}$, $\tilde{\nu}_H > \tilde{\nu}_L$; note we are not assuming $L < H$ so there is not loss of generality. We will perturb $c_L(\cdot)$ by one "unit" in the direction of $\tilde{\beta}_L(\cdot)$ (defined above) thus making L pay more coinsurance, and simultaneously perturb $c_H(\cdot)$ by $\frac{f_L}{f_H}$ units in the direction of $-\tilde{\beta}_H(\cdot)$, thus making H pay less coinsurance. The overall effect on the ex-ante budget is $f_L \cdot (1) - f_H \frac{f_L}{f_H} \cdot (1) = 0$ since both perturbations keep treatment choices constant at the margin and $\int \beta^{c,\pi}_t\left(m(s)\right) dG_t(s) = 1 \ \ \forall t$ (again, by Lemma D.1 part i). The total effect on ex-ante expected utility is

$$\begin{aligned} f_L \frac{\partial}{\partial \varepsilon} V_L\left(\tilde{c}_L + \varepsilon\tilde{\beta}_L, \tilde{\pi}_L\right) + f_H \frac{\partial}{\partial \varepsilon} V_H\left(\tilde{c}_H - \varepsilon\frac{f_L}{f_H}\tilde{\beta}_H, \tilde{\pi}_H\right) \\ = -f_L\tilde{\nu}_L + \frac{f_L}{f_H} f_H \tilde{\nu}_H = f_L\left[-\nu_L + \nu_H\right] > 0, \end{aligned}$$

where the first equality is from Lemma D.1 part ii and the inequality follows by the assumption that $\nu_H > \nu_L$. This shows the perturbation is profitable, which is a contradiction, thus completing part i.

[92] If $c \in C^2$, $\beta^{c,\pi}_t(m)$ will be C^1 in m, thus making (G.34) applicable. This can be seen by simply unpacking definitions.

Part ii) Under separable, $D(\hat{m}) = -\frac{U_{11}c'}{U_1} = \frac{\partial}{\partial \hat{m}} \log U_1$. Dropping subscripts on $m_{c,\pi}(\cdot)$, this implies $\int_{m(\underline{s})}^{m(s)} D(\hat{m})d\hat{m} = \int_{m(\underline{s})}^{m(s)} \frac{\partial}{\partial \hat{m}} \log U_1 d\hat{m} = [\log U_1]\,|_{m(\underline{s})}^{m(s)}$, so $L_{c,\pi,\underline{s}}(s) \equiv \exp\left(-\int_{m(\underline{s})}^{m(s)} D(\hat{m})d\hat{m}\right) = \frac{U_1|_{m(\underline{s})}}{U_1|_{m(s)}} = \frac{W_1(\underline{s})}{W_1(s)}$ (note the negative inside the exp). Then,

$$\int W_1^{c,\pi}(s) L_{c,\pi,\underline{s}}(s) dG_t(s) = [W_1^{c,\pi}(\underline{s})] \int W_1^{c,\pi}(s) \frac{1}{W_1^{c,\pi}(s)} dG_t(s) = [W_1^{c,\pi}(\underline{s})].$$

It follows that $\nu_t^{c,\pi} = W_1^{c,\pi}(\underline{s}) \left[\int \frac{W_1^{c,\pi}(\underline{s})}{W_1^{c,\pi}(s)} dG_t(s)\right]^{-1} = \left[\int \frac{1}{W_1^{c,\pi}(s)} dG_t(s)\right]^{-1}$, and so $\nu_t^{c,\pi} = \theta_t^{c,\pi}$ as desired.

Part iii) Given the definition of $\nu_t^{c,\pi}$ in (D.2), this follows immediately from the fact that, under money-loss, $D(m) = 0$ (since $\eta_{NG}^{c,\pi}(m) = 0$). This completes the proof. □

G.4 Single Crossing Lemmas: Proofs

Lemma D.4 (Willingness-to-Pay: Restated)**.** *Let $L < H \in \mathcal{T}$, and let (x, π) be a linear contract with $x \geq 0$ and (unless utility is money-loss) $x \leq 1$.*[93] *Then*

$$\frac{\frac{\partial}{\partial x} V_H(x,\pi)}{\lambda_H^{x,\pi}} > \frac{\frac{\partial}{\partial x} V_L(x,\pi)}{\lambda_L^{x,\pi}}.$$

Proof. Note that $m_{x,\pi}(s)$ is weakly positive by definition and strictly increasing in s for $s \geq s_\pi^*$ by Lem. G.4 iv which applies since, by hypothesis, $x \geq 0$ (i.e., $c' \leq 1$) and either utility is money-loss or $x \leq 1$ (i.e., $c' \geq 0$). Further note that G_t has positive measure above s_π^* since $s_\pi^* < \bar{s}$ and G_t has common support on $[\underline{s}, \bar{s}]$ across t.

By the envelope theorem, we can see that for $t = L, H$, $\frac{\frac{\partial}{\partial x} V_t(x,\pi)}{\lambda_t^{x,\pi}} = \int m_{x,\pi}(s) \frac{g_t(s) W_1^{x,\pi}(s)}{\lambda_t^{x,\pi}} ds$, where (recall) $W_1^{x,\pi}(s)$ is the ex-post marginal utility defined in Definition 6. Thus our goal is then to show

$$\int m_{x,\pi}(s) \frac{g_L(s) W_1^{x,\pi}(s)}{\lambda_L^{x,\pi}} ds < \int m_{x,\pi}(s) \frac{g_H(s) W_1^{x,\pi}(s)}{\lambda_H^{x,\pi}} ds. \tag{G.36}$$

Clearly, $\frac{g_t(s) W_1^{x,\pi}(s)}{\lambda_t^{x,\pi}} \geq 0$ and $\int \frac{g_t(s) W_1^{x,\pi}(s)}{\lambda_t^{x,\pi}} ds = 1$ for each t, so (G.36) is comparing the expectation of $m_{x,\pi}(s)$ with respect to two different probability densities: $\frac{g_t(s) W_1^{x,\pi}(s)}{\lambda_t^{x,\pi}}$ for $t = L, H$. Then by the first paragraph it suffices to show $\frac{g_L(s) W_1^{x,\pi}(s)}{\lambda_L^{x,\pi}} \prec_{\text{MLRP}} \frac{g_H(s) W_1^{x,\pi}(s)}{\lambda_H^{x,\pi}}$. For

[93] For brevity, the main-text statement of this lemma (Lem. 4) didn't mention that, under money-loss, we can dispense with the assumption $x \leq 1$. The same goes for the next two lemmas (Lemmas 5 and 6), which are based on this lemma.

all $s \in \mathbb{R}$,

$$\frac{g_H(s)W_1^{x,\pi}(s)/\lambda_H^{x,\pi}}{g_L(s)W_1^{x,\pi}(s)/\lambda_L^{x,\pi}} = \frac{g_H(s)}{g_L(s)}\frac{\lambda_L^{x,\pi}}{\lambda_H^{x,\pi}},$$

which is strictly increasing in s since $\frac{g_H(s)}{g_L(s)}$ is assumed to be. This concludes the proof.

□

Lemma D.5 (Single Crossing: Restated)**.** *Let $L < H \in \mathcal{T}$. Let $x, x' \in \mathbb{R}_+$ and (unless utility is money-loss) $\max\{x, x'\} \leq 1$. If $x' < x$ and $V_L(x', \pi') \leq V_L(x, \pi)$ for some $\pi', \pi \in \mathbb{R}$, then $V_H(x', \pi') < V_H(x, \pi)$.*

Proof. Define $\check{\pi}_L$ to be the amount such that $V_L(x, \pi - \check{\pi}_L) = V_L(x', \pi')$. [Intuitively, we can think of $\check{\pi}_L$ as compensation.]

We will argue that:

Claim 1: $V_H(x, \pi - \check{\pi}_L) > V_H(x, \pi')$.

Since $V_H(x, \pi)$ is decreasing in π[94], Claim 1 will give us $V_H(x, \pi) \geq V_H(x, \pi - \check{\pi}_L) > V_H(x, \pi')$, and thus conclude the proof.

For each $\ell \in [0, 1]$, let $\pi_{L,\ell}$ be the unique number that preserves L's utility as we vary the coinsurance between x and x' — that is, the unique number $\pi_{L,\ell}$ such that the first indifference holds in:

$$V_L(\ell x + (1-\ell)x', \pi_{L,\ell}) = V_L(x, \pi - \check{\pi}_L) = V_L(x', \pi') \tag{G.37}$$

By the implicit function theorem, $\pi_{L,\ell}$ exists, is locally unique, and is a differentiable function of ℓ [Need conditions for IFT and differentiability! Specifically also show differentiability of V_L]. Observe that $\pi_{L,1} = \pi - \check{\pi}_L$ and $\pi_{L,0} = \pi'$.

By the implicit function theorem applied to the first equation in (G.37),

$$\begin{aligned}\frac{d\pi_{L,\ell}}{d\ell} &= -\frac{(x - x')\frac{\partial}{\partial \text{ first arg}}V_L(\ell x + (1-\ell)x', \pi_{L,\ell})}{\frac{\partial}{\partial \text{ second arg}}V_L(\ell x + (1-\ell)x', \pi_{L,\ell})} \\ &= \frac{(x - x')\frac{\partial}{\partial \text{ first arg}}V_L(\ell x + (1-\ell)x', \pi_{L,\ell})}{\lambda_L^{\ell x + (1-\ell)x', \pi_{L,\ell}}},\end{aligned} \tag{G.38}$$

where the second equality is from G.10.

For $t = L, H$ and $\ell \in [0, 1]$, define $B_t(\ell)$ by

$$B_t(\ell) \equiv V_t(\ell x + (1-\ell)x', \pi_{L,\ell}). \tag{G.39}$$

[94] By Lemma G.10, $\frac{\partial}{\partial \pi}V_H(x, \pi) = -\lambda_H^{x,\pi} < 0$.

By (G.37), $B_L(\ell)$ is constant in $\ell \in [0,1]$, so $B'_L(\ell) = 0$ for $\ell \in [0,1]$.

To show Claim 1, we wish to show

Claim 1.a: $B_H(1) - B_H(0) > 0$, and for that we wish to show

Claim 1.a.a: $\forall \ell \in [0,1], B'_H(\ell) > 0$.

To ease notation, for any t and any contract (x,π) — not to be confused with the specific x and π in the current proof — let $A_t^{x,\pi} \equiv \frac{\partial}{\partial \text{first arg}} V_t(x,\pi)$. For each risk-type t, by the chain rule (and the envelope theorem when introducing λ),

$$B'_t(\ell) = (x-x') \frac{\partial}{\partial \text{ first arg}} V_t\left(\ell x + (1-\ell)x', \pi_{L,\ell}\right) + \frac{d\pi_{L,\ell}}{d\ell} \frac{\partial}{\partial \text{ second arg}} V_t\left(\ell x + (1-\ell)x', \pi_{L,\ell}\right) \tag{G.40}$$

$$= (x-x') A_t^{\ell x+(1-\ell)x', \pi_{L\ell}} + \left[\frac{(x-x') A_L^{\ell x+(1-\ell)x', \pi_{L,\ell}}}{\lambda_L^{\ell x+(1-\ell)x', \pi_{L,\ell}}}\right] \left[-\lambda_t^{\ell x+(1-\ell)x', \pi_{L\ell}}\right]$$

$$= (x-x') \cdot \left[A_t^{\ell x+(1-\ell)x', \pi_{L\ell}} - \frac{A_L^{\ell x+(1-\ell)x', \pi_{L,\ell}}}{\lambda_L^{\ell x+(1-\ell)x', \pi_{L\ell}}} \lambda_t^{\ell x+(1-\ell)x', \pi_{L\ell}}\right],$$

where the second equality uses Lemma G.10 and (G.38).

[Though it was already shown above, I note that setting $t = L$ shows $B'_L(\ell) = 0$ for $\ell \in [0,1]$]. Setting $t = H$ gives

$$B'_H(\ell) = (x-x') \left[A_H^{\ell x+(1-\ell)x', \pi_{L\ell}} - \frac{A_L^{\ell x+(1-\ell)x', \pi_{L,\ell}}}{\lambda_L^{\ell x+(1-\ell)x', \pi_{L\ell}}} \lambda_H^{\ell x+(1-\ell)x', \pi_{L\ell}}\right].$$

Thus since $x > x'$, simple algebra on the previous expression shows we need the following relation in order to show $B'_H(\ell) > 0$:

$$\frac{A_L^{\ell x+(1-\ell)x', \pi_{L,\ell}}}{\lambda_L^{\ell x+(1-\ell)x', \pi_{L,\ell}}} < \frac{A_H^{\ell x+(1-\ell)x', \pi_{L,\ell}}}{\lambda_H^{\ell x+(1-\ell)x', \pi_{L,\ell}}},$$

but Lemma D.4 gives us exactly that; in order to make that lemma apply, we assumed that $x, x' \in \mathbb{R}_+$ and that either utility is money-loss or $x, x' \leq 1$. Thus we have, for all $\ell \in [0,1]$, $B'_H(\ell) > 0$, proving Claim 1a.a.

Now by the fundamental theorem of calculus and Claim 1.a.a, $B_H(1) - B_H(0) = \int_0^1 B'_t(\ell) d\ell > 0$, so $B_H(1) > B_H(0)$, establishing Claim 1.a.

Finally, Claim 1.a and the definition of $B_t(\ell)$, imply that we have the inequality relation

in

$$V_H\left(x, \pi - \check{\pi}_L\right) = V_H\left(x, \pi_{L,1}\right) = V_H\left(\ell x + (1-\ell)x', \pi_{L,\ell}\right)|_{\ell=1} > V_H\left(\ell x + (1-\ell)x', \pi_{L,\ell}\right)|_{\ell=0} = V_H\left(x', \pi_{L,0}\right) = V_H\left(x', \pi'\right),$$

where the outermost equalities come from the observation above that $\pi_{L,1} = \pi - \pi_{LC}$ and $\pi_{L,0} = \pi'$ respectively. Relating the outer expressions, we conclude that $V_H\left(x, \pi - \pi_{LC}\right) > V_H\left(x, \pi'\right)$, proving Claim 1.

As stated above, since $V_H(x, \pi)$ is decreasing in π (and $\check{\pi}_L \geq 0$), we have the first inequality in $V_H\left(x, \pi\right) \geq V_H\left(x, \pi - \check{\pi}_L\right) > V_H\left(x, \pi'\right)$, which concludes the proof. □

Lemma D.6 (Implementability: Restated)**.** *Suppose $(x_t, \pi_t)_{t=1}^T$ satisfies all IC constraints (5.1b) of the full problem (5.1) with $x_t \geq 0$ and (unless utility is money-loss) $x_t \leq 1$ for all t. Then $(x_t)_{t=1}^T$ is weakly increasing in t, and so is $(\pi_t)_{t=1}^T$. Further, $x_t < x_{t+1}$ if and only if $\pi_t < \pi_{t+1}$.*

Proof. Suppose by contradiction that $x_t > x_{t+1}$ for some $t \in \mathcal{T}$. By upward-adjacent IC, $V_t(x_t, \pi_t) \geq V_t(x_{t+1}, \pi_{t+1})$. Then by single crossing (Lem 5) — which applies since $x_\tau \geq 0$ and (outside money-loss) $x_\tau < 1$ for all $\tau \in \mathcal{T}$ — we have that $V_{t+1}(x_t, \pi_t) > V_{t+1}(x_{t+1}, \pi_{t+1})$ since $x_t > x_{t+1}$. But this violates downwards-adjacent IC. This establishes that $(x_t)_{t=1}^T$ is weakly increasing in t. The remaining claims then follow from the fact that each $V_t(\cdot, \cdot)$ is strictly increasing in its first argument and strictly decreasing in its second, as shown in Lemma G.10. Specifically that if one of these claims failed, one contract would be unambiguously worse than another thus violating the IC constraint of the type that's supposed to choose it. This completes the proof.

□

G.5 Proof of Theorem 1/1′

G.5.1 Lemmas for Theorem 1/1′

Proof of Lemma 2

Lemma 2 (Expected Marginal Utility Increasing in t: Restated)**.** *Let utility be money-loss CARA. Suppose $(c_t, \pi_t)_{t=1}^T$ satisfies the upward-adjacent IC constraints (5.2) of the private-risk problem. Then, $\lambda_t^{c_t, \pi_t}$ is strictly increasing in t. In particular, λ_t^* and λ_t^{**} are both strictly increasing in t.*

Proof. Fix $t \in \mathcal{T}$ with $t < T$. Since upward-adjacent IC is satisfied, $V_t(c_t, \pi_t) \geq V_t(c_{t+1}, \pi_{t+1})$. We show $\lambda_{t+1}^{c_{t+1}\pi_{t+1}} > \lambda_t^{c_t\pi_t}$. By Lemma G.5 part i, endogenous ex-post utility $W^{c,\pi}(s)$ is strictly decreasing in s for all c, π.[95] Thus, since $V_t(c,\pi) \equiv \int W^{c,\pi}(s) dG_t(s)$ for all t, the standing assumption that $G_t \prec_{\text{MLRP}} G_{t+1}$ (given that $t < t+1$) implies $V_t(c_{t+1}, \pi_{t+1}) > V_{t+1}(c_{t+1}, \pi_{t+1})$. By transitivity, we have $V_t(c_t, \pi_t) > V_{t+1}(c_{t+1}, \pi_{t+1})$. For any CARA utility function $u(x)$, we have $u'(x) = -\psi u(x)$ for all x where $\psi > 0$ is the coefficient of absolute risk aversion, so under money-loss CARA utility, $W_1^{c,\pi}(s) = -\psi W^{c,\pi}(s)$ for all s, c, π. Thus $\lambda_\tau^{c,\pi} = -\psi V_\tau(c,\pi)$ for all $\tau \in \mathcal{T}$ since $\lambda_\tau^{c,\pi} \equiv \int W_1^{c,\pi}(s) dG_t(s)$. The desired conclusion $\lambda_{t+1}^{c_{t+1}\pi_{t+1}} > \lambda_t^{c_t\pi_t}$ then follows from $V_t(c_t, \pi_t) > V_{t+1}(c_{t+1}, \pi_{t+1})$. □

Proof of Lemma 3/3′

Lemma 3′ (Generalized Upward IC Binds: Restated)**.** *Let utility be general. Suppose* $(x_t^{**}, \pi_t^{**})_{t=1}^T$ *is an optimal solution to the private-risk problem with* $x_t^{**} \geq 0$ *for all* t*. Let* $\mathcal{P}$ *be the set of pools, i.e., the set of equivalence classes of types defined by the relation* $t \sim t'$ *iff* $x_t^{**} = x_{t'}^{**}$*.*

Suppose either utility is money-loss CARA, or $\left[\sum_{t\in p} f_t \lambda_t^{**}\right] \cdot \left[\sum_{t\in p} f_t \frac{\partial}{\partial \pi} P_t(x_t^{**}, \pi_t^{**})\right]^{-1}$ *is strictly increasing in* p *over the set of pools* $\mathcal{P}$ *and* $x_t^{**} \leq 1$ *for all* t*.*[96] *Then all upward-adjacent ICs bind at* $(x_t^*, \pi_t^*)_{t=1}^T$*:* $V_{t-1}(x_{t-1}^*, \pi_{t-1}^*) = V_{t-1}(x_t^*, \pi_t^*)$*.*

Further, the same result holds if $(x_t^{**}, \pi_t^{**})_{t=1}^T$ *is replaced by any optimal solution* $(x_t^*, \pi_t^*)_{t=1}^T$ *to the FOA-relaxed problem.*

Finally, in the nonlinear coinsurance case, upward-adjacent IC also binds at the FOA-relaxed solution under money-loss CARA, as asserted in Lemma 3.[97]

Proof. For the FOA-relaxed problem, if $(t-1)$'s upward-adjacent IC doesn't bind, we can show a contradiction by simply conducting a budget-neutral premium transfer from $t-1$ to t. By assumption, this preserves IC at the margin (given continuity of V_{t-1}). It also increases ex-ante expected utility by increasingness of λ_t^* (Lem 2), which holds whenever upward-adjacent IC is satisfied and thus in both the full and relaxed FOA-problems. This argument goes through not only for linear coinsurance — as in the other parts of the present lemma — but also for the *nonlinear*-coinsurance problem since Lemma 2 does as well.[98]

[95] Lemma G.5 part i is where the standing assumption B2 that $b_2(m,s) < 0$ is used.

[96] Under money-loss CARA, the increasingness condition reduces to strict increasingness of λ_t^{**} which holds there by Lem 2. $x_t^{**} \leq 1$ is not needed under money-loss utility (CARA or otherwise) to use Lemmas 5 and 6 which is the only place it will be used.

[97] For nonlinear coinsurance, I have not extended this result to the full problem since with nonlinear, we lack the single crossing property established for linear in Lem. 5. However, the proof shows it does extend to the intermediate problem that only ignores non-adjacent upward ICs while still considering all downward ICs.

[98] With non-linear coinsurance the argument does not obviously extend to the full problem since without

Now for the full-problem part. In what follows, I will omit discussion the role of the hypothesis $x_t^* \leq 1$; as discussed in Footnote 96, it is needed, outside money-loss utility, to invoke Single Crossing (Lem 5) and its corollary Lemma 6.

Although the full-problem part of the result has a disjunctive hypothesis, it suffices to prove only the general-utility disjunct for the following reason: the extra hypotheses required by that case have been shown to hold under money-loss CARA in Lemma 2; see Footnote 96 for more detail.

For all pools $p \in \mathcal{P}$, let t_p^+ and t_p^- denote the highest and lowest risk-types in p. By Lemma 6 (which applies because we are in the full-problem case), $(x_t^*)_{t\in\mathcal{T}}$ is weakly increasing so $t_p^+ < t_{p+1}^-$ for each p for which these objects exist. Clearly, if t is in a pool p but not the highest type therein (i.e. $t \in p$ but $t < t_p^+$), t's upward-adjacent binds IC trivially. It remains to address the case $t = t_p^+$ for some pool p.

We prove the result by induction on p starting from the base case of $p = |\mathcal{P}| - 1$ so that $p + 1$ is the highest pool. Suppose for contradiction that upward-adjacent IC does not bind at $p = |\mathcal{P}| - 1$. Let q be the highest pool that is *weakly* lower than p such that downward-adjacent IC doesn't bind at t_q^-. q must exist since, at a minimum, "downward IC" doesn't bind at $t = 1$ since it doesn't exist. Single crossing (Lem. 5) implies q's non-adjacent downward ICs don't bind either. For $r \in \{q, ..., p\}$, I *claim* t_r^+'s upward ICs also don't bind. Clearly if $q = p$, this is true by hypothesis. If instead $q < p$, then, for $r \in \{q, ..., p\}$, t_{r+1}^-'s downward IC binds by definition of q, so single crossing and $x_{t_r^+}^{**} < x_{t_{r+1}^-}^{**}$ imply t_r^+'s upward-adjacent IC doesn't bind; non-bindingness of t_r^+'s non-adjacent upward ICs then follows (a fortiori) by single crossing, establishing the claim.

We now use the premiums to uniformly transfer money in a budget-neutral way from each type in q to each type in $p + 1$, keeping the pools intact.[99] Before describing the deviation precisely, note that it is IC preserving because of the following factors: $p + 1 = |\mathcal{P}|$ (using the *base-case assumption*), q and $p + 1$ are (respectively) made unambiguously worse- and better-off, t_q^-'s downward IC doesn't bind, and we just argued t_r^+'s upward ICs don't bind for $r \in \{q, ..., p\}$. Now to specify the deviation, in the space of coordinates (π_q, π_{p+1}), consider

linearity, we lack single crossing (Lem 5). However, the argument can be extended if we ignore only the non-adjacent upward ICs while still considering all downward ICs. In that case, the deviation is slightly more complex. Specifically, the deviation sends money uniformly to risk-types $\{t, ..., T\}$ through the premiums. This costs $1 - F_{t-1}$ dollars in expectation across risk-types. The ex-ante welfare effect is $\sum_{\tau=t}^T f_\tau \lambda_t^*$. To pay for this, draw $1 - F_{t-1}$ dollars uniformly from all risk-types through the premiums. This saves $1 - F_{t-1}$ dollars in expectation. The ex-ante welfare effect of that is $[1 - F_{t-1}] \sum_{\tau=1}^T f_\tau \lambda_\tau^*$. Overall, this is budget neutral and has ex-ante welfare effect $[1 - F_{t-1}] \left[-\sum_{\tau=1}^T f_\tau \lambda_\tau^* + \sum_{\tau=t}^T \frac{f_\tau}{1-F_{t-1}} \lambda_t^* \right]$, which can be seen to be strictly positive since λ_τ^* is strictly increasing in τ by Lemma 2. It's easy to see that under money-loss CARA, the uniform premium transfer from all types doesn't effect incentives for plan choice. Thus, the deviation marginally preserves all ICs considered.

[99] The deviation is essentially the same as that in the proof of Prop. 1 part ii, just now allowing pooling.

the direction

$$\left(\left[\sum_{t\in q} f_t \frac{\partial}{\partial \pi} P_t(x_t^{**}, \pi_t^{**})\right]^{-1}, -\left[\sum_{t\in p+1} f_t \frac{\partial}{\partial \pi} P_t(x_t^{**}, \pi_t^{**})\right]^{-1}\right), \tag{G.41}$$

so that we are increasing π_q and decreasing π_p, keeping x_q^{**}, x_{p+1}^{**} constant and leaving constant $(x_\tau^{**}, \pi_\tau^{**})$ for all $\tau \in \mathcal{P}\backslash\{q, p+1\}$. This strictly increases ex-ante expected utility by the assumption that $\left[\sum_{t\in p} f_t \lambda_t^{**}\right] \cdot \left[\sum_{t\in p} f_t \frac{\partial}{\partial \pi} P_t\left(x_t^{**}, \pi_t^{**}\right)\right]^{-1}$ is increasing in t. That it's budget neutral is easy to see and was shown (for the non-pooling case) in the proof of Prop. 1 part ii. Thus the deviation is feasible and profitable, which contradicts optimality of $(x_t^{**}, \pi_t^{**})_{t\in\mathcal{T}}$, completing the base case $p = |\mathcal{P}| - 1$.

Now for the inductive step, let $p < |\mathcal{P}| - 1$ and suppose upward IC binds at t_r^+ for all pools $r > p$. By single crossing then, downward IC doesn't bind at pools strictly above $p+1$. Since the base-case assumption ($p = |\mathcal{P}| - 1$) was only used to justify that we didn't need to worry about downward IC at pools strictly above $p+1$ (since there were no such pools), the same argument-by-contradiction used in the base-case can now be used to show upward-adjacent IC binds at t_p^+. This concludes the proof. □

The next lemma supports the proof of Theorem 1/ 1′. by showing $x_t^* \geq \hat{x}_t$ (and more generally $x_t^* \geq \hat{x}_{tS_t^*}$ for general utility). The proof of Theorem 1/1′ will then show that this inequality is also strict. Note that unlike some results in this appendix, it is not a restatement of a result from a prior section.

Lemma G.16 (Weak Distortion)**.** *Let utility be general and suppose $(x_t^{**}, \pi_t^{**})_{t=1}^T$ is an optimal private-risk solution with $x_t^{**} \geq 0$ and — if utility is not money-loss — $x_t^{**} \leq 1$. Suppose that, for all t and S, the static problem is strictly concave with solution $\hat{x}_{tS}$. or utility is money-loss.*

i) Suppose upward-adjacent ICs bind for all t.[100] *For all $t > 1$, if $t = \max\{\tau \in \mathcal{T} | x_\tau^{**} = x_t^{**}\}$, then $x_t^{**} \geq \hat{x}_{tS_t^{**}}$. Further, if $(x_t^*, \pi_t^*)_{t=1}^T$ is an optimal solution to the FOA-relaxed problem or an optimal solution to the relaxed private-risk problem where only upward ICs (including non-adjacent) are considered (in either case with $x_{tS_t}^* \leq 1$), then $x_t^* \geq \hat{x}_{tS_t^*}$ for all $t > 1$.*

ii) Again, suppose upward-adjacent ICs bind for all t. For any $p \subset \mathcal{T}$, let $G_p(s) = \sum_{t\in p} \frac{f_t}{\sum_{\tau\in p} f_\tau} G_t(s)$ and suppose the static problem corresponding to G_p is strictly concave and has a solution $\hat{x}_{pS} \leq 1$ for every S. Then for each pool $p \in \mathcal{P}\backslash\{1\}$ (where $\mathcal{P}$ is the set of pools), $x_t^{**} \geq \hat{x}_{pS_p^{**}}$ for each $t \in p$, where $S_p^{**} \equiv -\sum_{t\in p} \frac{f_t}{\sum_{\tau\in p} f_\tau} P_t(x_t^{**}, \pi_t^{**})$.

[100] Sufficient conditions for this were given in Lem. 3′.

iii) $x_1^{**} \leq \hat{x}_{1S_1^{**}}$ *for the lowest risk-type* $t = 1$*. If all downward ICs pointing to* $t = 1$ *are slack, then* $x_1^{**} = \hat{x}_{1S_1^{**}}$*. Further, if* $(x_t^*, \pi_t^*)_{t=1}^T$ *is optimal solution to the FOA-relaxed solution, then* $x_1^* = \hat{x}_{1S_1^*}$*.*

Proof. i) The proof for the FOA-relaxed problem — as well as the only upward ICs problem — is basically the same argument as in the full problem but we no longer need the hypotheses $t = \max\{\tau \in \mathcal{T} | x_\tau = x_t\}$ and bindingness of upward IC. I omit the details, and instead only prove the part about the fully optimal solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$.

Suppose for contradiction that $x_t^{**} < \hat{x}_{tS_t^{**}}$ for some t such that $t = \max\{\tau \in \mathcal{T} | x_\tau^{**} = x_t^{**}\}$. This implies that either $t = T$ or $x_t^{**} < x_{t+1}^{**}$. If $t = T$, clearly we don't need to worry about downward IC. If $x_t^{**} < x_{t+1}^{**}$, our hypothesis that t's upward-adjacent IC binds together with single crossing (Lem 5) implies $(t+1)$'s downward IC doesn't bind. Then, using our hypothesis that (indeed) all upward-adjacent ICs bind allows us to iteratively apply single crossing (Lem. 5) and transitivity to show non-adjacent downward ICs pointing towards t bind don't bind either.[101] As a result, in either case ($t = T$ or $x_t^{**} < x_{t+1}^{**}$), local deviations will not violate downward IC, given continuity of the value functions V_τ for $\tau \in \mathcal{T}$.

Recall $A_t^{x,\pi} \equiv \frac{\partial}{\partial x} V_t(x, \pi)$ (from (G.18)) and π_{txS} is the unique premium such that $V_t(x, \pi_{txS}) = -S$ (Def. (7)). Further, let $A_\tau^t \equiv A_\tau^{x_t^{**}, \pi_t^{**}}$ and $A_t \equiv A_t^t$ for any risk-types τ and t, and similarly let $\lambda_\tau^t \equiv \lambda_\tau^{x_t^{**}, \pi_t^{**}}$ and $\lambda_t \equiv \lambda_t^t$. Additionally, to ease notational burden, I sometimes omit the $**$ superscript on S_t^{**} which will be always intended for the rest of the proof of part i.

There are two cases: $\frac{A_{t-1}^t}{\lambda_{t-1}^t} < \frac{\partial \pi_{txS_t}}{\partial x}|_{x=x_t^{**}}$ and $\frac{A_{t-1}^t}{\lambda_{t-1}^t} > \frac{\partial \pi_{txS_t}}{\partial x}|_{x=x_t^{**}}$. In either case, $\frac{A_{t-1}^t}{\lambda_{t-1}^t} < \frac{A_t}{\lambda_t}$ by Lemma 4.

Case 1: $\frac{A_{t-1}^t}{\lambda_{t-1}^t} < \frac{\partial \pi_{txS_t}}{\partial x}|_{x=x_t^{**}}$. Starting at (x_t^{**}, π_t^{**}), consider marginally increasing t's coverage x_t and simultaneously increasing t's premium π_t at a rate that preserves the budget. Formally, for each unit of increase in x_t, increase π_t by $\frac{\partial \pi_{txS_t}}{\partial x}|_{x=x_t^{**}}$. By definition of π_{txS_t} (Definition 7), the budget is preserved. To see $t-1$'s upward IC is preserved, note that the assumption of the case guarantees this. Further, single crossing then implies the same for lower types than $t-1$.[102] That t is made strictly better off follows from the following factors: $x_t^{**} < \hat{x}_{tS_t}$, the strict concavity of the static problem, and the definitions of $\frac{\partial \pi_{txS}}{\partial x}$ and $\hat{x}_{tS_t}$. As argued above, we don't need to worry about downward IC. Thus this is a profitable marginal deviation in this case.

[101] For a more explicit argument, see the proof of Prop. F.1 i.

[102] To see this formally for lower types, suppose for contradiction, the deviation strictly attracted some type $\tau < t-1$. Then τ would strictly prefer t's perturbed contract over not only her own but (by transitivity) over $t-1$'s. But then by single crossing (given $x_\tau^{**} \leq x_{t-1}^{**}$ which holds by Lem. (6)), $t-1$ must also be strictly attracted to t's perturbed contract. But this contradicts that $t-1$ is not strictly attracted thereto which we just established before this footnote.

Case 2: $\frac{A^t_{t-1}}{\lambda^t_{t-1}} > \frac{\partial \pi_{txS_t}}{\partial x}|_{x=x^{**}_t}$. Starting at (x^{**}_t, π^{**}_t), consider marginally increasing t's coverage x_t and simultaneously increasing π_t at a rate that leaves $(t-1)$'s value of t's contract the same. Formally, for each unit of increase in x_t, increase π_t by $\frac{A^t_{t-1}}{\lambda^t_{t-1}}$. Clearly, $t-1$'s upward IC is preserved, and single crossing implies the same for lower types than $t-1$.[103] By the assumption of case 2, the budget is unaffected since (recall) $\frac{\partial \pi_{txS_t}}{\partial x}|_{x=x^{**}_t}$ is the rate of increase of the premium needed to preserve the budget. Since $\frac{A^t_{t-1}}{\lambda^t_{t-1}} < \frac{A_t}{\lambda_t}$, the deviation makes t better off. Since now downward ICs bound in the first place, continuity implies none are violated by the local deviation. As argued above, we don't need to worry about downward IC. Thus this is a profitable marginal deviation in this case.

Altogether, we've shown there is a profitable marginal deviation, a contradiction. This concludes the proof. Finally, it's easy to see that the two extra hypotheses of part i are not needed for the above argument in the case of $t = T$ since there is no type above to worry about attracting to t.

ii) The argument is very similar to that of part i — I merely mention the key differences. Instead of just perturbing one type's contract, here we uniformly perturb (x^{**}_t, π^{**}_t) for every type t in the pool p. We use the fact that for pools $p < q$, the MLRP order on individual types implies $G_p(s) \equiv \sum_{t\in p} \frac{f_t}{\sum_{\tau\in A} f_\tau} G_t(s) \prec_{\text{MLRP}} \sum_{t\in q} \frac{f_t}{\sum_{\tau\in q} f_\tau} G_t(s) \equiv G_q(s)$. Lastly, we need to define $\frac{\partial \pi_{pxS^{**}_t}}{\partial x}$ for a pool p of risk-types in the obvious way by analogy to the definition of $\frac{\partial \pi_{txS^{**}_t}}{\partial x}$ (Definition 7).

iii) For the fully optimal $(x^{**}_t, \pi^{**}_t)_{t=1}^T$, the argument for $x^{**}_1 \leq \hat{x}_{1S^{**}_1}$ is exactly analogous to that for part i above, specifically for the case $t = T$ mentioned in the first paragraph thereof (since $t = 1$ and T are extremal types). Here, the deviation will move x^{**}_1 downward instead of upward. I omit further details, which are, again, analogous. For the FOA-relaxed solution $(x^*_t, \pi^*_t)_{t=1}^T$, the equality $x^*_1 = \hat{x}_{1S^*_1}$ is obvious given the definitions of $\hat{x}_{1S^*_1}$ and S^*_1 since under the FOA-relaxed problem, there are no ICs pointing to $t = 1$; see the proof of Prop. 1 i for a more formal argument to this effect. □

G.5.2 Proof of Theorem 1/1′

Theorem 1′ (Generalized Main Result: Restated)**.** *Let utility be general. Suppose that the static problem is strictly concave for every t and S.*

Suppose $(x^*_t, \pi^*_t)_{t=1}^T$ is an optimal solution to either the FOA-relaxed problem or the problem with only upward ICs (including non-adjacent ones) with $x_t \geq 0$.

i) Suppose that either utility is money-loss CARA, or $\lambda_t^{x^*_t,\pi^*_t} \left[\frac{\partial}{\partial \pi} P_t(x^*_t, \pi^*_t)\right]^{-1}$ is strictly

[103] The argument is the same as in Case 1 — see Footnote 102 there.

increasing in t and $x_t^* \leq 1$ for all t.[104] *Then $x_t^* > \hat{x}_{tS_t^*}$ for all $t > 1$.*
For parts ii and iii, suppose $(x_t^{**}, \pi_t^{**})_{t=1}^T$ is an optimal solution to the full private-risk problem with $x_t^{**} \leq 1$ for all t. Let $\mathcal{P}$ be the set of pools — i.e., the set of equivalence classes of types defined by the relation $t \sim t'$ iff $x_t^{**} = x_{t'}^{**}$. Since all pools consist of an interval of risk-types, let $\mathcal{P}$ inherit the order from the set of risk-types $\mathcal{T}$. For all pools $p \in \mathcal{P}$, let t_p^+ denote the highest risk-type in p.

ii) Suppose that either utility is money-loss CARA or the following hypotheses hold: $\left[\sum_{t\in p} f_t \lambda_t^{**}\right] \cdot \left[\sum_{t\in p} f_t \frac{\partial}{\partial \pi} P_t\left(x_t^{**}, \pi_t^{**}\right)\right]^{-1}$ is strictly increasing in p over the set of pools $\mathcal{P}$, $x_t^{**} \leq 1$ for all t, and for all $p \in \mathcal{P}$, $\left[\sum_{t\in p\setminus\{t_p^+\}} f_t \lambda_t^{**}\right] \cdot \left[\sum_{t\in p\setminus\{t_p^+\}} f_t \frac{\partial}{\partial \pi} P_t\left(x_t^{**}, \pi_t^{**}\right)\right]^{-1} < \lambda_{t_p^+}^{**}\left[\frac{\partial}{\partial \pi} P_{t_p^+}\left(x_{t_p^+}^{**}, \pi_{t_p^+}^{**}\right)\right]^{-1}$. [Note that a sufficient condition for the latter is that $|\frac{\partial}{\partial \pi} m_{x_t^{**}, \pi_t^{**}}(s)|$ is weakly decreasing in s.] Then $x_{t_p^+}^{**} > \hat{x}_{t_p^+ S_{t_p^+}^{**}}$ for all $p \in \mathcal{P}$ such that $t_p^+ > 1$.

iii) Suppose that either utility is money-loss CARA or the following hypotheses hold: $\left[\sum_{t\in p} f_t \lambda_t^{**}\right] \cdot \left[\sum_{t\in p} f_t \frac{\partial}{\partial \pi} P_t\left(x_t^{**}, \pi_t^{**}\right)\right]^{-1}$ is strictly increasing in p over the set of pools $\mathcal{P}$ and $x_t^{**} \leq 1$ for all t. Then $x_p^{**} > \hat{x}_{pS_p^{**}}$ for all $p \in \mathcal{P}$ satisfying $p \cap \{1\} = \emptyset$, where $(x_p^{**}, \pi_p^{**}) \equiv (x_t^{**}, \pi_t^{**})$ for all $t \in p$, $S_p^{**} \equiv -\sum_{t\in p} \frac{f_t}{\sum_{\tau\in p} f_\tau} P_t(x_t^{**}, \pi_t^{**})$, and $\hat{x}_{pS_p^{**}}$ is defined as follows. $\hat{x}_{pS_p^{**}}$ denotes the unique solution to the static problem (4.4) with distribution $G_p(s) \equiv \sum_{t\in p} \frac{f_t}{\sum_{\tau\in p} f_\tau} G_t(s)$ and subsidy S_p^{**}.

Proof. Note: an informal argument for this result is given after the statement Theorem 1 in the main text (Section 6.2).

I give a full proof only for part ii but mention the key differences for parts i and iii, which are very similar.

i) The argument is very similar to that for part ii, though simpler since we only need to worry about upward IC (or even only upward-adjacent IC). As in part ii below, we apply Lemma G.16 part i but the sentence thereof concerning the FOA-relaxed solution. Note that under money-loss CARA, Lem. 2 applies even at the relaxed-optimal solution, so that $\lambda_t^{x_t^*, \pi_t^*}\left[\frac{\partial}{\partial \pi} P_t\left(x_t^*, \pi_t^*\right)\right]^{-1}$ is strictly increasing in t given that (under money-loss) $\frac{\partial}{\partial \pi} P_t\left(x_t^*, \pi_t^*\right) = 1$. I omit further details.

ii) Under money-loss CARA, it follows from Lemma 2 that the extra hypotheses that are assumed outside money-loss CARA hold, in both the relaxed and full problems; under money-loss CARA, $x_t^{**} \leq 1$ is not needed to apply Lemmas 5 and 6 which are the only place it will be used outside money-loss.

Suppose for contradiction that $x_{t_p^+}^{**} \leq \hat{x}_{t_p^+ S_t^{**}}$ for some p with $t_p^+ > 1$. The first step is to

[104] Under money-loss CARA, the increasingness condition holds by Lem. 2 (and by $\frac{\partial}{\partial \pi} P_t = 1$) at the optimal solutions to both the FOA-relaxed and full problems. The latter condition ($x_t^{**} \leq 1$) is not needed under money-loss in order to apply Lemmas 5 and 6, which are the only places it is otherwise used.

apply Lemma G.16 (part i) — whose hypotheses we satisfy for t_p^+ by Lem. 3′, which applies given our hypothesis that $\left[\sum_{t\in p} f_t\lambda_t^{**}\right]\cdot\left[\sum_{t\in p} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^{**},\pi_t^{**}\right)\right]^{-1}$ is strictly increasing in p (and that $x_t^*\leq 1$). This shows $\hat{x}_{t_p^+ S^{**}_{t_p^+}}\leq x^{**}_{t_p^+}$ (since $t_p^+>1$), thus leaving only the case $x^{**}_{t_p^+}=\hat{x}_{t_p^+ S^{**}_{t_p^+}}$, which obviously also implies $\pi^{**}_{t_p^+}=\hat{\pi}_{t_p^+ S^{**}_{t_p^+}}$. The remainder of the proof is dedicated to deriving a contradiction from this case, so suppose $x^{**}_{t_p^+}=\hat{x}_{t_p^+ S^{**}_{t_p^+}}$.

Recall $A_t^{x,\pi}\equiv\frac{\partial}{\partial x}V_t(x,\pi)$ (from (G.18)) and π_{txS} is the unique premium such that $V_t(x,\pi_{txS})=-S$ (Def. (7)). Further, let $A_\tau^t\equiv A_\tau^{x_t^{**},\pi_t^{**}}$ and $A_t\equiv A_t^t$ for any risk-types τ and t, and similarly let $\lambda_\tau^t\equiv\lambda_\tau^{x_t^{**},\pi_t^{**}}$ and $\lambda_t\equiv\lambda_t^t$.

Consider perturbing t_p^+'s contract in the direction $\left(1,\frac{\partial\pi_{txS_t^{**}}}{\partial x}|_{x=x^{**}_{t_p^+}}\right)$ in the coordinate space $(x_{t_p^+},\pi_{t_p^+})$, starting at $(x^{**}_{t_p^+},\pi^{**}_{t_p^+})$. By definition of the notation π_{txS}, $\lim_{\varepsilon\downarrow 0}P_{t_p^+}\left(x^{**}_{t_p^+}+\varepsilon,\pi^{**}_{t_p^+}+\varepsilon\frac{\partial\pi_{txS_t^{**}}}{\partial x}|_{x=x^{**}_{t_p^+}}\right)=$ 0. Since by assumption $x^{**}_{t_p^+}=\hat{x}_{t_p^+ S^{**}_{t_p^+}}$, the definition of $\hat{x}_{t_p^+ S^{**}_{t_p^+}}$ implies $\lim_{\varepsilon\downarrow 0}V_{t_p^+}\left(x^{**}_{t_p^+}+\varepsilon,\pi^{**}_{t_p^+}+\varepsilon\frac{\partial\pi_{txS_t^{**}}}{\partial x}|_{x=x^{**}_{t_p^+}}\right)=$ 0. But from Lemma G.11 — and the equality $\left(x^{**}_{t_p^+},\pi^{**}_{t_p^+}\right)=\left(\hat{x}_{t_p^+ S^{**}_{t_p^+}},\hat{\pi}_{t_p^+ S^{**}_{t_p^+}}\right)$, — $\frac{\partial\pi_{txS_t^{**}}}{\partial x}|_{x=x^{**}_{t_p^+}}=\frac{A_{t_p^+}}{\lambda_{t_p^+}}$ so that the deviation can be written $\left(1,\frac{A_{t_p^+}}{\lambda_{t_p^+}}\right)$.

By Lem. 4 we have $\frac{A_{t_p^+}}{\lambda_{t_p^+}}>\frac{A_\tau^{t_p^+}}{\lambda_\tau^{t_p^+}}$ for all risk-types $\tau<t_p^+$, so deviating t_p^+'s contract marginally as described above strictly relaxes all upward ICs $\tau\to t_p^+$ for such τ. What about downward IC's pointing towards t_p^+? By our hypotheses, Lem. 3′ implies that for all pools $q\in\mathcal{P}$, t_q^+'s upward-adjacent ICs bind (except $q=|\mathcal{P}|$). By definition of pools, for all $q\in\mathcal{P}$, $x^{**}_{t_q^+}<x^{**}_{t_q^++1}=x^{**}_{t^-_{q+1}}$, so iterative application of single crossing and transitivity implies slack in all downward ICs pointing from one pool to a *distinct* lower pool.[105] By continuity of V_t, such slack will be preserved by local deviations.

Thus no IC will be locally violated by a sufficiently small, further marginal budget-preserving premium transfer to t_p^+ from the group of types just below it (which exists since $t_p^+>1$). More precisely, IC will not be violated by a sufficiently small marginal budget-preserving transfer to t_p^+ from *either* the pool $p\backslash\{t_p^+\}$ — in the event it's non-empty — *or* the pool $p-1$ otherwise, i.e., if p is a trivial pool. Formally, the deviation is similar to the one in the proof of Lem. 3′ (Eqn. G.41) and also to the one in the proof of Prop. 1 ii (Eqn. G.33). Such a transfer strictly increases ex-ante expected utility due to — in the former case — the hypothesis $\left[\sum_{t\in p\backslash\{t_p^+\}} f_t\lambda_t^{**}\right]\cdot\left[\sum_{t\in p\backslash\{t_p^+\}} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^{**},\pi_t^{**}\right)\right]^{-1}<\lambda^{**}_{t_p^+}\left[\frac{\partial}{\partial\pi}P_{t_p^+}\left(x^{**}_{t_p^+},\pi^{**}_{t_p^+}\right)\right]^{-1}$ and — in the latter case — the hypothesis that $\left[\sum_{t\in p} f_t\lambda_t^{**}\right]\cdot\left[\sum_{t\in p} f_t\frac{\partial}{\partial\pi}P_t\left(x_t^{**},\pi_t^{**}\right)\right]^{-1}$ is

[105] For an explicit example of this sort of argument, see the proof of Prop. F.1 i.

strictly increasing in p.[106] In the statement of part ii, a sufficient condition for the former hypothesis is given and its sufficiency alleged. Sufficiency thereof can be shown by an obvious extension of Lemma G.9 part v. Altogether, we've shown a joint deviation — the premium-compensated increase in coverage together with the budget-neutral premium transfer — that is profitable and feasible, thus contradicting optimality of $(x_t^{**}, \pi_t^{**})_{t=1}^T$. This concludes the proof.

iii) The argument is very similar to that for part ii — I note only some key differences. Here, we apply Lemma G.16 part ii instead of part i thereof. The initial direction of perturbation is now $\left(1, \frac{\partial \pi_{px S_p^{**}}}{\partial x}|_{x=x_p^{**}}\right)$, where for each $p \in \mathcal{P}$ (and any x, S), the notation π_{pxS} is defined in the obvious way given $G_p(s)$. It is applied uniformly to the entire pool of types p. The budget-neutral premium transfer is now a uniform transfer from $p-1$ to p.

□

G.6 Proofs of Proposition E.1 and Fact E.1

Fact E.1. *Let L, H be any two risk-types in $\mathcal{T}$ with $L < H$. $V_L(\hat{c}_L, \hat{\pi}_L) > V_L(\hat{c}_H, \hat{\pi}_H)$.*

Proof. Let $\pi_{L\hat{c}_H}$ be L's unique actuarial fair premium for coverage $\hat{c}_H$, as defined formally in Definition 7 of Appendix G.1.2. Since $G_L \prec_{\text{MLRP}} G_H$ and $V_H(c, \pi)$ is strictly decreasing in π (Lemma G.10), it's easy to see that $\pi_{L\hat{c}_H} < \hat{\pi}_H$. Then

$$V_L(\hat{c}_L, \hat{\pi}_L) \geq V_L(\hat{c}_H, \pi_{L\hat{c}_H}) > V_L(\hat{c}_H, \hat{\pi}_H),$$

where the first inequality is by definitions of $(\hat{c}_L, \hat{\pi}_L)$ and $\pi_{L\hat{c}_H}$ — in particular, note $(\hat{c}_H, \pi_{L\hat{c}_H})$ is feasible in L's unsubsidized static problem (4.4) — and the second inequality is by $\pi_{L\hat{c}_H} < \hat{\pi}_H$. □

Proposition E.1. *Suppose $\mathcal{T} = \{L, H\}$ with $L < H$. Suppose $(x_t^*, \pi_t^*)_{t=L,H}$ is an optimal private-risk solution with $x_t \geq 0$ for all t. Suppose either utility is CARA money-loss, or $\lambda_H^* \left[\frac{\partial}{\partial \pi} P_H\left(x_H^*, \pi_H^*\right)\right]^{-1} > \lambda_L^* \left[\frac{\partial}{\partial \pi} P_L\left(x_L^*, \pi_L^*\right)\right]^{-1}$ and (outside money-loss) $x_L^*, x_H^* \leq 1$. Suppose also that there exist a unique solution $\hat{x}_L$ to L's static problem (4.4). Then $S_H^* > 0 > S_L^*$.*

Proof. It should be obvious that $S_L^* + S_H^* = 0$ since $V_t(x, \pi)$ is strictly decreasing in π (Lemma

[106] More explicitly, the transfer is, in the former case, $\left(\left[\sum_{\tau \in t_p \setminus \{t_p^+\}} f_\tau \frac{\partial}{\partial \pi} P_\tau(x_\tau^{**}, \pi_\tau^{**})\right]^{-1}, -f_{t_p^+} \frac{\partial}{\partial \pi} P_{t_p^+}(x_{t_p^+}^{**}, \pi_{t_p^+}^{**})^{-1}\right)$ in the space of premiums $\left(\pi_{t_p \setminus \{t_p^+\}}, \pi_{t_p^+}\right)$ and $\left(\left[\sum_{\tau \in t_{p-1}} f_\tau \frac{\partial}{\partial \pi} P_\tau(x_\tau^{**}, \pi_\tau^{**})\right]^{-1}, -f_{t_p^+} \frac{\partial}{\partial \pi} P_{t_p^+}(x_{t_p^+}^{**}, \pi_{t_p^+}^{**})^{-1}\right)$ in the space $\left(\pi_{t_{p-1}}, \pi_{t_p^+}\right)$, and in the latter case.

G.10).[107]

Thus it suffices to show $S_L^* < S_H^*$. So suppose for contradiction that $S_H^* \leq S_L^*$. Then since $S_L^* + S_H^* = 0$, $S_H^* \leq 0 \leq S_L^*$.

First consider the pooling case: $(x^p, \pi^p) \equiv (x_L^*, \pi_L^*) = (x_H^*, \pi_H^*)$. The fact that $L \prec_{\text{MLRP}} H$ implies the inequality in $-S_L^* = P_L(x^p, \pi^p) > P_H(x^p, \pi^p) = -S_H^*$ by definition of P_t (Eqn. 4.3) and Lemma G.4 iv.[108] Thus $S_L^* < S_H^*$, a contradiction.

Now consider the separation case: $x_L^* < x_H^*$; recall $x_H^* < x_L^*$ is ruled out by Lemma 6.

By Lemma 3′, the disjunctive hypothesis of the present result implies upward IC binds. Our goal is to use Lemma E.1 to show upward IC doesn't bind, which will be a contradiction.

Since upward IC binds and $x_L^* < x_H^*$, single crossing (Lem 5) implies — given that $x_t^* \leq 1$ for money-loss and $x_t \geq 0$ — that downward IC doesn't bind. Then $x_L^* = \hat{x}_{LS_L^*}$ by Lemma G.16 iii — specifically the part about the fully optimal solution (x_t^*, π_t^*) when all downward ICs pointing to $t = 1$ (i.e. $t = L$) are slack. We observe one more fact for use below.

Observation: if $S' \leq S$, then $V_t(\hat{x}_{tS'}, \hat{\pi}_{tS'}) \leq V_t(\hat{x}_{tS}, \hat{\pi}_{tS})$ for all t, since $P_t(\hat{x}_{tS'}, \hat{\pi}_{tS'}) \leq S' \leq S$ by definition of $(\hat{x}_{tS'}, \hat{\pi}_{tS'})$.

Now we show upward IC doesn't bind ($V_L(x_L^*, \pi_L^*) > V_L(x_H^*, \pi_H^*)$) and thus a contradiction (since we showed it does bind above), from the following series of relations (which wraps and is justified afterwards):

$$\begin{aligned} V_L(x_L^*, \pi_L^*) = V_L(\hat{x}_{LS_L^*}, \hat{\pi}_{LS_L^*}) \geq V_L(\hat{x}_L, \hat{\pi}_L) \\ > V_L(\hat{x}_H, \hat{\pi}_H) \geq V_L(\hat{x}_{HS_H^*}, \hat{\pi}_{HS_H^*}) \geq V_L(x_H^*, \pi_H^*). \end{aligned}$$

The equality is from $x_L^* = \hat{x}_{LS_L^*}$, as was shown above, and from the fact that the latter equality implies $\pi_L^* = \hat{\pi}_{LS_L^*}$ given the definitions thereof. The first inequality is by the "observation" and $0 \leq S_L^*$. The second inequality is by Lemma E.1. To see the third inequality, note that $V_H(\hat{x}_H, \hat{\pi}_H) \geq V_H(\hat{x}_{HS_H^*}, \hat{\pi}_{HS_H^*})$ by the "observation" and $S_H^* \leq 0$; then single crossing (Lem 5) implies $V_L(\hat{x}_H, \hat{\pi}_H) \geq V_L(\hat{x}_{HS_H^*}, \hat{\pi}_{HS_H^*})$, indeed strictly. The last inequality is by definition of $(\hat{x}_{HS_H^*}, \hat{\pi}_{HS_H^*})$ and of S_H^*. This concludes the proof. □

[107] For general utility, this is slightly less obvious since lowering the premium increases spending, but this can be dealt with by lowering the premium even more. See Lemma G.9.

[108] The latter shows $m_{c,\pi}(s)$ is strictly increasing in m for s in the positive measure interval $[s_{\pi^p}^*, \bar{s}]$ — where s_π^* was defined (for each π) in Assumption B5 in Section 3 — if c is C^1 with $c' \geq 0$ (which holds since $x^p \leq 1$ by hypothesis) and $c' \leq 1$ (which also holds by assumption).

G.7 Proofs for FOA-Validity and Separation-vs-Pooling (Prop. F.1 and F.2)

Proposition F.1 (FOA-Validity: Restated)**.** *Let utility be general. Suppose $(x_t, \pi_t)_{t=1}^T$ is feasible in the FOA-relaxed problem (henceforth just "relaxed problem"). If utility is not money-loss CARA, suppose further that all upward-adjacent ICs bind*[109] *and that $x_t \leq 1$ for all t.*

i) If $(x_t)_t$ is weakly increasing (in t) and $\pi_t = \pi_{t+1}$ whenever $x_t = x_{t+1}$, then $(x_t, \pi_t)_{t=1}^T$ is feasible in the full problem. If further, $x_t < x_{t+1}$ for some t, then all downward ICs pointing to t hold with slack, i.e., $V_{t+i}(x_{t+i}, \pi_{t+i}) > V_{t+i}(x_t, \pi_t)$ for all $i \geq 1$. Additionally, if $(x_t)_t$ is strictly increasing (in t) throughout, then all extra ICs not imposed by the relaxed problem hold with slack at $(x_t, \pi_t)_{t=1}^T$.

ii) Suppose $(x_t, \pi_t)_{t=1}^T$ is also optimal in the relaxed problem. The FOA is valid if and only if both $(x_t)_t$ is weakly increasing and $\pi_t = \pi_{t+1}$ whenever $x_t = x_{t+1}$. Thus, strict increasingness of $(x_t)_t$ implies strict increasingness of $(x_t)_t$ is optimal in the full problem.

iii) Suppose the relaxed problem is strictly concave (by which I include the assumption that its feasible set is convex). Then the FOA will be valid if the following holds: there exists an optimal solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$ to the full problem with $x_t^{**} \leq 1$ (for all t) such that $(x_t^{**})_t$ is strictly increasing and all upward-adjacent ICs bind there.

iv) Again suppose the relaxed problem is strictly concave (by which I include the assumption that its feasible set is convex). If the optimal solution $(x_t^*, \pi_t^*)_{t=1}^T$ to the relaxed problem has $x_t^* > x_{t+1}^*$ for some t, then $x_\tau^{**} = x_{\tau+1}^{**}$ (and $\pi_\tau^{**} = \pi_{\tau+1}^{**}$) for some $\tau \in \mathcal{T}$ in any optimal solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$ to the full problem.

Proof. i) [Note: The hypothesis that upwards-adjacent IC binds will only be used to show downward IC is satisfied (and in fact doesn't bind).]

We use an integrated argument to prove all three statements of part i. Suppose that $(x_t)_t$ is weakly (strictly) increasing (with $\pi_t = \pi_{t+1}$ whenever $x_t = x_{t+1}$) and upward-adjacent IC binds.

First consider non-adjacent upward ICs. We need to show that for any t and for all $i \geq 2$, $V_t(x_t, \pi_t) \geq (>) V_t(x_{t+i}, \pi_{t+i})$ in the weakly (strictly) increasing case. Fix t. By feasibility in the FOA-relaxed problem, upward adjacent IC is satisfied i.e. for any $i \geq 2$, $V_{t+i-1}(x_{t+i-1}, \pi_{t+i-1}) \geq V_{t+i-1}(x_{t+i}, \pi_{t+i})$ If $x_{t+1} > x_t$, then by single crossing (Lem 5) $V_t(x_{t+i-1}, \pi_{t+i-1}) > V_t(x_{t+i}, \pi_{t+i})$; if alternatively $x_{t+1} = x_t$, then $\pi_{t+1} = \pi_t$ by assumption so $V_t(x_{t+i-1}, \pi_{t+i-1}) \geq V_t(x_{t+i}, \pi_{t+i})$ in fact with equality. But if by an inductive hypothesis $V_t(x_t, \pi_t) \geq V_t(x_{t+i-1}, \pi_{t+i-1})$ — and note the base case $i = 2$ holds by t's upward-adjacent IC

[109] This is guaranteed for money-loss CARA by Lem. 2.

— then $V_t(x_t, \pi_t) \geq (>) V_t(x_{t+i}, \pi_{t+i})$ by transitivity. Thus, by induction, we've thus shown $V_t(x_t, \pi_t) \geq (>) V_t(x_{t+i}, \pi_{t+i})$ for all $i \geq 2$.

Now consider downward ICs. We need to show that for any t and for all $i \geq 1$, $V_t(x_t, \pi_t) \geq (>) V_t(x_{t-i}, \pi_{t-i})$ in the weakly (strictly) increasing case. [The second assertion of part i of the result statement will be addressed at the end of the is paragraph.] Recall that upward adjacent IC binds as part of our hypothesis or as an implication of money-loss CARA (Lem. 3). In the case of $i = 1$, the desired conclusion then follows by single crossing in the strict case (when $x_t < x_{t+1}$) and in the weak case (when just $x_t \leq x_{t+1}$) by the assumption that pooled x implies pooled π. For $i \geq 2$, we again use an inductive argument. Suppose as inductive hypothesis that $V_t(x_t, \pi_t) \geq V_t(x_{t-i+1}, \pi_{t-i+1})$. Letting $t - i + 1$ play the role of t momentarily, we have just argued that $(t - i + 1)$'s downward-adjacent IC satisfies $V_{t-i+1}(x_{t-i+1}, \pi_{t-i+1}) \geq (>) V_{t-i+1}(x_{t-i}, \pi_{t-i})$ when, respectively, $x_{t-i} = x_{t-i+1}$ or $x_{t-i} < x_{t-i+1}$. Now, in the latter case, single crossing implies $V_t(x_{t-i+1}, \pi_{t-i+1}) > V_t(x_{t-i}, \pi_{t-i})$[110] and in the former case (where $\pi_{t-i+1} = \pi_{t-i}$ by hypothesis of the proposition given that case), $V_t(x_{t-i+1}, \pi_{t-i+1}) \geq V_t(x_{t-i}, \pi_{t-i})$ in fact with equality. But then $V_t(x_t, \pi_t) \geq (>) V_t(x_{t-i}, \pi_{t-i})$ by transitivity and the induction hypothesis. This shows $V_t(x_t, \pi_t) \geq (>) V_t(x_{t-i}, \pi_{t-i})$ for all $i \geq 1$ by induction; note that to show the strict inequality $V_t(x_t, \pi_t) > V_t(x_{t-i}, \pi_{t-i})$, we only used $x_{t-i} < x_{t-i+1}$ thus justifying the second assertion in the result statement. Note that although that assertion was indexed slightly differently as $V_{t+i}(x_{t+i,}\pi_{t+i}) > V_{t+i}(x_{t,}\pi_t)$ for all $i \geq 1$ whenever $x_t < x_{t+1}$, this is clearly equivalent to what we showed by simply reindexing.

Altogether, this completes the proof of part i.

ii) The background assumption of this part is that $(x_t, \pi_t)_{t=1}^T$ is relaxed-optimal. Suppose x_t is weakly increasing in t and that if $x_t = x_{t+1}$ for some t then $\pi_t = \pi_{t+1}$. Then by part i, $(x_t, \pi_t)_{t=1}^T$ satisfies all ICs in the full problem. Thus, since $(x_t, \pi_t)_{t=1}^T$ is relaxed-optimal, it is full-optimal a fortiori. This shows that the FOA is valid.

Conversely, suppose the FOA is valid. Then, by definition of FOA, $(x_t, \pi_t)_{t=1}^T$ is full-optimal since it's relaxed-optimal and in particular is feasible in the full problem. Thus, by Lemma 6, $x_t \leq x_{t+1}$ for all $t < T$ and $\pi_t = \pi_{t+1}$ whenever $x_t = x_{t+1}$. This proves the converse. Finally, the last statement about "strict increasingness of x_t" is obvious given the first assertion of part ii.

iii) Assume the hypotheses and suppose by contradiction that the FOA fails. Then for some optimal solution $(x_t^{**}, \pi_t^{**})_{t=1}^T$ to the *full* problem, $(x_t^{**}, \pi_t^{**})_{t=1}^T$ gives strictly less ex-ante expected utility than the optimal solution $(x_t^*, \pi_t^*)_{t=1}^T$ to the *FOA-relaxed* problem. Consider a direction of deviation moving the $2T$-dimensional vector $(x_t^{**}, \pi_t^{**})_{t=1}^T$ in a straight

[110] Technically, to get this we don't need the strict relation in $V_{t-i+1}(x_{t-i+1}, \pi_{t-i+1}) \geq (>) V_{t-i+1}(x_{t-i}, \pi_{t-i})$ from the previous sentence.

line towards $(x_t^*, \pi_t^*)_{t=1}^T$. Since the objective is strictly higher at the latter, the hypothesis of strict concavity of relaxed problem implies that this deviation increases the objective at the margin. We now check the deviation also preserves feasibility at the margin. By hypothesis the relaxed problem is convex so its feasible region is convex. Thus, the deviation preserves feasibility in the *FOA-relaxed* problem given that both $(x_t^*, \pi_t^*)_{t=1}^T$ and $(x_t^{**}, \pi_t^{**})_{t=1}^T$ are feasible therein. Further, by hypothesis $(x_t^{**})_t$ is strictly increasing and upward-adjacent IC binds, so since $(x_t^{**}, \pi_t^{**})_{t=1}^T$ is relaxed-feasible, part i of the present result implies all extra ICs hold with slack. This means that the deviation locally preserves all extra ICs in the *full* problem by continuity of the indirect utility functions V_t. Thus the deviation is locally feasible and profitable, which contradicts optimality of $(x_t^{**}, \pi_t^{**})_{t=1}^T$ in the full problem. This establishes that the FOA is valid.

iv) Suppose has $x_t^* > x_{t+1}^*$ for some t. Then by part ii, the FOA must not be valid. By Lemma 6, $(x_t^{**})_t$ must be weakly increasing since it is feasible in the full problem, but part iii (which applies by the strict concavity hypothesis), it must not be the case that it's strictly increasing throughout. Thus, $x_\tau^{**} = x_{\tau+1}^{**}$ for some $\tau \in \mathcal{T}$;[111] Lemma 6 then implies $\pi_\tau^{**} = \pi_{\tau+1}^{**}$ for that τ as well. This completes the proof. □

Proposition F.2 (Static Monotonicity drives FOA-Validity, two types: Restated)**.** *Let $\mathcal{T} = \{L, H\}$ with $L < H$. Suppose $(x_t, \pi_t)_{t=L,H}$ is an optimal solution to the FOA-relaxed problem with $x_t \geq 0$ for all t. Suppose also that the static problem is strictly concave and $\hat{x}_{tS}$ exists for all t, S.*

i) Let utility be money-loss CARA. If $\hat{x}_1 \leq \hat{x}_2$, then $x_1 < x_2$ and the FOA is valid, so $(x_t, \pi_t)_{t=1}^T$ is optimal in the full problem.

ii) Now let utility be general and suppose $\lambda_t^{x_t,\pi_t} \left[\frac{\partial}{\partial \pi} P_t(x_t, \pi_t)\right]^{-1}$ is strictly increasing in t and $x_t \leq 1$. Let $S_t \equiv -P_t(x_t, \pi_t)$ for all t. If $\hat{x}_{1S_1} \leq \hat{x}_{2S_2}$, then $x_1 < x_2$ and the FOA is valid so $(x_t, \pi_t)_{t=1}^T$ is optimal in the full problem.

Proof. Let $(x_t, \pi_t)_{t=L,H}$ be as in the statement. Consider the following chain of inequalities.

$$x_L = \hat{x}_{LS_L} \leq \hat{x}_{HS_H} < x_H. \tag{G.42}$$

The first equality $x_L = \hat{x}_{LS_L}$ by Lemma G.16 part iii, specifically the part about the FOA-relaxed solution. The inequality $\hat{x}_{LS_L} \leq \hat{x}_{HS_H}$ is by hypothesis. The last inequality is by Theorem 1′ which applies here under money-loss CARA for part i (of the present result) and under general utility by increasingness of $\lambda_t^{x_t,\pi_t} \left[\frac{\partial}{\partial \pi} P_t(x_t, \pi_t)\right]^{-1}$ for part ii (of the present result).

[111] Technically, we've only shown this for *some* τ, not necessarily for $\tau = t$ (and don't need to show that), though it seems very likely.

Thus, $x_L < x_H$ by (G.42). It then follows by Prop. F.1 ii that the FOA is valid so that $(x_t, \pi_t)_{t=L,H}$ is optimal in the full problem. This concludes the proof. □

H Nonlinear Coinsurance: Extra Analysis

H.1 Direct Formulation

As discussed in Section 7.1, when nonlinear contracts are allowed, it is useful to reformulate the private-risk problem as a direct revelation mechanism. To formulate the direct mechanism, recall that I normalize baseline wealth to zero and that $r_t(s)$ is to be interpreted as the "reimbursement" the consumer gets from insurance when spending $m_t(s)$ (i.e. when announcing s). The direct mechanism is:[112]

$$\max_{(r_t(\cdot),m_t(\cdot))_{t\in\mathcal{T}}} \sum_{t=1}^{T} f_t \int U\left(-m_t(s)+r_t(s),\, h\left(m_t(s),s\right)\right) dG_t(s)$$

$$\text{s.t.} \quad \text{(IC1)} \quad t \in \arg\max_{\hat{t}\in\mathcal{T}} \int U\left(-m_{\hat{t}}(s)+r_{\hat{t}}(s),\, h\left(m_{\hat{t}}(s),s\right)\right) dG_t(s) \tag{H.1a}$$

$$\text{(IC2)} \quad s \in \arg\max_{\hat{s}} U\left(-m_t(\hat{s})+r_t(\hat{s}),\, h\left(m_t(\hat{s}),s\right)\right) \tag{H.1b}$$

$$\text{(IR)} \quad \sum_{t=1}^{T} f_t \int r_t(s)\, dG_t(s) \leq 0 \tag{H.1c}$$

H.2 Relation between Direct and Indirect Mechanisms and Formulations

Equivalence of Mechanisms (Direct and Indirect) Clearly, any indirect mechanism of the form (5.1) can be implemented as a direct mechanism of the form (H.1) using the change of variables $m_t(s) = m_{c_t,\pi_t}(s)$ and $r_t(s) = -\pi_t + \left(m_{c_t,\pi_t}(s) - c_t\left(m_{c_t,\pi_t}(s)\right)\right)$. Conversely, given a solution $(m_t, r_t)_{t\in\mathcal{T}}$ to the direct formulation (H.1), we can define π_t and c_t as mentioned in Section 7.1. To repeat in more detail, first, let $\pi_t = -r_t(\underline{s}) + m_t(\underline{s})$ where (recall) $\underline{s}$ is the lowest severity in the support of G_t (common across t's). Second, let $c_t(m) = \left[m - r_t\left(s_t^{-1}(m)\right)\right] - \pi_t$ for those m such that $m = m_t(s)$ for some s in the support of G_t, and $c_t(m) = M_m$ for some sufficiently large function M_m of m.[113]

[112] Double deviations are ruled out by the fact that t is not payoff relevant and the common support assumption.

[113] $s_t^{-1}(\cdot)$ is the (potentially) set-valued inverse function of $m_t(s)$. $r_t\left(s_t^{-1}(m)\right)$ is well-defined because if two types s_1, s_2 select the same treatment level m — in the sense of reporting truthfully when $m_t(s_1) = m_t(s_2)$ — then they must have been paying the same out-of-pocket i.e. $r_t(s_1) = r_t(s_2)$ since otherwise the one paying more wouldn't be optimally reporting their ex-post type.

To see that the induced indirect mechanism preserves incentives at the ex-post stage, consider a type s who had reported $\hat{t}$ at the interim stage. Under the direct mechanism, their total monetary consumption if they report $\hat{s}$ in the ex-post stage is $-m_{\hat{t}}(\hat{s}) + r_{\hat{t}}(\hat{s})$ (and obviously their treatment is $m_{\hat{t}}(\hat{s})$), whereas under the indirect its $-\pi_{\hat{t}} - c_{\hat{t}}\left(m_{\hat{t}}(\hat{s})\right)$ when choosing $m_{\hat{t}}(\hat{s})$.[114] After substituting for $c_{\hat{t}}\left(m_{\hat{t}}(\hat{s})\right)$, the latter equals $-m_{\hat{t}}(\hat{s}) + r_{\hat{t}}(\hat{s})$ as desired.

Equivalence of Formulations (Direct and Indirect) The above verified that any given direct mechanism can be implemented by an indirect mechanism (and vice-versa). We now ask: when are the direct and indirect optimization problems (a.k.a formulations) equivalent ((5.1) and (H.1)) in the sense of having both the same optimal value and having the respective maximizers be related by the above conversion formulas. Specifically for what choices of $\mathcal{A}$ do we have this equivalence? For example, we may want to set $\mathcal{A}$ as the set of either continuous functions, Lipschitz functions, or functions with Lipschitz constant 1 — the latter being quite natural as mentioned in 4.1. Recalling the formula $c_t(m) = \left[m - r_t\left(s_t^{-1}(m)\right)\right] - \pi_t$, it is easy to see that (Lipschitz) continuity of the induced c_t would be assured if the optimal $m_t(s)$ and $r_t(s)$ are (Lipschitz) continuous. To assure the rightward derivative c' is ≤ 1, it's not hard to see it's sufficient that the optimal $r_t(s)$ be weakly increasing (since the optimal $m_t(\cdot)$ will be weakly increasing). Thus, if these (respective) properties hold in the optimal static solution, we will have an equivalence between the direct and indirect problems under the respective choice of admissible coinsurance schedules $\mathcal{A}$.

Gaps in $\{m_t(s)\}$: a More Natural Construction The *above* approach to dealing with "unsupported" treatment levels m — namely, setting $c_t(m) = M_m$ — would be quite unnatural if there are *gaps* in the interior of the set of supported treatment levels $\{m_t(s)|s \in [\underline{s}, \bar{s}]\}$ under the optimal $m_t(\cdot)$. Such gaps or, equivalently, discontinuities in the optimal $m_t(\cdot)$ seem unlikely, but I have not formally ruled them out.[115,116] In any case, I now describe a more natural approach to defining $c(m)$ for m in such gaps. First let m^+ the spending level on the right endpoint of the gap: $m^+ \equiv \inf\{z|z \geq m,\ z = m_t(s) \text{ for some } s \text{ in supp } G_t\}$. Then set $c_t(m) = \lim_{z \downarrow m^+} c_t(w)$. It remains to check that this construction will not (strictly) attract types to choose (in the induced mechanism) spending levels that were unsupported in the direct mechanism.

[114] By construction of M_m, "unsupported" m's will never be attractive to choose under the induced indirect mechanism.

[115] See [Lollivier and Rochet 1983] for discussion of irregular solutions in optimal income taxation.

[116] For a result showing regularity of control variables (continuity and, separately, Lipschitz) under certain conditions, see the math textbook [Clarke 2013] (Theorem 23.3); in the FOA-relaxed direct formulation, $m_t(s)$ and $r_t(s)$ will be control variables.

I sketch that argument here, leaving a full proof to the reader. It is easy to see that unsupported spending levels m on the interior or left-endpoint of a gap will not attract any types (to choose them) since those bundles are effectively dominated by $(m^+, c_t(m^+))$ which offers strictly more treatment for the same price.[117] It remains to deal with a right-discontinuity in $m_t(\cdot)$ by showing a newly available spending level at the right endpoint of the gap will not attract anyone. This will only be true up to weak preference. Specifically, if s is the point at which the right-discontinuity occurs, I claim s will at least *weakly* prefer her original bundle $(m_t(s), -m_t(s) + r_t(s))$ to the newly available $(m^+, c_t(m^+))$. The reason is that if s strictly preferred $(m^+, c_t(m^+))$, then since preferences are continuous and $c_t(m^+) \equiv \lim_{z \downarrow m^+} c_t(w)$, s would have (under the direct mechanism) also strictly preferred some spending level above but sufficiently close to m^+. Since only weak preference is assured, we will need to assume (for this construction to implement the direct mechanism) that if a tie arises, s's choice will resolve it in favor of the original bundle $(m_t(s), -m_t(s) + r_t(s))$.

H.3 Nonlinear FOA-Validity

The FOA can fail, but a sufficient condition for validity is known.[118] In his slightly different context, [Kapička 2013]'s Lemma 12 provides such a sufficient condition, namely increasingness of (what I'll call) marginal rents in the relaxed-optimal solution. In the present context, that means increasingness of $V_t(c_\tau, \pi_\tau) - V_{t-1}(c_\tau, \pi_\tau)$ — in intended-risk-type τ — at the relaxed solution.[119] Below, in Proposition H.1 part i, I adapt [Kapička 2013]'s sufficient condition result — which is in the context of an income-taxation model — to this context. Both results allow nonlinear contracts, but part ii of my proposition newly specializes to linear contracts and connects with indirect mechanisms.

Specifically, Proposition H.1 part ii shows that the sufficient condition from part i (increasing marginal rents) is implied by increasingness of coverage $(x_t)_t$ in the relaxed solution; the latter was shown to be sufficient for FOA-validity (under linear contracts) in Prop. F.1 ii (in Appendix F), as previewed at the end of Section 6.2 of the main text. The intuition for part ii is that marginal rent can be thought of as a proxy for — the lack of — expected

[117] A caveat is that sufficiently excessive treatment may be harmful to health ($h_1 < 0$), thus casting doubt on the term "dominated". However, in that case, the optimal solution would not have made $m_t(\cdot)$ discontinuously jump up (at some s) to a level m^+ that is harmful to s and types near it. I leave a rigorous argument to the reader.

[118] Whereas for linear contracts, validity of the FOA at the interim stage is effectively equivalent to separation — as is shown in Prop. (F.1) ii and iii of Appendix F and surrounding discussion — with nonlinear contracts it can happen that both upward- and downward-adjacent ICs bind (i.e. FOA fails) despite separation occurring. Indeed, I strongly conjecture that under nonlinear contracts separation will be at least a generic property of the optimal menu if not guaranteed altogether.

[119] More precisely, $(c_\tau, \pi_\tau)_{\tau=1}^T$ here is the indirect solution induced by the relaxed-optimal direct solution.

marginal coverage since the benefit from being lower-risk is higher when expected marginal coverage is low.

Proposition H.1. *i) Suppose $(c_t, \pi_t)_{t=1}^T$ is the indirect solution induced by a feasible direct solution to the interim-FOA-relaxed problem (7.1). The following conditions are sufficient — under discrete and continuous risk-types respectively — to guarantee that it is also a feasible solution to the full problem (H.1) — and thus fully optimal if it was optimal in the relaxed problem.*

Discrete: upward-adjacent IC binds and $V_t(c_\tau, \pi_\tau) - V_{t-1}(c_\tau, \pi_\tau)$ is weakly increasing in τ for $\tau \geq 1$ and $t \geq 2$.

Continuous: $\frac{\partial}{\partial \tau} U_\tau^t|_{\tau=t}$ is weakly increasing in t where

$$V_\tau^t \equiv \int U\left(-\pi_t - c_t\left(m_t(s)\right), h\left(m_t(s), s\right)\right) dG_\tau(s).$$

ii) Suppose $c_t(m) = (1 - x_t)m$ for $x_t > 0$ and that $(x_t)_t$ is strictly increasing in t while $(\pi_t)_t$ is weakly increasing in t. Further, suppose either that U is money-loss utility or that the following hold: $U_{12} \leq 0$, $c \in C^1$, and both $c' > (\geq) 0$ and $h\left(m_{c,\pi}(s), s\right)$ is weakly (strictly) decreasing in s for all t; the parentheticals indicate an alternate pair of hypotheses.[120] *Then the "discrete" condition above holds.*

H.4 Nonlinear Analog of $x_t^* > \hat{x}_t$?

For context, I repeat the last paragraph of Section (7.2) (with paraphrase): If the dynamic term in (7.3) is ignored, that expression reduces to the analogous formula for the "dual" of t's *static* moral hazard problem:[121]

$$1 - c_t'\left(m_t(s)\right) = \frac{1 - G_t(s)}{g_t(s)} h_{12} \underbrace{\left[-1 + \frac{\psi_t(s)}{\lambda_t}\right]}_{\text{static term}}, \tag{H.2}$$

where (c_t, π_t) (as induced by m_t, r_t) are meant to appear inside the unpacked formulas for $\psi_t(s)$ and λ_t (recall (7.2)).

Although I cannot formally state a nonlinear analog of the inequality $x_t^* > \hat{x}_t$ from the linear case (Thm 1), I contend that the spirit of that result does extend to the nonlinear case — again, under the FOA. To see this informally, continue to assume money-loss CARA. First note (without proof) that the solution to t's static problem $(\hat{c}_t, \hat{\pi}_t)$ satisfies (7.4), and because

[120] These hypotheses guarantee that $W_1^{x_t, \pi_t}(s)$ is strictly increasing in s by Lemma G.5 ii and iii.

[121] "Dual" in the sense of minimizing (expected) insurer losses subject to achieving a fixed expected utility for the consumer.

of CARA, this actually holds with $\hat{\pi}_t$ replaced by any π_t (Lem. 1).[122] Thus, if we set $c_t = \hat{c}_t$ on both sides of (7.3), then for any π_t, the LHS will be lower than the RHS given the presence of the dynamic term, which is strictly positive by Lemmas 7 and 2. Accordingly, in order to make the two sides of the equation balance, the marginal coverage $1 - c_t'(m_{c_t}(s))$ on the LHS must be increased relative to $1 - \hat{c}_t'(m_{\hat{c}_t}(s))$.[123] This observation suggests the profitability (at any s) of an upward deviation of marginal coverage $1 - c_t'(m_{c_t}(s))$ starting from any putatively optimal menu $(c_t^*, \pi_t^*)_{t=1}^T$ with $c_t^* = \hat{c}_t$ for some $t > 1$. I keep this argument informal because it's not clear that this implies the full pointwise inequality $1 - c_t^{*\prime}(m) \geq 1 - \hat{c}_t'(m)$ for all m simultaneously.

The discussion in [Kapička 2013] suggests it may be possible to show roughly that marginal coverage is distorted upwards *in expectation*. More specifically, the result would be expressed as an upward distortion of the *marginal rent*. As mentioned above, such marginal rent can be thought of as a proxy for (lack of) expected marginal coverage, as I verify for the linear case in Prop. H.1 ii (in Appendix H.3).[124]

I Nonlinear Coinsurance: Wealth Effects

This section extends the money-loss CARA analysis of section 7.2 to other functional forms and type space assumptions. I first extend to non-CARA money, then give the general-utility formula — of which the other forms are a special case — and then specialize to separable utility formula. Then I turn to continuous risk-types (with all functional forms). All of this is still under the first-order approach.

I.1 Discrete Types

I.1.1 General Utility

The next definition introduces several pieces of notation which are interpreted afterwards.

[122] Again, π_t appears inside the terms on both sides of (7.4).

[123] Although c_t' also appears on the RHS (inside $\psi_t(s)$ and λ_t), this does not disrupt the argument since $\frac{\psi_t(s)}{\lambda_t}$ decreases when we increase c_t' (and correspondingly increase π_t).

[124] To further develop the argument for upward distortion of marginal rent, it would be easiest (following [Kapička 2013]) to work with the continuous risk-type case as formulated in problem S1.1 in Supplementary Appendix S1.1 There, the object corresponding to marginal rent is $\int W_t(s) \frac{\partial}{\partial t} g_t(s) ds$ where $W_t(s)$ represents ex-post utility (as in (7.1b) from the main text). That object appears there in the envelope condition (S1.1b) for interim IC. To complete the argument, it would be necessary to sign the co-state multiplier on that constraint in order to confirm the benefit of relaxing interim IC via the distortion of marginal rent. For reference, that multiplier is denoted $\gamma(t)$, and is introduced in Supplementary Appendix S2 in constraint S2.3 during the proof of Prop. S1.2 — the continuous risk-type analog of Prop. 3 (with wealth effects).

Definition 10. Let $U_{1,t,s} \equiv U_{1,t,s}\left(-\pi_t - c_t(m_t(s)), h\left(m_t(s), s\right)\right)$ and similarly $U_{12,t,s} \equiv U_{12}\left(-\pi_t - c_t(m_t(s)), h\left(m_t(s), s\right)\right)$.[125] [I sometimes omit t and s, and the commas in the subscripts.]

$$B_{t,s}(\hat{s}) \equiv \exp\left(\int_s^{\hat{s}} \left[\frac{U_{21,t,\tilde{s}}}{U_{1,t,\tilde{s}}} h_2\left(m_t(\tilde{s}), \tilde{s}\right)\right] d\tilde{s}\right). \tag{I.1}$$

Let $N_t^\tau(s) \equiv \int_s^{\bar{s}} B_{\tau,s}(\hat{s}) \frac{g_t(\hat{s})}{1-G_\tau(s)} d\hat{s}$ and $N_t(s) \equiv N_t^t(s)$. Let $\chi_t(s) \equiv \int_s^{\bar{s}} B_{t,s}(\hat{s}) \frac{1}{U_{1,t,\hat{s}}} \frac{g_t(\hat{s})}{1-G_t(s)} d\hat{s}$. Let $J_t^\tau \equiv N_t^\tau(\underline{s})$, $J_t \equiv J_t^t$, and $J_0^t \equiv 1$ for all t, and let $\rho_t \equiv \chi_t(\underline{s})$. Let $\nu_t^\tau \equiv J_t^\tau/\rho_\tau$ and $\nu_t \equiv \nu_t^t$.

$B_{t,s}(\hat{s})$ is the number of utils given to $(t, \hat{s})$ for $\hat{s} > s$ by a *treatment-choice-preserving* transfer to t that gives 1 util to s. When we transfer money to ex-post types under general utility, we must do so in a non-uniform fashion in order to preserve ex-post IC. We also want to so in a way that prioritizes the highest severities s as much as possible. Towards this end, if we transfer money to (t, s) which is worth one util to (t, s), we can and should also transfer $B_{t,s}(\hat{s})$ utils worth of money to all the types $(t, \hat{s})$ with $\hat{s} > s$.[126] $N_t^\tau(s)$ is the $[1 - G_\tau(s)]$-normalized expected number of utils received by t due to such a transfer to τ's plan when t buys τ's plan. I refer to such a transfer (to types $[s, \bar{s}]$ in t) as a $N(s)$-transfer to τ, regardless of whether $\tau = t$. $\chi_t(s)$ is the cost of an $N(s)$-transfer to t on path. J_τ^t and ρ_t are the special case of $N_\tau^t(s)$ and $\chi_t(s)$ when $s = \underline{s}$, thus corresponding to a transfer to all severities within t. It follows from the definition of $\nu_t^\tau \equiv J_t^\tau/\rho_\tau$ that ν_t^τ is the amount of expected utils (under G_t) generated per dollar of $N(s)$-transfer to τ when the cost is evaluated on path. Actually, $\nu_t \equiv \nu_t^t$ was already introduced in Appendix D.1 using very-slightly different notation $\nu_t^{c_t,\pi_t}$. There I described it as the utils per dollar created by what I called a ν-transfer (to t) which is simply a dollars worth of treatment-preserving transfer to t on path. A ν-transfer is simply one dollar worth of $N(\underline{s})$-transfer to t.

The next lemma extends Lemma 7 from money-loss to general utility.

Lemma I.1 (Generalizing Lem. 7)**.** *Under general utility,* $\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t} > 0$.[127]

The next proposition extends Proposition 3 from money-loss CARA to general utility and verifies that the latter is indeed a special case. Then in the proposition following it, I show how this specializes to separable utility.

[125] Equivalently, we could write the first arguments as $-m_t(s) + r_t(s)$ rather than $-\pi_t - c_t(m_t(s))$.

[126] To be clear, if other types $\tilde{s} \neq \hat{s}$ misreport as $\hat{s}$, the transfer to $\hat{s}$ will not be worth $B_{t,s}(\hat{s})$ utils to them and indeed will be worth strictly less than that for $\tilde{s} > \hat{s}$.

[127] This result is proved as part of Lem. J.1 in Appendix J.2.

Proposition I.1 (Optimal Menu: General Utility)**.** *Suppose* $(c_t, \pi_t)_{t\in\mathcal{T}}$ *is the indirect solution induced by an optimal solution to FOA-relaxed direct problem (7.1) under general utility. For each* $t \in \mathcal{T}$ *and* $m \in \mathbb{R}$*, if there exists* s *such that* $m_t(s) = m$ *and the derivative* $c_t'(m)$ *exists (at* m*), then the latter satisfies the formula (I.2) just below.*

$$1 - c_t'\left(m_t(s)\right) = \frac{\left[\left(-c_t'U_{21} + U_{22}h_1\right)h_2 + U_2h_{12}\right]}{\lambda} \cdot \left[\frac{1-G_t(s)}{g_t(s)}\right] \cdot \Bigg\{ \lambda \underbrace{\left[-\chi_t(s) + \frac{N_t(s)}{\nu_t}\right]}_{\textit{static term}} + \frac{1-\check{F}_{t-1}}{f_t} \underbrace{\left[\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t}\right]}_{\textit{slackening}} \underbrace{\left[-\lambda + \sum_{k=t}^{T} \frac{f_k Q_t^{k-1}}{1-\check{F}_{t-1}} \nu_k\right]}_{\textit{redistr. force}} \Bigg\} \tag{I.2}$$

$$\textit{dynamic term}$$

where

$$Q_\ell^\tau \equiv \Pi_{x=\ell}^{\tau} \frac{\nu_x}{\nu_x^{x+1}} \;\; \ell \leq \tau = 1; \quad Q_t^{t-1} \equiv 1 \;\forall t, \tag{I.3}$$

$$1 - \check{F}_{t-1} \equiv \sum_{k=t}^{T} f_k Q_t^{k-1} \quad \textit{i.e.} \;\; \check{F}_{t-1} \equiv 1 - \sum_{k=t}^{T} f_k Q_t^{k-1} \;\; \forall t \geq 1, \quad \textit{and} \;\; \check{F}_0 \equiv 1, \tag{I.4}$$

$$\lambda \equiv \frac{1}{Q} \sum_{k=1}^{T} f_k Q_1^{k-1} \nu_k \quad \textit{where} \; Q \equiv \sum_{k=1}^{T} f_k Q_1^{k-1}. \tag{I.5}$$

The interpretation of (I.2) is similar to that of (7.3) in the main text.[128] The formula (I.2) can be decomposed into two parts capturing the static and dynamic forces. The static force, which captures the risk-incentive tradeoff, is represented by the LHS and the term outside the curly braces on the RHS alongside and first term in curly braces, which is labeled "static term". The latter represents the money saved by a $N_t(s)$-transfer and a compensating ν_t-transfer to t. This amount is strictly positive because such a compound transfer redistributes money to higher marginal utility consumers — mathematically, $\frac{N_t(s)}{\chi_t(s)} > \nu_t \equiv \frac{J_t}{\rho_t}$. The dynamic force is represented by the second term in curly braces (excluding $\frac{1-\check{F}_{t-1}}{f_t}$). It captures the value of redistribution to type t through T (from the budget) via ν-transfers, which is enabled by the IC-slack (which is strictly positive by Lem. I.1) resulting from the $N_t(s)$-transfer and compensating ν_t-transfer to t. As with the premium transfers in the money-loss case, the ν-transfers are weighted by the weights P_t^{k-1} in order to prioritize the highest-risk types while preserving interim IC.[129] The inverse hazard rate $\frac{1-\check{F}_{t-1}}{f_t}$ captures the relative importance

[128] As there, (I.2) is an *implicit* characterization in that c_t appears inside several objects on the RHS.

[129] Note that $-c_t'U_{12}h_2 + U_{22}h_2h_1 + U_2h_{12}$ is the *partial* derivative of the treatment FOC with respect to s.

of the dynamic and static terms conditional on the size of each.[130] Again, this reflects the principle of no distortion at the top under the FOA.

λ, defined in (I.5), represents the value of money that is to be distributed in a way that both preserves IC and maximally prioritizes the highest marginal utility consumers. When transferring to a given risk-type, this is done via ν-transfers. Thus, the question becomes how to conduct ν-transfers in an upward-IC-preserving way that maximally prioritizes the highest expected marginal utilities per-dollar ν_t (i.e. return). For t's upward IC to be preserved in a joint ν-transfer to t and $t+1$, we need the per-person transfers to $t+1$ (i.e., per f_{t+1}) to be no more than $\frac{\nu_t}{\nu_t^{t+1}}$ times those to t. Thus, if $\nu_t < \nu_{t+1}$ for all t, the optimal IC-preserving joint ν-transfer to $t+1$ will be exactly $\frac{\nu_t}{\nu_t^{t+1}}$ times the transfer to t. And indeed, $\nu_t < \nu_{t+1}$ is guaranteed for money-loss CARA (Lem 2) — where it will turn out that $\nu_\tau^t = \lambda_\tau^t$ — and is informally expected more generally.

Outside, money-loss CARA — where we'll see $\frac{\lambda_t}{\lambda_t^{t+1}} = 1$ (Lemma I.3 below) — the above analysis implies that the optimal ν-transfers to each risk-type must be *iteratively* "bumped-up" by the ratios $\frac{\nu_t}{\nu_t^{t+1}} \neq 1$.[131] Formally, a one-dollar transfer to risk-type 1 will be paired with a $Q_1^{t-1} \equiv \Pi_{x=1}^{t-1} \frac{\nu_x}{\nu_x^{x+1}}$-dollar transfer to each risk-type $t > 1$. This explains the formula for λ (I.8); the term $\left[\sum_{k=1}^T f_k Q_1^{k-1}\right]^{-1}$ is just a normalization so that units work out. This explains the domino redistribution term (the last term in (I.2)), thus accounting for the appearance of $Q_t^{\ell-1}$ and $\check{F}_{t-1}$ there.

I.1.2 Non-CARA Money-Loss

In order to specialize the general utility formula to money-loss and thus prove Proposition 3, we need to manipulate the expression $B_{t,s}(\hat{s})$ which appears in $N_t(s)$,$\chi_t(s)$, and ν_t; recall it's definition in equation (I.1).

Toward this end, define the following notation which is just a slight notational variational on a definition from Appendix D.1 (Eqn. (D.1) there)

$$L_s(\hat{s}) \equiv \exp\left(-\int_s^{\hat{s}} \left\{\left[-U_{11}c'(m_t(\tilde{s})) + U_{21}h_1\left(m_t(\tilde{s}),\tilde{s}\right)\right] / \left[U_1\right]\right\} \dot{m}_t(\tilde{s}) d\tilde{s}\right) \tag{I.6}$$

where the arguments of the derivatives of U are evaluated at $(-\pi_t - c_t(m_t(\tilde{s})), h\left(m_t(\tilde{s}),\tilde{s}\right))$. We have the following lemma, which will help specialize to the money-loss case.

[130] As noted in the main text (for money-loss CARA), for $t = 1$, the dynamic term drops out because $1 - \check{F}_{t-1} = 0$ since $\check{F}_0 \equiv 1$.The convention $\check{F}_0 \equiv 1$ works here because $\sum_{k=1}^T f_k Q_t^{k-1}\left[\nu_k - \lambda\right] = 0$ as can be seen easily from (I.5).

[131] At the end of section (I.1.2) below, I argue that the "up" in "bumped up" is literal in the sense that $\frac{\lambda_t}{\lambda_t^{t+1}} > 1$ for money-loss DARA.

Lemma I.2. *For all $s \leq \hat{s}$, $B_{t,s}(\hat{s}) = \frac{U_{1,t,\hat{s}}}{U_{1,t,s}} L_{t,s}(\hat{s})$. In the special case of money-loss, $L_{t,s}(\hat{s}) = 1$, so in particular $B_{t,s}(\hat{s}) = \frac{U_{1,t,\hat{s}}}{U_{1,t,s}}$ there.*

The details of specializing to the money-loss case are in Appendix J.3.3 where Proposition 3 is proved.

Proposition I.2 (Optimal Menu: Non-CARA Money-Loss)**.** *Under money-loss, the formulas from Proposition I.1 specialize to*

$$1 - c_t'\left(m_t(s)\right) = \frac{1 - G_t(s)}{g_t(s)} \frac{h_{12}}{\lambda} \cdot \Bigg\{ \underbrace{\lambda \left[-1 + \frac{\psi_t(s)}{\lambda_t} \right]}_{\text{static term}} + \underbrace{\frac{1 - \check{F}_{t-1}}{f_t} \underbrace{\left[\frac{\psi_t(s)}{\lambda_t} - \frac{\psi_{t-1}^t(s)}{\lambda_{t-1}^t} \right]}_{\text{slackening}} \underbrace{\left[-\lambda + \sum_{\ell=t}^{T} \frac{f_\ell Q_t^{\ell-1}}{1 - \check{F}_{t-1}} \lambda_\ell \right]}_{\text{redistr. force}}}_{\text{dynamic term}} \Bigg\}, \tag{I.7}$$

where,

$$Q_\ell^\tau \equiv \Pi_{x=\ell}^\tau \frac{\lambda_x}{\lambda_x^{x+1}} \;\; \ell \leq \tau = 1, \quad Q_t^{t-1} \equiv 1 \;\forall t, \tag{I.8}$$

$$\lambda \equiv \frac{1}{Q} \sum_{k=1}^{T} f_k Q_1^{k-1} \lambda_k, \quad \text{where } Q \equiv \sum_{k=1}^{T} f_k Q_1^{k-1}, \tag{I.9}$$

$$1 - \check{F}_{t-1} \equiv \sum_{k=t}^{T} f_k Q_t^{k-1} \quad \text{i.e. } \check{F}_{t-1} \equiv 1 - \sum_{k=t}^{T} f_k Q_t^{k-1} \;\; \forall t \geq 1, \quad \check{F}_0 \equiv 1 \text{ by convention}, \tag{I.10}$$

and for CARA in particular, $Q_\ell^\tau = 1$ for all $\ell \leq \tau$, so (I.7) reduces to (7.3) (in Prop. 3) from the main text.

λ, defined in (I.9), represents the value of money that is to be distributed in a way that both preserves IC and maximally prioritizes the highest marginal utility consumers. When transferring to a given risk-type, this is done via premium transfers. Thus, the question becomes how to conduct premium-transfers in an upward-IC-preserving way that maximally prioritizes the highest *expected* marginal utilities λ_t. For t's upward IC to be preserved in a joint premium-transfer to t and $t+1$, we need the per-person transfers to $t+1$ to be no more than $\frac{\lambda_t}{\lambda_t^{t+1}}$ times those to t.[132] Thus, if $\lambda_t < \lambda_{t+1}$ for all t, the optimal IC-preserving joint

[132] When I say "per-person" transfers to a type τ, I mean per f_τ.

premium-transfer to $t+1$ will be exactly $\frac{\lambda_t}{\lambda_t^{t+1}}$ times the transfer to t; and indeed, $\lambda_t < \lambda_{t+1}$ is guaranteed for CARA (Lem. 2) and is informally expected more generally. The next lemma shows that $\frac{\lambda_t}{\lambda_t^{t+1}} = 1$ under CARA.

Finally, I informally justify the term "bumped-*up*" for DARA money-loss. When u exhibits strict DARA, we can expect $\frac{\lambda_t}{\lambda_t^{t+1}} > 1$, so that $Q_\ell^\tau \equiv \Pi_{x=\ell}^{\tau} \frac{\lambda_x}{\lambda_x^{x+1}} > 1$ for $\ell \leq \tau$. [The opposite will hold for IARA.] To argue this for DARA, I need to make two assumptions, which though very weak, are not technically guaranteed and the latter of which is not even precisely defined. First, I assume upward IC binds, which (despite the FOA assumption) could in principle fail under non-CARA since I have not ruled out the case where neither direction binds (though this seems pathological). Second, I assume that plans intended for lower-risk types will have lower coverage "overall", thus making them riskier "overall".[133] With these assumptions in place, we can reason as follows. Since t is indifferent between her own plan and a less risky plan intended for $t+1$, when the premiums of both plans are reduced by one dollar, t will come to strictly prefer her own plan (given DARA) since her baseline wealth has been effectively increased. It follows that t's expected marginal utility under her own plan must have been strictly greater than under $t+1$'s plan. That is, $\frac{\lambda_t}{\lambda_t^{t+1}} > 1$. The same argument establishes the opposite for IARA. For CARA, $\frac{\lambda_t}{\lambda_t^{t+1}} = 1$, as the next lemma confirms.

Lemma I.3. *Under money-loss CARA with T discrete risk-types, $\lambda_{t-1}^t = \lambda_{t-1}$ for all $t \geq 2$.*

I.1.3 Separable Utility

I now specialize Proposition I.1 to separable utility. I defer the proof until Appendix J.3.4, but the key step is to observe that $B_{t,s}(\hat{s}) = 1$ since $U_{21} = 0$ when utility takes the separable form: $U(y,h) = u(y) + v(h)$, where y is monetary consumption and h is health.

Proposition I.3 (Optimal Menu: Separable Utility)**.** *Under* separable *utility, the formulas from the general case (Prop. I.1) simplify to (I.11) (and (I.12)) just below.*

$$\begin{aligned} 1 - c_t'\left(m_t(s)\right) = \left[U_{22}h_1h_2 + U_2h_{12}\right] \Bigg[& \\ \frac{1-G_t(s)}{g_t(s)} \left[\theta_t^{-1} - \int_s^{\bar{s}} \frac{1}{U_{1,t,\hat{s}}} \frac{g_t(\hat{s})}{1-G_t(s)} d\hat{s}\right] & \\ + \frac{1-F_{t-1}}{f_t} \left[\left[\frac{1-G_t(s)}{g_t(s)}\right] - \left[\frac{1-G_{t-1}(s)}{g_t(s)}\right]\right] \left[\lambda^{-1} - \sum_{k=t}^{T} \frac{f_k}{1-F_{t-1}} \theta_t^{-1}\right] \Bigg], & \end{aligned} \tag{I.11}$$

[133] We saw this formally for linear contracts (Lem. 6), but it would be harder to precisely formulate and show for nonlinear contracts since they are not totally ordered.

where $F_0 \equiv 1$ *and*

$$\theta_t \equiv \left[\int_{\underline{s}}^{\bar{s}} \frac{1}{U_{1,t,s}} g_t(s) ds\right]^{-1}; \qquad \lambda \equiv \left[\sum_{k=t}^{T} f_k \theta_t^{-1}\right]^{-1}. \tag{I.12}$$

Recall that θ_t was already defined at the end of Section 4.2, using the slightly different notation $\theta_t^{c,\pi} \equiv \left[\int 1/\left[W_1^{c,\pi}(s)\right] g_t(s) ds\right]^{-1}$ right before Proposition D.1. Its reciprocal θ_t^{-1} is the expected cost of performing what I called a θ-transfer to t in section 4.2 which is simply the special case of a ν-transfer under separable. In the proof, it is shown that, for separable, $\theta_t = \nu_t = \frac{1}{\rho_t}$. Unlike ν-transfers in general, θ-transfers distribute utility uniformly with a risk-type; this contrasts premium-transfers which distribute money uniformly.

Here (with separable utility), the slackening term, which was $\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t}$ for general utility, takes the form $\left[\frac{1-G_t(s)}{g_t(s)}\right] - \left[\frac{1-G_{t-1}(s)}{g_t(s)}\right]$ (after dividing by $g_t(s)$). In the continuous separable case below, we will see the continuous analog of this expression, $\frac{\partial}{\partial t}\left[\frac{1-G_t(s)}{g_\tau(s)}\right]|_{t=\tau}$, which has appeared in prior papers starting at least from [Courty and Li 2000] and is called the "impulse response function" by [Pavan et al. 2014].

As shown in the proof, arriving at (I.11) from (I.2) requires some manipulations including distributing the $\frac{1}{\lambda}$ out front in the latter.

J Proofs for Nonlinear Coinsurance

J.1 Proof of Prop H.1

Proposition G.1. *i) Suppose* $(c_t, \pi_t)_{t=1}^T$ *is the indirect solution induced by a feasible direct solution to the interim-FOA-relaxed problem (7.1). The following conditions are sufficient — under discrete and continuous risk-types respectively — to guarantee that it is also a feasible solution to the full problem (H.1) — and thus fully optimal if it was optimal in the relaxed problem.*

Discrete: upward-adjacent IC binds and $V_t(c_\tau, \pi_\tau) - V_{t-1}(c_\tau, \pi_\tau)$ *is weakly increasing in* τ *for* $\tau \geq 1$ *and* $t \geq 2$.

Continuous: $\frac{\partial}{\partial \tau} U_\tau^t|_{\tau=t}$ *is weakly increasing in* t *where*

$$V_\tau^t \equiv \int U\left(-\pi_t - c_t\left(m_t(s)\right), h\left(m_t(s), s\right)\right) dG_\tau(s).$$

ii) Suppose $c_t(m) = (1 - x_t)m$ *for* $x_t > 0$ *and that* $(x_t)_t$ *is strictly increasing in* t *while* $(\pi_t)_t$ *is weakly increasing in* t. *Further, suppose either that* U *is money-loss utility or that the*

following hold: $U_{12} \leq 0$, $c \in C^1$, *and both* $c' > (\geq) 0$ *and* $h\left(m_{c,\pi}(s), s\right)$ *is weakly (strictly) decreasing in* s *for all* t*; the parentheticals indicate an alternate pair of hypotheses.*[134] *Then the "discrete" condition above holds.*

Proof. i) First consider the discrete case. Without loss, let $t > t'$. We have the chain of reasoning below. The first equality is since upward IC binds (by hypothesis). The inequality is by also by hypothesis, given that $t > t'$. The second equality is by regrouping terms. The penultimate equality is by telescoping.

$$\begin{aligned}
V_t(c_t, \pi_t) =& V_t(c_t, \pi_t) - \sum_{\tau=t'}^{t-1} \left[\underbrace{V_\tau(c_{\tau+1}, \pi_{\tau+1}) - V_\tau(c_\tau, \pi_\tau)}_{0} \right] \\
=& \sum_{\tau=t'}^{t-1} \left[V_{\tau+1}(c_{\tau+1}, \pi_{\tau+1}) - V_\tau(c_{\tau+1}, \pi_{\tau+1}) \right] + V_{t'}(c_{t'}, \pi_{t'}) \\
\geq& \sum_{\tau=t'}^{t-1} \left[V_{\tau+1}(c_{t'}, \pi_{t'}) - V_\tau(c_{t'}, \pi_{t'}) \right] + V_{t'}(c_{t'}, \pi_{t'}) \\
=& V_t(c_{t'}, \pi_{t'}) - V_{t'}(c_{t'}, \pi_{t'}) + V_{t'}(c_{t'}, \pi_{t'}) = V_t(c_{t'}, \pi_{t'})
\end{aligned}$$

which establishes $V_t(c_t, \pi_t) \geq V_t(c_{t'}, \pi_{t'})$,as desired.

Now consider the continuous case. This part is only very slightly modified from [Kapička 2013] Lemma 12.

$$\begin{aligned}
V_t(c_t, \pi_t) - V_{t'}(c_{t'}, \pi_{t'}) =& \int_{t'}^{t} \frac{\partial}{\partial \tau'} V_{\tau'}^{\tau} |_{\tau'=\tau} d\tau \\
\geq& \int_{t'}^{t} \frac{\partial}{\partial \tau'} V_{\tau'}^{t'} |_{\tau'=\tau} d\tau \\
=& V_t(c_{t'}, \pi_{t'}) - V_{t'}(c_{t'}, \pi_{t'}).
\end{aligned}$$

where the equalities use the envelope theorem and the fact that the solution is a solution to the FOA-relaxed problem. Canceling the common term $V_{t'}(c_{t'}, \pi_{t'})$ shows $V_t(c_t, \pi_t) \geq V_t(c_{t'}, \pi_{t'})$ as desired.

ii) Suppose risk-types $L, H \in \mathcal{T}$ are such that $L < H$ and suppose U is either money-loss utility or the following: $U_{12} \leq 0$, $x \geq 0$, and $h\left(m_{c,\pi}(s), s\right)$ is weakly decreasing in s. Lemma G.5 parts ii and iii then show that $W_1^{x,\pi}(s)$ is strictly increasing in s. Consider two contracts (x_+, π) and (x_-, π) with the same π and $x_+ > x_-$. Then

[134] These guarantee that $W_1^{x_t, \pi_t}(s)$ is strictly increasing in s by Lemma G.5 ii and iii.

$$
\begin{aligned}
&[V_H(x_+,\pi) - V_L(x_+,\pi)] - [V_H(x_-,\pi) - V_L(x_-,\pi)] \\
&= \int_{x_-}^{x_+} \frac{\partial}{\partial x} \left[V_H(x,\pi) - V_L(x,\pi)\right] dx \\
&= \int_{x_-}^{x_+} \left[\int m_{x,\pi}(s) W_1^{x,\pi}(s) dG_H(s) - \int m_{x,\pi}(s) W_1^{x,\pi}(s) dG_L(s) \right] dx \\
&= \int_{x_-}^{x_+} \int \underbrace{m_{x,\pi}(s)}_{\text{incr. in } s,\, >0} \underbrace{W_1^{x,\pi}(s)}_{\text{incr. in } s,\, >0} \left[g_H(s) - g_L(s)\right] ds dx > 0
\end{aligned}
$$

where the second equality is by the envelope theorem and the last inequality is by $g_H \succ_{\text{MLRP}} g_L$ and the annotation on the LHS.

Altogether, this shows $[V_H(x_+,\pi) - V_L(x_+,\pi)] - [V_H(x_-,\pi) - V_L(x_-,\pi)] > 0$.

Now consider two contracts (x,π_+) and (x,π_-) with the same x and $\pi_+ \geq \pi_-$. The argument is similar to the one for x above (rather than π here):

$$
\begin{aligned}
&[V_H(x,\pi_+) - V_L(x,\pi_+)] - [V_H(x,\pi_-) - V_L(x,\pi_-)] \\
&= \int_{\pi_-}^{\pi_+} \frac{\partial}{\partial \hat{\pi}} \left[V_H(x,\hat{\pi}) - V_L(x,\hat{\pi})\right] d\hat{\pi} = \int_{\pi_-}^{\pi_+} \left[\lambda_H^{x,\hat{\pi}} - \lambda_L^{x,\hat{\pi}}\right] d\hat{\pi} \\
&= \int_{\pi_-}^{\pi_+} \left[\int W_1^{x,\hat{\pi}}(s) dG_H(s) - \int W_1^{x,\hat{\pi}}(s) dG_L(s) \right] d\hat{\pi} \\
&= \int_{\pi_-}^{\pi_+} \int \underbrace{W_1^{x,\hat{\pi}}(s)}_{\text{incr. in } s,\, >0} \left[g_H(s) - g_L(s)\right] ds d\hat{\pi} \geq 0
\end{aligned}
$$

Finally, consider two contracts (x_+,π_+) and (x_-,π_-) with $x_+ > x_-$ and $\pi_+ > \pi_-$. The two arguments above allow us to reason:

$$
\begin{aligned}
&[V_H(x_+,\pi_+) - V_L(x_+,\pi_+)] - [V_H(x_-,\pi_-) - V_L(x_-,\pi_-)] \\
&= \underbrace{\left\{[V_H(x_+,\pi_+) - V_L(x_+,\pi_+)] - [V_H(x_+,\pi_-) - V_L(x_+,\pi_-)]\right\}}_{\geq 0} \\
&+ \underbrace{\left\{[V_H(x_+,\pi_-) - V_L(x_+,\pi_-)] - [V_H(x_-,\pi_-) - V_L(x_-,\pi_-)]\right\}}_{0} \quad > 0
\end{aligned}
$$

This completes the proof. □

J.2 Proofs of Lemmas 7 and I.1

The next result restates and slightly generalizes Lemmas 7 and I.1 — the latter being itself the generalization of the former for general utility.

Lemma J.1 (Combining and Extending Lem. 7 and I.1)**.** *Let utility be general. Let* $U_{1,\tau,s}$ *be shorthand for* $U_1\left(-\pi_\tau - c_\tau\left(m_{c_\tau,\pi_\tau}(s)\right), h\left(m_{c_\tau,\pi_\tau}(s), s\right)\right)$.[135] *Let* $\psi_t^\tau(s)$ *and* λ_t^τ *be defined as in (7.2b) and (7.2a), and let* $N_t^\tau(s)$ *and* $\nu_{t'}^\tau$ *be defined as in Definition 10. Let* $t' < t \in \mathcal{T}$ *so that* $G_{t'} \prec_{MLRP} G_t$.

Then for any $\tau \in \mathcal{T}$ *and all* $s \in (\underline{s}, \bar{s})$*, we have* $\frac{N_{t'}^\tau(s)}{\nu_{t'}^\tau} < \frac{N_t^\tau(s)}{\nu_t^\tau}$ *and* $\frac{\psi_{t'}^\tau(s)}{\lambda_{t'}^\tau} < \frac{\psi_t^\tau(s)}{\lambda_t^\tau}$*. [Note that whereas under money-loss utility, objects involving* N *and* ν *reduce to those involving* ψ *and* λ *— as can be seen from the analysis of Appendix J.3.3 — so that the two conditions coincide, under general utility, these are two distinct conditions that are simultaneously true.]*

Proof. I show $\frac{N_{t'}^\tau(s)}{\nu_{t'}^\tau} < \frac{N_t^\tau(s)}{\nu_t^\tau}$ by showing $\frac{N_{t'}^\tau(s)}{J_{t'}^\tau} < \frac{N_t^\tau(s)}{J_t^\tau}$ which is equivalent to $\frac{N_{t'}^\tau(s)}{\nu_{t'}^\tau} < \frac{N_t^\tau(s)}{\nu_t^\tau}$ since $\nu_t^\tau \equiv J_t^\tau/\rho_\tau$ and $\rho_\tau > 0$ so that the ρ_τ's on both sides cancel. I omit a full proof of $\frac{\psi_{t'}^\tau(s)}{\lambda_{t'}^\tau} < \frac{\psi_t^\tau(s)}{\lambda_t^\tau}$ which is similar to that of $\frac{N_{t'}^\tau(s)}{J_{t'}^\tau} < \frac{N_t^\tau(s)}{J_t^\tau}$, but I comment on how it is different at the end of the proof.

From the definition of $B_{\tau,s}(\hat{s})$ in (I.1), its easy to see that $B_{\tau,\underline{s}}(\hat{s}) = B_{\tau,\underline{s}}(s)B_{\tau,s}(\hat{s})$.Observe that

$$\frac{\left[1-G_\tau(s)\right]N_{\tau'}^\tau(s)}{J_{\tau'}^\tau} = \int_{\underline{s}}^{\bar{s}} \frac{B_{\tau,s}(\hat{s})g_{\tau'}(\hat{s})}{J_{\tau'}^\tau}\mathbf{1}_{\hat{s}\geq s}d\hat{s} = \frac{1}{B_{\tau,\underline{s}}(s)}\int_{\underline{s}}^{\bar{s}} \frac{B_{\tau,\underline{s}}(\hat{s})g_{\tau'}(\hat{s})}{J_{\tau'}^\tau}\mathbf{1}_{\hat{s}\geq s}d\hat{s}, \tag{J.1}$$

where the first equality is by definition of $N_{\tau'}^\tau(s)$ while the second is by the multiplicative property of B just mentioned. Also, by definition of $J_{\tau'}^\tau$,

$$\int_{\underline{s}}^{\bar{s}} \frac{B_{\tau,\underline{s}}(\hat{s})g_{\tau'}(\hat{s})}{J_{\tau'}^\tau}d\hat{s} = \frac{1}{J_{\tau'}^\tau}\int_{\underline{s}}^{\bar{s}} B_{\tau,\underline{s}}(\hat{s})g_{\tau'}(\hat{s})d\hat{s} = 1,$$

and it's easy to see that $\frac{B_{\tau,\underline{s}}(\hat{s})g_{\tau'}(\hat{s})}{J_{\tau'}^\tau} > 0$ by definitions of $B_{\tau,s}(\hat{s})$ and $J_{\tau'}^\tau$ (Def. 10).

Thus, $\frac{[1-G_\tau(s)]N_{\tau'}^\tau(s)}{J_{\tau'}^\tau}$ is an expectation — with respect to the density $\frac{B_{\tau,s}(\hat{s})g_{\tau'}(\hat{s})}{J_{\tau'}^\tau}$ — of the function $\mathbf{1}_{\hat{s}\geq s}$ viewed as a function of $\hat{s}$. As a result, we can show $\frac{[1-G_\tau(s)]N_{t'}^\tau(s)}{J_{t'}^\tau} < \frac{[1-G_\tau(s)]N_t^\tau(s)}{J_t^\tau}$ — or equivalently $\frac{N_{t'}^\tau(s)}{J_{t'}^\tau} < \frac{N_t^\tau(s)}{J_t^\tau}$ — by showing that $\frac{B_{\tau,\underline{s}}(\hat{s})g_{t'}(\hat{s})}{J_{t'}^\tau} \prec_{\text{MLRP}} \frac{B_{\tau,\underline{s}}(\hat{s})g_t(\hat{s})}{J_t^\tau}$ since the MLRP order implies the first-order stochastic dominance order which is preserved by expectations of increasing functions (such as $\mathbf{1}_{\hat{s}\geq s}$) that are non-constant (such as $\mathbf{1}_{\hat{s}\geq s}$).

[135] In the notation of the direct mechanism (H.1), this is $U_1\left(-m_\tau(s) + r_\tau(s), h\left(m_\tau(s), s\right)\right)$.

Toward that end, observe

$$\frac{B_{\tau,\underline{s}}(\hat{s})g_t(\hat{s})/J^\tau_t}{B_{\tau,\underline{s}}(\hat{s})g_{t'}(\hat{s})/J^\tau_{t'}} = \frac{g_t(\hat{s})}{g_{t'}(\hat{s})}\frac{J^\tau_{t'}}{J^\tau_t},$$

which is strictly increasing in $\hat{s}$ since $\frac{g_t(\hat{s})}{g_{t'}(\hat{s})}$ is by assumption. This completes the proof.

The proof of $\frac{\psi^\tau_{t'}(s)}{\lambda^\tau_{t'}} < \frac{\psi^\tau_t(s)}{\lambda^\tau_t}$ is similar, but we would instead work with $\frac{u'_{\tau s}g_{\tau'}(s)}{\lambda^\tau_{\tau'}}$ instead of $\frac{B_{\tau,s}(\hat{s})g_{\tau'}(\hat{s})}{J^\tau_{\tau'}}$ and use the fact that $\frac{1}{\lambda^\tau_{\tau'}}\int_{\underline{s}}^{\bar{s}} U_{1,\tau,\hat{s}}g_{\tau'}(\hat{s})d\hat{s} = 1$ instead of that $\frac{1}{J^\tau_{\tau'}}\int_{\underline{s}}^{\bar{s}} B_{\tau,\underline{s}}(\hat{s})g_{\tau'}(\hat{s})d\hat{s} = 1$. □

J.3 Discrete Types: Proof of Propositions 3, I.1, I.2, I.3

J.3.1 General Utility, Discrete: Proof of Proposition I.1

To apply the first-order approach to the direct formulation (H.1), we replace the ex-post IC constraints (H.1b) with the corresponding envelope condition ((J.3) below) and ignore all interim IC constraints (H.1a) except the upward adjacent ones ((J.4) below). The resulting relaxed problem can be formulated as follows. [I indicate Lagrange multipliers with the word "mult".]

$$\max_{m_t(\cdot),r_t(\cdot),W_t(\cdot)} \sum_{t=1}^{T} f_t \int U\left(-m_t(s)+r_t(s), h\left(m_t(s),s\right)\right) dG_t(s)$$

$$\text{s.t.} \quad \sum_{t=1}^{T} f_t \int r_t(s)G_t(s) = 0 \quad \text{mult: } -\lambda$$

$$W_t(s) = U\left(-m_t(s)+r_t(s), h\left(m_t(s),s\right)\right), \ \forall t \quad \text{mult: } \alpha_t(s) \tag{J.2}$$

$$\int W_t(s)dG_t(s) \geq \int W_{t+1}(s)dG_t(s) \ \forall t \leq T-1 \quad \text{mult: } \gamma_t \tag{J.3}$$

$$W'_t(s) = U_2\left(-m_t(s)+r_t(s), h\left(m_t(s),s\right)\right)h_2\left(m_t(s),s\right) \ \forall t \quad \text{mult: } \mu_t(s) \tag{J.4}$$

Note: Since T is the highest risk-type, there is no upward IC for $t = T$ and thus no such multiplier as γ_T there-on. However, I define $\gamma_T \equiv 0$ by convention to ease notation.

By the Pontryagin maximum principle [Jahn 2020] (Thm. 5.19), we have the following necessary FOC's, where the last condition (J.6) follows from the fact that the control

variables m_t and r_t are unrestricted at the endpoints $\underline{s}$ and $\bar{s}$ for all t.

$$\begin{aligned}
[m_t(s)] \quad & f_t g_t(s)\left[-U_1 + U_2 h_1\right] - \alpha_t(s)\left[-U_1 + U_2 h_1\right] \\
& - \mu_t(s)\left[-U_{21} h_2 + U_{22} h_1 h_2 + U_2 h_{12}\right] = 0 \qquad \text{(J.5)} \\
[r_t(s)] \quad & f_t g_t(s) U_1 - \alpha_t(s) U_1 - f_t g_t(s)\lambda - \mu_t(s) U_{21} h_2 = 0 \\
[W_t(s)] \quad & \mu_t'(s) = -\alpha_t(s) - g_t(s)\gamma_t + g_{t-1}(s)\gamma_{t-1} \\
[\text{Boundary Condition}] \quad & \mu_t(\underline{s}) = 0 = \mu_t(\bar{s}) \quad \text{since } W_t(s) \text{ has free endpoints.} \qquad \text{(J.6)}
\end{aligned}$$

Using $[r_t(s)]$ to solve for $\alpha_t(s)$,

$$\alpha_t(s) = f_t g_t(s) - f_t g_t(s)\frac{\lambda}{U_1} - \mu_t(s)\left[\frac{U_{21}}{U_1} h_2\right].$$

Then plugging in for $\alpha_t(s)$ in $[W_t(s)]$,

$$\mu_t'(s) = -f_t g_t(s) + f_t g_t(s)\frac{\lambda}{U_1} + \mu_t(s)\left[\frac{U_{21}}{U_1} h_2\right] - g_t(s)\gamma_t + g_{t-1}(s)\gamma_{t-1}. \qquad \text{(J.7)}$$

Plugging in for $\alpha_t(s)$ in $[m_t(s)]$ and successively canceling terms:

$$\begin{aligned}
f_t g_t(s)\frac{\lambda}{U_1}\left[-U_1 + U_2 h_1\right] + \mu_t(s)\left[\frac{U_{21}}{U_1} h_2\right]\left[-U_1 + U_2 h_1\right] - \mu_t(s)\left[-U_{21} h_2 + U_{22} h_1 h_2 + U_2 h_{12}\right] = 0 \\
f_t g_t(s)\frac{\lambda}{U_1}\left[-U_1 + U_2 h_1\right] + \mu_t(s)\left[\frac{U_{21}}{U_1} U_2 h_1 h_2 - U_{22} h_1 h_2 - U_2 h_{12}\right] = 0 \qquad \text{(J.8)}
\end{aligned}$$

Let $(c_t, \pi_t)_{t\in\mathcal{T}}$ be the indirect solution induced by the optimal direct solution as defined in sections 7.1 or H.1. By the ex-post IC (H.1b) for money-loss, if $c_t(\cdot)$ is differentiable $-c_t' U_1 + U_2 h_1 = 0$. Thus we can write (J.8) as

$$\begin{aligned}
f_t g_t(s)\frac{\lambda}{U_1}\left[U_1\left[1 - c_t'\left(m_t(s)\right)\right]\right] + \mu_t(s)\left[\frac{U_{21}}{U_1} U_2 h_1 h_2 - U_{22} h_1 h_2 - U_2 h_{12}\right] = 0 \\
1 - c_t'\left(m_t(s)\right) = -\frac{\mu_t(s)\left[\frac{U_2 h_1}{U_1} U_{21} h_2 - U_{22} h_1 h_2 - U_2 h_{12}\right]}{f_t g_t(s)\lambda} \\
1 - c_t'\left(m_t(s)\right) = \frac{\mu_t(s)\left[-c_t' U_{21} h_2 + U_{22} h_1 h_2 + U_2 h_{12}\right]}{f_t g_t(s)\lambda} \qquad \text{(J.9)}
\end{aligned}$$

Solving the ODE J.7 using the boundary condition $\mu_t(\bar{s}) = 0$,

$$\mu_t(s) = -\int_s^{\bar{s}} \exp\left(\int_s^{\hat{s}} \left[\frac{U_{21}}{U_1} h_2\right]_x dx\right) \left\{ f_t g_t(\hat{s}) \left[-1 + \frac{\lambda}{U_1}\right] - \gamma_t g_t(\hat{s}) + \gamma_{t-1} g_{t-1}(\hat{s}) \right\} d\hat{s}. \tag{J.10}$$

Using the notation introduced in Definition 10 from the beginning of section I.1, we can rewrite (J.10) as:

$$\mu_t(s) = [1 - G_t(s)] \left\{ f_t \left[N_t(s) - \lambda \cdot \chi_t(s) \right] + \gamma_t N_t(s) - \gamma_{t-1} N_{t-1}^t(s) \right\}. \tag{J.11}$$

Plugging in $\mu(\underline{s}) = 0$ to (J.11), we have (for $t = 2, ..., T$):

$$0 = f_t J_t \left[1 - \lambda \frac{\rho_t}{J_t}\right] + \gamma_t J_t - \gamma_{t-1} J_{t-1}^t. \tag{J.12}$$

$$\therefore \; 0 = f_t \nu_t \left[1 - \frac{\lambda}{\nu_t}\right] + \gamma_t \nu_t - \gamma_{t-1} \nu_{t-1}^t, \tag{J.13}$$

where we have divided both sides of (J.12) by ρ_t to get (J.13). Thus we get the following recursive relation for $t = 2, ..., T$:

$$\gamma_{t-1} = \gamma_t \frac{\nu_t}{\nu_{t-1}^t} + f_t \frac{\nu_t}{\nu_{t-1}^t} \left[1 - \frac{\lambda}{\nu_t}\right]. \tag{J.14}$$

Let $R_\ell^\tau \equiv \Pi_{x=\ell}^{\tau} \frac{\nu_{x+1}}{\nu_x^{x+1}}$ for $\ell \leq \tau = 1, \ldots, T-1$, and $R_t^{t-1} \equiv 1 \; \forall t$ by convention.

Claim: the solution has the following form:

$$\gamma_{n-j} = \sum_{k=(n-j)+1}^{n} f_k R_{n-j}^{k-1} \left[1 - \frac{\lambda}{\nu_k}\right] \quad \text{for all } j \in \{1, ..., n-1\}. \tag{J.15}$$

This can be proved by induction. Recalling $\gamma_T \equiv 0$ by convention (since T is the highest type), so $\gamma_{T-1} = R_{T-1}^{T-1} f_T \left[1 - \frac{\lambda}{\nu_T}\right]$ by (J.14), which verifies the base case of $j = 1$. Now suppose

$$\gamma_{T-j} = \sum_{k=T-j+1}^{T} R_{T-j}^{k-1} f_k \left[1 - \frac{\lambda}{\nu_k}\right]$$

for all $j \in \{1, ..., i-1\}$ for some $i \leq T-1$. Then by (J.14),

$$\gamma_{T-i} = \gamma_{T-i+1}\frac{\nu_{T-i+1}}{\nu_{T-i}^{T-i+1}} + f_{T-i+1}\frac{\nu_{T-i+1}}{\nu_{T-i}^{T-i+1}}\left[1 - \frac{\lambda}{\nu_{T-i+1}}\right],$$

but then plugging in for γ_i by induction,

$$\gamma_{T-i} = \left\{\sum_{k=(T-i+1)+1}^{T} R_{T-i+1}^{k-1} f_k \left[1 - \frac{\lambda}{\nu_k}\right]\right\}\frac{\nu_{T-i+1}}{\nu_{T-i}^{T-i+1}} + f_{T-i+1}\frac{\nu_{T-i+1}}{\nu_{T-i}^{T-i+1}}\left[1 - \frac{\lambda}{\nu_{T-i+1}}\right],$$

$$\therefore\ \gamma_{T-i} = \left\{\sum_{k=(T-i+1)+1}^{T} \frac{\nu_{T-i+1}}{\nu_{T-i}^{T-i+1}} R_{T-i+1}^{k-1} f_k \left[1 - \frac{\lambda}{\nu_k}\right]\right\} + f_{T-i+1}\frac{\nu_{T-i+1}}{\nu_{T-i}^{T-i+1}}\left[1 - \frac{\lambda}{\nu_{T-i+1}}\right],$$

$$\therefore\ \gamma_{T-i} = \left\{\sum_{k=(T-i+1)+1}^{T} R_{T-i}^{k-1} f_k \left[1 - \frac{\lambda}{\nu_k}\right]\right\} + f_{T-i+1} R_{T-i}^{T-i}\left[1 - \frac{\lambda}{\nu_{T-i+1}}\right],$$

$$\therefore\ \gamma_{T-i} = \sum_{k=T-i+1}^{T} R_{T-i}^{k-1} f_k \left[1 - \frac{\lambda}{\nu_k}\right].$$

This shows the claim by induction.

At this point, I break up the analysis into two cases: $t = 1$ and $t = 2, ..., T$. First consider the case $t = 1$ so that $\gamma_{t-1} = \gamma_0 \equiv 0$. Plugging γ_t from (J.15) into (J.11),

$$\mu_t(s) = [1 - G_t(s)]\left\{f_t\left[N_t(s) - \lambda\chi_t(s)\right] + \gamma_t N_t(s) - \gamma_{t-1} N_{t-1}^t(s)\right\}$$

$$\mu_t(s) = [1 - G_t(s)]\left\{f_t\left[N_t(s) - \lambda\chi_t(s)\right] + N_t(s)\left[\sum_{k=t+1}^{T} f_k R_t^{k-1}\left[1 - \frac{\lambda}{\nu_k}\right]\right]\right\}$$

By adding and subtracting $N_t(s) f_t\left[1 - \frac{\lambda}{\nu_t}\right]$:

$$\mu_t(s) = [1 - G_t(s)]\left\{f_t\lambda\left[-\chi_t(s) + \frac{N_t(s)}{\nu_t}\right] + N_t(s)\left[\sum_{k=t}^{T} f_k R_t^{k-1}\left[1 - \frac{\lambda}{\nu_k}\right]\right]\right\}, \tag{J.16}$$

where recall $R_t^{t-1} \equiv 1$. Let $Q_\ell^\tau \equiv \Pi_{x=\ell}^{\tau}\frac{\nu_x}{\nu_x^{x+1}}$ for $\ell \leq \tau$ and $Q_\ell^{\ell-1} \equiv 1$.

For $\ell \leq \tau$, we have $R_\ell^\tau \equiv \Pi_{x=\ell}^{\tau}\frac{\nu_{x+1}}{\nu_x^{x+1}} = [\nu_\ell]^{-1}\left[\Pi_{x=\ell}^{\tau}\frac{\nu_x}{\nu_x^{x+1}}\right]\nu_{\tau+1} = [\nu_\ell]^{-1} Q_\ell^\tau \nu_{\tau+1}$ so that,

$$R_\ell^\tau = [\nu_\ell]^{-1} Q_\ell^\tau \nu_{\tau+1}. \tag{J.17}$$

Then (J.16) becomes (again, for $t = 1$ only)

$$\mu_t(s) = [1 - G_t(s)] \left\{ f_t \lambda \left[-\chi_t(s) + \frac{N_t(s)}{\nu_t} \right] + \frac{N_t(s)}{\nu_t} \left[\sum_{k=t}^{T} f_k Q_t^{k-1} [\nu_k - \lambda] \right] \right\}. \tag{J.18}$$

Eventually, when we solve for λ in (J.27), it will become clear that the last term in $\sum_{k=1}^{T} f_k Q_t^{k-1} [\nu_k - \lambda] = 0$ so that

$$\mu_t(s) = [1 - G_t(s)] f_t \lambda \left[-\chi_t(s) + \frac{N_t(s)}{\nu_t} \right] \quad \text{for} \quad t = 1 \quad \text{only.} \tag{J.19}$$

Now turn to the case $t = 2, ..., T$. Plugging into (J.11) for γ_t *and* γ_{t-1} using (J.15) gives

$$\mu_t(s) = [1 - G_t(s)] \left\{ f_t \left[N_t(s) - \lambda \chi_t(s) \right] \right.$$
$$\left. + N_t(s) \left[\sum_{k=t+1}^{T} f_k R_t^{k-1} \left[1 - \frac{\lambda}{\nu_k} \right] \right] - N_{t-1}^t(s) \left[\sum_{k=t}^{T} f_k R_{t-1}^{k-1} \left[1 - \frac{\lambda}{\nu_k} \right] \right] \right\}.$$

By adding and subtracting $N_t(s) f_t \underbrace{R_t^{t-1}}_{1} \left[1 - \frac{\lambda}{\nu_t} \right]$:[136]

$$\mu_t(s) = [1 - G_t(s)] \left\{ f_t \lambda \left[-\chi_t(s) + \frac{N_t(s)}{\nu_t} \right] + \left[\frac{N_t(s)}{R_{t-1}^{t-1}} - N_{t-1}^t(s) \right] \left[\sum_{k=t}^{T} f_k R_{t-1}^{k-1} \left[1 - \frac{\lambda}{\nu_k} \right] \right] \right\}. \tag{J.20}$$

By (J.17), we can rewrite the second term in curly braces in (J.20) as

$$\left[\frac{N_t(s)}{[\nu_{t-1}]^{-1} Q_{t-1}^{t-1} \nu_t} - N_{t-1}^t(s) \right] \left[\sum_{k=t}^{T} f_k [\nu_{t-1}]^{-1} Q_{t-1}^{k-1} \nu_k \left[1 - \frac{\lambda}{\nu_k} \right] \right] \tag{J.21}$$

$$= \left[\frac{N_t(s)}{[\nu_{t-1}]^{-1} \frac{\nu_{t-1}}{\nu_{t-1}^t} \nu_t} - N_{t-1}^t(s) \right] [\nu_{t-1}]^{-1} \left[\sum_{k=t}^{T} f_k Q_{t-1}^{k-1} [\nu_k - \lambda] \right] \tag{J.22}$$

$$= \left[\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t} \right] \frac{\nu_{t-1}^t}{\nu_{t-1}} \left[\sum_{k=t}^{T} f_k Q_{t-1}^{k-1} [\nu_k - \lambda] \right] \tag{J.23}$$

$$= \left[\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t} \right] \left[\sum_{k=t}^{T} f_k Q_t^{k-1} [\nu_k - \lambda] \right]. \tag{J.24}$$

[136] Alternatively, could break up the γ_{t-1} term to get the sums to go from $t+1$ up instead of t. This would make things in terms of a perturbation that makes $t-1$ indifferent rather than t indifferent.

Now, plugging (J.24) in for the second term in curly braces in (J.20) and plugging the latter back into (J.9) (and simplifying) gives (J.25) just below for $t = 2, ..., T$:

$$1 - c_t'\left(m_t(s)\right) = \frac{\left[-c_t' U_{21} h_2 + U_{22} h_1 h_2 + U_2 h_{12}\right]}{f_t \lambda} \frac{1 - G_t(s)}{g_t(s)} \cdot \left\{ f_t \lambda \left[-\chi_t(s) + \frac{N_t(s)}{\nu_t} \right] + \left[\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t} \right] \left[\sum_{k=t}^{T} f_k Q_t^{k-1} \left[-\lambda + \nu_k \right] \right] \right\}. \tag{J.25}$$

Introducing the following notation,

$$1 - \check{F}_{t-1} \equiv \sum_{k=t}^{T} f_k Q_t^{k-1} \quad \text{i.e. } \check{F}_{t-1} \equiv 1 - \sum_{k=t}^{T} f_k Q_t^{k-1} \ \ \forall t \geq 1, \quad \text{and } \check{F}_0 \equiv 1 \text{ by convention,}$$

we make (J.25) actually cover the case $t = 1$ (Eqn. (J.19)) too: The last square-bracketed term in (J.25) becomes $\left[1 - \check{F}_{t-1}\right] \left[\sum_{k=t}^{T} \frac{f_k Q_t^{k-1}}{1 - \check{F}_{t-1}} \left[-\lambda + \nu_k\right]\right]$, which equals zero for $t = 1$ given the convention $\check{F}_0 \equiv 1$. That convention is valid because $\sum_{k=1}^{T} f_k Q_t^{k-1} \left[-\lambda + \nu_k\right] = 0$, as will be shown below (and as was mentioned before (J.19)).

From (J.25), we can arrive at the desired formula (I.2) (in the Proposition) by using the new form of the last square-bracketed term and distributing the $\frac{1}{f_t}$.

Now, we solve for λ. Evaluating (J.15) at $j = T - 1$,

$$\begin{aligned} \gamma_1 = \gamma_{T-[T-1]} &= \sum_{k=(T-[T-1])+1}^{T} f_k R_{T-[T-1]}^{k-1} \left[1 - \frac{\lambda}{\nu_k}\right], \\ \therefore \gamma_1 &= \sum_{k=2}^{T} f_k \left[\nu_1\right]^{-1} Q_1^{k-1} \nu_k \left[1 - \frac{\lambda}{\nu_k}\right], \\ \therefore \gamma_1 &= \left[\nu_1\right]^{-1} \sum_{k=2}^{T} f_k Q_1^{k-1} \left[\nu_k - \lambda\right], \end{aligned} \tag{J.26}$$

where the third line uses (J.17). From (J.13) evaluated at $t = 1$, and subsequently plugging in (J.26) for γ_1,

$$\begin{aligned} 0 &= f_1 \nu_1 \left[1 - \frac{\lambda}{\nu_1}\right] + \gamma_1 \nu_1, \\ \therefore 0 &= f_1 \left[\nu_1 - \lambda\right] + \sum_{k=2}^{T} f_k Q_1^{k-1} \left[\nu_k - \lambda\right]. \end{aligned}$$

Solving for λ gives

$$\lambda = \left[\sum_{k=1}^{T} f_k Q_1^{k-1}\right]^{-1} \left[\sum_{k=1}^{T} f_k Q_1^{k-1} \nu_k\right]. \tag{J.27}$$

which is the desired formula (I.5) from the proposition. As foreshadowed above, (J.27) makes it easy to see that $\sum_{k=1}^{T} f_k Q_t^{k-1} [\nu_k - \lambda] = 0$, thus justifying the $t = 1$ form of $\mu_t(s)$ (Eqn. (J.19)) and, ultimately, the use of the convention $\check{F}_0 \equiv 1$. This completes the proof of Proposition I.1.

J.3.2 Proof of Lemma I.2

Recall

$$L_s(\hat{s}) \equiv \exp\left(-\int_s^{\hat{s}} \left\{\left[-U_{11} c'(m_t(\tilde{s})) + U_{21} h_1\left(m_t(\tilde{s}), \tilde{s}\right)\right] / \left[U_1\right]\right\} \dot{m}_t(\tilde{s}) d\tilde{s}\right) \tag{J.28}$$

where the arguments of the derivatives of U are evaluated at $(-\pi_t - c_t(m_t(\tilde{s})), h\left(m_t(\tilde{s}), \tilde{s}\right))$.

Lemma I.2 (Restated)**.** *For all $s \leq \hat{s}$, $B_{t,s}(\hat{s}) = \frac{U_{1,t,\hat{s}}}{U_{1,t,s}} L_{t,s}(\hat{s})$. In the special case of money-loss, $L_{t,s}(\hat{s}) = 1$, so in particular $B_{t,s}(\hat{s}) = \frac{U_{1t\hat{s}}}{U_{1ts}}$ there.*

Proof. Omitting t subscripts and some arguments, we have the chain of reasoning below. The first equality is by adding and subtracting and adding and subtracting. The second equality (not including $\equiv$) is simply carrying out the differentiation $\frac{d}{d\tilde{s}}$ in the 3rd step to get the expression in the second step. The third equality is by the fundamental theorem of calculus, where I use the notation $[f(x)]_s^{\hat{s}} \equiv f(\hat{s}) - f(s)$ for a generic function f. The last

equality is by algebra and the definition of $L_{t,s}(\hat{s})$.

$$
\begin{aligned}
B_{t,s}(\hat{s}) &\equiv \exp\left(\int_s^{\hat{s}} \left[\frac{U_{21,\tilde{s}}}{U_{1,\tilde{s}}} h_2\left(m_t(\tilde{s}),\tilde{s}\right)\right] d\tilde{s}\right) \\
&= \exp\left(\int_s^{\hat{s}} \left[\frac{U_{21}}{U_1} h_2\left(m(\tilde{s}),\tilde{s}\right) + \left[-U_{11}c'(m(\tilde{s})) + U_{21}h_1\right]\dot{m}(\tilde{s})/\left[U_1\right]\right] d\tilde{s}\right. \\
&\qquad \left. - \int_s^{\hat{s}} \left\{\left[-U_{11}c'(m(\tilde{s})) + U_{21}h_1\right]/\left[U_1\right]\right\}\dot{m}(\tilde{s}) d\tilde{s}\right) \\
&= \exp\left(\int_s^{\hat{s}} \frac{d}{d\tilde{s}} \log U_1\left(-\pi - c(m(\tilde{s})), h\left(m(\tilde{s}),\tilde{s}\right)\right) d\tilde{s}\right. \\
&\qquad \left. - \int_s^{\hat{s}} \left\{\left[-U_{11}c_t'(m(\tilde{s})) + U_{21}h_1\right]/\left[U_1\right]\right\}\dot{m}(\tilde{s}) d\tilde{s}\right) \\
&= \exp\left(\left[\log U_1\left(-\pi - c(m(\tilde{s})), h\left(m(\tilde{s}),\tilde{s}\right)\right)\right]_s^{\hat{s}}\right) \\
&\quad \cdot \exp\left(-\int_s^{\hat{s}} \left\{\left[-U_{11}c'(m(\tilde{s})) + U_{21}h_1\right]/\left[U_1\right]\right\}\dot{m}(\tilde{s}) d\tilde{s}\right) \\
&= \frac{U_{1\hat{s}}}{U_{1s}} L_{t,s}\left(\hat{s}\right).
\end{aligned}
$$

The above chain shows $B_{t,s}(\hat{s}) = \frac{U_{1,t,\hat{s}}}{U_{1,t,s}} L_{t,m(s)}(\hat{s})$ as desired. Now consider the money-loss case. $\left[-U_{11}c'(\tilde{m}) + U_{21}h_1\right] = -u''c'(\tilde{m}) + u''h_1 = u''\left[-c'(\tilde{m}) + h_1\right] = 0$, so $L_{t,s}(\hat{s}) = 1$ and thus $B_{t,s}(\hat{s}) = \frac{U_{1t\hat{s}}}{U_{1ts}}$. □

J.3.3 Money-Loss as Special Case of General: Proof of Proposition 3

In this subsection, I prove Proposition 3 by showing that the general utility formula (I.2) from Proposition I.1 (proved in Appendix J.3.1) specializes to the money-loss formula (7.3) from Proposition 3 in the main text.

We now use Lemma I.2 to simplify the notation from Definition 10 as follows.

$N_t^\tau(s) \equiv \int_s^{\bar{s}} B_{\tau,s}(\hat{s}) \frac{g_t(\hat{s})}{1-G_\tau(s)} d\hat{s} = \frac{1}{U_{1,\tau,s}} \int_s^{\bar{s}} U_{1,\tau,\hat{s}} \frac{g_t(\hat{s})}{1-G_\tau(s)} d\hat{s} = \frac{1}{U_{1,\tau,s}} \psi_t^\tau(s)$.

$\chi_t(s) \equiv \int_s^{\bar{s}} B_{t,s}(\hat{s}) \frac{1}{U_{1,t,\hat{s}}} \frac{g_t(\hat{s})}{1-G_t(s)} d\hat{s} = \frac{1}{U_{1,t,s}}$,

$\rho_t \equiv \chi_t(\underline{s}) = \frac{1}{U_{1,t,\underline{s}}}$.

$J_t^\tau \equiv N_t^\tau(\underline{s}) = \frac{1}{U_{1,\tau,\underline{s}}} \int_{\underline{s}}^{\bar{s}} U_{1,\tau,\hat{s}} g_t(\hat{s}) d\hat{s} = \frac{1}{U_{1,\tau,\underline{s}}} \lambda_t^\tau$, where recall $\lambda_t^\tau \equiv \int_{\underline{s}}^{\bar{s}} U_{1,\tau,\hat{s}} g_t(\hat{s}) d\hat{s}$ (Eqn. (7.2a)).

Thus $\nu_t^\tau \equiv \frac{J_t^\tau}{\rho_t} = \lambda_t^\tau$.

Thus $Q_\ell^\tau \equiv \Pi_{x=\ell}^\tau \frac{\nu_x}{\nu_x^{x+1}} = \Pi_{x=\ell}^\tau \frac{\lambda_x}{\lambda_x^{x+1}}$ for $\ell \leq \tau$ and $Q_\ell^{\ell-1} \equiv 1$. Thus, our definition of Q_ℓ^τ for general utility form the previous sentence — which was officially given equation (I.3) of

Proposition I.1 — specializes to the definition we gave for money-loss in (I.8) of Proposition 3, namely $\Pi_{x=\ell}^{\tau}\frac{\lambda_x}{\lambda_x^{x+1}}$. The same then applies automatically to the notation $1-\check{F}_{t-1}\equiv\sum_{k=t}^{T}f_kQ_t^{k-1}$, which was also used in both the general utility and money-loss formulas. The convention $\check{F}_0\equiv 1$ needs no translation.

With this notation in place, we turn to simplifying the main expressions obtained for general utility.[137]

Looking at the expression (J.25), the first term in curly braces is

$$f_t\lambda\left[-\chi_t(s)+\frac{N_t(s)}{\nu_t}\right]=f_t\lambda\frac{1}{U_{1,t,s}}\left[-1+\frac{\psi_t(s)}{\lambda_t}\right],$$

and the second term in curly braces is

$$\begin{aligned}&\rho_t\left[\frac{N_t(s)}{\nu_t}-\frac{N_{t-1}^t(s)}{\nu_{t-1}^t}\right]\left[\sum_{k=t}^{T}f_kQ_t^{k-1}\left[-\lambda+\nu_k\right]\right]\\=&\frac{1}{U_{1,t,\underline{s}}}\left[\frac{\frac{1}{U_{1,t,s}}\psi_t(s)}{\frac{1}{U_{1,t,\underline{s}}}\lambda_t}-\frac{\frac{1}{U_{1,t,s}}\psi_{t-1}^t(s)}{\frac{1}{U_{1,t,\underline{s}}}\lambda_{t-1}^t}\right]\left[\sum_{k=t}^{T}f_kQ_t^{k-1}\left[-\lambda+\lambda_k\right]\right]\\=&\frac{1}{U_{1,t,s}}\left[\frac{\psi_t(s)}{\lambda_t}-\frac{\psi_{t-1}^t(s)}{\lambda_{t-1}^t}\right]\left[\sum_{k=t}^{T}f_kQ_t^{k-1}\left[-\lambda+\lambda_k\right]\right].\end{aligned}$$

Thus for money-loss, the whole curly braced expression (in (J.25)) is

$$f_t\lambda\frac{1}{U_{1,t,s}}\left[-1+\frac{\psi_t(s)}{\lambda_t}\right]+\frac{1}{U_{1,t,s}}\left[\frac{\psi_t(s)}{\lambda_t}-\frac{\psi_{t-1}^t(s)}{\lambda_{t-1}^t}\right]\left[\sum_{k=t}^{T}f_kQ_t^{k-1}\left[-\lambda+\lambda_k\right]\right].$$

Thus overall, (J.25) becomes

$$\begin{aligned}1-c_t'\left(m_t(s)\right)=&\frac{\left[-c_t'U_{21}h_2+U_{22}h_1h_2+U_2h_{12}\right]}{\lambda U_{1,t,s}}\frac{1-G_t(s)}{g_t(s)}\cdot\Bigg\{\\&\lambda\left[-1+\frac{\psi_t(s)}{\lambda_t}\right]+\frac{1}{f_t}\left[\frac{\psi_t(s)}{\lambda_t}-\frac{\psi_{t-1}^t(s)}{\lambda_{t-1}^t}\right]\left[-\lambda+\sum_{k=t}^{T}f_kQ_t^{k-1}\lambda_k\right]\Bigg\}.\end{aligned}$$

[137] Again, I work with the expression (J.25) from the end of the proof of Proposition I.1 rather than the expression (I.2) since the latter is in terms of J_τ^t and ρ_t rather than $\nu_\tau^t\equiv\frac{J_\tau^t}{\rho_t}$.

But

$$\frac{1}{U_{1,t,s}}\left[-c_t' U_{21} h_2 + U_{22} h_1 h_2 + U_2 h_{12}\right] = \frac{1}{u_{t,s}'}\left[u_{t,s}'' h_2 \underbrace{\left[-c_t' + h_1\right]}_{0} + u_{t,s}' h_{12}\right] = h_{12},$$

where $-c_t' + h_1 = 0$ by the ex-post FOC under money-loss. Thus we have

$$1 - c_t'\left(m_t(s)\right) = \frac{h_{12}}{\lambda}\frac{1 - G_t(s)}{g_t(s)} \cdot \left\{\lambda\left[-1 + \frac{\psi_t(s)}{\lambda_t}\right]\right.$$
$$\left. + \frac{1}{f_t}\left[\frac{\psi_t(s)}{\lambda_t} - \frac{\psi_{t-1}^t(s)}{\lambda_{t-1}^t}\right]\left[-\lambda + \sum_{k=t}^{T} f_k Q_t^{k-1} \lambda_k\right]\right\}.$$

This is the same formula (I.7) we got in the main text, after using the notation $1 - \check{F}_{t-1} \equiv \sum_{k=t}^{T} f_k Q_t^{k-1}$.

Finally, we simplify λ. From (J.27)

$$\lambda = \left[\sum_{k=t}^{T} f_k Q_t^{k-1}\right]^{-1}\left[\sum_{k=t}^{T} f_k Q_t^{k-1}\frac{J_k}{\rho_k}\right]$$
$$= \left[\sum_{k=t}^{T} f_k Q_t^{k-1}\right]^{-1}\left[\sum_{k=t}^{T} f_k Q_t^{k-1}\lambda_k\right],$$

which is the desired formula (I.9) from Proposition I.2.

We now derive the CARA-specialized formulas.

Lemma I.3 (Restated)**.** *Under money-loss CARA, $\lambda_{t-1}^t = \lambda_{t-1}$ for all $t \geq 2$.*

Proof. When u is CARA, marginal utility u_{ts}' is proportional to utility u, so expected marginal utility is proportional to expected utility. Under money-loss CARA, upward IC binds by Lemma 3, which means that $t-1$'s expected utility is equal whether they choose their own plan or the one intended for t. Proportionality then implies $\lambda_{t-1}^t = \lambda_{t-1}$. □

Recalling $Q_\ell^\tau = \Pi_{x=\ell}^{\tau} \frac{\lambda_x}{\lambda_x^{x+1}}$, this then implies $Q_\ell^\tau = 1$ under CARA. This in turn allows us to specialize $\check{F} = F$ and $\lambda = \sum_{k=t}^{T} f_k \lambda_k$ under CARA. Finally, for CARA, we rewrite the convention $\check{F}_0 \equiv 1$ as $F_0 \equiv 1$. This concludes the proof of Proposition I.2 (money-loss without CARA), and Proposition 3 (Money-Loss with CARA) in the main text.

J.3.4 Separable as Special Case of General : Proof of Proposition I.3

This subsection proves Proposition I.3 by specializing the general utility result of Proposition I.1 which was proved at the beginning of this appendix (subsection J.3.1). For separable, $U_{21} = 0$, so that $B_{t,s}(\hat{s}) \equiv \exp\left(\int_s^{\hat{s}} \left[\frac{U_{21}}{U_1} h_2\right]_{t,\tilde{s}} d\tilde{s}\right) = 1$. Then the notation from Definition 10, simplifies as follows.

$N_t^\tau(s) = \frac{1-G_t(s)}{1-G_\tau(s)}$, $N_t(s) = 1$. $\chi_t(s) \equiv \int_s^{\bar{s}} \frac{1}{U_{1,t,s}} \frac{g_t(\hat{s})}{1-G_t(s)} d\hat{s}$.

$J_t^\tau \equiv N_t^\tau(\underline{s}) = 1$, and let $\rho_t \equiv \chi_t(\underline{s}) = \int_{\underline{s}}^{\bar{s}} \frac{1}{U_{1,t,s}} g_t(s) ds$, so that $\nu_t^\tau \equiv \frac{J_t^\tau}{\rho_\tau} = \frac{1}{\rho_\tau} = \left[\int_{\underline{s}}^{\bar{s}} \frac{1}{U_{1,t,s}} g_t(s) ds\right]^{-1}$ (and $\nu_t \equiv \nu_t^t$ as always).

I claim we will also have $Q_\ell^\tau = \frac{\rho_{\tau+1}}{\rho_\ell}$ for $\ell \leq \tau$ (and $Q_\ell^{\ell-1} \equiv 1$). To see this, observe $Q_\ell^\tau \equiv \Pi_{x=\ell}^\tau \frac{\nu_x}{\nu_x^{x+1}} = \frac{\nu_\ell}{\nu_\tau^{\tau+1}} \Pi_{x=\ell+1}^\tau \frac{\nu_x}{\nu_{x-1}^x}$, which equals $\frac{J_\ell/\rho_\ell}{J_\tau^{\tau+1}/\rho_{\tau+1}} \Pi_{x=\ell+1}^\tau \frac{J_x/\rho_x}{J_{x-1}^x/\rho_x} = \frac{\rho_{\tau+1}}{\rho_\ell} \Pi_{x=\ell}^\tau \frac{J_x}{J_x^{x+1}}$ which equals $\frac{\rho_{\tau+1}}{\rho_\ell}$ since $J_t^\tau = 1$.

We now apply these "translations" to the main expressions obtained for general utility. I work with the expression (J.25) from the end of the proof of Proposition I.1 rather than the expression (I.2) since the latter hasn't yet implemented the notation $1 - \check{F}_{t-1} \equiv \sum_{k=t}^T f_k Q_t^{k-1}$.

The first term in curly brackets in (J.25) is

$$f_t \lambda \left[-\chi_t(s) + \frac{N_t(s)}{\nu_t}\right] = f_t \lambda \left[\int_{\underline{s}}^{\bar{s}} \frac{1}{U_{1,t,s}} g_t(s) ds - \int_s^{\bar{s}} \frac{1}{U_{1,t,\hat{s}}} \frac{g_t(\hat{s})}{1-G_t(s)} d\hat{s}\right].$$

And the second term in curly brackets in (J.25) is

$$\begin{aligned}
&\left[\frac{N_t(s)}{\nu_t} - \frac{N_{t-1}^t(s)}{\nu_{t-1}^t}\right] \left[\sum_{k=t}^T f_k Q_t^{k-1} \left[-\lambda + \nu_k\right]\right] \\
&= \rho_t \left[\frac{N_t(s)}{J_t} - \frac{N_{t-1}^t(s)}{J_{t-1}^t}\right] \left[\sum_{k=t}^T f_k \frac{\rho_k}{\rho_t} \left[-\lambda + \frac{J_k}{\rho_k}\right]\right] \\
&= \left[\frac{N_t(s)}{J_t} - \frac{N_{t-1}^t(s)}{J_{t-1}^t}\right] \left[\sum_{k=t}^T f_k \left[-\rho_k \lambda + J_k\right]\right] \\
&= \left[\frac{1}{1-G_t(s)}\right] \left[\left[1-G_t(s)\right] - \left[1-G_{t-1}(s)\right]\right] \left[\sum_{k=t}^T f_k \left[1 - \lambda \int_{\underline{s}}^{\bar{s}} \frac{1}{U_{1,t,s}} g_t(s) ds\right]\right].
\end{aligned}$$

Overall we have

$$1-c_t'\left(m_t(s)\right)=\frac{\left[-c_t'\underbrace{U_{21}}_{0}h_2+U_{22}h_1h_2+U_2h_{12}\right]}{f_tg_t(s)\lambda}\left[1-G_t(s)\right]$$
$$\cdot\left\{f_t\lambda\left[\int_{\underline{s}}^{\bar{s}}\frac{1}{U_{1,t,s}}g_t(s)ds-\int_{s}^{\bar{s}}\frac{1}{U_{1,t,\hat{s}}}\frac{g_t(\hat{s})}{1-G_t(s)}d\hat{s}\right]\right.$$
$$\left.+\left[\frac{1}{1-G_t(s)}\right]\left[\left[1-G_t(s)\right]-\left[1-G_{t-1}(s)\right]\right]\left[\sum_{k=t}^{T}f_k\left[1-\lambda\int_{\underline{s}}^{\bar{s}}\frac{1}{U_{1,t,s}}g_t(s)ds\right]\right]\right\}.$$

Letting $F_0\equiv 1$ and simplifying gives the final answer for $t=1,...,T$:

$$1-c_t'\left(m_t(s)\right)=\left[U_{22}h_1h_2+U_2h_{12}\right]\left\{\right.$$
$$\frac{1-G_t(s)}{g_t}\left[\int_{\underline{s}}^{\bar{s}}\frac{1}{U_{1,t,s}}g_t(s)ds-\int_{s}^{\bar{s}}\frac{1}{U_{1,t,\hat{s}}}\frac{g_t(\hat{s})}{1-G_t(s)}d\hat{s}\right]$$
$$\left.+\frac{1-F_{t-1}}{f_t}\left[\left[\frac{1-G_t(s)}{g_t(s)}\right]-\left[\frac{1-G_{t-1}(s)}{g_t(s)}\right]\right]\left[\frac{1}{\lambda}-\sum_{k=t}^{T}\left[\frac{f_k}{1-F_{t-1}}\int_{\underline{s}}^{\bar{s}}\frac{1}{U_{1,k,s}}g_k(s)ds\right]\right]\right\}.$$

This is (I.11), as desired.

Finally, we simplify λ. Recall we have $\nu_t^\tau=\frac{1}{\rho_t}=\left[\int_{\underline{s}}^{\bar{s}}\frac{1}{U_{1,t,s}}g_t(s)ds\right]^{-1}$ for all t,τ, and $Q_\ell^\tau=\frac{\rho_{\tau+1}}{\rho_\ell}$ have for $\ell\leq\tau$.

From (J.27) — or equivalently (I.5) in the general-utility result (Proposition I.3) — we have

$$\begin{aligned}\lambda&=\left[\sum_{k=1}^{n}f_kQ_1^{k-1}\right]^{-1}\left[\sum_{k=1}^{n}f_kQ_1^{k-1}\nu_k\right]\\&=\left[\sum_{k=1}^{n}f_k\frac{\rho_k}{\rho_1}\right]^{-1}\left[\sum_{k=1}^{n}f_k\frac{\rho_k}{\rho_1}[\rho_k]^{-1}\right]\\&=\left[\sum_{k=1}^{n}f_k\rho_k\right]^{-1}\\&=\left[\sum_{k=1}^{n}f_k\left[\int_{\underline{s}}^{\bar{s}}\frac{1}{U_{1,k,s}}g_k(s)ds\right]\right]^{-1},\end{aligned}\tag{J.29}$$

which is the desired formula (I.12) in the separable-utility result (Proposition I.3). This concludes the proof of Proposition I.3.