

Optimal Health Insurance for Multiple Goods and Time Periods

Randall P. Ellis^{a,*}, Shenyi Jiang^b, Willard G. Manning^c

^a Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215, USA.

<ellisrp@bu.edu>

^b Renmin University of China, Beijing, P.R. China. <syijiang@gmail.com>

^c Harris School of Public Policy Studies, The University of Chicago, Chicago IL USA. <w-manning@uchicago.edu>

Acknowledgements: We have benefited from comments during presentations at Boston University, Harvard University, Renmin University, Peking University, The University of Chicago, the AEA annual meeting, and the American Society of Health Economists conference. We also thank David Bradford, Kate Bundorf, Albert Ma, Tom McGuire, David Meltzer, Joe Newhouse, Kosali Simon, Frank Sloan, Wenija Zhu, and two anonymous referees for their useful comments. The opinions expressed are those of the authors, and not those of Boston University, Renmin University, or The University of Chicago.

Abstract

We examine the efficiency-based arguments for second-best optimal health insurance with multiple treatment goods and multiple time periods. Correlated shocks across health care goods and over time interact with complementarity and substitutability to affect optimal cost sharing. Health care goods that are substitutes or have positively correlated demand shocks should have lower optimal patient cost sharing. Positive serial correlations of demand shocks and uncompensated losses that are positively correlated with covered health services also reduce optimal cost sharing. Our results rationalize covering pharmaceuticals and outpatient spending more fully than is implied by static, one good, or one period models.

1. Introduction

A central theme in health economics is the behavior of patients and providers in the presence of partial health insurance that covers some but not all of the cost of health care. The economic motivation for such coverage is that risk-averse individuals can reduce their financial risk by pooling risks with others through insurance that effectively shifts funds from healthy to sick states of the world to reduce the financial risk from uncertain health. Partial, rather than complete, insurance can be desirable to offset the adverse effects of moral hazard when the marginal payment paid by the consumer for a service is less than the social cost of producing that service. Given that patients respond to lower out-of-pocket prices of health care, health insurance increases the amount of the care that they purchase, generating a deadweight loss from the increased use. There is a great deal of empirical support for the law of demand applying to health care in the literature, coming from observational studies, natural experiments, and randomized experiments (Newhouse et al., 1981; Manning et al., 1987; Zweifel and Manning, 2000; Baicker and Finkelstein, 2011).

We focus here on the normative issues of what the second-best coinsurance policies should be over multiple treatment goods and multiple periods. This topic has taken on new importance in the recent American debate on health care reform, where demand-side cost sharing has reemerged as a central cost containment strategy. Cost sharing plays a central role in the health insurance exchanges, which promote a series of plan tiers (platinum, gold, silver, bronze) that differ primarily in cost sharing. This issue is also central in Value-Based Insurance, which promotes higher coverage for more effective services (Fendrick et al, 2001; Choudhry et al, 2010). Key issues in this policy discussion are what “essential benefits” should be covered, and

how to justify higher or lower consumer cost sharing. In an earlier paper, we examined optimal insurance where the tradeoff is between one treatment good and one preventive service (Ellis and Manning, 2007); in this paper, we focus on the new issues that arise from more than one dimension of uncertainty coming from multiple treatment goods and multiple time periods.

Although we are fully aware of the new insights for insurance from behavioral economics (e.g., Kunreuther, Pauly and McMorow, 2012), we nonetheless adopt a neoclassical approach and derive results based on a model of expected utility maximization in which health care demand curves reflect true marginal benefits, consumers are risk averse, and only demand side cost sharing is available to influence costs. Even if consumers make the wrong choices because of misperceptions, inertia, or other behavioral frictions, it can still be of interest to model optimal insurance more formally as a gold standard with which to contrast imperfect choices. Even under alternative assumptions, there will always be a tradeoff between deadweight losses from moral hazard and welfare gains from the risk protection offered by incomplete insurance.

Our work is not the first to carefully consider multiple health treatment goods (Goldman and Philipson, 2007), but to our knowledge it is the first to examine the issues that arise from multiple, correlated sources of uncertainty, whether the correlations arise due to correlated demands across different health care goods at a point in time, correlated demand shocks over time, or the correlated losses that arise from uncompensated losses (e.g. job loss) that accompany the treatment cost shocks that can be covered by the insurance plan. We derive four new findings about second-best optimal insurance:

1. Holding variances constant, positive contemporaneous correlations among goods justify lower coinsurance rates for at least one good, while negative correlations justify higher coinsurance rates. This finding modifies previous findings on the role of complements vs. substitutes.
2. Positive serial correlation over multiple periods (such as with chronic care or prescriptions filled) justifies lower coinsurance rates for at least one good. Static models of only one period miss the implied accumulation of risk over time.
3. Health-related losses that are not fully compensated by insurance (such as lost income or time costs of treatment) justify lower coinsurance rates.
4. Savings behavior modifies optimal coinsurance rates and its effect depends on the nature of serial correlation.

In the following sections of the paper, we provide a systematic review of the existing literature, highlighting that our model generalizes two limiting cases that are already in the literature. After deriving comparative statics for our model, we numerically evaluate the relevance of our approach using empirical estimates from a set of large employment based insurance plans. As we proceed, we show how more traditional rules for coinsurance rates change as a result of our extensions beyond the one good, one period model.

2. Literature Review

Much of the economic literature on optimal health insurance focuses on “the fundamental tradeoff of risk spreading and appropriate incentives” (Cutler and Zeckhauser, 2000, p. 576). Specifically, it examines either the dead weight losses from moral hazard, the tradeoff between moral hazard and the gains from insuring against financial risk, or the differential coverage of multiple goods with varying degrees of risk. Much of this work employs a one-period, one health care good model with uncertainty about health states or uncertainty about levels of health care expenditure [Arrow, 1963; Besley, 1988; Cutler and Zeckhauser, 2000; Dardanoni and Wagstaff, 1990; Pauly, 1968; Spence and Zeckhauser, 1971; Zeckhauser, 1970]. Papers that derive the optimal insurance contracts using this framework have employed variants of the tradeoff between the risk premium (as reflected by the Arrow-Pratt approximation) and the deadweight loss from moral hazard (as reflected in Harberger’s loss or related measures) or the compensating variation (Marquis and Manning, 1999).¹

Almost all of this literature has been based on either a one-period model or a two-period model where the consumer selects the coinsurance before knowing her realized state of health, but chooses her health care expenditures conditional on the state of the world that occurs. The common result in this literature is that one should select coverage in a plan with a constant coinsurance rate such that the marginal gains from risk reduction from a change in the coinsurance rate just equals the marginal costs of increasing moral hazard. The consensus of the empirical strain of this literature is that optimal levels of cost sharing usually involve neither full

¹ See Feldstein (1973), Feldstein and Freedman (1977), Buchanan and Keeler (1991), Manning and Marquis (1996), Newhouse et al. (1993), Feldman and Manning (1997) for other theoretical and empirically-based examinations of optimal insurance. Blomqvist (1997) extends the theory to nonlinear insurance schedules.

insurance (zero out-of-pocket cost), nor no insurance. Depending on the formal model approach and the data employed, optimal coinsurance rates range from the 50-60 percent range (Feldstein and Friedman, 1977; Manning and Marquis, 1996) down to values that are in the mid 20 percent range or lower, possibly with a deductible and/or stop-loss (Blomqvist, 1997; Buchanan et al., 1991; Feldman and Dowd, 1991; Feldman and Manning, 1997; Newhouse et al., 1993).

In this paper, we reconsider the optimal insurance for health care in markets with one nonhealth good and two or more health care goods. We do this in the context of both one and two period models. We are aware of only three papers that have considered multiple health goods. Besley (1988) provides a multi-good extension to this literature in which demands for health care goods are stochastic, but does not model either complementarity or substitutability between goods or consider correlated shocks in demand. Goldman and Philipson (2007) model two goods in one period to show how complementarity and substitutability of health care services affects optimal cost sharing in an expected utility format. However, they do not model multiple dimensions of uncertainty nor do they distinguish the separate roles of correlated risks and complementarity vs. substitutes in their expected utility framework. Ellis and Manning (2007) model the case of one treatment good and one prevention good to highlight how optimal insurance rules differ for the two types of goods, but they include only one health demand shock and do not model multiple treatment goods, correlations over time, or savings behavior.

To keep the model tractible, we take a step backwards from Ellis and Manning and do not examine the preventive effects of health care spending. In doing so we also differentiate our work from Hall and Jones (2007) in which health care affects mortality and life expectancy, but not utility, directly. Using the same utility structure as Ellis and Manning, we examine how optimal cost sharing is affected by the correlation structure of random shocks affecting demand for health care both across goods and over time. We also examine how savings and uncompensated health losses affect the optimal insurance calculations.

The important insight here is that, all other things being equal, health care goods that are positively correlated should be more generously insured than those that are negatively correlated or uncorrelated. This holds both for contemporaneously correlated health care treatment goods and shocks that are serially correlated over time. The basic logic is that if the demand for two goods or over two periods is uncertain, then the *ex ante* combined variance is larger if they are

positively correlated than if there is no or negative correlation.² Risk averse individuals will prefer more generous insurance (lower coinsurance rates) to reduce their financial risk when the health shocks affecting health goods are positively correlated than if demand shocks are independent or negatively correlated.

We first study the optimal demand-side insurance coverage for health care treatment when there are two health care treatment goods. After developing a general analytical model as our “base case,” we do comparative statics on a series of special cases. A key attraction of our specification is that we not only solve for the optimal cost sharing rate as a closed-form solution to recreate the results from the basic model but also address a richer class of economic behavior. These include new results involving uncompensated health care losses, correlated health care shocks, and cross price elasticities of demand with multiple goods.

Our second set of analytical results considers a two-period model in which health care shocks persist over time due to chronic conditions. In a multi-period context, if a consumer’s savings react to healthcare shocks, then both the cost of risk and the optimal cost sharing rates change. Positively serially correlated shocks (e.g. chronic illness) imply that healthcare should be more generously covered (lower cost sharing) than when shocks are independent or negatively correlated across periods.

Our next set of analytical results considers the two good, two period model, with both contemporaneous and serial correlations, which gives insights into optimal coverage for acute versus chronic conditions. Optimal copayments with two periods are higher unless shocks are perfectly correlated over time.

The concluding section of the paper discusses a few empirical results that have a bearing on our analytical findings. We briefly discuss empirical estimates of the variance of three broad sets of services and the magnitudes of contemporaneous and intertemporal correlations that shed light on the empirical relevance of our findings.

3. Model

Following Arrow (1963), Pauly (1968), and Zeckhauser (1970), we focus only on the demand-side while examining optimal cost sharing. We consider risk aversion, moral hazard,

²As shown below, the details are more complicated because optimal coinsurance also depends on the variability in demand, own and cross price effects, correlated error shocks, and (in the cases of multiple periods) the interest and discount rates.

and uncompensated losses without attempting to model other concerns that influence optimal insurance coverage such as the following: correcting for externalities, such as those that can occur with communicable diseases (Hofmann, 2007); altruism or public good arguments for insurance coverage (Coate, 1995; Rask and Rask, 2000); effects of insurance on technological change (Goddeeris, 1984); distributional concerns including goals like the elimination of poverty or achieving social solidarity (Andrulis, 1998; Atim, 1999; Maarse and Paulus, 2003; Sin et al., 2003), corrections of informational problems (i.e. consumers make the wrong decisions) (Doherty and Thistle, 1996; Choudhry et al, 2010); or insurance that fosters more complete coordination among health care providers (Duggan, 2004; Lichtenberg, 2002; and Newhouse, 2006). Without denying the relevance of these other factors affecting optimal insurance, we set them aside and reexamine the efficiency-based arguments for insurance with multiple goods and periods and derive new results which refine our understanding of the value of generous insurance coverage from the consumer's point of view.

We assume the individual's utility function involves one non-health consumption good, Y , and two health treatment goods, $X = \{X_1, X_2\}$, first within one period and then over two periods.³ Prices of the three goods are P_Y, P_1 , and P_2 , and the consumer's income in each period is I . In the underlying behavioral model, there is a health production function that transforms health care X into health status.

Following much of the literature (for example, Cutler and Zeckhauser, 2000), we examine only health insurance plans with a constant coinsurance rate $0 \leq c \leq 1$ for both health treatment goods. There are no new insights gained from introducing loading costs (as this is well covered in Arrow, 1963, and Ellis and Manning, 2007); therefore, we simplify our model and assume loading costs to be zero throughout this paper. Insurance premiums, π , are competitively determined and depend on the coinsurance rates and the demand structure, but do not vary across individuals. The insurance policy is thus a pure coinsurance plan with no deductible, stop-loss, or limit on the maximum expenditure or level(s) of covered services.

We assume the following sequence of steps in our full model:

³In an earlier version of the paper, we used more general notation that allowed for an arbitrary number of health care goods, but we found that all of the relevant intuition is obtained using only two health care goods.

1. The insurer chooses the premium π and coinsurance rates $\{c_1, c_2\}$ for health care treatment for goods X_1 and X_2 , respectively.
2. Nature decides the consumer's state of illness as a vector of random health shocks, $\{\theta_i\}$, that affect the demand for the vector of health care goods, X .
3. The consumer chooses quantities X and Y to maximize period 1 utility.
4. If a two-period model, repeat steps 2 and 3 for Period 2.

Conditional on the available technology, the demand for medical care services has been shown by many empirical studies to be highly income inelastic, with $\mu_I = .1$ or $.2$, (Chernew and Newhouse, 2012).⁴ For simplicity, we assume that X is perfectly income inelastic: $\partial X / \partial I = 0$. While this income elasticity assumption is strong, it buys us a great deal of simplicity that enables many closed form solutions for cases with multiple health treatment goods. While we make this strong assumption about income elasticities, we make weaker assumptions about other parameters below. We avoid concerns about corner solutions by further assuming that income is always sufficient to pay for at least some of all other goods Y after paying for X .

Utility in every period is separable in health status and the utility of consumption. A corollary of this is that health care shocks do not have any effect on the marginal utility of income, other than through their effect on medical expenditures. Health shocks affect the marginal utility of income through spending on medical care, but do not directly affect the marginal utility of income. This is a common theoretical assumption, used in many empirical studies (e.g., Arrow, 1963; Zeckhauser, 1970; Manning and Marquis, 1996).

4. Single period model with two health goods

Demand curves for each of two goods are assumed to be linear.

⁴Across countries or over time, the income elasticity is greater than one (Newhouse, 1993; Hall and Jones, 2012). Most likely because the available technology grows with income, but for an efficiency analysis of optimal cost sharing, the relevant elasticity is the short run income elasticity, which is very low. Holding social, demographic, environmental and supply side attributes constant, low and high income individuals demand very similar levels of health care even across dramatically different income levels in generous health plans. By assuming zero income effect, we also separate the moral hazard loss from the price effect and the income transfer effect. Nyman (1999) points out the income transfer effect on medical care consumption is eliminated when income transfers are eliminated, and this would only occur if the probability of illness were 1, which is the case we consider as the source of our *ex post* efficiency problem.

$$(1) \quad \begin{aligned} X_1 &= A_1 - B_1 P_1 / P_Y + G_{12} P_2 / P_Y \\ X_2 &= A_2 - B_2 P_2 / P_Y + G_{12} P_1 / P_Y \end{aligned}$$

For simplicity, we normalize the marginal costs of all goods to be one, and express all prices in terms of the share of this marginal cost paid by the consumer. Hence, $P_Y = 1$ and $P_i = c_i$, and the cost share is the consumer price of the i^{th} health care good X_i . Note the presence of the (symmetric) cross price coefficient, G_{12} , which will play a key role later. As in Ellis and Manning (2007), we also incorporate the idea that illness may involve other uninsured (uncompensated) losses. These losses can be of two types: uncompensated out-of-pocket costs (which we assume are proportional to covered costs), and L_i , the uncompensated health shock losses, $L_{\theta_i}(\theta_i)$.⁵

The marginal benefit function for each medical service X_i is assumed to be a linear demand curve with a constant slope B_i on the price regardless of the realization of the random health shock θ_i . Stochastic health treatment demand is introduced by letting $A_i = \mu_i + \theta_i$, where $\theta \sim F(\theta)$, with $E(\theta) = 0$.

Assumptions about the variance/covariance vector of θ are made below but throughout we assume that the distribution does not depend on the out-of-pocket price or income. This corresponds to the horizontal intercept of the demand curves having a mean of μ_i when the full consumer price $(c_i + L_i)$ is zero. Using this convention, a single consumer's demand curve for each medical service X_i can be written as

$$(2) \quad X_i = \mu_i - B_i(c_i + L_i) + G_{ij}(c_j + L_j) + \theta_i$$

⁵ Arrow (1963) was the first to raise the issue of uncompensated losses in thinking about health insurance. Doherty and Schlesinger (1983) show results similar to our findings below, while Gravelle and Rees (2004, Chapter 19) reach similar conclusions. We include these losses because earlier results are based largely on findings that assume risks are independent of the health risk. Section 4.3 discusses the sensitivity of the usual second best results to this issue.

In order to introduce risk aversion, we apply a monotonically increasing concave function $V(\dots)$ with the property of constant absolute risk aversion to the indirect utility function consistent with the demand function in Equation (2). Using this notation, we write the one period indirect utility function with two treatment goods as

$$(3) \quad V(I, C, \theta) = V \left[J + \frac{B_1(c_1 + L_1)^2}{2} + \frac{B_2(c_2 + L_2)^2}{2} - (\mu_1 + \theta_1)(c_1 + L_1) \right. \\ \left. - (\mu_2 + \theta_2)(c_2 + L_2) - G_{12}(c_1 + L_1)(c_2 + L_2) \right] - L_{\theta_1}(\theta_1) - L_{\theta_2}(\theta_2)$$

where $J = I - \pi$,

That this utility function corresponds to the above demand functions (including their symmetric cross price terms G_{12}) can be verified by applying Roy's identity. Using the linear demand equation for X , the insurer's break-even condition for the insurance premium is

$$(4) \quad \pi = (1 - c_1)[\mu_1 - B_1(c_1 + L_1) + G_{12}(c_2 + L_2)] + (1 - c_2)[\mu_2 - B_2(c_2 + L_2) + G_{12}(c_1 + L_1)].$$

In a two-period model, we assume that the same premium is charged in both periods.

While we have used the somewhat restrictive assumptions of linear demand, additive errors, and zero income effects, this specification has two attractive features. The error terms only interact with cost sharing (c_i) in a simple multiplicative form. This facilitates introducing multiple goods and multiple periods. The linear specification also allows us to consider cross price elasticities in a natural way.

We now turn to the social planner's problem of choosing the optimal coinsurance rates when there are two health care treatment goods (X_1 and X_2) and a composite all-other-goods commodity, Y . We develop the model using a general specification and then derive various cases of interest as special cases. The optimal cost sharing rates c_i for health care treatment will maximize the expectation of Equation (3). Taking its partial derivative with respect to c_i and setting equal to zero yields an equation that characterizes the social optimum. Since this expression will not in general have a simple closed form solution, we take a Taylor series approximation of the partial derivative V_J , evaluated around the nonstochastic arguments of the utility function. This solution can be written as

$$(5) \quad 0 = \frac{\partial E_\theta V}{\partial c_1} = E_\theta \left\{ \left[V_I [J - K] - V_{II} [J - K] [\theta_1 (c_1 + L_1) + \theta_2 (c_2 + L_2)] \right] \right. \\ \left. \left(-\frac{\partial \pi}{\partial c_1} - \mu_1 + B_1 (c_1 + L_1) - G_{12} (c_2 + L_2) - \theta_1 \right) \right\}$$

where

$$J = I - \pi,$$

$$K = (c_1 + L_1)\mu_1 + (c_2 + L_2)\mu_2 - B_1(c_1 + L_1)^2 / 2 - B_2(c_2 + L_2)^2 / 2 \\ + G_{12}(c_1 + L_1)(c_2 + L_2),$$

$$\pi = (1 - c_1)[\mu_1 - B_1(c_1 + L_1) + G_{12}(c_2 + L_2)] + (1 - c_2)[\mu_2 - B_2(c_2 + L_2) + G_{12}(c_1 + L_1)],$$

$$\frac{\partial \pi}{\partial c_1} = [-\mu_1 - B_1 + 2B_1c_1 + B_1L_1 + G_{12} - 2c_2G_{12} - G_{12}L_2]$$

Define the absolute risk aversion parameter $R^A \equiv -V_{II}/V_I$, let

$\sigma_1^2 \equiv E(\theta_1)^2$, $\sigma_2^2 \equiv E(\theta_2)^2$, $\sigma_{12} \equiv E(\theta_1\theta_2)$, and use $E(\theta_1) = E(\theta_2) = 0$. We show in the appendix that this result can be rearranged to obtain first order conditions characterizing the optimal coinsurance rates c_1 and c_2 as follows.

$$(6) \quad [-\mu_1 - B_1 + 2B_1c_1 + B_1L_1 + G_{12} - 2c_2G_{12} - G_{12}L_2] + \mu_1 - B_1(c_1 + L_1) + G_{12}c_2 \\ + R^A [\sigma_1^2 (c_1 + L_1) + \sigma_{12} (c_2 + L_2)] = 0$$

$$(7) \quad [-\mu_2 - B_2 + 2B_2c_2 + B_2L_2 + G_{12} - 2c_1G_{12} - G_{12}L_1] + \mu_2 - B_2(c_2 + L_2) + G_{12}c_1 \\ + R^A [\sigma_2^2 (c_2 + L_2) + \sigma_{12} (c_1 + L_1)] = 0$$

We consider a variety of ways of interpreting these two equations below. In broad terms, the first line of each equation gives the marginal cost of increasing the cost-sharing rate, while the second line, which involves R^A gives the marginal benefit in terms of reduced cost of risk.

4.1 One health care good, base case ($L_1 = L_2 = \mu_2 = \sigma_2^2 = \sigma_{12} = B_2 = G_{12} = 0$)

The Base Case of a single health care good in a single period with no uncompensated costs corresponds to $L_1 = L_2 = \mu_2 = \sigma_2^2 = \sigma_{12} = B_2 = G_{12} = 0$. This yields the well-known result from the literature characterizing the second-best optimal insurance contract for constant coinsurance rate plans when there is a simple tradeoff between moral hazard and the cost of risk:

$$(8) \quad -B_1(1-c_1) + R^A \sigma_1^2 c_1^* = 0$$

Where $B_1 \geq 0$ and the first term (the marginal costs due to moral hazard) is increasing in the price response $|B_1|$. The gains from risk pooling are increasing in the variance of health care demand. Solving for the optimal coinsurance rate yields Equation (9), where the optimal base case coinsurance rate, c_1 , is increasing in the price response, B_1 , and decreasing in the underlying variance in demand, σ_1^2 .

$$(9) \quad c_1^* = \frac{B_1}{B_1 + R^A \sigma_1^2} \equiv c_1^{BaseCase}$$

Intuitively, there is a tradeoff in changing c to decrease risk protection as the patient pays more of the price, and to reduce moral hazard. The optimal coinsurance rate is the ratio of the loss from moral hazard to the net loss from changing risk protection plus moral hazard. If demand is perfectly inelastic ($B_1 = 0$), and there is any risk at all ($\sigma_1^2 > 0$) then the optimal coinsurance rate is zero. If there is no variance or the consumer is risk neutral ($R^A = 0$), then the optimal coinsurance rate should be 1. For all interesting cases the optimal coinsurance rate lies between 0 and 1: $0 < c_1^{BaseCase} < 1$.

4.2 Adding uncompensated health-related losses in the one good case ($L_1 > 0$)

Incorporating uncompensated health losses so that $L_1 > 0$ also affects optimal cost sharing for covered treatment goods, increasing the desired coverage (decreasing the cost sharing rate c_1) as considered in the literature on incomplete insurance markets from Doherty and Schlesinger (1983), Gravelle and Rees (2004), and Ellis and Manning (2007). The expression for the optimal cost sharing rate for health good X_I becomes

$$(10) \quad c_1^* = c_1^{BaseCase} - \frac{R^A \sigma_1^2 L_1}{B_1 + R^A \sigma_1^2}.$$

The optimal coinsurance rate c_1 is decreasing in the size of the uncompensated loss L_1 because all of the terms after the minus sign are positive. With uncompensated losses, it is also possible for c_1^* to be negative or to reach a corner solution where $c_1^* = 0$ for either large L_1 or

small B_1 . Equation (10) provides an efficiency-based rationale for why full insurance can be second-best optimal: the absence of complete insurance markets to fully transfer income into high-cost health states of the world means that coinsurance rates are set at or closer to zero than they would have been if the alternative insurance markets were complete and consumers were able to insure against all health care losses. There are many health services and conditions which have substantial uncompensated health care related losses. This is particularly true in developing countries where disability and unemployment insurance are rare and productivity losses from ill health are often large.⁶ Thus, incomplete insurance markets provide a rationale for more generous insurance coverage of health care treatment, even when welfare losses due to moral hazard and insurance loading may be large.

4.3 Multiple health care treatment goods ($\mu_i > 0, \sigma_i^2 > 0, B_i > 0, L_1 = L_2 = 0$)

An important motivation for modeling two rather than one health care treatment goods is to be able to examine the role of cross price elasticities and correlated health shocks. If $L_1 = L_2 = 0$, then (6) and (7) above can be solved for the optimal cost share for health good 1 and simplified to:

$$(11) \quad c_1^* = c_1^{Base Case} - \frac{G_{12}}{B_1 + R^A \sigma_1^2} (1 - c_2^*) - \frac{R^A \sigma_{12}}{B_1 + R^A \sigma_1^2} c_2^*$$

with a similar expression for the second health care good. Because all of the terms multiplying σ_{12} and G_{12} are non-negative, the negative sign means that the first order effect of increasing either term is to reduce the optimal cost share c_1^* relative to the base case. In this case,

we have $\frac{\partial c_1^*}{\partial G_{12}} = \frac{c_2^* - 1}{B_1 + R^A \sigma_1^2} < 0$, reducing c_2^* reinforces the fact that $\frac{\partial c_1^*}{\partial G_{12}} < 0$. Hence, when two

health services become stronger gross substitutes in the sense that $G_{12} = \partial X_1 / \partial c_2$ is increased, then both services should have lower cost sharing rates. This finding replicates the finding of Goldman and Philipson (2007) that as goods become stronger complements ($G_{12} < 0$) they

⁶ Wagstaff (2007) provides documentation of the large magnitudes of income losses from illness in Vietnam.

should have higher cost sharing. We now refine this conclusion in light of allowing for second order changes that affect demand for these two health goods.

The implications of the covariance (or correlation) between shocks affecting health goods on optimal insurance can be complex. As a result of the 2nd order effect of σ_{12} on c_2^* , we cannot unambiguously sign $\frac{\partial c_1^*}{\partial \sigma_{12}}$ for all possible values of G_{12} and σ_{12} . We can, however evaluate this

derivative for various special cases. In the Appendix we show that $\frac{\partial c_1^*}{\partial \sigma_{12}}$ and $\frac{\partial c_1^*}{\partial G_{12}}$ can be

unambiguously signed to be negative when either of the following holds true: 1) both health care goods have the same demand curve slope and variance ($B_1 = B_2$, $\sigma_1 = \sigma_2$) or 2) as the limiting case where σ_{12} and G_{12} both approach zero. The ambiguous cases, as discussed below, occur when the two parameters differ in sign. We use a graphical approach to contrast optimal insurance as G_{12} and σ_{12} change by examining how they change relative to the Base Case when G_{12} and σ_{12} are both equal to zero. In each of Figures 1 through 3, the equilibrium expression (11) and its counterpart for c_2^* are shown as solid lines, and are used to examine the net effect of changing σ_{12} (Figure 1) and changing G_{12} (Figures 2), and changing both (Figure 3). The Base Case illustrates the commonly cited result for optimal coinsurance that $0 < c_i^* < 1$ when health care shocks for uncorrelated goods are neither substitutes nor complements (i.e., $G_{12}=0$). Figure 1 illustrates how the solution lines pivot as σ_{12} increases, which forces both optimal coinsurance rates to decrease, as the covariance term σ_{12} increases from zero to a positive level. If goods 1 and 2 are gross substitutes ($G_{12} > 0$), and there is no covariance ($\sigma_{12} = 0$), then as in Figure 2, there is an unambiguous decrease in the optimal coinsurance rates for both goods. Figure 3 illustrates the corresponding case where the two goods are substitutes and positively correlated, that is, $G_{12} > 0$, and $\sigma_{12} > 0$. The substitution effect and a positive correlation reinforce each other to further cut the optimal cost sharing. (The counterpart figures for complements are shown in Figure A1 (a)-(b).) If both G_{12} and σ_{12} are of the same sign it is possible that the partial derivatives of either parameter on c_i^* can be ambiguous because of second order effects, which

means the second order derivatives can be positive or negative. As shown in Figure A2 (a)-(c), both equilibrium lines are rotating and shifting because of the logic illustrated in the Figures 1 (non-zero covariance) and 2 (substitutes). It is possible that the second order effect of the two lines moving will leave one c_i^* higher while the other is made lower.

Our model is parameterized in terms of variance and covariance terms. However, since in the preceding exposition we were holding the variance terms constant, this implies that an increase in the covariance is directly interpretable as an increase in the contemporaneous correlation. Restating our main result in terms of correlations: for demand structures where goods are either substitutes or neither substitutes nor complements, increasing correlation always warrants more generous insurance coverage. If two goods are sufficiently strong complements, ($G_{12} < 0$), then increasing correlation does not necessarily warrant more generous insurance coverage and it is possible that optimal insurance rates can increase, so that $c_i^* > c_i^{base}$. Conversely, for sufficiently negatively correlated shocks, optimal coverage can also decrease when goods are strong substitutes.

It is straightforward to show that other elements of the conventional results for insurance are also validated in this model, even if the overall level depends on complementarity in demand or covariance information. It is shown in the appendix that

$$\frac{\partial c_1^*}{\partial \sigma_1^2} < 0, \frac{\partial c_1^*}{\partial \sigma_2^2} < 0, \frac{\partial c_1^*}{\partial B_1} > 0, \text{ and } \frac{\partial c_1^*}{\partial B_2} < 0.$$

We can now add that the increase in the covariance, or more specifically the correlation ρ_{12} will lead to the reduction of at least one optimal coinsurance rate, possibly both. It is that ambiguity about which is the affected health care good, that leads us to put the results in parentheses in Table 1

5. Multiple period model

Our framework can also address the case of multiple periods with correlated health care demands. In the Appendix, we derive optimal savings rules for models with $T > 2$ periods, while here we focus on the case where there are only two periods (indexed by 1 and 2). In this section we focus on the case where there is only one good in each period, before turning to two goods in two periods in the next section. We assume that the cost share is the same in both periods, and

hence in this section use $c_1 = c_2 = c$. We also assume there are no uncompensated losses, so that $L_1 = L_2 = 0$, and we focus on the case where demand is the same in each period except for the health shocks θ_i , $X_i = A - Bc + \theta_i$. Additionally, we focus on the case where the parameters and price structure are constant over time: $I_1 = I_2 = I$, $\pi_1 = \pi_2 = \pi$, and $\sigma_1^2 = \sigma_2^2 = \sigma^2$. To allow for the possibility that the health care demands in the two periods are correlated, we assume that the second period health shock is $\theta_2 = \rho\theta_1 + \varepsilon_2$, where $-1 \leq \rho \leq 1$. This assumption significantly differentiates our paper from Gollier (2003), which examines the demand for insurance in a lifecycle model with no serial correlation in the insurable risk.

In a dynamic model, we need to introduce savings, which we assume to be optimally chosen. In our two-period model, net saving is decided in period 1 after θ_1 is known and all income and savings are spent in period 2 (i.e., there are no bequests). Optimal savings will depend on all of the parameters of the model, but of special interest is that savings will depend on the cost sharing rate c and the first period health shock θ_1 , $S_1^*(\theta_1, c)$ with $\partial S_1^* / \partial \theta_1 < 0$. In the Appendix, we show that the objective function to be maximized through the choice of c_i can be written as follows:

$$EV = E_{\theta_1, \theta_2} \left\{ \left[V^1(J - K - S_1^*(\theta_1, c) - c\theta_1) \right] + \varphi E_{\theta_2|\theta_1} \left\{ \left[V^2(J - K + (1+r)S_1^*(\theta_1, c) - c\theta_2) \right] \right\} \right\}$$

where

$$J = I - \pi$$

$$(12) \quad \theta_2 = \rho\theta_1 + \varepsilon_2$$

$$K = c\mu - \frac{Bc^2}{2}$$

$$\pi = (1-c)(\mu - Bc)$$

Except for the savings function and the introduction of discounting, φ , this formulation is very similar in structure to what was used above for the case with one period with multiple states of the world. The solution for the optimal choice of c is derived in the appendix. We make the following three further assumptions in deriving our solution:

- Savings is optimal so that for all θ_l , $V_l^1(\dots) = (1+r)\varphi E_{\theta_2|\theta_1} \{V_l^2(\dots)\}$ where r is the interest rate.
- The utility function can be approximated using a second order approximation with constant absolute risk aversion (CARA).
- Consumers can earn a return on savings $(1+r)$ that is the inverse of their discount rate φ , so that $\varphi(1+r) = 1$.

In using a second order approximation to the utility function with CARA, we rule out third order terms that might warrant a precautionary saving motive. In our framework, an optimal insurance program should never include catastrophic levels of health care spending, so the precautionary, self-insurance motive should be less important.

For our CARA utility function specification, the optimal savings is approximated by

$S_1^*(\theta_1, c) = \bar{S}_1 - s_1 c \theta_1$ (shown in appendix), where the expected (*ex ante*) savings are

$\bar{S}_1 = \frac{\varphi(1+r)-1}{R^A [1+\varphi(1+r)^2]}$ and the optimal (ex post) savings are reduced by the proportion

$s_1 = \frac{1-\varphi(1+r)\rho}{1+\varphi(1+r)^2}$ multiplied by the out-of-pocket health payments in time period 1. The term s_1 is

the marginal propensity to save. In particular, if $\varphi(1+r) = 1$ (as assumed), then $S_1^*(\theta_1, c)$ is

reduced to a simple functional form $S_1^* = -\frac{1-\rho}{2+r} c \theta_1$ ⁷. Under these assumptions, when $n = 2$ the

optimal cost share is

$$(13) c^* = \frac{B}{B + R^A \sigma^2 \left[1 - s_1 \frac{1-\rho}{1+\varphi} \right]} = \frac{B}{B + R^A \sigma^2 \left[1 - \frac{(1-\rho)^2}{(2+r)(1+\varphi)} \right]}.$$

This result is very similar to the Base Case equation (9) for the case of a single period and one health care good without any uncompensated losses ($L_i = 0$). Equation (13) differs from the Base Case by the addition of a new savings-related term in the denominator,

⁷ In the more general case of any $n > 2$ periods, where the autocorrelation terms are allowed to have an arbitrary pattern rather than a first order autocorrelation, we show in the appendix (A44) that the optimal saving function still has a closed form.

$$\Sigma \equiv \left[1 - \frac{(1-\rho)^2}{(2+r)(1+\phi)} \right].$$

Σ is a function of the correlation coefficient between period 1 and period 2 health shocks, ρ , the interest rate, r , and the consumer discount rate, ϕ . Note that Σ is non-decreasing in ρ , r , and ϕ , and that c^* is decreasing in Σ .

As before, we interpret the optimal cost sharing result for a variety of special cases. First, consider the case where the marginal propensity to save is zero, $s_1 = 0$, so that savings does not respond to health shocks. In this case, $\Sigma = 1$ and the one period model results remain correct even with two periods. The consumer must absorb all health shocks fully in the first period, so there is no difference between the static and dynamic choices of c^* . However when $s_1 > 0$, optimal copayments are higher unless shocks are perfectly correlated. Saving is a sort of self-insurance that smooths risk over the both periods and affects optimal insurance by reducing the need for generous insurance when shocks are not perfectly correlated in the presence of second best insurance markets.

Second, consider the case where the period 1 and period 2 shocks are perfectly correlated, so that $\rho = 1$. Once again $\Sigma = 1$ and the one period model results hold. Although savings is possible, there are none if the consumer knows exactly what the shock will be in period 2. There is no diversification across periods in the burden of health shocks. In this limiting case, insurance should be the same as with no savings.

Third, consider the case where health shocks are uncorrelated over time, so that $\rho = 0$. The discount rate ϕ is a number close to one, and it is convenient to consider the case where $\phi = 1$ and $r = 0$ so that there is no interest or discounting. The plausible result in this two period model is that s_1 will be close to one half, and half of the burden of a health shock is born in period 1 and half is deferred to period 2. Since the cost of risk goes up with the square of the deviation from certainty, the savings reduces riskiness in the first period to one-quarter of the one period value and, hence, the cost of first period risk (proportional to the variance) is reduced to one quarter of the one period model results. Since this burden is shared between two periods, the net reduction in risk is by one half of the variance. The reason that Σ only declines to 0.75 is that in a two period model there is no opportunity to reduce the burden of shocks in the second

period. So while savings can reduce the burden of first period shocks to one quarter of their uncertainty cost, savings cannot reduce the burden of second period shocks.

This result with no discounting or interest rates closely approximates the result with discounting, since the two terms $(2+r)(1+\varphi)$ will be approximately 4 if consumers use the market interest rate for discounting.

6. Two-periods with two-goods

So far, we have addressed the cases of two treatment goods and two periods separately. It is also important to consider the case of multiple treatment goods with more than one period so that we can address issues such as the differential coverage of acute and chronic health care. We focus on the case where there are two goods and two periods.

The consumption of two goods in two periods is indexed by $\{(X_1^1, X_2^1), (X_1^2, X_2^2)\}$ with subscripts indicating goods, and superscripts time periods. Following the notation in the one-good multiple period case, we assume $c_{X_i^1} = c_{X_i^2} = c_i$, $i = 1, 2$, $I_1 = I_2 = I$ and $\pi_1 = \pi_2 = \pi$. In the appendix we allow for $L_1 = L_2 = L > 0$, but here in the main text we focus on the case of no other uncompensated costs, so that $L_1 = L_2 = 0$.

To make the results more concrete, we consider the case where X_1 is a medical treatment good for an acute condition and X_2 as a medical treatment good for a chronic condition. Then, the acute good shock in the first period is not correlated with its own future amount, while the shocks affecting the chronic condition good two in both periods are positively correlated, where $\theta_2^2 = \rho\theta_2^1 + \varepsilon_2$, $0 < \rho \leq 1$. We also assume that $E(\theta_1^1) = E(\theta_2^1) = E(\theta_1^2) = E(\theta_2^2) = 0$ and

$$VAR \begin{pmatrix} \theta_1^1 \\ \theta_2^1 \\ \theta_1^2 \\ \theta_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 \\ \sigma_{12} & \sigma_2^2 & 0 & \rho\sigma_2^2 \\ 0 & 0 & \sigma_1^2 & \sigma_{12} \\ 0 & \rho\sigma_2^2 & \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

The objective function and first order conditions characterizing the optimum are shown in the appendix. The main new result of interest relates to how consumers optimally change savings in response to acute rather than chronic conditions. If discounting exactly equals $1+r$, so that $\varphi(1+r) = 1$. Then, we show in the appendix that the optimal saving function is approximated by

$$(14) \quad S_1^*(\theta_1^l, \theta_2^l, c_1, c_2) = -\frac{1}{2+r}c_1\theta_1^l - \frac{1-\rho}{2+r}c_2\theta_2^l,$$

From this expression one can see that the marginal savings out of the unexpected health spending from the two types of shocks differ, with $s_1 = \frac{1}{2+r}$ and $s_2 = \frac{1-\rho}{2+r}$. The expected (*ex ante*) savings \bar{S}_1 are the same with the two-period one-good case. The marginal propensity to save s_2 is lower than s_1 because good 2 is chronic care and patients who receive a large shock for chronic care anticipate a correlated shock the following period and, hence, do not adjust their savings as much as for an acute care shock.

Results for the general case are presented in the appendix. For the special case where demand shocks have zero covariance in each period and the goods are neither complements nor substitutes, so that $\sigma_{12} = 0$, $G_{12} = 0$, and $L = 0$, the two-good, two-period case reduces to the one-good two-period case. The optimal cost sharing is approximated by

$$(15) \quad c_1^* = \frac{B_1}{B_1 + R^A \sigma_1^2 \left[1 - \frac{1}{(2+r)(1+\phi)} \right]}$$

$$(16) \quad c_2^* = \frac{B_2}{B_2 + R^A \sigma_2^2 \left[1 - \frac{(1-\rho)}{(2+r)(1+\phi)} \right]}$$

The optimal cost sharing rates are identical to equation (13) with $\rho = 0$ for acute care. Thus, the correlation of spending for chronic care in two periods does not affect the optimal cost sharing of the acute case. Thus, our findings indicate that chronic care should have better coverage than acute, all other things equal, given that chronic care is positively serially correlated, while acute care is uncorrelated over time.

In the appendix we also examine the case where the two goods are substitutes or complements and where the two goods are positively or negatively correlated contemporaneously, with derived expressions (A56) and (A57). While savings and positively serially correlated errors makes optimal cost sharing more sensitive to these two parameters, and consumer discounting makes them less sensitive, the arguments that substitutes and positively contemporaneously

correlated shocks deserve better insurance coverage than implied by the one period model continue to hold with two periods. For the special case where $\sigma_{12} = 0$ and $G_{12} = 0$, expressions (15) and (16) reveal that optimal cost sharing satisfies $\frac{\partial c_i^*}{\partial r} < 0$ and $\frac{\partial c_i^*}{\partial \varphi} < 0$. Although we expect these first order effects to dominate, because the real interest rate r and φ have a complex effect on savings and consumption decisions, we could not unambiguously sign the effect of these variables on optimal insurance in the general case.

7. Theoretical Summary

We have extended the theoretical literature on efficiency-based models of optimal insurance to address issues that arise from correlated sources of uncertainty, whether the source of the correlation is due to correlated demands across different health care goods at a point in time, correlated demands over time, or the correlated losses that arise from uncompensated losses that accompany the losses covered by the insurance plan. By using a quadratic indirect utility function and, hence a linear demand specification with zero income effects on the demand for health care treatment, we have been able to derive closed form expressions characterizing optimal cost sharing for health care treatment when there are two health care goods and two time periods.

Table 1 summarizes the comparative statics findings in this paper for the various parameters considered for health care treatment goods and multiple time periods. In some of the cases that we have considered, we can only sign the effects of a parameter on optimal cost sharing rates for certain parameter values. These cases have the comparative static results in parentheses.

The first four rows of Table 1 reaffirm conventional results found in the previous literature, while the bottom nine rows reflect new or at least newly nuanced results that extend the previous literature. It is well established that optimal cost sharing on health care treatment should be higher as demand becomes more elastic, consumers become less risk averse, or the variance of spending decreases. These first four findings are consistent with the findings from Besley (1988) and others. Our theoretical findings are also consistent with those of Goldman and Philipson (2007) on complements and substitutes – that all other things equal, cost sharing should be higher for complements than substitutes. Our main new finding are: (1) positively

correlated losses across health care goods or over time should almost always lead to more generous coverage (lower cost-sharing) than uncorrelated or negatively correlated losses; and (2) optimal savings behavior shifts optimal insurance coverage depending on the nature of the shocks and correlation of each good over time. Chronic conditions, which by their nature display positive serial correlation, deserve greater coverage than acute conditions that are not serially correlated since chronic conditions impose more risk and savings cannot make up for this additional risk as easily as when shocks are uncorrelated.

8. Empirical Relevance

In this section, we briefly examine the empirical magnitudes of the two innovations of our model: the role of contemporaneous correlations across multiple goods and of serial correlation over time. We use data from the Thomson-Reuters (now Truven Analytics) MarketScan database from the 2000-2004 period on a population of non-elderly (age <65) enrollees in employment-based commercial plans. We have restricted the sample to FFS, HMO, PPO and POS plans which covered outpatient pharmacy services in addition to outpatient physician and inpatient services. We included only those individuals who were continuously enrolled for the full five year period. This yields a sample of 1,335,448 individuals. Besides the large sample size, these data have two major advantages. The first is that the enrollees are followed for several years, allowing us to study correlations by type of health care over time. Second, all of the enrollees had pharmacy coverage, allowing us to contrast pharmacy expenditure patterns with those of both inpatient and outpatient care.

Table 2 summarizes key means, standard deviations and correlations from our five year sample, decomposed into three broad types of services – inpatient (facility, not inpatient physician) services, outpatient services, and pharmacy services.⁸ The first two columns in the top half of Table 2 reaffirm that inpatient spending, while not the largest expected cost, is by far the most risky in a one year framework. Table 2 also shows that the serial correlation coefficients for spending for each of the three services differ meaningfully, with pharmaceutical spending having

⁸Some researchers may expect a higher proportion of spending to be on inpatient care. The MEDSTAT commercial claims do not include Medicaid or Medicare enrollees, who have higher hospitalization rates. MEDSTAT claims have 24% of all covered charges for inpatient care in 2004. Our sampling frame of using only people with five consecutive years of insurance coverage has somewhat lower proportion of 22 percent of spending in inpatient services.

the highest serial correlations and inpatient spending the lowest. The serial correlations also reveal that following a health shock spending returns to normal levels much more slowly than a simple autoregressive, AR(1), pattern would suggest. Chronic conditions obviously explain this pattern. The larger correlations for pharmacy versus outpatient care and for outpatient versus inpatient care suggest a larger correction for pharmacy than outpatient from the usual results for a one period model. The final column summarizes the implications of using multiple years of spending to calculate financial risk by presenting the standard deviation of five year sums of spending rather than one year spending. Whereas inpatient spending has nearly four times as much variation as pharmacy using a one year horizon, it is less than twice as variable if a five year horizon is used.

Our theoretical model shows that contemporaneous correlations between multiple health care goods can also be important. Table 3 presents two correlation matrices for spending on inpatient, outpatient, pharmacy and total spending. The top half is for one year (2004) while the bottom half corresponds to correlations of five year sums of each of the three components and total spending. The top correlation matrix shows that one-year inpatient spending is the most closely correlated with total spending and that correlations among other services are relatively modest – the conventional view. Taking into account the autocorrelation effect, we generate the bottom half of the table. It reveals that the pattern for inpatient spending is weaker using five year total spending. Outpatient spending becomes the category most closely correlated with five year total spending, and pharmacy spending has a significantly higher correlation with total spending when compared to the one-period case. Considering the contemporaneous correlation and the autocorrelation, our multi-period model suggests better coverage for outpatient and pharmacy spending than suggested by conventional rules for insurance coverage.

Needless to say, the MarketScan data do not have information on the range of uncompensated losses. Thus, we are unable to comment on the magnitude of the correction for correlated uncompensated losses. Moreover, in the absence of estimates of the underlying demand elasticities for these three services, or, even more challenging, estimates of the degree of complementarity among them, it is difficult to determine how different the coinsurance rates under our rules would be from those based on Besley's formulation or older approaches. But the direction is clear. With their higher positive correlation over time, pharmaceuticals should have a lower coinsurance rate than would occur under the traditional rules for one-period models.

Spending on inpatient services, which is less serially correlated than outpatient and pharmaceuticals, would change the least.

9. Discussion

The new results that we find most interesting are those that (1) focus on the roles of uncompensated losses differentially over health care goods and time periods, and (2) those that address the role of correlations across goods and time. As in our earlier work (Ellis and Manning (2007), as well as Gravelle and Rees (2004) and Doherty and Schlesinger (1983)), uncompensated health losses that are related to insured services should influence the level of cost sharing for correlated health care goods. These losses provide a rationale for both reducing out-of-pocket costs for those goods which tend to have larger uncompensated losses, such as time lost due to hospitalization and recovery, time lost by going for a physician visit, or copayments. The intuition is clear. If consumers face uncertain income losses which are correlated with health care spending shocks on certain treatment goods, then overinsuring those treatment goods is a second best response to reduce this combined risk from the compensated and uncompensated elements. In the tradeoff of moral hazard against risk protection, the key term in uncertainty in the demand for health goods in the single period, two-good model is $\left[\theta_1 (c_1 + L_1) + \theta_2 (c_2 + L_2) \right]$ where the θ 's are uncertain *ex ante*. The risk premium depends on the variance of this whole expression, which in turn depends on the size of the uncompensated losses (the L 's) compared to the out-of-pocket copayments (the c 's). Uncompensated losses increase both the benefits from risk reduction (the variance term from above) and the costs of insurance (the demand / moral hazard term.).

Our finding that optimal treatment cost sharing should be lower for positively correlated treatment goods and goods with positive cross-price effects augments the findings of Besley (1988) who focuses on multiple goods with uncorrelated demand shocks. It makes strong intuitive sense that positively correlated variables cause greater variance which needs to be partially offset by lower cost sharing.⁹ The empirical significance of these results is difficult to

⁹Besley (1988, 329-330) indicates that cross-elasticities and covariances jointly affect which good is more generously covered. But covariances are the product of rho and the two standard deviations. Thus his result is actually applicable to correlation ρ_{12} , because he only uses whether covariance is zero or not, which is equivalent to $\rho_{12} = 0$.

assess, since relatively little research has focused on estimating covariance and substitution parameters. They may nonetheless provide guidance on coverage for goods such as brand name drugs, specialty curative goods, or the coverage of serious chronic illnesses, which may be strong substitutes or complements. This framework also provides a rationale for more generously covering services that are often jointly provided, such as specialty services and the lab tests or drugs that specialists prescribe, since such goods will be positively contemporaneously correlated.

Finally, our multi-period model shows the key role that savings decisions and correlated errors play in setting optimal cost sharing. We are not aware of any papers in the health economics literature that have emphasized this topic, although there is a sizeable literature on how large uncovered health losses can lead to dissaving and bankruptcy. While there is a literature on how consumers respond to health spending shocks, the implications for optimal health insurance design deserves reexamination. Expensive, chronic conditions, which exhibit strong positive serial correlations over time, provide an economic rationale for more generous insurance coverage because consumers cannot use intertemporal savings to reduce the burdens of such spending. Thus, in a world where chronic and acute care look to be otherwise equivalent in terms of price responses and variability, the stronger correlation in health care spending for the chronically ill would lead to better coverage than the standard one period model would suggest. Given that so much of pharmaceutical use exhibits the same property, this finding also supports better pharmaceutical coverage.

It is worth highlighting the limitations of our study. As is common in this literature, we develop all of our models using a linear demand structure, with additive errors that have constant variance, and our demand for health care treatment is perfectly income inelastic.¹⁰ By making these simplifications, we assume away income effects and corner solutions, which are particularly relevant in equity discussions of optimal health insurance. In our model, increasing insurance coverage does not subsidize low income people but rather subsidizes those with poor health. Obviously if poor health and low income are correlated, then transfers are implied. We acknowledge that we use strong assumptions, but believe our results make intuitive sense and are

¹⁰ The empirical literature finds that demand is income inelastic overall, especially in the absence of adverse selection on insurance coverage (Manning et al, 1987; Newhouse et al, 1993). But demand for specific health treatment services may be more highly income elastic and yield different results.

more general than many other models that use only consumer surplus, or assume only two health states or one health care good.

We have repeatedly used a second order approximation of the utility function, which is consistent with approximating the utility function with a constant absolute risk aversion function. We are not especially troubled by this assumption because our results should hold as an approximation for any arbitrary function, as long as the absolute risk aversion parameter does not vary too much across states of the world. Our uncompensated loss function and optimal savings function were also approximated using linear functions. Again, our results should hold as an approximation for more general nonlinear functions.

The other restrictive assumption that we have made for tractability's sake is that the variance in health care expenditure is a constant, conditional on the health state. Specifically we have assumed that the variance and the other higher order moments in healthcare treatment do not depend on the level of cost sharing, that is $\partial \sigma_\theta^2 / \partial c_i = 0$, or other observable factors in the demand function. An extension of the current work would allow for the common observation that the variability in health care expenditures is an increasing function of the mean or expected value of expenditures given the covariates in the model.¹¹

Our models highlight the empirical significance of demand and cost parameters. Of all of our parameters, the cost sharing demand responsiveness of various health care treatment goods has been studied the most. The RAND Health Insurance Experiment and other studies have established that spending on inpatient care is less responsive to cost sharing than spending on outpatient care, which is less responsive than spending on pharmaceuticals. The variances and means of spending on different types of treatment goods are also well understood. Inpatient spending is much more variable than outpatient spending and drug spending, justifying greater

¹¹ When the variance becomes an increasing function of the mean, we pick up an extra term in the cost-of-risk part of the first order condition that did not exist if demand for health care treatment $X(c_x)$ had constant variance, conditional on health status. As we increase the coinsurance rate, we have the usual increase in the term related to the variance of out-of-pocket expenses. With increasing coinsurance also decreasing the mean, it also decreases the variance. Thus, we have a partially offsetting term to include in the cost of risk. The magnitude of this reduction depends on how price responsive demand is. As long as the demand for health care treatment is inelastic with respect to cost sharing, the qualitative pattern described earlier in this section prevails. See Feldman and Manning (1997) for such an extension to the basic model that we considered briefly in Ellis and Manning (2007) that allowed for a constant coefficient of variation for health care expenditures instead of a constant variance assumption.

coverage for inpatient care than other health services. These conclusions follow from the previous literature as well as our framework.

Less well studied are the implications of cross price effects, contemporaneous correlations, and serial correlations over time of specific treatment goods. Drug spending and outpatient spending have higher contemporaneous correlations with other types of spending than inpatient care, suggesting they may deserve greater insurance coverage than would otherwise be the case. Cross elasticities of demand for drug and outpatient care are higher than for inpatient care, justifying greater coverage. Some evidence on this is provided in Meyerhoefer and Zuvekas (2006) who demonstrate that cross price elasticities between non-mental health drugs and spending on treatment for physical health are moderately large and statistically significant.

Relatively little work has explored the intertemporal correlations of specific treatment services. We do know that spending on some drugs and outpatient care is much more highly correlated over time than spending on inpatient care. Ellis, Jiang and Kuo (2013) provide recent evidence on this issue in examining more than 30 different medical services defined by type of service, provider specialty and place of service. Our evidence from the Thomson-Reuters MarketScan data also suggests that these correlations are sizeable and important for both outpatient and for pharmacy, but stronger for pharmacy than outpatient.

Perhaps the area most in need of empirical work is to document the magnitude of uncompensated health losses that are correlated with health care spending. Significant uncompensated costs provide a rationale for zero or even negative cost sharing on treatment goods. It would be interesting to know how much the conventional model results must be adjusted given their magnitudes.

REFERENCES

- Atim, C., 1999. Social movements and health insurance: a critical evaluation of voluntary, non-profit insurance schemes with case studies from Ghana and Cameroon. *Social Science & Medicine*; 48; 881-896.
- Andrulis, D. P., 1998. Access to care is the centerpiece in the elimination of socioeconomic disparities in health. *Annals of Internal Medicine* 129; 412-416.
- Arrow, K. J., 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* 53; 941-973.
- Baicker, K., Finkelstein, A., 2011. The Effects of Medicaid Coverage — Learning from the Oregon Experiment. *New England Journal of Medicine* 365; 683-685.
- Blomqvist, A., 1997. Optimal non-linear health insurance. *Journal of Health Economics* 16; 303-321.
- Besley, T., 1988. Optimal reimbursement health insurance and the theory of Ramsey taxation. *Journal of Health Economics* 7; 321-336.
- Buchanan, J. L., Keeler, E.B., Rolph, J.E., Holmer M.R., 1991. Simulating health expenditures under alternative insurance plans. *Management Science* 37; 1067-1090.
- Choudhry, N.K., Rosenthal, M.B., Milstein, A., 2010. Assessing The Evidence For Value-Based Insurance Design. *Health Affairs* 29;1988-1994.
- Chernew, M.E., Newhouse, J.P. 2012. Health Care Spending Growth. In: Pauly MV, McGuire TG, Barros PP. (Eds), *Handbook of Health Economics*, vol. 2. North-Holland: Amsterdam,. p. 1-1126.
- Coate, S., 1995. Altruism, the Samaritan's dilemma, and government transfer policy. *The American Economic Review* 85; 46-57.
- Cutler, D.M., Zeckhauser, R.J. 2000. The anatomy of health insurance. In: Culyer A, Newhouse JP (Eds), *Handbook of Health Economics*, vol. 1. North Holland: Amsterdam; p. 563-643.
- Dardanoni, V., Wagstaff, A., 1990. Uncertainty and the demand for medical care. *Journal of Health Economics* 9; 23-38.
- Doherty, N.A. and Schlesinger. 1983. Optimal insurance in incomplete markets. *Journal of Political Economy*, 91(6): 1045-1054.
- Doherty, N.A., Thistle, P.D., 1996. Adverse selection with endogenous information in insurance markets. *Journal of Public Economics* 63; 83-102.
- Duggan, M., 2004. Does contracting out increase the efficiency of government programs? Evidence from Medicaid HMOs. *Quarterly Journal of Economics*, 2549-2572

- Ellis, R.P., Jiang, S, Kuo, T., 2013. Does service-level spending show evidence of selection across health plan types? *Applied Economics* 45; 1801-1712.
- Ellis R.P., Manning W.G. 2007. Optimal health insurance for prevention and treatment. *Journal of Health Economics* 26; 1128-1150.
- Feldstein, M., 1973. The welfare loss of excess health insurance. *Journal of Political Economy*, University of Chicago Press, vol. 81(2), pages 251-80, Part I, M.
- Feldstein, M, Friedman B., 1977. Tax subsidies, the rational demand for health insurance, and the health care crisis. *Journal of Public Economics* 7; 155-178.
- Feldman, R, Dowd, B., 1991. A new estimate of the welfare loss of excess health insurance. *The American Economic Review* 81; 297-301
- Feldman, R., Manning, W.G., 1997. Une Formule Simple du Taux Optimal pour Une Police D-Assurance Maladie (English Translation: A simple formula for the optimal coinsurance rate in a health insurance policy). In: Jacobzone S(Ed), *Economie de la Sante: Trajectoires du Futur*, InseeMethodes, vol. 64-65. French National Ministry of Health. p. 133-144.
- Fendrick A.M., Smith D.G., Chernew M.E., Shah S.N., 2001. A benefit-based copay for prescription drugs: patient contribution based on total benefits, not drug acquisition cost. *American Journal of Managed Care* 7; 861
- Goddeeris, J.H., 1984, Medical insurance, technological change, and welfare, *Economic Enquiry* 22: 56-67.
- Goldman, D., Philipson, T.J. 2007. Integrated insurance design in the presence of multiple medical technologies. *The American Economic Review* 97; 427-32.
- Gollier, C., 2003. To insure or not to insure?: An insurance puzzle. *The Geneva Papers on Risks and Insurance Theory*, 28: 5-24.
- Gravelle, H. and Rees, Ray. 2004. *Microeconomics*, Third Edition. Prentice Hall: Harlow, England.
- Hall, R.E., Jones, C.I., 2007. The value of life and the rise in health spending. *Quarterly Journal of Economics* 122; 39–72.
- Hofmann, A., 2007. Internalizing externalities of loss prevention through insurance monopoly: an analysis of interdependent risks. *The Geneva Risk and Insurance Review* 32; 91-111.
- Lichtenberg, F., 2002. The effects of Medicare on health care utilization and outcomes. *Frontiers in Health Policy Research*, Vol. 5, ed. By Alan Garber (MIT Press)
- Kunreuther, H.C., Pauly, M.V., McMorro, S. 2012. *Insurance and behavioral economics*. Cambridge: Cambridge University Press .
- Maarse, H, Paulus, A., 2003. Has solidarity survived? A comparative analysis of the effects of social health insurance reform in four European countries. *Journal of Health Politics, Policy and Law* 28; 585-614.

- Manning, W.G., Newhouse, J.P., Duan, N. et al., 1987. Health insurance and the demand for medical care: Evidence from a randomized experiment. *American Economic Review* 77; 251-277.
- Manning, W.G., Marquis, S.M., 1996. Health insurance: the trade-off between risk pooling and moral hazard. *Journal of Health Economics* 15; 609-639.
- Marquis, S.M., Manning, W.G., 1999. Lifetime costs and compensation for injuries. *Inquiry* 36; 244-254.
- Meyerhoefer, C., Zuvekas, S., 2006, Coverage for mental health treatment: Do the gaps still persist? *Journal of Mental Health Policy and Economics*, Sep. 9(3): 155-163
- Newhouse, J.P., Manning, W.G., Morris, C.N., et al., 1981, Some interim results from a controlled trial of cost sharing in health insurance. *New England Journal of Medicine*; 305; 1501-1507.
- Newhouse J.P. and the Insurance Experiment Group, 1993. *Free-for-all: Health insurance, medical Costs, and health outcomes: the results of the health insurance experiment.* Cambridge: Harvard University Press.
- Newhouse, J.P., 2006. Reconsidering the moral hazard-risk avoidance tradeoff. *Journal of Health Economics* 25; 1005–1014.
- Nyman, J.A., 1999. The economics of moral hazard revisited. *Journal of Health Economics* 18; 811-824.
- Pauly, M.V., 1968. The economics of moral hazard: Comment. *American Economic Review* 58; 531-37.
- Rask, K.N., Rask, K.J., 2000. Public insurance substituting for private insurance: new evidence regarding public hospitals, uncompensated care funds, and Medicaid. *Journal of Health Economics* 19; 1-31.
- Spence, M, Zeckhauser, R., 1971. Insurance, information, and individual action. *American Economic Review* 61; 380-387.
- Sin, D.D., Svenson, L.W., Cowie, R.L., Man, P.S., 2003. Can universal access to health care eliminate health inequities between children of poor and nonpoor families? A case study of childhood asthma in Alberta. *Chest* 124; 51-56.
- Wagstaff, A., 2007. Health insurance for the poor: initial impacts of Vietnam's health care fund for the poor. *Policy Research Working Paper Series 4134, The World Bank*
- Zeckhauser, R., 1970. Medical insurance: a case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory* 2; 10–26.
- Zweifel, P.J., Manning, W.G. 2000. Consumer incentives in health care. In: Culyer A, Newhouse JP (Eds), *Handbook of Health Economics*, vol. 1. North Holland: Amsterdam p.409-59.

Table 1. Comparative Statics on Optimal Coinsurance Rate c_i^*

Effect of		on c_i^*
Own demand slope	$ B_i $	+
Other good demand slope	$ B_j $	-
Risk aversion parameter	R^A	-
Variance of spending on own i	σ_i^2	-
Uncompensated losses affecting income	L_i	-
Uncompensated losses affecting utility directly	$L(\theta)$	0
Variance of spending on other good j	σ_j^2	-
Covariance of spending on i and j	σ_{ij}	(-)
Correlation between spending on goods i and j	ρ_{ij}	(-)
Cross price term for other good	G_{ij}	(-)
Real interest rate	r	(-)
Consumer discount factor	φ	(-)
Correlation with next period error	ρ	(-)

Note: results in parentheses only hold for specific values of key parameters.

Table 2

Means, standard deviations, five year autocorrelation rates of inpatient, outpatient and pharmacy spending

	<u>2004</u>	2004	<u>Autocorrelation with spending in year</u>					<u>Standard</u>
	<u>Mean</u>	<u>Standard</u>	<u>2004</u>	<u>2003</u>	<u>2002</u>	<u>2001</u>	<u>2000</u>	<u>deviation of</u>
		<u>deviation</u>						<u>5 year sum</u>
Using unadjusted spending								
2004 inpatient spending	\$ 871	\$ 8,116	1.000	0.111	0.082	0.062	0.055	\$ 16,288
2004 outpatient spending	\$ 2,150	\$ 7,018	1.000	0.460	0.332	0.271	0.240	\$ 17,702
2004 pharmacy spending	\$ 977	\$ 2,321	1.000	0.825	0.622	0.585	0.528	\$ 7,700
2004 total spending	\$ 3,999	\$12,856	1.000	0.418	0.304	0.248	0.223	\$ 31,875

Note: Results are based on healthcare services for MEDSTAT Marketscan data for individuals aged < 65, who were continuously enrolled during 2000-2004. N= 1,335,448.

Table 3
Correlations across spending on healthcare services

2004 unadjusted spending

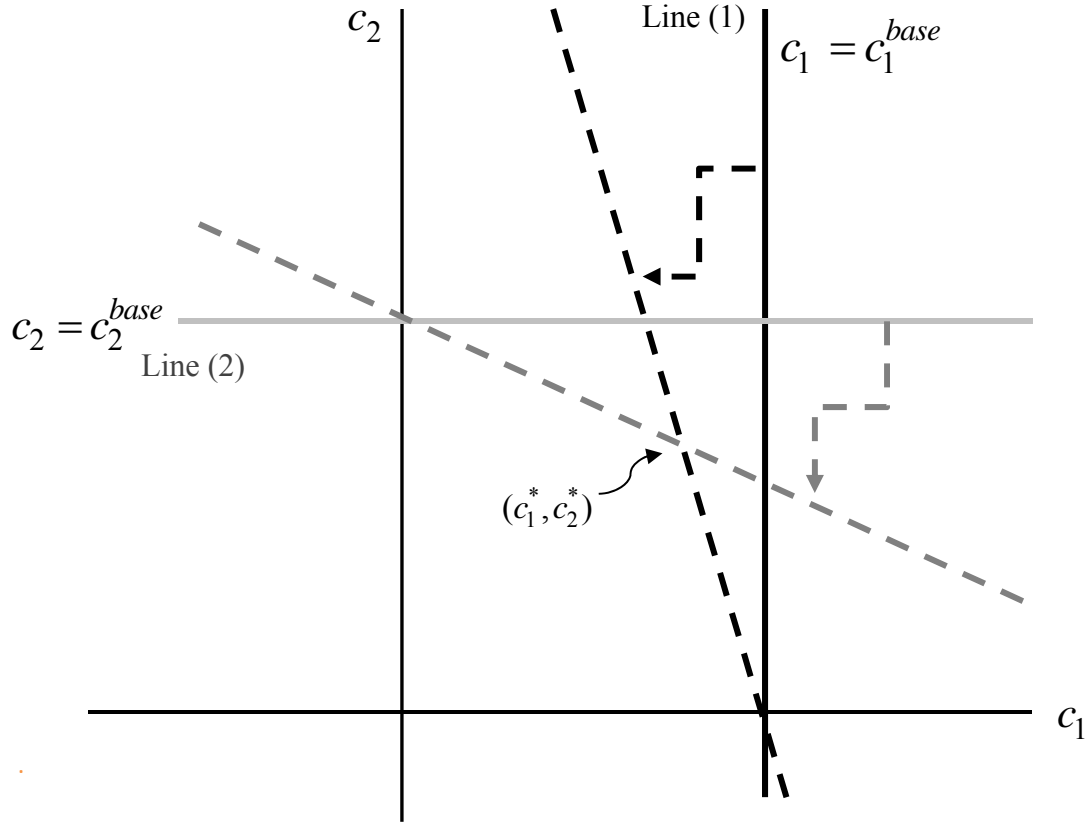
	<u>Inpatient</u>	<u>Outpatient</u>	<u>Pharmacy</u>	<u>Total</u>
Inpatient	1.000	0.278	0.129	0.807
Outpatient		1.000	0.251	0.767
Pharmacy			1.000	0.399
Total	(symmetric)			1.000

2000-2004 Five year total spending

	<u>Inpatient</u>	<u>Outpatient</u>	<u>Pharmacy</u>	<u>Total</u>
Inpatient	1.000	0.395	0.206	0.793
Outpatient		1.000	0.346	0.836
Pharmacy			1.000	0.529
Total	(symmetric)			1.000

Note: Results are based on healthcare services for MEDSTAT MarketScan data for individuals aged < 65 in 2004, who were continuously enrolled during 2000-2004. N= 1,482,232.

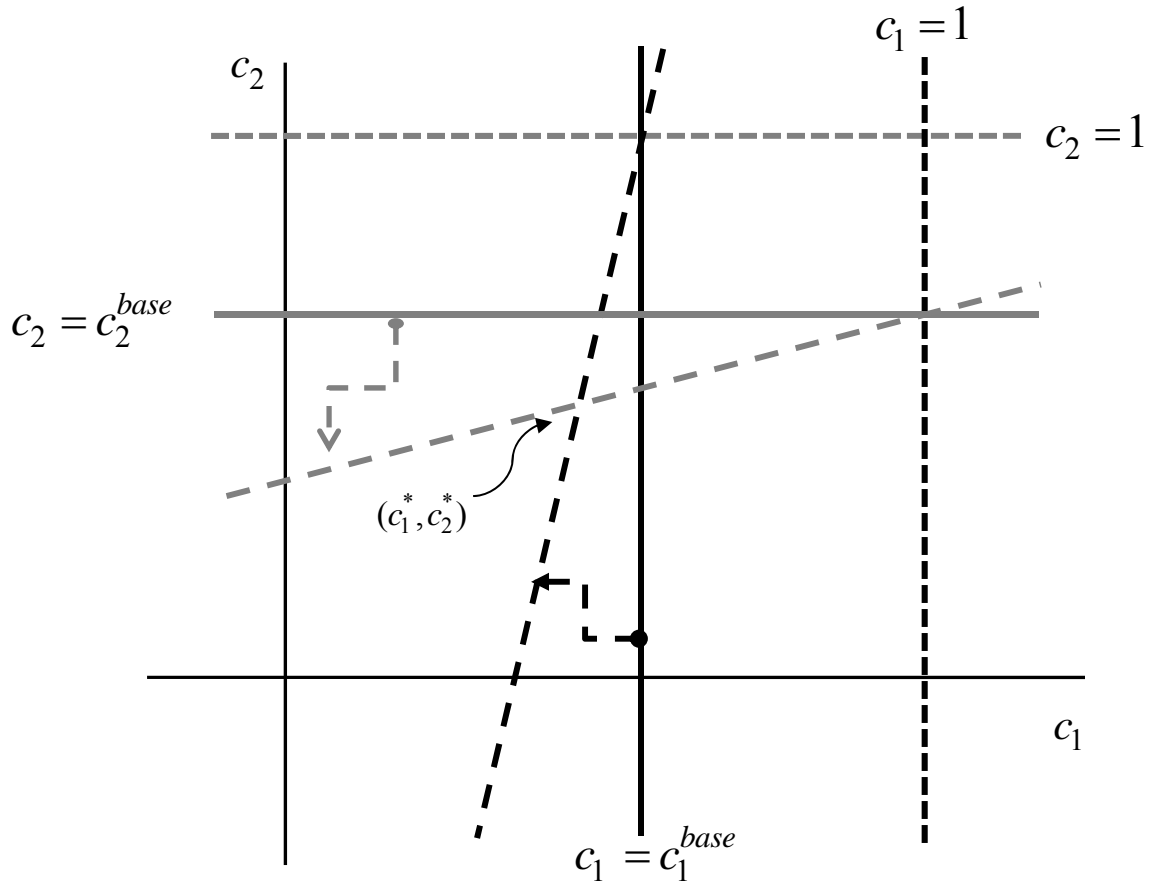
Figure 1
Optimal coinsurance rates c_i^* for $G_{12} = 0$ as $\sigma_{12} \uparrow$
Two Health Care Goods, One Period: Increasing correlation



Line (1) and Line (2) correspond to the solutions of Equation 11, when there is no covariance between the two goods and they are neither substitutes nor complements. With $G_{12} = \sigma_{12} = 0$, each optimal coinsurance rate equals the Base Case value and the optimal coinsurance rates do not depend on each other.

Holding the variances terms σ_1^2 and σ_2^2 constant while increasing the covariance σ_{12} implies increasing the correlation, ρ_{12} . Both Line (1) and Line (2) rotate inward away from the Base Case. Hence, $c_i^* < c_i^{base}$ as $\rho_{12} \uparrow$. Case shown is for $G_{12} = 0$, but the pivots happen in the same direction for arbitrary $G_{12} < R^A \sigma_{12}$ as $\sigma_{12} \uparrow$

Figure 2
 Optimal coinsurance rates c_i^* when $(G_{12} \uparrow \text{from } 0)$ & $\sigma_{12} = 0$
 Two Health Care Goods, One Period: Substitutes



As goods become stronger substitutes, intercepts decrease while slopes increase, such that intersection points are below base case. Thus, substitutes unambiguously lead to lower c_i^* than the Base Case. Similar results can be shown for general σ_{12} , since σ_{12} only affects the slopes but not the intercepts.

The dash-dot-dash lines have $(G_{12} > 0 \text{ \& } \sigma_{12} > 0)$ while the solid line is the base case $(G_{12} = \sigma_{12} = 0)$ and the dashed line corresponds to $(G_{12} > 0 \text{ \& } \sigma_{12} = 0)$. In this case, the dash-dot-dash intersection is southwest of the dashed intersection, which means the two effects reinforce each other.

Appendix

This appendix derives selected analytical results in the main paper. For convenience, Table A-1 presents our notation. Equation numbers shown without an A suffix correspond to numbering in the main text.

Table A-1 Notation

$X = \{X_i\}$ = quantity of health care treatment of service i

Y = quantity of all other consumption goods

I = consumers total income

J = disposable income after premiums

π = premium paid by consumer

P_i, P_Y = demand prices of X_i and Y

c_i = cost sharing rate on treatment X_i (share paid by consumer)

$\theta = \{\theta_i\}$ = random shock affecting health and demand for X_i

μ_i = mean of health care spending on X_i

B_i = slope of demand curve when written in the form $X_i = \mu_i - B_i c_i + \theta_i$

G_{ij} = cross price effect of c_j on X_i and also c_i on X_j

φ = discount rate used by consumer

r = interest rate received on savings by consumer

S_1 = savings in period 1

σ_i^2 = variance of θ_i and also variance of health care spending X_i

σ_{ij} = covariance of θ_i and θ_j

V = the consumer's utility function

R^A = absolute risk aversion constant $= -V_{\pi\pi}/V_\pi$

L_i = per unit uncompensated costs of treatment X_i

$L_i(\theta_i)$ = uncompensated health losses that reduce effective income from random shocks θ_i

A-2. Optimal cost sharing rates for health care treatment

Assume there is no preventive good, and that treatment goods X_1 and X_2 have linear demand curves of the form

$$(A1) \quad \begin{aligned} X_1 &= A_1 - B_1 P_1 / P_Y + G_{12} P_2 / P_Y \\ X_2 &= A_2 - B_2 P_2 / P_Y + G_{12} P_1 / P_Y \end{aligned}$$

These demands are consistent with a risk neutral indirect utility function of

$$(A2) \quad \tilde{V}^S = \frac{I}{P_Y} + \frac{B_1 P_1^2}{2 P_Y^2} + \frac{B_2 P_2^2}{2 P_Y^2} - A_1 \frac{P_1}{P_Y} - A_2 \frac{P_2}{P_Y} - G_{12} \frac{P_1 P_2}{P_Y^2}$$

Using the normalizations $P_Y = 1$, $P_1 = c_1 + L_1$, $P_2 = c_2 + L_2$, $A_i = \mu_i + \theta_i$, letting $J = I - \pi$, applying the concave transformation $V^S(\dots)$,¹² this yields an indirect utility function that can be written as

$$(A3) \quad V^S = V^S \left[J + \frac{B_1 (c_1 + L_1)^2}{2} + \frac{B_2 (c_2 + L_2)^2}{2} - (\mu_1 + \theta_1)(c_1 + L_1) \right. \\ \left. - (\mu_2 + \theta_2)(c_2 + L_2) - G_{12} (c_1 + L_1)(c_2 + L_2) \right] - L_1(\theta_1) - L_2(\theta_2)$$

Taking partial derivatives with respect to c_1 yields

$$0 = \frac{\partial V^S}{\partial c_1} = E_\theta \left\{ V_I^S \left[J - K - \theta_1 (c_1 + L_1) - \theta_2 (c_2 + L_2) \right] \right. \\ \left. \times \left(-\frac{\partial \pi}{\partial c_1} - \mu_1 + B_1 (c_1 + L_1) - G_{12} (c_2 + L_2) - \theta_1 \right) \right\}$$

where

$$(A4) \quad J = I - \pi,$$

$$K = (c_1 + L_1)\mu_1 + (c_2 + L_2)\mu_2 - \frac{1}{2} B_1 (c_1 + L_1)^2 - \frac{1}{2} B_2 (c_2 + L_2)^2 + G_{12} (c_1 + L_1)(c_2 + L_2),$$

$$\pi = (1 - c_1)[\mu_1 - B_1 (c_1 + L_1) + G_{12} (c_2 + L_2)] + (1 - c_2)[\mu_2 - B_2 (c_2 + L_2) + G_{12} (c_1 + L_1)],$$

$$\frac{\partial \pi}{\partial c_1} = -\mu_1 - B_1 + 2B_1 c_1 + B_1 L_1 + G_{12} - 2c_2 G_{12} - G_{12} L_2$$

¹² This model is convex quadratic in health care goods and we apply a risk adverse (concave) function to spending on all other goods. Hence objective function is strictly quasiconcave, and the constraints are all linear or strictly concave. A unique maximum is assured.

Taking a first order approximation of V_I^s around J-K, we can write

$$\begin{aligned}
\frac{\partial V^s}{\partial c_1} &\approx E_\theta \left\{ \left[V_I^s [J-K] - V_{II}^s [J-K] [\theta_1 (c_1 + L_1) + \theta_2 (c_2 + L_2)] \right] \right. \\
&\quad \left. \left(-\frac{\partial \pi}{\partial c_1} - \mu_{x_1} + B_1(c_1 + L_1) - G_{12}(c_2 + L_2) - \theta_1 \right) \right\} \\
(A5) \quad &= E_\theta \left\{ V_I^s [J-K] \left[-\frac{\partial \pi}{\partial c_1} - \mu_1 + B_1(c_1 + L_1) - G_{12}(c_2 + L_2) - \theta_1 \right] \right. \\
&\quad \left. - V_{II}^s [J-K] [\theta_1 (c_1 + L_1) + \theta_2 (c_2 + L_2)] \right. \\
&\quad \left. \times \left[-\frac{\partial \pi}{\partial c_1} - \mu_1 + B_1(c_1 + L_1) - G_{12}(c_2 + L_2) - \theta_1 \right] \right\}
\end{aligned}$$

Defining $R^A \equiv -V_{II}^s / V_I^s$, $\sigma_1^2 \equiv E_{\theta_1}(\theta_1)^2$, $\sigma_2^2 \equiv E_{\theta_2}(\theta_2)^2$, $\sigma_{12} \equiv E_\theta(\theta_1 \theta_2)$, using $E_\theta(\theta_i) = 0$, the first order conditions for the maximum, once divided through by the (nonstochastic) $V_I^s(J-K)$ can be approximated as follows.

$$\begin{aligned}
(A6) \quad 0 &= E_\theta \left\{ \left[-\frac{\partial \pi}{\partial c_1} - \mu_{x_1} + B_1(c_1 + L_1) - G_{12}(c_2 + L_2) - \theta_1 \right] \right. \\
&\quad \left. + R^A [\theta_1 (c_1 + L_1) + \theta_2 (c_2 + L_2)] \left[-\frac{\partial \pi}{\partial c_1} - \mu_1 + B_1(c_1 + L_1) - G_{12}(c_2 + L_2) - \theta_1 \right] \right\} \\
&= \left[-\frac{\partial \pi}{\partial c_1} \right] - [\mu_1 - B_1(c_1 + L_1) + G_{12}(c_2 + L_2)] - R^A [\sigma_1^2(c_1 + L_1) + \sigma_{12}(c_2 + L_2)]
\end{aligned}$$

Finally replacing $\frac{\partial \pi}{\partial c_1}$ and rearranging slightly, we get expressions that are linear in c_1 and

c_2 .

$$\begin{aligned}
(A7) \quad &[-\mu_1 - B_1 + 2B_1c_1 + B_1L_1 + G_{12} - 2c_2G_{12} - G_{12}L_2] \\
&+ \mu_1 - B_1(c_1 + L_1) + G_{12}c_2 + R^A [\sigma_1^2(c_1 + L_1) + \sigma_{12}(c_2 + L_2)] = 0
\end{aligned}$$

$$\begin{aligned}
(A8) \quad &[-\mu_2 - B_2 + 2B_2c_2 + B_2L_2 + G_{12} - 2c_1G_{12} - G_{12}L_1] \\
&+ \mu_2 - B_2(c_2 + L_{x_2}) + G_{12}c_1 + R^A [\sigma_2^2(c_2 + L_2) + \sigma_{12}(c_1 + L_1)] = 0
\end{aligned}$$

In the main text we examine a number of special cases.

A-2.1 One good, no uncompensated losses $\mu_2 = B_2 = G_{12} = L_1 = L_2 = \sigma_2^2 = \sigma_{12} = 0$

$$(A9) \quad [-\mu_1 - B_1 + 2B_1c_1] + \mu_1 - B_1c_1 + R^A\sigma_1^2c_1 = 0$$

$$(A10) \quad c_1^* = \frac{B_1}{B_1 + R^A\sigma_1^2}$$

A-2.2 One good, uncompensated losses. $\mu_2 = B_2 = G_{12} = L_2 = \sigma_2^2 = \sigma_{12} = 0$, but $L_1 > 0$,

FOC (A7) simplifies to

$$(A11) \quad [-\mu_1 - B_1 + 2B_1c_1] + \mu_1 - B_1c_1 + R^A\sigma_1^2(c_1 + L_1) = 0$$

$$(A12) \quad c_1^* = \frac{B_1 - R^A\sigma_1^2L_1}{B_1 + R^A\sigma_1^2}$$

A-2-3 Two goods no uncompensated losses, general demand structure

If $L_1 = L_2 = 0$, then equations (A7) and (A8) can be solved for c_1^* as

$$(A13) \quad c_1^* = \frac{(G_{12} - B_2)(R^A\sigma_{12} - G_{12}) - (G_{12} - B_1)(R^A\sigma_2^2 + B_2)}{(R^A\sigma_1^2 + B_1)(R^A\sigma_2^2 + B_2) - (R^A\sigma_{12} - G_{12})^2}$$

It is straightforward to show from (A13) that

$$(A14) \quad \frac{\partial c_1^*}{\partial \sigma_1^2} < 0, \quad \frac{\partial c_1^*}{\partial \sigma_2^2} < 0, \quad \frac{\partial c_1^*}{\partial B_1} > 0, \quad \frac{\partial c_1^*}{\partial B_2} < 0,$$

The expressions for $\frac{\partial c_1^*}{\partial \sigma_{12}}$ and $\frac{\partial c_1^*}{\partial G_{12}}$ cannot be signed for all possible values of G_{12} and σ_{12} ,

however their signs are unambiguous for certain special cases. We consider two special cases.

A-2-3-1 Symmetric demand slopes and variances

If $B_1 = B_2$ and $\sigma_1^2 = \sigma_2^2$ then by factoring the denominator and simplifying, it is straightforward to show from (A13) that

$$c_1^* = c_2^* = \frac{B_i - G_{12}}{B_i + R^A\sigma_i^2 - G_{12} + R^A\sigma_{12}}$$

For cases where $0 < c_i^* < 1$, it is straightforward to verify that $\frac{\partial c_i^*}{\partial \sigma_{12}} < 0$ and $\frac{\partial c_i^*}{\partial G_{12}} < 0$ for all σ_{12} and G_{12} .

A-2-3-1 Limiting cases where G_{12} and σ_{12} approach zero

These partial derivatives can also be signed for the limiting case where G_{12} and σ_{12} both approach zero. In this limiting case the two partial derivatives become

$$(A15) \quad \left. \frac{\partial c_1^*}{\partial \sigma_{12}} \right|_{\sigma_{12}=0, G_{12}=0} = \frac{-R^A B_2}{(R^A \sigma_1^2 + B_1)(R^A \sigma_2^2 + B_2)} < 0$$

$$(A16) \quad \left. \frac{\partial c_1^*}{\partial G_{12}} \right|_{\sigma_{12}=0, G_{12}=0} = \frac{-R^A \sigma_2^2}{(R^A \sigma_1^2 + B_1)(R^A \sigma_2^2 + B_2)} < 0$$

This means that for sufficiently small G_{12} and σ_{12} , as the covariance of the errors between two services increases (becomes more positive), then the optimal coinsurance rate decreases.

The more general relationship for various combinations of σ_{12} and G_{12} shows that these derivatives cannot be signed unambiguously, as discussed in the main text.

A-3. Two period model

It is well known that there are close parallels between models with multiple states of the world and models with multiple periods. One key difference is that there is the possibility of correlated outcomes in different periods, either because health shocks are serially correlated or because the two periods are linked by savings. We examine here a two-period model with one health care treatment good in each period, allowing both savings and correlated errors. The demand structure and premiums are assumed to be the same in both periods. We use the direct utility function which is the dual to the indirect utility function used thus far. We also focus on the case where $I_1 = I_2 = I$ and $L_1 = L_2 = L$, hence $\pi_1 = \pi_2 = \pi$. We also assume constant variances over time $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and $\varphi(1+r) = 1$.

$$\begin{aligned} \text{Max}_{X_1, Y_1, S_1} \quad EV &= E_{\theta_1, \theta_2} \left\{ V^1(X_1, Y_1, \theta_1) - L_{\theta_1}(\theta_1) \right. \\ &\quad \left. + \varphi E_{\theta_2 | \theta_1} \left\{ \left[V^2(X_2, Y_2, \theta_2) - L_{\theta_2}(\theta_2) \right] \right\} \right\} \end{aligned}$$

where

$$\begin{aligned} (A17) \quad V^1(X_1, Y_1, \theta_1) &= V^s(Y_1 - \frac{(\mu + \theta_1 - X_1)^2}{2B}) \\ V^2(X_2, Y_2, \theta_2) &= V^s(Y_2 - \frac{(\mu + \theta_2 - X_2)^2}{2B}) \\ \theta_2 &= \rho\theta_1 + \varepsilon_2 \\ \varphi &= \text{consumers discount factor} \\ K_1 = K_2 &= (c+L)\mu - \frac{B(c+L)^2}{2} \\ \pi &= (1-c)(\mu - B(c+L)) \end{aligned}$$

As long as income is sufficient to always buy the optimal amount of X_i , then the same amount of X_i will be purchased as in the one period case.

Hence we can use:

$$\begin{aligned} (A18) \quad X_1^* &= \mu + \theta_1 - B(c+L) \\ X_2^* &= \mu + \theta_2 - B(c+L) \end{aligned}$$

Savings will equilibrate the expected marginal utility of income in period 2 with the marginal utility in period 1. Hence we have

$$(A19) \quad V_t^1(\theta_1) = (1+r)\varphi E_{\theta_2|\theta_1} \left\{ \left[V_t^2(\theta_2) \right] \right\}$$

Let the optimal savings function be $S_1^*(\theta_1, c_x)$ (derived below) and write the problem using indirect utility as

$$(A20) \quad EV^* = E_{\theta_1, \theta_2} \left\{ \left[V^1 \left(J - K - S_1^*(\theta_1, c) - (c+L)\theta_1 \right) - L_{\theta_1}(\theta_1) \right] \right. \\ \left. + \varphi E_{\theta_2|\theta_1} \left\{ \left[V^2 \left(J - K + (1+r)S_1^*(\theta_1, c) - (c+L)\theta_2 \right) - L_{\theta_2}(\theta_2) \right] \right\} \right\}$$

where

$$J_1 = J_2 = J = I - \pi$$

$$\theta_2 = \rho\theta_1 + \varepsilon_2$$

φ = consumers discount factor

$$K = (c+L)\mu - \frac{B(c+L)^2}{2}$$

$$\pi = (1-c)(\mu - B(c+L))$$

Except for the savings function, this formulation is very similar in structure to that used for multiple states of the world. Differentiating with regard to c yields

$$(A21) \quad \frac{\partial EV^*}{\partial c} = E_{\theta_1, \theta_2} \left\{ \left[\left[V_t^1 \left(J - K - S_1^*(\theta_1, c) - (c+L)\theta_1 \right) \right. \right. \right. \\ \left. \left. \times \left(-\frac{\partial S_1^*(\theta_1, c)}{\partial c} - \frac{\partial \pi}{\partial c} - \mu + B(c+L) - \theta_1 \right) \right] \right. \\ \left. + \varphi E_{\theta_2|\theta_1} \left[\left[V_t^2 \left(J - K + (1+r)S_1^*(\theta_1, c) - (c+L)\theta_2 \right) \right. \right. \right. \\ \left. \left. \times \left((1+r)\frac{\partial S_1^*(\theta_1, c)}{\partial c} - \frac{\partial \pi}{\partial c} - \mu + B(c+L) - \theta_2 \right) \right] \right] \right\}$$

By the envelope theorem, the terms involving $\frac{\partial S_1^*(\theta_1, c)}{\partial c}$ in the above expression will cancel out due to the assumption of optimal savings. Hence we can rewrite this as:

$$(A22) \quad \frac{\partial EV^*}{\partial c} = E_{\theta_1} \left\{ \left[\left[V_t^1 \left(J - K - S_1^*(\theta_1, c) - (c+L)\theta_1 \right) \times \left(-\frac{\partial \pi}{\partial c} - \mu + B(c+L) - \theta_1 \right) \right] \right. \right. \\ \left. + \varphi E_{\theta_2|\theta_1} \left[\left[V_t^2 \left(J - K + (1+r)S_1^*(\theta_1, c) - (c+L)\theta_2 \right) \right. \right. \right. \\ \left. \left. \times \left(-\frac{\partial \pi}{\partial c} - \mu + B(c+L) - \theta_2 \right) \right] \right] \right\}$$

Except for the fact that two different period utility functions are used, and the appearance of the savings function as an argument of the V_t^1 this is identical to the earlier specification. Taking a

first order Taylor series approximation of the V_I^1 and V_I^2 functions, as before, we obtain the similar expansions

$$(A23) \quad 0 = E_{\theta_1} \left\{ \left[1 - R^A \left(-S_1^*(\theta_1, c) - (c + L)\theta_1 \right) \right] \left(-\frac{\partial \pi}{\partial c} - \mu + B(c + L) - \theta_1 \right) \right. \\ \left. + \varphi E_{\theta_2|\theta_1} \left\{ \left[1 - R^A \left((1 + r)S_1^*(\theta_1, c) - (c + L)\theta_2 \right) \right] \left(-\frac{\partial \pi}{\partial c} - \mu + B(c + L) - \theta_2 \right) \right\} \right\}$$

(A24)

$$\frac{\partial EV^*}{\partial c} \approx E_{\theta_1} \left\{ \left[V_I^1(J - K) + V_I^1(J - K) \left(-S_1^*(\theta_1, c) - (c + L)\theta_1 \right) \right] \left(-\frac{\partial \pi}{\partial c} - \mu + B(c + L) - \theta_1 \right) \right. \\ \left. + \varphi E_{\theta_2|\theta_1} \left\{ \left[V_I^2(J - K) + V_I^2(J - K) \left((1 + r)S_1^*(\theta_1, c) - (c + L)\theta_2 \right) \right] \right. \right. \\ \left. \left. \times \left(-\frac{\partial \pi}{\partial c} - \mu + B(c + L) - \theta_2 \right) \right\} \right\} \\ = (1 + \varphi) \left(-\mu + Bc - \frac{\partial \pi}{\partial c} + BL \right) - R^A(c + L)\sigma^2(1 + \varphi) \\ - R^A \left\{ E_{\theta_1} \left[\begin{aligned} & - \left(S_1^*(\theta_1, c) \right) \left(-\frac{\partial \pi}{\partial c} - \mu + B(c + L) \right) \\ & + \varphi \left((1 + r)S_1^*(\theta_1, c) \right) \left(-\frac{\partial \pi}{\partial c} - \mu + B(c + L) \right) \end{aligned} \right] \right. \\ \left. + E_{\theta_1, \theta_2} \left[\left(S_1^*(\theta_1, c) \right) (\theta_1 - \varphi(1 + r)\theta_2) \right] \right\}$$

If consumers use the same discount rate as that is implied by their real interest rate on savings, then $\varphi(1 + r) = 1$ and the top expression in brackets will be zero, and we can further simplify

$$(A25) \quad 0 = (1 + \varphi) \left(-\mu + Bc - \frac{\partial \pi}{\partial c} + BL \right) - R^A\sigma^2(c + L)(1 + \varphi) \\ - R^A E_{\theta_1, \theta_2} \left[\left(S_1^*(\theta_1, c) \right) (\theta_1 - \theta_2) \right]$$

Using $\frac{\partial \pi}{\partial c} = -\mu + 2Bc + BL - B$.

$$(A26) \quad 0 = (1 + \varphi)(B - Bc) - R^A\sigma^2(c + L)(1 + \varphi) - R^A E_{\theta_1, \theta_2} \left\{ \left(S_1^*(\theta_1, c) \right) (\theta_1 - \theta_2) \right\}$$

This expression cannot be simplified further without explicit savings function form $S_1^*(\theta_1, c_X)$. As is showed below, optimal savings rule can be approximated by the following linear function

$$(A27) \quad S_1^*(\theta_1, c) = \bar{S}_1 - s_1(c + L)\theta_1$$

$$\text{Where } \bar{S}_1 = \frac{\varphi(1+r)-1}{R^A[1+\varphi(1+r)^2]}, \quad s_1 = \frac{\varphi(1+r)\rho-1}{1+\varphi(1+r)^2},$$

This implies that there is an average savings level \bar{S}_1 but that savings is reduced by proportion s_1 for all losses (compensated or uncompensated). Since \bar{S}_1 will be uncorrelated with $(\theta_1 - \theta_2)$, it will drop out once expectations are taken and we can write

$$(A28) \quad 0 = (1+\varphi)B(1-c) - R^A\sigma^2(c+L)[(1+\varphi) - s_1(1-\rho)]$$

Rearranging yields the following condition for optimal cost sharing with multiple periods,

$$(A29) \quad c^* = \frac{B - R^A L \sigma^2 \left[1 - s_1 \frac{1-\rho}{1+\varphi} \right]}{B + R^A \sigma^2 \left[1 - s_1 \frac{1-\rho}{1+\varphi} \right]}$$

If we plug in function form of $s_1 = \frac{1-\rho}{2+r}$, then

$$(A30) \quad c^* = \frac{B - R^A L \sigma^2 \left[1 - \frac{(1-\rho)^2}{(2+r)(1+\varphi)} \right]}{B + R^A \sigma^2 \left[1 - \frac{(1-\rho)^2}{(2+r)(1+\varphi)} \right]}.$$

Now we turn to deriving optimal saving function. As our model set up, saving is determined after health shock is revealed in the 1st period. Assuming individuals are not budget constrained, the optimal amount of X_i will be purchased in both periods and will not be affected by the optimal saving decision, however saving does affect the amount of Y_i . Due to this feature, we solve our optimal saving function after taking optimal choices of X_i as given. This approach brings us the same solution for the optimal saving function as used in the choice of X_1 , Y_1 and S_1 simultaneously in the first period.

$$\begin{aligned}
(A31) \quad \text{Max}_{S_1} \quad EV &= \left[V(J - K - S_1 - (c + L)\theta_1) - L_{\theta_1}(\theta_1) \right] \\
&\quad + \varphi E_{\theta_2} \left[V(J - K + (1 + r)S_1 - (c + L)\theta_2) - L_{\theta_2}(\theta_2) \right] \\
&\quad \text{where} \\
&\quad J = I - \pi \\
&\quad \theta_2 = \rho\theta_1 + \varepsilon_2 \\
&\quad \varphi = \text{consumers discount factor} \\
&\quad K = (c + L)\mu - \frac{B(c + L)^2}{2} \\
&\quad \pi = (1 - c)(\mu - B(c + L))
\end{aligned}$$

The optimal saving function S_1^* satisfies

$$(A32) \quad -V_I(J - K - S_1^* - (c + L)\theta_1) + \varphi(1 + r)E_{\theta_2|\theta_1} \left[V_I(J - K + (1 + r)S_1^* - (c + L)\theta_2) \right] = 0$$

Taking a first order Taylor series approximation of V_I function at $J - K$,

$$\begin{aligned}
(A33) \quad & \\
(A34) \quad & V_I(J - K) - [S_1^* + (c + L)\theta_1]V_{II}(J - K) = \varphi(1 + r)E_{\theta_2|\theta_1} \left\{ V_I(J - K) \right. \\
&\quad \left. + [(1 + r)S_1^* - (c + L)\theta_2]V_{II}(J - K) \right\} \\
& V_I(J - K) - [S_1^* + (c + L)\theta_1]V_{II}(J - K) = \varphi(1 + r)V_I(J - K) + \varphi(1 + r)^2 S_1^* V_{II}(J - K) \\
&\quad - \varphi(c + L)(1 + r)V_{II}(J - K)E_{\theta_2|\theta_1}(\theta_2)
\end{aligned}$$

As is assumed $\theta_2 = \rho\theta_1 + \varepsilon_2$, $E_{\theta_2|\theta_1}(\theta_2) = \rho\theta_1$, we obtain

$$\begin{aligned}
(A35) \quad & V_I(J - K) - [S_1^* + (c + L)\theta_1]V_{II}(J - K) = \varphi(1 + r)V_I(J - K) + \varphi(1 + r)^2 S_1^* V_{II}(J - K) \\
&\quad - \varphi(1 + r)(c + L)V_{II}(J - K)\rho\theta_1
\end{aligned}$$

$$(A36) \quad 1 + [S_1^* + (c + L)\theta_1]R^A = \varphi(1 + r) - \varphi(1 + r)^2 S_1^* R^A + \varphi(c + L)(1 + r)R^A \rho\theta_1$$

Rearranging the above equation, we solve the optimal saving function as following

$$(A37) \quad S_1^* = \frac{\varphi(1 + r) - 1}{R^A [1 + \varphi(1 + r)^2]} - \frac{1 - \varphi(1 + r)\rho}{1 + \varphi(1 + r)^2} (c + L)\theta_1, \text{ which implies that}$$

$$(A38) \quad \bar{S}_1 = \frac{\varphi(1 + r) - 1}{R^A [1 + \varphi(1 + r)^2]}, \quad s_1 = \frac{1 - \varphi(1 + r)\rho}{1 + \varphi(1 + r)^2}$$

Given our assumption $\varphi(1+r) = 1$, S_1^* can be simplified further as

$$(A39) S_1^* = -\frac{1-\rho}{2+r}(c+L)\theta_1 \text{ and } \bar{S}_1 = 0, s_1 = \frac{1-\rho}{2+r}$$

A-4 Optimal saving for model with multiple periods $T \geq 2$

Now we allow more general cases. We assume that $E_{\theta_t}(\theta_{t+i}) = \rho_i \theta_t$, for any t , without making any restrictions on the stochastic process for θ_t .

It can readily be shown that

$$(A40) S_1^* = -\frac{(1+r)(1-\rho_1) + (1-\rho_2)}{1 + (1+r) + (1+r)^2} \theta_1 \text{ for } T=3.$$

$$(A41) S_1^* = -\frac{(1+r)^2(1-\rho_1) + (1+r)(1-\rho_2) + (1-\rho_3)}{1 + (1+r) + (1+r)^2 + (1+r)^3} \theta_1 \text{ for } T=4.$$

$$(A42) S_1^* = -\frac{(1+r)^3(1-\rho_1) + (1+r)^2(1-\rho_2) + (1+r)(1-\rho_3) + (1-\rho_4)}{1 + (1+r) + (1+r)^2 + (1+r)^3 + (1+r)^4} \theta_1 \text{ for } T=5$$

By induction, it is easy to see that the optimal saving result for any $T \geq 2$ is

$$(A43) S_1^* = -\frac{\sum_{t=1}^{T-1} (1+r)^{t-1} (1-\rho_{T-t+1})}{\sum_{t=1}^T (1+r)^{t-1}} \theta_1.$$

We do not try to derive optimal cost sharing rules for multiple periods, other than to note that these more general savings rules imply that risk spreading of health shocks over multiple periods will depend in a natural way on interest and discount rates.

A-5 Optimal saving for model with two periods and two goods

The consumption of two goods in two periods is indexed by subscripts for goods and superscripts for periods $\{(X_1^1, X_2^1), (X_1^2, X_2^2)\}$. Following the notation for the multiple period

case, we assume $c_{X_i^1} = c_{X_i^2} = c_i$, $i=1, 2$; $I_1 = I_2 = I$, $L_{X_1^1} = L_{X_1^2} = L_1$, $L_{X_2^1} = L_{X_2^2} = L_2$ and

$$\pi_1 = \pi_2 = \pi.$$

Consider X_1 as an acute medical treatment good which is not correlated with its own future and

X_2 as a chronic medical treatment good that is positively correlated: $\theta_2^2 = \rho\theta_2^1 + \varepsilon_2$, $0 < \rho \leq 1$.

Assuming that $E(\theta_1^1) = E(\theta_2^1) = E(\theta_1^2) = E(\theta_2^2) = 0$ we have

$$VAR \begin{pmatrix} \theta_1^1 \\ \theta_2^1 \\ \theta_1^2 \\ \theta_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 \\ \sigma_{12} & \sigma_2^2 & 0 & \rho\sigma_2^2 \\ 0 & 0 & \sigma_1^2 & \sigma_{12} \\ 0 & \rho\sigma_2^2 & \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

As is assumed that income is sufficient to support the optimal amounts of medical service, the demands of two treatment goods in period 1 are

$$(A44) \quad \begin{aligned} X_1^1 &= \mu_1 + \theta_1^1 - B_1(c_1 + L_1) + G_{12}(c_2 + L_2) \\ X_2^1 &= \mu_2 + \theta_2^1 - B_2(c_2 + L_2) + G_{12}(c_1 + L_1) \end{aligned}$$

Similarly, the demands in period 2 are

$$(A45) \quad \begin{aligned} X_1^2 &= \mu_1 + \theta_1^2 - B_1(c_1 + L_1) + G_{12}(c_2 + L_2) \\ X_2^2 &= \mu_2 + \theta_2^2 - B_2(c_1 + L_2) + G_{12}(c_1 + L_1) \end{aligned}$$

If we let the optimal savings function be $S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)$, then we can write the objective function using the expected indirect utility as

$$EV^* = E_{\theta_1^1, \theta_2^1} \left\{ \left[V^1 \left(J - K - S_1^* - (c_1 + L_1)\theta_1^1 - (c_2 + L_2)\theta_2^1 \right) - L_{\theta_1^1}(\theta_1^1) - L_{\theta_2^1}(\theta_2^1) \right] \right. \\ \left. + \varphi E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} \left\{ \left[V^2 \left(J - K + (1+r)S_1^* - (c_1 + L_1)\theta_1^2 - (c_2 + L_2)\theta_2^2 \right) \right] \right. \right. \\ \left. \left. - L_{\theta_1^2}(\theta_1^2) - L_{\theta_2^2}(\theta_2^2) \right] \right\} \right\}$$

where

$$(A46) \quad J_1 = J_2 = J = I - \pi$$

φ = consumers discount factor

$$\begin{aligned} K &= (c_1 + L_1)\mu_1 + (c_2 + L_2)\mu_2 - \frac{B(c_1 + L_1)^2}{2} - \frac{B(c_2 + L_2)^2}{2} \\ &\quad + G_{XY}(c_1 + L_1)(c_2 + L_2) \\ \pi &= (1 - c_1)[\mu_1 - B_1(c_1 + L_1) + G_{12}(c_2 + L_2)] \\ &\quad + (1 - c_2)[\mu_2 - B_2(c_2 + L_2) + G_{12}(c_1 + L_1)] \end{aligned}$$

The optimal saving function S_1^* satisfies

$$(A47) \quad -V_I (J - K - S_1^* - (c_1 + L_1)\theta_1^1 - (c_2 + L_2)\theta_2^1) + \varphi(1+r)E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} \left[V_I (J - K + (1+r)S_1^* - (c_1 + L_1)\theta_1^2 - (c_2 + L_2)\theta_2^2) \right] = 0$$

Taking a first order Taylor series approximation of V_I function at $J - K$,

$$(A48) \quad V_I (J - K) - \left[S_1^* + (c_1 + L_1)\theta_1^1 + (c_2 + L_2)\theta_2^1 \right] V_{II} (J - K) = \varphi(1+r)E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} \left\{ V_I (J - K) + \left[(1+r)S_1^* + (c_1 + L_1)\theta_1^2 + (c_2 + L_2)\theta_2^2 \right] V_{II} (J - K) \right\}$$

$$(A49) \quad V_I (J - K) - \left[S_1^* + (c_1 + L_1)\theta_1^1 + (c_2 + L_2)\theta_2^1 \right] V_{II} (J - K) = \varphi(1+r)V_I (J - K) + \varphi(1+r)^2 S_1^* V_{II} (J - K) - \varphi(c_1 + L_1)(1+r)V_{II} (J - K) E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} (\theta_1^2) - \varphi(c_2 + L_2)(1+r)V_{II} (J - K) E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} (\theta_2^2)$$

Having assumed $\theta_2^2 = \rho\theta_2^1 + \varepsilon_2$, $0 < \rho \leq 1$, $E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} (\theta_2^2) = \rho\theta_2^1$, we obtain

$$(A50) \quad V_I (J - K) - \left[S_1^* + (c_1 + L_1)\theta_1^1 + (c_2 + L_2)\theta_2^1 \right] V_{II} (J - K) = \varphi(1+r)V_I (J - K) + \varphi(1+r)^2 S_1^* V_{II} (J - K) - \varphi(c_2 + L_2)(1+r)V_{II} (J - K) \rho\theta_2^1$$

$$(A51) \quad 1 + \left[S_1^* + (c_1 + L_1)\theta_1^1 + (c_2 + L_2)\theta_2^1 \right] R^A = \varphi(1+r) - \varphi(1+r)^2 S_1^* R^A + \varphi(c_2 + L_2)(1+r)R^A \rho\theta_2^1$$

Rearranging the above equation, we solve the optimal saving function as follows:

$S_1^*(\theta_1^1, \theta_2^1, c_1, c_2) = \bar{S} - s_1(c_1 + L_1)\theta_1^1 - s_2(c_2 + L_2)\theta_2^1$, where the savings and overall savings rates

are $\bar{S}_1 = \frac{\varphi(1+r)-1}{R^A [1 + \varphi(1+r)^2]}$, $s_1 = \frac{1}{1 + \varphi(1+r)^2}$ and $s_2 = \frac{1 - \varphi(1+r)\rho}{1 + \varphi(1+r)^2}$. The expected (*ex ante*)

savings \bar{S}_1 are same with the two-period one-good case. The marginal propensity to save s_2 is

lower than s_1 because good 2 is chronic care and patients reduce the uncertainty of expected

spending of good 2 by treating it better in period 1. In particular, if $\varphi(1+r) = 1$, $S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)$

is reduced to a simple functional form $S_1^* = -\frac{1}{2+r}(c_1 + L_1)\theta_1^1 - \frac{1-\rho}{2+r}(c_2 + L_2)\theta_2^1$.

Differentiating with regard to c_i yields

$$(A52) \frac{\partial EV^*}{\partial c_1} = E_{\theta_1^1, \theta_2^1} \left\{ \left[V_I^1 \times \left(-\frac{\partial S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)}{\partial c_1} - \frac{\partial \pi}{\partial c_1} - \mu_1 - \mu_2 \right) \right] \right. \\ \left. + \varphi E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} \left[V_I^2 \times \left((1+r) \frac{\partial S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)}{\partial c_1} - \frac{\partial \pi}{\partial c_1} - \mu_1 - \mu_2 \right) \right] \right\}$$

$$(A53) \frac{\partial EV^*}{\partial c_2} = E_{\theta_1^1, \theta_2^1} \left\{ \left[V_I^1 \times \left(-\frac{\partial S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)}{\partial c_2} - \frac{\partial \pi}{\partial c_2} - \mu_1 - \mu_2 \right) \right] \right. \\ \left. + \varphi E_{\theta_1^2, \theta_2^2 | \theta_1^1, \theta_2^1} \left[V_I^2 \times \left((1+r) \frac{\partial S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)}{\partial c_2} - \frac{\partial \pi}{\partial c_2} - \mu_1 - \mu_2 \right) \right] \right\}$$

By the envelope theorem, the terms involving $\frac{\partial S_1^*(\theta_1^1, \theta_2^1, c_1, c_2)}{\partial c_2}$ in the above expression will cancel out due to the assumption of optimal savings.

The optimal cost sharing rates c_1^* and c_2^* must satisfy

(A54)

$$(1+\varphi) \left[B_1(1-c_1^*) - G_{12}(1-c_2^*) \right] - R^A \sigma_1^2 (c_1^* + L_1) \left[(1+\varphi) - s_1 \right] - R^A \sigma_{12} (c_2^* + L_2) \left[(1+\varphi) - s_2 \right] = 0$$

(A55)

$$(1+\varphi) \left[B_2(1-c_2^*) - G_{12}(1-c_1^*) \right] - R^A \sigma_2^2 (c_2^* + L_2) \left[(1+\varphi) - s_2(1-\rho) \right] - R^A \sigma_{12} (c_1^* + L_1) \left[(1+\varphi) - s_1 \right] = 0$$

If we simplify the problem by assuming that there are no uncompensated cost related to health care (by assuming $L_1 = L_2 = 0$), the optimal conditions become

$$(A56) c_1^* = \frac{(1+\varphi)(B_1 - G_{12})}{(1+\varphi)(B_1 + R^A \sigma_1^2) - R^A \sigma_1^2 s_1} + \frac{(1+\varphi)(G_{12} - R^A \sigma_{12}) + R^A \sigma_{12} s_2}{(1+\varphi)(B_1 + R^A \sigma_1^2) - R^A \sigma_1^2 s_1} c_2^*$$

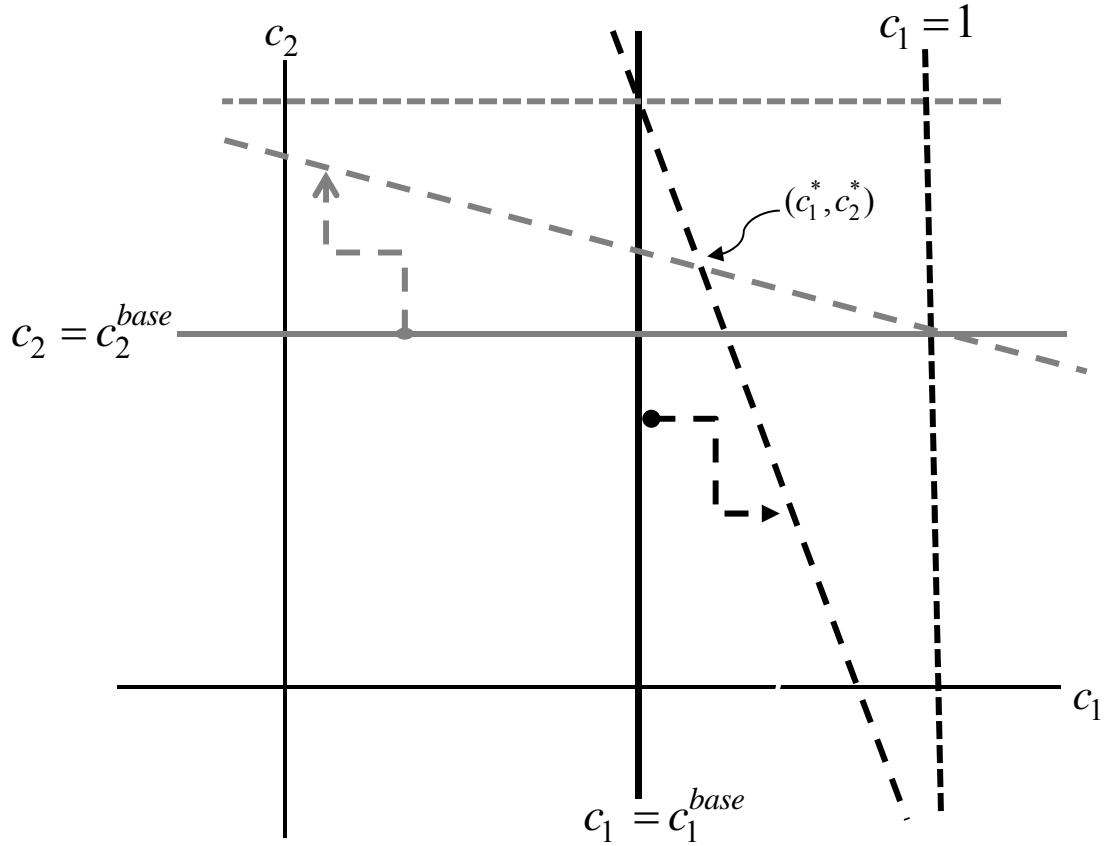
$$(A57) c_2^* = \frac{(1+\varphi)(B_2 - G_{12})}{(1+\varphi)(B_2 + R^A \sigma_2^2) - R^A \sigma_2^2 s_2(1-\rho)} + \frac{(1+\varphi)(G_{12} - R^A \sigma_{12}) + R^A \sigma_{12} s_1}{(1+\varphi)(B_2 + R^A \sigma_2^2) - R^A \sigma_2^2 s_2(1-\rho)} c_1^*$$

Paralleling the analysis in section A.2.3 above for two goods in one period, we can sign the effects of G_{12} and σ_{12} on c_1^* and c_2^* only for special cases, such as when they are both in the neighborhood of zero. A graphical analysis can be used to show results similar to those in Table A2. Savings tends for both acute and chronic conditions raises optimal cost sharing, but has a larger effect on acute conditions which are serially uncorrelated.

Figure A1 (a)

Optimal coinsurance rates c_i^* when $(G_{12} \downarrow \text{from } 0)$ & $\sigma_{12} = 0$

Two Health Care Goods, One Period: Complements

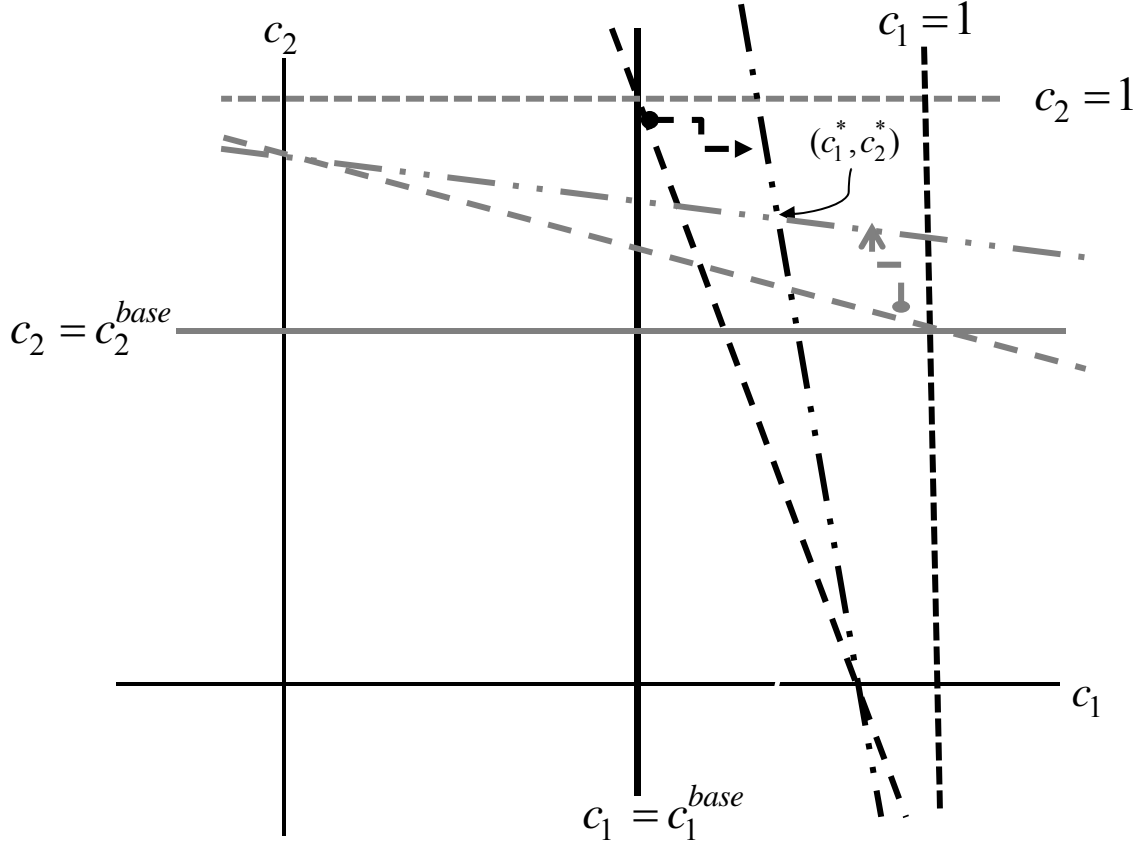


As two goods become stronger complements, intercepts increase while slopes decrease, so that equilibrium lines shift outward such that new intersection point lies between c_i^{base} and 1. Thus, complements unambiguously lead to higher c_i^* than Base Case when there is no correlation.

Figure A1 (b)

Optimal coinsurance rates c_i^* when $G_{12} < 0$ & $\sigma_{12} < 0$

Two Health Care Goods, One Period: Complements

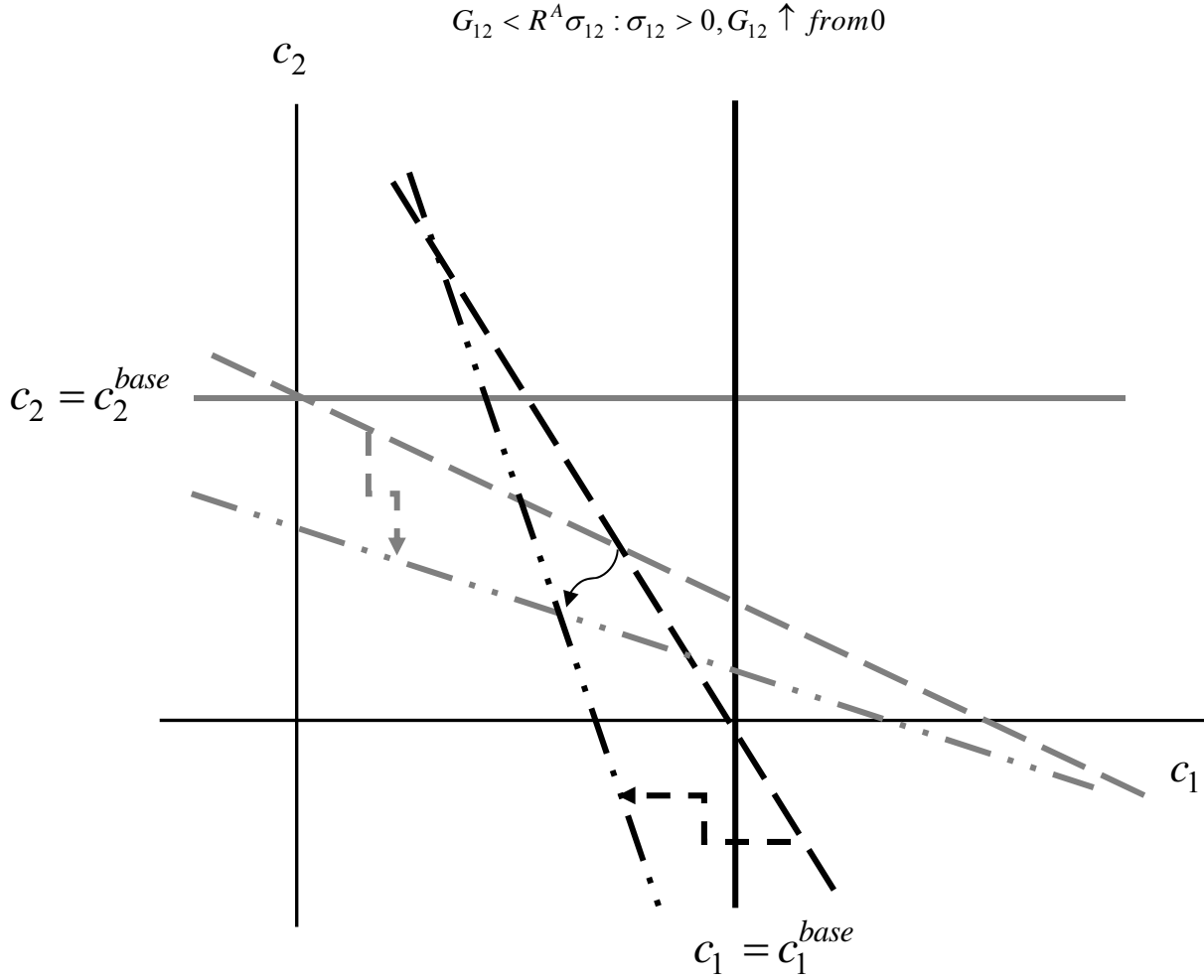


Starting from Figure A1 and decreasing the correlation from zero holding σ_1^2, σ_2^2 constant to preserve the Base-Case comparison. Here we have an unambiguous case where the effect of G_{12} and σ_{12} reinforce each other, pushing both c_i^* away from zero.

The dash-dot-dash lines have ($G_{12} < 0$ & $\sigma_{12} < 0$) while the solid line is the base case ($G_{12} = \sigma_{12} = 0$) and the dashed line corresponds to ($G_{12} < 0$ & $\sigma_{12} = 0$). In this case, the dash-dot-dash intersection is northeast of the dashed intersection, which means the two effects reinforce each other.

Note that the two slopes of the dash-dot-dash lines must be both positive/zero/negative at the same time.

Figure A2 (a)



The sign of $\frac{\partial c_i^*}{\partial G_{12}}$ when $\sigma_{12} > 0$ and $G_{12} \uparrow \text{from } 0$ depends on whether G_{12} remains lower than/exceeds $R^A \sigma_{12}$ (i.e. depending on whether the sign of slope switches when G_{12} goes up from zero) as well as the relative movement of the two lines determining whether resulting c_1^* and c_2^* move in the same direction. The figure A3(a)-A3(C) correspond to the cases when $G_{12} < R^A \sigma_{12}$, $G_{12} = R^A \sigma_{12}$, and $G_{12} > R^A \sigma_{12}$.

Note: these figures do not exhaust all the possibilities of changes in c_1^* and c_2^* .

Figure A2 (b)

$$G_{12} = R^A \sigma_{12} : \sigma_{12} > 0, G_{12} \uparrow \text{ from } 0$$

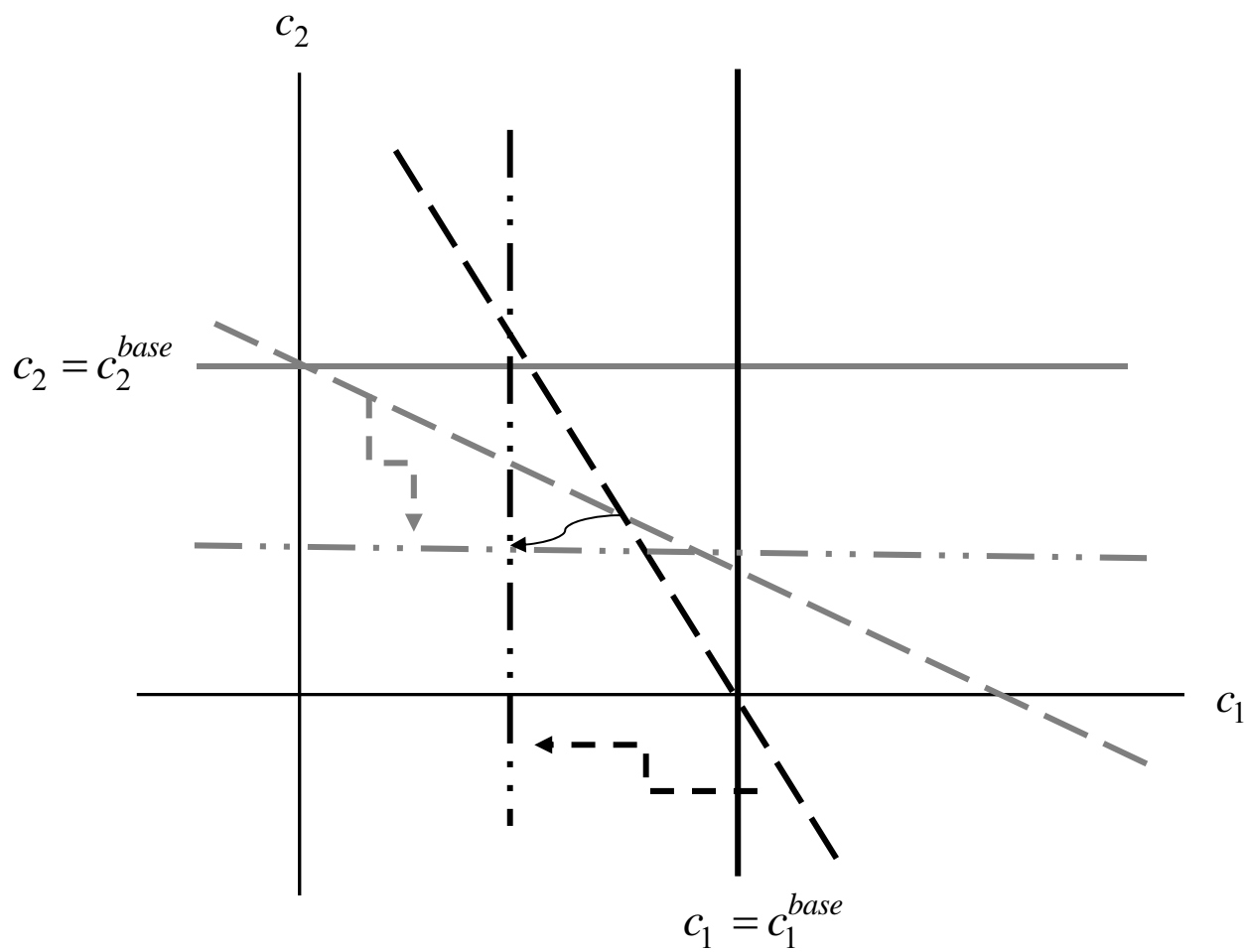


Figure A2 (c)

$$G_{12} > R^A \sigma_{12} : \sigma_{12} > 0, G_{12} \uparrow \text{ from } 0$$

