Multidimensional Screening in Insurance Markets with Adverse Selection

Keith J. Crocker  
Smeal College of Business  
The Pennsylvania State University  
University Park PA 16802

and

Arthur Snow  
Department of Economics  
University of Georgia  
Athens, GA 30602

August 2006

Abstract

Bundled coverage of different losses and distinct perils, along with differential deductibles and policy limits, are common elements of insurance contracts offered in homeowners, automobile, and commercial insurance markets. We show that these practices allow insurers to implement multidimensional screening of insurance applicants who possess hidden knowledge of their risks of loss, and thereby to reduce the externality cost of adverse selection. Competitive forces drive insurers to exploit multidimensional screening and enhance the efficiency of insurance contracting. We also find that efficient multidimensional screening allows competitive insurance markets to attain pure strategy Nash equilibria over a wider range of applicant pools, providing an avenue for insurers to circumvent the nonexistence problem identified by Rothschild and Stiglitz (1976).

JEL classifications: D82, G22, L10  
Keywords: Efficient screening, coverage limitations, signaling, existence of equilibrium
I. Introduction

When insurance applicants possess hidden knowledge of their risks of incurring a loss, as in the seminal model of insurance contracting with adverse selection analyzed by Rothschild and Stiglitz (1976), the market is driven by competition for those in the low risk class. These applicants wish to distinguish themselves from high-risk applicants in order to obtain insurance coverage at terms more favorable than would be made available if they were pooled with the high risks. The desired separation is attained through a screening mechanism whereby insurers offer applicants their choice from a menu of contracts, one of which provides high risks with full and fair insurance, while the other incorporates a lower average price for coverage but also includes a deductible that is unacceptable to high risks.

In this paper we show that separation of this sort is accomplished most effectively by multidimensional screening implemented through the contractual bundling of different perils and different losses. Bundled coverage, with differential deductibles and policy limits, allows insurers to screen applicants in several dimensions thereby reducing the externality cost that low risk applicants must bear to distinguish themselves from high risks. As a consequence, insurers can compete most effectively for the low risk applicants by offering contracts that exploit efficient multidimensional screening.1

The prediction that insurance markets will offer contracts with differential coverage for alternative losses and distinct perils is confirmed by commonly observed practices. For example, a typical homeowner’s policy provides bundled indemnification for both liability and property losses, with different policy limits for these losses. Further, within each category of loss, a homeowner faces different levels of coverage for
distinct perils that cause the same damage, or different losses caused by the same peril. Generally, there is a deductible for the home being destroyed by fire, but if the same home is destroyed by flood the deductible is essentially 100%, as no coverage is provided for this peril. Moreover, losses of specific types of property, such as jewelry, rugs or antiques, are also subject to specific (lower) sublimits.

Similarly, automobile insurance typically bundles different levels of coverage, such as losses resulting from collision, personal injury, uninsured motorists, and medical bills, into the same policy with different sublimits depending on the peril or category of loss. It is also common to observe different deductibles for alternative perils causing the same loss. For example, an automobile lost through theft or fire is typically subject to a lower deductible through the “comprehensive” portion of the policy than would be the case were the loss a result of an accident and therefore covered by the “collision” part of the policy. Coverage limitations are also commonly observed in commercial insurance, where sublimits, coinsurance provisions, deductibles and exclusions often vary according to the peril or type of loss.2

Our analysis not only shows that multidimensional screening is efficient, but also shows that multidimensional screening provides an avenue for insurance markets to circumvent the nonexistence problem identified by Rothschild and Stiglitz. They established that a pure strategy Nash equilibrium fails to exist if the proportion of high risks in the applicant pool is sufficiently low.3 We demonstrate that, by reducing the externality cost of adverse selection, efficient multidimensional screening allows competitive insurance markets to attain pure strategy Nash equilibria over a wider range of applicant pools.4 Indeed, in the limit efficient multidimensional screening eliminates
adverse selection entirely. We conclude that market forces compel insurers to offer contracts with differential coverage for alternative losses and distinct perils that implement multidimensional screening, thereby enhancing both the efficiency of insurance contracting and the prospects for existence of Nash equilibrium.

The paper proceeds as follows. In the next section, we describe the Rothschild-Stiglitz insurance market, the externality cost of adverse selection, and the potential nonexistence problem. Multidimensional screening is introduced in section three through differential levels of insurance coverage, and necessary and sufficient conditions are identified under which multidimensional screening is efficient. In section four we show that efficient multidimensional screening enhances the prospects for existence of a pure strategy Nash equilibrium. In section five we consider the limiting case in which increasing the dimensionality of the screening space leads to perfect discrimination among applicants by their risk types, thus eliminating the adverse selection externality. The final section contains concluding remarks.

II. The One-Dimensional Screening Model of Rothschild and Stiglitz

In the one-dimensional screening model of insurance contracting with adverse selection developed by Rothschild and Stiglitz, competing insurers face two large classes of insurance applicants presenting independent risks. Applicants are expected utility maximizers, each endowed with an exogenous wealth $W$, a risk averse von Neumann-Morgenstern utility function $U(\cdot)$, and the possibility of incurring a loss with damage $D$. They differ only with respect to the probability $p^t \in (0,1)$ of incurring the loss, and each applicant belongs to one of two risk classes $t \in \{H, L\}$ with $p^H$ greater than $p^L$. 


Although insurers know these probabilities, they do not know the risk class to which a particular applicant belongs.

An insurance contract \( C \equiv (\alpha, S) \) consists of a premium \( \alpha \) and a deductible \( S \) that together determine wealth levels \( W_1 = W - \alpha - S \) if the damage is incurred and \( W_o = W - \alpha \) otherwise. Since the amount of loss \( D \) is fixed, a deductible is simply a policy indemnification that is less than the amount of loss, and this coverage limitation could instead be implemented through policy limits or coinsurance provisions. Therefore, we shall use the term “deductible” to refer to any type of coverage restriction.

The expected utility of a contract \( C \) to an insurance applicant of risk type \( t \) is defined by

\[
V^t(C) \equiv (1 - p^t)U(W_o) + p^tU(W_1) .
\] (1)

The expected profit earned by an insurer selling contract \( C \) to an applicant of risk type \( t \) is defined by

\[
\pi^t(C) \equiv \alpha - p^t(D - S) .
\] (2)

Insurers are assumed to be risk neutral, and to incur no costs other than indemnified claims.

In the absence of adverse selection, insurers could perfectly discriminate among applicants by risk type, and an applicant of type \( t \) would receive the contract \( (\alpha^t, S^t) \) that entailed full and fair insurance, with no deductible \( (S^t = 0) \) and a premium equal to the expected loss \( (\alpha^t = p^tD) \). The wealth allocations that would result from these contracts are depicted in Figure 1 as \( H^* \) and \( L^* \) for the high-risk and low-risk applicants, respectively. Given the assumed informational asymmetry, however, the contractual
pair resulting in \((H^*, L^*)\) cannot be implemented, since insurers cannot distinguish between the two types of applicants and \(H\)-types prefer the \(L\)-type contract to their own. As a consequence of this adverse selection, the only informationally feasible contracts are those that satisfy the incentive compatibility constraints

\[
V^t(C^t) \geq V^{t'}(C^{t'}) \text{ for } t, t' \in \{H, L\}. \tag{3}
\]

Rothschild and Stiglitz define contracts to be *informationally consistent* if they satisfy these incentive constraints and also individually earn zero profit.

The informationally consistent contractual pairs include zero-profit pooling contracts that result in wealth positions along the pooled fair-odds line \(FE\) in Figure 1, as well as incentive compatible, separating pairs, in which each contract \(C^t\) satisfies the zero-profit condition

\[
\pi^t(C^t) = 0 \text{ for } t \in \{H, L\}. \tag{4}
\]

Examples of *separating* informationally consistent contracts include pairs that result in \(H^*\) for \(H\)-types and a point between \(E\) and \(A\) along the \(L\)-type fair-odds line \(EL^*\) in Figure 1.

A *pure strategy Nash equilibrium* in this contracting environment is a pair of informationally consistent contracts with the property that no other pair would jointly earn positive expected profit when offered in competition for customers against the equilibrium pair. As Rothschild and Stiglitz show, no pooling contract can be sustained as a Nash equilibrium, since \(L\)-types could always be attracted to a profitable “cream skimming” defection rendering the original pooling contract unprofitable, as it would be purchased only by \(H\)-types. They also demonstrate that the only candidate for
equilibrium is the Pareto dominant member of the class of separating informationally consistent contractual pairs, depicted as \((H^*, A)\) in Figure 1. We shall refer to these as the Rothschild-Stiglitz contracts.

Competition for \(L\)-types ensures that they receive their most preferred contract amongst those that are informationally consistent. Thus, the Rothschild-Stiglitz contracts provide \(H\)-types with full and fair insurance, allowing insurers to offer a fair contract for \(L\)-types that incorporates the lowest deductible consistent with incentive compatibility for \(H\)-types. As a consequence, \(L\)-types bear an externality cost since adverse selection by \(H\)-types prevents insurers from offering a contract that provides \(L\)-types with a lower deductible at a fair price. Nonetheless, competition ensures that this externality cost is minimized subject to the zero-profit and incentive compatibility constraints.

We shall demonstrate in the next section conditions under which multidimensional screening reduces the externality cost of adverse selection by bundling different losses and perils and incorporating differential deductibles that are advantageous to \(L\)-types but disadvantageous to \(H\)-types. Competition for \(L\)-types then drives insurers to offer bundled contracts that efficiently exploit multidimensional screening.

The Rothschild-Stiglitz contracts are sustained as a pure strategy Nash equilibrium if and only if the proportion of \(H\)-types in the applicant pool, \(\lambda\), is greater than or equal to a critical value, \(\lambda^*\). When \(\lambda\) is less than \(\lambda^*\), no pure strategy Nash equilibrium exists. To understand the reason for this nonexistence result, it is useful to consider the class of \textit{interim incentive efficient} contracts that solve the problem of maximizing the expected utility of \(L\)-types, \(V^L(C^L)\), subject to the incentive compatibility constraints (3), the resource constraint for the economy.
and a voluntary participation constraint for $H$-types,

$$V^H(C^H) \geq \bar{V}^H,$$  

that specifies an arbitrary lower bound $\bar{V}^H$ for their expected utilities. The class of interim incentive efficient contracts is generated by varying the participation constraint utility level.

A solution to an interim efficiency problem, which is formally characterized by Crocker and Snow (1985, 1986), can be described with reference to Figure 2. At a solution, the resource constraint (5) holds with equality, as does the incentive constraint (3) for $H$-type applicants. The contingent wealth position for the $H$-types entails full insurance, and so lies along the $45^\circ$ line. The corresponding contingent wealth for $L$-types provides less than full insurance, and is located on the (dotted) locus $FA$ in Figure 2. One member of the class of interim efficient contracts is depicted in Figure 2 by the wealth allocations $(M, Y)$, which is a solution to the efficiency problem when the participation constraint utility for $H$-types is set at or below the level associated with the indifference curve labeled $\bar{V}^H$. Higher levels of reservation utility for the $H$-types result in efficient wealth allocation pairs closer to point $F$, which corresponds to the first-best pooling contract and is itself a solution to the efficiency problem when $\bar{V}^H$ is set equal to $U(W - \bar{p}D)$, where $\bar{p} \equiv \lambda p^H + (1 - \lambda)p^L$.

The case depicted in Figure 2 illustrates the nonexistence problem, which arises when $(H^*, A)$ is not in the set of interim incentive efficient allocations. Since both types of insurance applicants prefer some interim efficient contracts, such as those resulting
in \((M, Y)\), to those resulting in \((H^*, A)\), there are potentially profitable defections that attract both types away from their Rothschild-Stiglitz contracts, which therefore cannot be sustained as a pure strategy Nash equilibrium. Note that all of the interim efficient contracts for \(L\)-types in Figure 2 lie between \(F\) and \(Y\) along locus \(FA\) and entail cross-subsidization, with losses accruing to the \(H\)-type contract offset by profits earned on the \(L\)-type contract. As the proportion of \(L\)-types in the population falls, however, the profitability of a defecting pair of contracts declines as the insurance pool tilts toward the losing \(H\)-type contract. In the Figure, point \(F\) moves down the 45º line and the slope of locus \(FA\) declines in absolute value. At a critical value \(\lambda^*\), the slope of locus \(FA\) at point \(A\) coincides with the marginal rate of substitution of the \(L\)-types at that point, and the Rothschild-Stiglitz contracts resulting in the wealth allocations \((H^*, A)\) are now a solution to an interim efficiency problem. In this event, there are no jointly profitable and incentive compatible defections that can attract applicants away from their Rothschild-Stiglitz contracts, which is now sustained as a pure strategy Nash equilibrium. This situation is depicted in Figure 3.

As the proportion of \(L\)-types declines further and \(\lambda\) rises above \(\lambda^*\), locus \(FA\) continues to pivot downward about point \(A\), and incentive compatible contracts in which \(H\)-types subsidize \(L\)-types become efficient, as illustrated by the efficient pair \((M, Y)\) in Figure 4. Since only \(L\)-types would be attracted to these defections, they do not overturn the Rothschild-Stiglitz contracts as a pure strategy Nash equilibrium. Hence, for all applicant pools in which \(\lambda\) is greater than or equal to \(\lambda^*\), the pure strategy Nash
equilibrium exists, and these are precisely the environments in which the Rothschild-Stiglitz contracts are interim incentive efficient.

We shall demonstrate in section four that the critical value $\lambda^*$ is lower when insurance contracts exploit efficient multidimensional screening. As a consequence, competitive insurance markets attain pure strategy Nash equilibria over a broader range of applicant pools when multidimensional screening is efficient. In this manner, insurance markets circumvent the nonexistence problem as firms are driven by competition for $L$-types to offer contracts that bundle different losses and perils with differential deductibles.

III. Efficient Multidimensional Screening

We introduce multidimensional screening by assuming that $D$ is the expected damage conditional on a loss occurring, and that the potential causes of this expected damage can be partitioned into $n > 1$ mutually exclusive and collectively exhaustive perils. We use $D_i$ to denote the loss associated with the $i$th peril and $\theta_i \in (0,1)$ to denote the conditional probability of this loss occurring for an applicant in risk class $t \in \{H, L\}$. Insurers categorize applicants on the basis of observable traits, and we assume that in partitioning the perils space insurers know the damage $D_i$ and conditional probability $\theta_i^t$ for each peril. The probability of a loss occurring, $p^t$, remains hidden knowledge, along with the conditional probabilities $\theta_i^t$, but insurers can now exploit $n$ signaling dimensions to screen insurance applicants.
In this context, a contract $C$ consists of a premium $\alpha$ and a vector of deductibles $S = (S_1, S_2, \ldots, S_n)$ that together determine wealth $W_i = W - \alpha - S_i$ if damage is incurred as a consequence of the $i$th peril, and $W_o = W - \alpha$ if no damage is incurred. The expected utility of a contract $C$ to an insurance applicant of risk type $t$ is now given by

$$V^t(C) \equiv (1 - p^t)U(W_o) + p^t \sum_{i=1}^{n} \theta^t_i U(W_i), \quad (7)$$

and the corresponding expected profit by

$$\pi^t(C) \equiv \alpha - p^t \sum_{i=1}^{n} \theta^t_i (D_i - S_i). \quad (8)$$

Hence, just as in the model of one-dimensional screening, there are only two types of insurance applicants and they signal their types through their deductible choices. Our innovation is to allow applicants to signal their hidden knowledge of risk type in more than one dimension through the choice of a vector of deductibles. To maintain strict comparability with the original Rothschild-Stiglitz framework, we assume that insurers are able to categorize applicants so that everyone in a given pool has an expected loss, conditional on a loss occurring, equal to $D$. That is, we require that

$$\sum_{i=1}^{n} \theta^H_i D_i = \sum_{i=1}^{n} \theta^L_i D_i \equiv D,$$

so that if we restrict the deductibles $S_i$ to be the same for all of the perils, the contracting environment reduces to the one-dimensional screening model considered by Rothschild and Stiglitz.

As in the case of one-dimensional screening where there is a single deductible, interim incentive efficient contracts maximize the expected utility of $L$-types subject to the incentive compatibility constraints (3), the resource constraint (5), and the voluntary participation constraint for $H$-types (6), but now using definitions (7) and (8) for expected utility and profit in place of (1) and (2). With coverage for all of the perils bundled into a
single insurance policy, there are only two incentive compatibility constraints, one for each risk class. If the perils were unbundled and covered by separate contracts, then additional incentive compatibility constraints would be required. Since these additional constraints cannot increase the value of efficient contracts, bundling of the perils is efficient, so there are only the two incentive compatibility constraints (3). Moreover, as in the one-dimensional screening environment, only the $H$-type incentive constraint is binding at a solution to an interim efficiency problem, and a solution provides $H$-types with full coverage and underinsures $L$-types. Thus, the introduction of multidimensional screening through a partitioning of the perils potentially responsible for damage does not fundamentally alter the structure of the screening environment set out by Rothschild and Stiglitz.

A pair of contracts is defined to be *incentive feasible* if it satisfies the incentive compatibility constraints (3) and the resource constraint (5). We define multidimensional screening to be *efficient* if, for any value of $\lambda \in (0,1)$ and any pair of incentive compatible one-dimensional screening contracts, $L$-types prefer, ceteris paribus, an incentive feasible contract with deductibles that vary across the perils. Our first result establishes that multidimensional screening is efficient whenever the two risk types differ with respect to the odds of the damage being caused by one peril rather than another. Moreover, it is always efficient to exploit these differences by setting a higher deductible in the $L$-type contract for the peril more likely for $H$-types. *Efficient* multidimensional screening fully exploits these differences to the advantage of $L$-types.
Theorem 1. Multidimensional screening is efficient if and only if for some pair of signaling dimensions $i$ and $j$, $\theta_i^L / \theta_j^L \neq \theta_i^H / \theta_j^H$. Moreover, efficient multidimensional screening requires $S_i^L < (>) S_j^L$ as $\theta_i^L / \theta_j^L > (<) \theta_i^H / \theta_j^H$.

Proof. Let $(\hat{C}^H, \hat{C}^L)$ be an incentive feasible pair of one-dimensional screening contracts associated with locus $FA$, providing $H$-types with the fully insured wealth $W^H < W - \bar{p}D$ and satisfying their incentive compatibility constraint with equality. Consider the problem of choosing an $L$-type contract $C^L$ to maximize their expected utility subject to the resource constraint (5) and the incentive compatibility constraint $V^H (C^L) \leq U(W^H)$. The solution to this problem, in conjunction with $\hat{C}^H$, is a pair of efficient multidimensional screening contracts. The Lagrangian expression associated with the problem can be written as

$$Z = V^L (C^L) + \xi [U(W^H) - V^H (C^L)] + \mu [\lambda \pi^H (\hat{C}^H) + (1 - \lambda) \pi^L (C^L)],$$

where $\xi$ and $\mu$ are the Lagrange multipliers for the binding incentive compatibility constraint and the resource constraint, respectively. The first-order necessary condition for the choice of the $L$-type deductible for the $k$th peril is\(^\,\,^12\)

$$-(p^L \theta_k^L - \xi p^H \theta_k^H) U'(W^L_k) + \mu (1 - \lambda) p^L \theta_k^L = 0,$$

where $U'(\cdot)$ denotes marginal utility. It follows that

$$(1 - \xi p^H \theta_k^H / p^L \theta_k^L) U'(W^L_k) = \mu (1 - \lambda),$$

implying that $U'(W^L_k)$ varies directly with $\theta_k^H / \theta_k^L$. We thus have

$$U'(W^L_k) > (<) \, U'(W^L_i)$$
when \( \theta_i^H / \theta_j^L > (=)[<] \theta_i^H / \theta_j^L \). Since \( U(\cdot) \) is strictly concave, reflecting risk aversion, we must then have \( S_i^L < (=)[>] S_j^L \) as \( \theta_i^L / \theta_j^L > (=)[<] \theta_i^H / \theta_j^H \). Finally, observe that this argument is independent of the \( L \)-type deductibles in dimensions other than \( i \) and \( j \), so that any difference in the conditional odds ratios for the two risk types is exploited by efficient multidimensional screening. \( \text{Q.E.D.} \)

Notice that efficient multidimensional screening always follows the pattern of comparative advantage enjoyed by \( L \)-types for signaling in one dimension rather than another. Moreover, this pattern depends solely on the exogenous conditional probabilities associated with the perils. While \( H \)-types are unaffected by the introduction of multidimensional screening, since their efficient contract entails full insurance, \( L \)-types benefit in two ways. First, they obtain relatively more coverage for the perils they are more likely to experience, and second, this shift in coverage makes their contract less attractive to \( H \)-types, relaxing the incentive constraint and thus further reducing the cost of adverse selection to the benefit of \( L \)-types. Hence, adopting efficient multidimensional screening is Pareto improving, and insurers competing for \( L \)-types will be led by market forces to implement efficient multidimensional screening.

**IV. Pure Strategy Nash Equilibrium with Multidimensional Screening**

The nonexistence of Nash equilibrium arises with multidimensional screening just as in the one-dimensional case. Indeed, our model reduces to the one-dimensional screening environment when insurance contracts are artificially constrained to require that the deductible be the same for each peril covered in a given contract, so that
for every pair of signaling dimensions \( i, j \in \{1, \ldots, n\} \) and each risk type \( t \). When these artificial constraints are not imposed, and coverage for the separate perils is bundled in a single contract for each risk type, competition through premium-deductible menus is simply elaborated by the potential for several different deductibles to appear in the same contract in order to screen the two types of applicants more effectively. Competition for \( H \)-types ensures that they receive full and fair insurance, just as in the one-dimensional setting. And, if it is possible to structure the \( L \)-type contract so that it breaks even and is preferred by them to their Rothschild-Stiglitz contract, while being disliked by \( H \)-types, then insurers competing for \( L \)-types will introduce multidimensional screening. Notice that this is precisely a circumstance in which the introduction of multidimensional screening is Pareto improving, so that market forces will compel insurers to implement efficient multidimensional screening.

We shall henceforth assume that multidimensional screening is efficient.\(^{13}\) As in the one-dimensional screening model, the only candidate for pure strategy Nash equilibrium is the pair of contracts that provides full and fair insurance for \( H \)-types and provides \( L \)-types with an incentive compatible, zero-profit contract. Now, however, the deductibles in the \( L \)-type contract vary across the perils in a fashion that maximizes the \( L \)-type’s expected utility.

In this section we establish that efficient multidimensional screening improves the prospects for existence of a pure strategy Nash equilibrium. It is important to recognize that this result is not a trivial consequence of the assumed efficiency of multidimensional screening, since, as an implication of Theorem 1, all of the second-best incentive efficient
contracts are more valuable to $L$-types with multidimensional screening. Thus, $L$-types are better off with multidimensional screening not only at the candidate for Nash equilibrium, but also at any potential defection. The key insight driving the analysis that follows is that incentive feasible contracts farther removed from the first-best pool, $F$, entail higher adverse selection costs for $L$-types, who then gain more from the introduction of multidimensional screening.

To establish our result, we utilize two incentive efficiency programs, program (A) for one-dimensional screening and program (B) for the corresponding multidimensional screening environment. In both programs, the contract $C^L$ is selected to maximize the $L$-type expected utility, $V^L(C^L)$, subject to a constraint on the expected utility of $H$-types, $V^H(C^L) \leq \bar{V}^H$, and a constraint on the $L$-type expected profit, $\pi^L(C^L) \geq \bar{\pi}^L$. Note that, by judicious selection of the constraint values $\bar{V}^H$ and $\bar{\pi}^L$, we can achieve as a solution to program (A) any desired contingent wealth position below the $45^\circ$ line. So, for example, in Figure 5 the choice set for program (A) consists of contingent wealth positions that lie simultaneously below $\bar{V}^H$ and $\bar{\pi}^L$, and the solution to program (A) is depicted as point $G$ for $L$-types, providing them with expected utility $\hat{V}^L$. The solution to program (B) is the corresponding efficient multidimensional screening contract.

We now introduce the concept of an *iso-utility curve for multidimensional screening*, which is defined to be the set of contingent wealth pairs $(W_0, W_1)$ associated with one-dimensional screening whose corresponding incentive efficient multidimensional screening contracts provide $L$-types with a constant level of expected utility. In the Lemma that follows, we will show that each iso-utility curve is downward
sloping and, at any point below the 45-degree line, is more steeply sloped than \( L \)-type indifference curves.

Formally, define the function \( \varphi(\hat{W}_o, \hat{W}_1) \) to be the indirect utility function for program (B) when the constraints are set by \((\hat{W}_o, \hat{W}_1)\), that is, when

\[
\overline{V}^H = (1 - p^H)U(\hat{W}_o) + p^H U(\hat{W}_1)
\]

and

\[
\pi^L = W - p^L D - (1 - p^L)\hat{W}_o - p^L \hat{W}_1.
\]

The corresponding iso-utility curve is \( \Psi(\hat{W}_o, \hat{W}_1) \equiv \{(W_o, W_1) | \varphi(W_o, W_1) = \varphi(\hat{W}_o, \hat{W}_1)\} \).

We now establish that each iso-utility curve is downward sloping and steeper than \( L \)-type indifference curves, as illustrated by the dashed line \( \Psi(\hat{W}_o, \hat{W}_1) \) passing through point \( G \) in Figure 5.

**Lemma.** Assume that the solutions to program (A) and the corresponding program (B) provide less than full coverage. If multidimensional screening is efficient, then, at a solution to program (A), the iso-utility curve is downward sloping and steeper than the \( L \)-type indifference curve.

**Proof.** See the Appendix

We are now in a position to prove that a pure strategy Nash equilibrium exists over a wider range of markets when insurers can exploit efficient multidimensional screening. Let \( \lambda^* \) and \( \lambda^{**} \) denote the critical values for the proportion \( \lambda \) of \( H \)-types in the applicant pool such that a pure strategy Nash equilibrium exists for all higher values
of $\lambda$ under one-dimensional and multidimensional screening, respectively. We demonstrate that $\lambda^{**}$ is less than $\lambda^*$ if multidimensional screening is efficient.

The intuition can be explained by reference to Figure 3. We have demonstrated that the iso-utility curve passing through point $A$ is steeper than the $L$-type indifference curve at that point. Thus, when $\lambda$ equals $\lambda^*$ and locus $FA$ is tangent to the $L$-type indifference curve at point $A$, the corresponding iso-utility curve has a steeper slope. As a consequence, there are values for $\lambda$ less than $\lambda^*$ such that locus $FA$ rises above the $L$-type indifference curve at $A$, but remains below the corresponding iso-utility curve. For these values of $\lambda$, profitable defections overturn the Rothschild-Stiglitz contracts as a Nash equilibrium with one-dimensional screening, but there are no profitable defections that overturn the multidimensional screening counterparts to the Rothschild-Stiglitz contracts. Thus, there is a range of values for $\lambda$ over which the Rothschild-Stiglitz one-dimensional screening contracts cannot be sustained as a Nash equilibrium, but their corresponding multidimensional screening contracts can.

**Theorem 2.** If multidimensional screening is efficient, then we have $\lambda^{**} < \lambda^*$.

**Proof.** Let $\Psi^{RS}$ denote the iso-utility curve passing through the $L$-type Rothschild-Stiglitz contract at point $A$ in the Figures, and let $\bar{\nu}^L$ denote the corresponding $L$-type indifference curve. Under the assumption that multidimensional screening is efficient, the Lemma demonstrates that $\Psi^{RS}$ is more steeply sloped than $\bar{\nu}^L$ at point $A$, implying the existence of a value $\lambda^c (< \lambda^*)$ such that, for all
\( \lambda \in [\lambda^c, \lambda^*], \) locus \( FA \) lies below \( \Psi^{RS} \) inside a neighborhood \( N \) of point \( A \), excluding point \( A \). Moreover, since \( \lambda < \lambda^* \) in this range, there is no pure strategy Nash equilibrium with one-dimensional screening for these values of \( \lambda \).

The Lemma also establishes that \( \Psi^{RS} \) lies strictly above \( \overline{V}^{L} \) between point \( A \) and the 45-degree line, and is more steeply sloped than \( L \)-type indifference curves. Hence, there exists a value \( \varepsilon > 0 \) such that the \( L \)-type indifference curve associated with expected utility \( \overline{V}^{L} + \varepsilon \) cuts \( \Psi^{RS} \) within the neighborhood \( N \) and lies below \( \Psi^{RS} \) between this intersection and the 45-degree line. Let \( \hat{\lambda} (< \lambda^* \) be the supremum value for \( \lambda \) such that there exists a profitable, one-dimensional screening defection from the Rothschild-Stiglitz contracts that yields expected utility \( \overline{V}^{L} + \varepsilon \) for \( L \)-types. Thus, for \( \lambda > \hat{\lambda} \), the locus \( FA \) necessarily lies below the \( \overline{V}^{L} + \varepsilon \) indifference curve. The situation when \( \lambda = \hat{\lambda} \) is depicted in Figure 6.

If \( \hat{\lambda} < \lambda^c \), then for all \( \lambda \in [\lambda^c, \lambda^*), \) locus \( FA \) must lie everywhere below the indifference curve for \( \overline{V}^{L} + \varepsilon \), which in turn lies below \( \Psi^{RS} \) outside neighborhood \( N \) and on its boundary, while locus \( FA \) lies below \( \Psi^{RS} \) inside neighborhood \( N \). Hence, if we have \( \hat{\lambda} < \lambda^c \), then for all \( \lambda \in [\lambda^c, \lambda^*), \) locus \( FA \) lies below \( \Psi^{RS} \) between point \( A \) and the 45-degree line, implying that there are no profitable defections from the Rothschild and Stiglitz contracts with multidimensional screening for these values of \( \lambda \). Therefore, we must have \( \lambda^{**} < \lambda^* \).
On the other hand, if we have $\hat{\lambda} \geq \lambda^c$, then for all $\lambda \in [\hat{\lambda}, \lambda^* )$, locus $FA$ must be tangent to or lie below the indifference curve for $\bar{V}^L + \varepsilon$ (recalling the definition of $\hat{\lambda}$), and therefore must once again lie below $\Psi^{RS}$ between point $A$ and the 45-degree line, implying that $\lambda^{**} < \lambda^*$. \textit{Q.E.D.}

V. Limit Cases

In this section, we investigate limit cases in which multidimensional screening ultimately eliminates adverse selection. We suppose that a finer partitioning of the $i$th signaling dimension increases the solution value of the interim incentive efficiency problem. Partitioning the $i$th dimension divides the conditional probability $\theta_i^I$ between two perils, one with conditional probability $\beta^I\theta_i^I$ and the other with probability $(1 - \beta^I)\theta_i^I$ for a fraction $\beta^I$ that depends on the applicant’s type. The following result shows that, as further partitioning increases the value of interim incentive efficiency problem, the smaller of the partitioned odds ratios, $\beta^L\theta_i^L / \beta^H\theta_i^H$ or $(1 - \beta^L)\theta_i^L / (1 - \beta^H)\theta_i^H$, declines toward zero. In the limit, where $\theta_i^L = 0 < \theta_i^H$ for some dimension $i$, adverse selection is eliminated as both types receive full and fair insurance with the deductible $S_i^L$ set high enough to deter the $H$-types from taking the $L$-type contract.\textsuperscript{16}
**Theorem 3.** If partitioning in the $i$th signaling dimension is efficient, then $L$-types have a stronger comparative advantage in accepting a deductible for one of the resulting perils than in the unpartitioned signaling dimension, so that

$$
\min\{\beta^L \theta_i^L / \beta^H \theta_i^H, (1-\beta^L) \theta_i^L / (1-\beta^H) \theta_i^H\} < \theta_i^L / \theta_i^H,
$$

and in the limit, as continued partitioning is valuable, the left-hand side approaches zero and adverse selection is eliminated.

**Proof.** Partitioning the $i$th signaling dimension would be of no value if $\beta^L$ were to equal $\beta^H$, for then the odds ratios for both parts would equal $\theta_i^L / \theta_i^H$ and, as shown by Theorem 1, multidimensional screening would have no additional value with such a partitioning. However, by hypothesis, the partitioning is valuable, so we may assume without loss of generality that $\beta^L$ is less than $\beta^H$, so that $\beta^L \theta_i^L / \beta^H \theta_i^H$ is less than $\theta_i^L / \theta_i^H$, establishing the inequality stated in the Theorem.

Now replace $\theta_i^L / \theta_i^H$ with $\beta^L \theta_i^L / \beta^H \theta_i^H$ throughout the inequality and repeat the preceding argument to conclude that the left-hand side is once again less than the right-hand side for an appropriate portioning. By proceeding in this manner, when partitioning of the signaling space is continually valuable, we can make the left-hand side of the inequality approach zero. Since the denominators of the two ratios on the left-hand side are bounded above by one, it must be that the smaller of the numerators equals zero in the limit. The Kuhn-Tucker condition for efficient signaling by $L$-types in the $i$th dimension requires

$$
-(p^L \theta_i^L - \xi p^H \theta_i^H)u'(w_i^L) + \mu(1-\lambda)p^L \theta_i^L \leq 0.
$$
With $\theta_i^L = 0$ and $\theta_i^H > 0$, we must have $\xi = 0$, indicating that the incentive constraint is not binding.\(^{17}\) Hence, in the limit, adverse selection is eliminated by efficient multidimensional screening. \textit{Q.E.D.}

In the preceding section, we established that, when costless partitioning of the signaling space enhances efficiency, the prospects improve for the existence of a pure strategy Nash equilibrium. We now see that, if costless finer partitioning continually enhances efficiency, then in the limit the existence of equilibrium is assured.

VI. Conclusions

Bundled coverage of different losses and distinct perils with differential deductibles and policy limits allows insurers to implement multidimensional screening of insurance applicants who possess hidden knowledge of their risks of loss. As in the case of one-dimensional screening examined by Rothschild and Stiglitz, competition for low risk applicants governs the level of signaling and the existence of Nash equilibrium, but is now also seen to determine the number of distinct screening dimensions exploited in insurance contracting. When the two types of applicants face different odds of incurring losses of distinct types or a loss caused by distinct perils, multidimensional screening reduces the externality cost of adverse selection that low-risk applicants must bear to distinguish themselves from high risks, as the two types can be more efficiently sorted by offering different coverage limitations for the alternative loses in a pattern advantageous to the low risks. Moreover, by reducing the externality cost of adverse selection, efficient
multidimensional screening allows competitive insurance markets to attain pure strategy Nash equilibria over a wider range of applicant pools.

These insights apply equally in other markets where adverse selection limits contracting possibilities. In the job market context examined by Spence (1973), where education serves as the signaling dimension, fields of specialization play the role of distinct perils, as those job applicants best suited for positions as engineers have an advantage in math and the natural sciences, while those best suited for positions as journalists find a more writing-intensive curriculum to their advantage. Similarly, in credit markets applicants can signal their creditworthiness by accepting a higher collateral requirement, as suggested by Bester (1985). By setting different requirements for different types of collateral, lenders can take advantage of any differences in the willingness of applicants to bear the risk of forfeiture, enhancing efficiency and the prospects for existence of a pure strategy Nash equilibrium in these markets.
Appendix

In this Appendix, we provide a proof of the Lemma. As a notational convention, we denote the Lagrange multipliers for the utility and profit constraints for the one-dimensional screening program (A) by $\hat{\xi}$ and $\hat{\mu}$, respectively, and those for the multidimensional screening program (B) by $\xi$ and $\mu$. The contingent wealth for $L$-types at a solution to program (A) is denoted by $(\hat{W}_o^L, \hat{W}_1^L)$ and the solution value for $L$-type expected utility by $\hat{V}^L$. The $L$-type contingent wealth at a solution to the corresponding program (B) is denoted by $(W_o^L, W_1^L)$ and the resulting expected utility by $V^L$.

We next establish two preliminary results. The first shows that, when multidimensional screening is efficient, the $L$-types pay a higher insurance premium than they do under the corresponding one-dimensional screening contract. The intuition behind this result is illustrated in Figure A-1, which refers to wealth levels for two perils $i$ and $j$. The indifference curves for a type $t$ applicant have the slope $\theta_i W_{UI} / \theta_j W_{UJ}$, while the slope of a corresponding isoprofit line is $-\theta_i / \theta_j$. The case depicted is that where $H_j H_i L_j L_i > \theta_i H / \theta_j H$, so that the $L$-type’s indifference curve is steeper than that of the $H$-type. Accordingly, multidimensional screening is efficient, and Theorem 1 implies that $S_i^L < S_j^L$ or, equivalently, $W_i^L > W_j^L$. As a result, starting from a point such as $K$ with equal wealth in both dimensions, the expected utility of $H$-types can be maintained while increasing the expected utility of $L$-types by moving to a position such as $J$. Since this change reduces profit earned off $L$-types, their
premium must increase to satisfy the \( L \)-type profit constraint. We now provide a formal proof of this result.

**Result 1.** If multidimensional screening is efficient, then we have \( W_o^L < \hat{W}_o^L \).

**Proof.** Let \( \delta \in [0,1] \) and define

\[
V(\delta) \equiv (1 - p^L)U(\hat{W}_o^L + \delta(W_o^L - \hat{W}_o^L)) + p^L \sum_i \theta_i^L U(\hat{W}_1^L + \delta(W_i^L - \hat{W}_1^L))
\]


to be the expected utility of an \( L \)-type under the one-dimensional screening contract when \( \delta = 0 \) and under the multidimensional screening contract when \( \delta = 1 \). If and only if multidimensional screening is efficient, \( V(1) > V(0) \) and then by construction \( V(\delta) \) has its minimum value at \( \delta = 0 \) and increases to its maximum value at \( \delta = 1 \). Thus, we have

\[
V'(\delta) = (1 - p^L)U'_o(W_o^L - \hat{W}_o^L) + p^L \sum_i \theta_i^L U'_i(W_i^L - \hat{W}_1^L) > 0
\]

for all \( \delta \in [0,1] \), where the subscripts on marginal utility indicate the state contingent wealth \( \hat{W}_s^L + \delta(W_s^L - \hat{W}_s^L) \) for \( s \in \{0,1,2,\ldots,n\} \). At \( \delta = 0 \) we have

\[
V'(0) = (1 - p^L)U'_o(W_o^L - \hat{W}_o^L) + p^L \sum_i \theta_i^L (W_i^L - \hat{W}_1^L) > 0.
\]

Since the expected profit for the \( L \)-types is the same under both \( L \)-type contracts for programs (A) and (B), we have from (2) and (8)

\[
\pi^L = W - p^L D - (1 - p^L)\hat{W}_o^L - p^L \hat{W}_1^L
= W - p^L D - (1 - p^L)W_o^L - p^L \sum_i \theta_i^L W_i^L,
\]

This implies

\[
p^L \sum_i \theta_i^L [W_i^L - \hat{W}_1^L] = -(1 - p^L) [W_o^L - \hat{W}_o^L].
\]

Substituting from this equation into the expression for \( V'(0) \), we arrive at
\[(1 - p^L) [U'(\hat{W}_o^L) - U'(\hat{W}_1^L)] (W_o^L - \hat{W}_o^L). \]

Since we have assumed multidimensional screening to be efficient, it must be the case that a solution to program (A) entails a positive deductible, so that \(U'(\hat{W}_o^L) < U'(\hat{W}_1^L), \)
from which the result \(W_o^L < \hat{W}_o^L\) follows immediately. \textbf{Q.E.D.}

Our second preliminary result shows that, near the 45-degree line where insurance coverage is almost full, the marginal cost of the adverse selection externality is lower with multidimensional screening.

\textbf{Result 2.} Assume that \(V^H\) and \(L^\pi\) are set so that the solutions to program (A) and the corresponding program (B) provide almost full coverage. If multidimensional screening is efficient, then we have \(\hat{\xi} > \xi.\)

\textit{Proof.} Given \(L^\pi\), suppose that \(V^H\) were set at a higher value such that the solutions to both programs (A) and (B) provided full coverage. In that case, multidimensional screening would have no value, so \(V^L - \hat{V}^L = 0,\) and the incentive compatibility constraints would impose no adverse selection costs, so we would necessarily have \(\hat{\xi} = \xi = 0.\) Now, holding \(L^\pi\) constant and having reduced \(V^H\) infinitesimally, a further marginal reduction in \(V^H\) must lead to an increase in \(V^L - \hat{V}^L\) if multidimensional screening is efficient. The Lagrangean expression associated with program (A) may be written as

\[L = V^L (\hat{C}^L) + \hat{\xi} [\bar{V}^H - V^H (\hat{C}^L)] + \hat{\mu} [\pi^L (\hat{C}^L) - \pi^L].\]
where \( \hat{C}^L = (\alpha^L, S^L) \), \( V' \) and \( \pi^L \) are defined as in (1) and (2), respectively, and the Lagrange multipliers are non-negative at a solution.\(^{19} \) The Lagrangean expression associated with program (B) is similar, only with \( \hat{C}^L \) replaced with \( C^L = (\alpha^L, S_1^L, \ldots, S_n^L) \), \( \hat{\xi} \) and \( \hat{\mu} \) replaced with \( \xi \) and \( \mu \), respectively, and \( V' \) and \( \pi^L \) are defined as in (7) and (8).

As envelope results from the Lagrangian formulations of programs (A) and (B), we have

\[
LHL dVdVd\pi\mu\xi\hat{\xi}\hat{\mu} = 0
\]

Thus, for the thought experiment just described, we have

\[
d\hat{\pi}^L = 0 > d\pi^H \quad \text{and} \quad d(V^L - \hat{V}^L) = (\xi - \hat{\xi})d\pi^H > 0,
\]

which together imply \( \hat{\xi} > \xi \). \textbf{Q.E.D.}

We now use Results 1 and 2 to prove the Lemma stated in the text.

\textit{Proof of the Lemma.} To prove that iso-utility curves are downward sloping and steeper than \( L \)-type indifference curves, we consider a thought experiment in which we perturb the \( L \)-type contingent wealth \((\hat{W}^L_o, \hat{W}^L_1)\) provided by a one-dimensional screening contract associated with program (A) in such a way that the expected utility of their corresponding multidimensional screening contract remains constant \((dV^L = 0)\). Hence, the perturbation entails movement along the iso-utility curve passing through \((\hat{W}^L_o, \hat{W}^L_1)\).
We first focus on perturbations in which the expected utility of $L$-types under one-dimensional screening declines $(d\hat{V}^L < 0)$. We will show that $d\hat{\pi}^L$ must be negative in this case, implying that this movement takes one below the $L$-type indifference curve and to the right of the iso-profit line, as illustrated by the move from point $G$ to point $X$ in Figure 5. We then demonstrate that, holding $V^L$ constant, positive perturbations in $d\hat{V}^L$ imply that $d\hat{\pi}^L$ must be positive, so that this movement takes one above the $L$-type indifference curve and to the left of the iso-profit line, as illustrated by the move from point $G$ to point $X'$ in Figure 5. It follows that, below the 45-degree line, iso-utility curves are downward sloping and steeper than $L$-type indifference curves.

To accommodate the perturbation $dV^L = 0$ with $d\hat{V}^L < 0$, the constraint parameters $\bar{\pi}^H$ and $\bar{\pi}^L$ in programs (A) and (B) must change. Recalling equations (10), implied by the Lagrangian functions for the two programs, $dV^L = 0$ implies $d\bar{\pi}^H = (\mu/\hat{\xi})d\bar{\pi}^L$, and $d\hat{V}^L < 0$ then implies

$$\left(\frac{\hat{\xi}}{\hat{\mu}} - \frac{\hat{\xi}}{\mu}\right)d\bar{\pi}^L < 0$$

(11)

since the multipliers are positive. It follows that we cannot have $d\bar{\pi}^L = 0$. Further, if we have $d\bar{\pi}^L > (\leq 0)$, then we must have $\hat{\xi}/\hat{\mu} < (\leq) \xi/\mu$. Since only one of this second pair of inequalities can hold, we must have either $d\bar{\pi}^L > 0$ or $d\bar{\pi}^L < 0$.

The first-order conditions for program (B) are, for each deductible $S^L_i$,

$$-p^L_i \theta_i^L U_i' + \hat{\xi} H^H_i (\theta_i^L + \mu p^L_i \theta_i^L) = 0$$

(12)

and that for the premium, after substitution of equations (12), may be written as
\[-(1 - p^L)U'_o + \xi(1 - p^H)U'_o + \mu(1 - p^L) = 0. \tag{13}\]

By solving equation (12) for $\xi$ with $i = 1$, substituting the result into equation (13), and then solving for $\mu$, we obtain

\[
\mu = \frac{1 - P\theta}{u - P\theta} U'(W_o^L),
\tag{14}
\]

where $u = U'(W_o^L)/U'(W_1^L)$, $P = p^H(1 - p^L)/p^L(1 - p^H)$, and $\theta = \theta_1^H/\theta_1^L$.

Substituting for $\mu$ from equation (14) into equation (13) and solving for $\xi$ yields

\[
\xi = \frac{1 - p^L}{1 - p^H} \left[ 1 - \frac{1 - P\theta}{u - P\theta} \right].
\tag{15}
\]

Doing the same for program (A), yields

\[
\hat{\mu} = \frac{1 - P}{\hat{u} - P} U'(\hat{W}_o^L) \text{ and } \hat{\xi} = \frac{1 - p^L}{1 - p^H} \left[ 1 - \frac{1 - P}{\hat{u} - P} \right],
\tag{16}
\]

where $\hat{u} = U'(\hat{W}_o^L)/U'(\hat{W}_1^L)$.

Recalling our stipulation that $V^L$ remains constant, we have

\[
0 = dV^L = (1 - p^L)U'(W_o^L)dW_o^L + p^L\Sigma^L,
\tag{17}
\]

where $\Sigma^L \equiv \sum \theta_i^L U'(W_i^L)dW_i^L$, along with

\[
d\pi^L = -(1 - p^L)dW_o^L - p^L\tilde{\Sigma},
\tag{18}
\]

where $\tilde{\Sigma} \equiv \sum \theta_i^L dW_i^L$, and

\[
d\overline{V}^H = (1 - p^H)U'(W_o^L)dW_o^L + p^H\Sigma^H,
\tag{19}
\]
where $\Sigma^H \equiv \sum_i \theta_i^H U'(W_i^L) dW_i^L$. Solving equation (17) for $dW_o^L$ and substituting the result into equations (18) and (19) yields

$$d\pi^L = p^L \left[ -\Sigma + \Sigma^L / U'(W_o^L) \right] \tag{20}$$

and

$$d\bar{\nu}^H = p^H \left[ \Sigma^H - \Sigma^L / P \right], \tag{21}$$

respectively.

From the Lagrangian for program (B), along with the constancy of $V^L$, we have

$$dV^L = \xi d\bar{\nu}^H - \mu d\pi^L$$

$$\quad = \frac{1 - p^L}{1 - p^H} d\bar{\nu}^H - \frac{1 - P\theta}{u - P\theta} \left[ \frac{1 - p^L}{1 - p^H} d\bar{\nu}^H + U'(W_o^L) d\pi^L \right] = 0 \tag{22}$$

after substituting for $\mu$ and $\xi$ from equations (14) and (15). Substituting in the second line of (22) for $d\pi^L$ from equation (20), and for the second $d\bar{\nu}^H$ from equation (21), and then rearranging terms, we arrive at

$$\frac{d\bar{\nu}^H}{p^H} u - P\theta = \Sigma^H - \frac{U'(W_o^L)}{P} \Sigma. \tag{23}$$

Following the same procedure with program (A), and recalling our stipulation that $\hat{\nu}^L$ declines, we have

$$d\hat{\nu}^L = \hat{\xi} d\bar{\nu}^H - \hat{\mu} d\pi^L$$

$$\quad = \frac{1 - p^L}{1 - p^H} d\bar{\nu}^H - \frac{1 - \hat{P}}{\hat{u} - P} \left[ \frac{1 - p^L}{1 - p^H} d\bar{\nu}^H + U'(\hat{W}_o^L) d\pi^L \right] < 0, \tag{24}$$

after substitution of $\hat{\xi}$ and $\hat{\mu}$ from (16). Substituting from (20) and (21) yields
\[
\frac{d\bar{\nu}^H}{p^H} \frac{\hat{u} - P}{1 - P} < \Sigma^H - \frac{U'(\hat{W}_o^L)}{P} \Sigma - \frac{1}{P} \left[ 1 - \frac{U'(\hat{W}_o^L)}{U'(W_o^L)} \right] \Sigma^L.
\]  

(25)

By solving equation (23) for \( \Sigma^H \), substituting the result into (25), and then combining like terms, we arrive at

\[
\frac{d\bar{\nu}^H}{p^H} \left[ \frac{\hat{u} - P}{1 - P} - \frac{u - P \theta}{1 - P \theta} \right] < \left[ \Sigma - \frac{\Sigma^L}{U'(W_o^L)} \right] \left[ \frac{U'(W_o^L) - U'(\hat{W}_o^L)}{P} \right].
\]

(26)

This inequality implies

\[
\frac{d\bar{\nu}^H}{p^H} \left[ \left( 1 - \frac{1 - p^H}{1 - p^L} \hat{\xi} \right)^{-1} - \left( 1 - \frac{1 - p^H}{1 - p^L} \xi \right)^{-1} \right] < \left[ \frac{U'(\hat{W}_o^L) - U'(W_o^L)}{P} \right] \frac{d\pi^L}{p^L}
\]

(27)

after substituting for the two terms in the bracketed expression on the left-hand side of (26) using equations (15) and (16) for \( \xi \) and \( \hat{\xi} \), and substituting for the first bracketed expression on the right-hand side of (26) using equation (20) for \( d\pi^L \).

Since \( \mu \) and \( \hat{\mu} \) are both positive, equations (14) and (16) for these multipliers reveal that the two terms in the bracketed expression on the left-hand side of (26) are both positive. It follows that the two bracketed expressions on the left-hand side of (27) are also positive, since each is equal to the corresponding positive ratio within brackets on the left-hand side of (26). As a result, when the two terms on the left-hand side of (27) are combined, the denominator is necessarily positive, and the numerator is equal to

\[
\frac{1 - p^H}{1 - p^L} (\hat{\xi} - \xi).
\]

(28)

Lemma 1 implies that the bracketed expression on the right-hand side of (27) is negative. Now suppose we have \( d\pi^L > 0 \). The right-hand side of (27) would then be negative.
Equation (22) shows that $dV^H$ must be positive if $d\pi^L$ is positive. Therefore, the expression at (28) would have to be negative, implying $\xi > \hat{\xi}$. However, from Lemma 2 we have $\hat{\xi} > \xi$ when coverage for the $L$-types is almost full. This contradiction implies that the supposition $d\pi^L > 0$ must be false, at least when coverage is almost full.

Finally, suppose that, for a given value of $\pi^L$, there exists a value for $V^H$ such that partial coverage for the $L$-types is far from full and the thought experiment $dV^L = 0 > d\hat{V}^L$ yields $d\pi^L > 0$. For the given value of $\pi^L$, there exist values for $V^H$ such that coverage for the $L$-types is almost full and the thought experiment $dV^L = 0 > d\hat{V}^L$ yields $d\pi^L < 0$. Since the environment is continuous, there would then have to exist a value for $V^H$ at which the thought experiment $dV^L = 0 > d\hat{V}^L$ yields $d\pi^L = 0$. However, we have established that we cannot have $d\pi^L = 0$. Hence, the thought experiment must always yield $d\pi^L < 0$.

Thus, when $dV^L = 0$ and $d\hat{V}^L < 0$, we must have $d\pi^L$ negative, so that the iso-utility curve through point $G$ in Figure 5 passes through a point such as $X$. The case in which $d\hat{V}^L > 0$ may be considered in a parallel fashion, the primary difference being that the inequality in (27) is then reversed. The result is that, when $dV^L = 0$ and $d\hat{V}^L > 0$, we must have $d\pi^L$ positive, so that the iso-utility curve through point $G$ passes through a point such as $X'$. It follows that the iso-utility curves are downward sloping and steeper than $L$-type indifference curves. \textbf{Q.E.D.}
Endnotes

1 It is worth noting that some authors have examined repeated contracting in adverse selection economies, which also introduces a kind of multidimensional screening wherein signaling dimensions correspond to distinct time periods, and experience rating is possible. Multidimensional screening is efficient in these contexts when principals and agents can commit to long-term contracts. Cooper and Hayes (1987) and Dionne and Doherty (1994) pursue analyses in this vein. Fluet (1992) considers a temporal version of partitioning perils, establishing that a time-dependent deductible is efficient when two risk types differ regarding the likely time of incurring the insured loss. None of these studies, however, consider the implications for existence of pure strategy Nash equilibrium.

2 See the CPCU Handbook of Insurance Policies (2001). Notable examples include the declarations of perils covered, and the associated amounts of indemnification, commonly encountered in commercial property policies (pp. 111-128), as well as in commercial general liability policies (pp. 340-354).

3 This finding of market failure has spawned a substantial literature aimed at resolving the existence question, including the non-Nash equilibrium concepts developed by Wilson (1977) and Riley (1979), and the Nash refinements in alternative three-stage games considered by Cho and Kreps (1987) and Hellwig (1987). Although each of these approaches admits a pure strategy Nash equilibrium, their predictions regarding the equilibrium do not agree. Puelz and Snow (1994) provide a summary of the contractual allocations supported by these alternative theories. In a related vein, Rosenthal and Weiss (1984) exhibit a mixed strategy Nash equilibrium that always exists given a fixed number of insurers. However, as they acknowledge, if existing firms are playing their equilibrium mixed strategies, then an entrant can play a different mixed strategy and earn positive profit. Hence, their solution to the existence problem does not apply in the environment studied by Rothschild and Stiglitz who assume (p. 631) that “[t]he market is competitive in that there is free entry.”

4 Only one study of multidimensional screening has considered the implications for existence of Nash equilibrium. Bond and Crocker (1991) introduce a second signaling dimension by assuming correlation between risk type and the intensity of preference for a hazardous good whose consumption can be monitored by the insurer. Their approach fundamentally alters the insurance contracting environment by introducing a moral hazard component through the hidden effects of the hazardous good on the probability of incurring the loss. In contrast, multidimensional screening is introduced in our model through optimal partitioning of the signaling space without otherwise altering the adverse selection environment.

5 Point E indicates the endowed contingent wealth \((W, W-D)\), and the loci \(H^*E\) and \(L^*E\) represent the fair-odds lines for high- and low-risk applicants, respectively.
6 The allocation \((M, Y)\) is also associated with the element of the set of interim incentive efficient contracts that is most preferred by \(L\)-type individuals.

7 Rothschild and Stiglitz restrict insurers to each offer only a single contract, and their counterexample demonstrating that a pure strategy Nash equilibrium may not exist concentrates on the existence of a pooling contract that, if offered, would be preferred by both types of customers to \((H^*, A)\). While the existence of such a pooling contract is sufficient to demonstrate non-existence, it is not necessary. Indeed, this is the case in Figure 2, where there are no profitable “pooling” defections of the type considered by Rothschild and Stiglitz (since the pooling fair-odds line \(FE\) lies everywhere below the indifference curve \(\bar{V}^L\)), yet the pair \((M, Y)\) is preferred by both types to \((H^*, A)\). The most attractive defections always involve separating pairs, such as \((M, Y)\) in which there is a cross-subsidy from \(L\)-types to the \(H\)-types. This issue is addressed in Crocker and Snow (1985).

8 To see that the locus \(FA\) becomes flatter as \(\lambda\) increases, totally differentiate the binding incentive compatibility constraint (3) for \(H\)-types (as an equality) and the resource constraint (5), and then eliminate \(dW^H\) to obtain

\[
-dW^L_1 / dW^L_o = \frac{(1-p^H)U'(W^L_o) + (1-p^L)U'(W^H)(1-\lambda)/\lambda}{p^H U'(W^L_o) + p^L U'(W^H)(1-\lambda)/\lambda}
\]

as the slope (in absolute value) of the locus \(FA\). Partially differentiating this expression with respect to \(\lambda\) yields

\[
\frac{\partial}{\partial \lambda}(-dW^L_1 / dW^L_o) = \frac{-U'(W^L_o)/(1-\lambda)\frac{\partial}{\partial \lambda}}{(1-p^H U'(W^L_o) + p^L U'(W^H)(1-\lambda)/\lambda)^2} T
\]

where \(T = p^H U'(W^L_o) - p^L (1-p^H) U'(W^L_o)/(1-p^L)\). This derivative is negative since we have \(p^H > p^L (1-p^H)/(1-p^L)\) and \(U'(W^L_o) > U'(W^L_o)\) when \(L\)-types are under insured, implying that \(T\) is positive.

9 Note that we are not introducing additional sources of risk, as in Dionne and Gollier (1992) but, rather, decomposing a given risk into its component parts.

10 Specifically, for peril \(i\) the insurer knows \(D_i\), and that the probability of this particular peril occurring is either \(\theta_i^H\) or \(\theta_i^L\), depending on whether the insurance applicant is a type \(H\) or \(L\) individual, respectively. The knowledge of type, of course, is still privately known by applicants.

11 Fluet and Pannequin (1997) show that bundling of independent risks in a single umbrella policy is efficient when two risk types present an insurer with a bundle of risks. We find that bundling is not only efficient, but is also part of any Nash equilibrium with multidimensional screening.
Corner solutions, in which \( L \)-types have a zero deductible in one or more dimensions, do not arise as long as \( \bar{V}^H \) is less than the utility realized in the first-best pooling contract, \( U(W - [\lambda p^H + (1 - \lambda) p^L] D) \), and in every peril \( i \), both \( \theta_i^L \) and \( \theta_i^H \) are positive. Corner solutions that arise when \( \theta_i^L \) or \( \theta_i^H \) is zero in the case of some perils are discussed in Section V.

Specifically, we assume that the conditional odds ratios differ, that is, for all \( j \), \( \theta_i^L / \theta_i^H \neq \theta_j^L / \theta_j^H \) for every \( i \neq j \), and so by Theorem 1 multidimensional screening is efficient and interim efficient \( L \)-type contracts contain \( n \) different deductibles.

Increased efficiency does not necessarily improve the prospects for existence of Nash equilibrium, as demonstrated by Hoy (1982), and Crocker and Snow (1986). The latter show that risk categorization by immutable observable traits is efficient when the proportion of \( H \)-types varies across the categories, so that categorization is informative. However, the implications for existence are ambiguous. The prospects for existence of Nash equilibrium are enhanced for the category in which \( H \)-types predominate, but are worsened for the category in which \( L \)-types are more prevalent.

To obtain this constraint level of profit, solve \( \hat{W}_o^L = W - \alpha \) for \( \alpha \) and \( \hat{W}_1^L = W - \alpha - S \) for \( S \), and then substitute the results into equation (2) for \( \pi^L \).

When \( \theta_i^L = 0 < \theta_i^H \) for some signaling dimension \( i \), the full-and-fair insurance contracts are incentive compatible if \( U(W - p^H D) \) is greater than or equal to \((1 - p^H \theta_i^H)U(W - p^L D) + p^H \theta_i^H U(W - p^L D - S_i^L)\), or equivalently, we may write \( U(W - p^L D) - U(W - p^L D - S_i^L) \geq [U(W - p^L D) - U(W - p^H D)] p^H \theta_i^H \), which can always be accomplished by setting the deductible \( S_i^L \) high enough.

We note that whenever there is a signaling dimension \( i \) for which \( \theta_i^L = 0 \) and \( \theta_i^H > 0 \), the incentive compatibility constraint is not binding, as the Kuhn-Tucker condition then implies \( \xi = 0 \), and in that event there is no adverse selection externality. We also note that whenever \( \theta_i^H = 0 \leq \theta_i^L \), \( L \)-types do no signaling in the \( i \)th dimension since, from Theorem 1, the first-order condition

\[
\frac{\partial Z}{\partial \alpha^L} = \left(\frac{\partial V^L}{\partial \alpha^L}\right) - \xi \left(\frac{\partial V^{HL}}{\partial \alpha^L}\right) + \mu(1 - \lambda) = 0
\]

implies \( \left(\frac{\partial V^L}{\partial \alpha^L}\right) + \mu(1 - \lambda) < 0 \) if \( \xi \) is positive, so that

\[
U'(W_i^L) \geq -\frac{\partial V^L}{\partial \alpha^L} > \mu(1 - \lambda)
\]
and the Kuhn-Tucker condition then implies $S^{L}_{\ell} = 0$.

18 These equalities are obtained in the manner described in footnote 19 using equations (2) and (8) for $\pi^{L}$ with one dimensional and multidimensional screening, respectively.

19 With reference to Figure 5, where a solution to program (A) is depicted as point $G$, note that an increase in $\pi^{L}$ (holding $V^{H}$ constant) reduces $\hat{V}^{L}$, while increasing $V^{H}$ (holding $\pi^{L}$ constant) decreases $\hat{V}^{L}$. 
References


Figure 1
\{H^*, A\} Not Efficient: \( \lambda < \lambda^* \)

Figure 2
\{H^*, A\} Just Efficient: \( \lambda = \lambda^* \)

**Figure 3**
{H*, A} Efficient: $\lambda > \lambda^*$

Figure 4
Figure 5
Figure 6
Figure A-1