ENDOGENOUS ADVERSE SELECTION: EVIDENCE FROM U.S. CROP INSURANCE

JOB MARKET PAPER

JAYASHREE SIL

NOVEMBER 2005

ABSTRACT: Adverse selection tests in the tradition of Chiappori and Salanie (2000) examine correlation between contract choice and risk, given variables observed by the insurer. Positive correlation is evidence that inefficiency is due to unobservable variables (which includes the hidden information about buyer heterogeneity). If the goal is to eliminate the inefficiency, such a conclusion falls short. I argue that the analytical framework must endogenize the adverse selection and postulate a possible source. The empirical analysis requires hypothesis tests that are motivated by predictions from a model of Endogenous Adverse Selection (EAS). Such analysis requires the analyst to posit the *presently unobserved, but potentially observable* source of adverse selection. The analysis requires a data set that includes more variables than observed by insurers. I use the Agricultural Resources Management Survey cross-sectional sample for 1996 to test if the U.S. crop yield insurance market is characterized by a zero-subsidy EAS equilibrium, caused by a producer's commitment to forward price contracts. Using data for corn and soybean producers, I estimate that use of forward contracts is associated with a 6% increase in risk in one state: Indiana. This is substantial when compared to the benchmark increase in risk of 1% to 2% for differentially rated producers who use higher risk crop production practices. The policy recommendation is to design a new policy at a higher coverage level, at the current or lower premium, for the low-risk insured Indiana producers that do not have forward contracts. Low-risk insured producers' welfare increases above the second-best level, while the high risk insured producers are already at their first-best level of welfare.

JEL: D820, G220, G280, Q180

Endogenous Adverse Selection: Evidence from U.S. Crop Insurance

Jayashree Sil
Agricultural & Resource Economics, UC Berkeley

Policy prescriptions emerging from tests of adverse selection in insurance markets typically do not state precisely how policy makers or firms can eliminate the inefficiency arising from the information problem. Eliminating some of the information constraints typically moves the market from second-best toward first-best.[1] Most tests in the tradition of Chiappori & Salanie (2000) examine the correlation between contract choice and risk, having accounted for variables observed by the insurer. Positive correlation conditional on these observables is taken as evidence of adverse selection. Studies conclude simply that the resulting inefficiency is due to unobservable variables. If one's goal is to eliminate the inefficiency, such a conclusion falls short.[2]

The goal of the present analysis is to provide a policy recommendation which can eliminate the inefficiency arising from adverse selection in an insurance market. We argue that the analytical framework must 'endogenize' the adverse selection and postulate a possible source of the hidden information problem. The source or characteristic must be *presently unobserved* by the insurer, but *potentially observable*. For our example, we consider the crop yield insurance market in which the buyer faces price as well as yield risk, can obtain a price insurance contract, and

[1] Choice (and associated welfare) in a second-best environment is said to be, at best, constrained-optimal. In the context here, the constraints arise due to hidden information about an insured's risk.

In a credit market with adverse selection and moral hazard, Vercammen (2002) shows how welfare is higher with hidden information. de Garidel-Thoron (2005) shows how hidden information (that is, presence of unobserved heterogeneity) improves welfare in a dynamic model in which an insurer does not know the early time period accident history or policy bought by the insured. Polborn, Hoy & Sadanand (2005) show that *regulatory adverse selection* (from information barriers imposed by disclosure laws) can improve welfare in life insurance markets.

[2] What exactly is the problem induced by adverse selection? Two basic cases can be considered. Rothschild & Stiglitz (1976) showed that a single policy (as would occur in a *pooling equilibrium*) would not be supported in a market with insurers who face unobserved heterogeneity in an insured's risk of accident. Since no one gets insurance, this is the problem of a lack of a market. They then showed that competitive insurers facing such information constraints would sell insurance if they could offer a menu of contracts, in a *separating equilibrium*. High and low risk buyers would select the policy that was optimal given their risk of accident. The low-risk would purchase less coverage (and at a higher price) than they would like. Insurers however are just fine with zero expected profit. Now, the problem is that some are under-insured.

The frequently-cited figure that some 40 million or so individuals in the United States do not have health insurance would be a stylized fact that is relevant for a market with *partial pooling*. A partial pooling equilibrium is considered by Demeza & Webb (2001) in which all low risk buyers are insured and some fraction of high risk potential buyers are excluded from the market.

Siegelman (2004) finds that while adverse selection frequently is cited as an argument for environmental and industrial legislation in the U.S., and for compensation awards in legal cases, evidence of adverse selection from econometric studies is mixed. Just, Calvin & Quiggin (1999) found that the U.S. crop insurance program was plagued by adverse selection that imposed a welfare cost on low-risk producers. Policies have subsequently been reformed in an attempt to remedy that situation, found to prevail in the 1990s.

uses a quasi-fixed factor of production[3]. Effort choice is a quasi-fixed factor choice. In this context, the relevant buyer is a crop producer who faces net revenue risk, where unit revenue is price times yield. Here, we posit the source of adverse selection to be a contract on a second source of risk (specifically, price risk), when the efficiency question is about the insurance market for yield risk. Given a broader notion of non-exclusive contracts, our test also serves as a test of whether non-exclusive contracts cause inefficiency in crop yield insurance.[4] A forward contract stipulated with a fixed (or flat) price is equivalent to a price insurance contract.[5,6]

We conduct a test of a zero-subsidy Endogenous Adverse Selection (EAS) separating equilibrium for the U.S. crop yield insurance market. While the conclusions of the present essay may reasonably apply to other types of insurance markets, we concentrate on the EAS model and associated test when the insurance is for crop yield risk, the seller is the crop yield insurer, and the buyer is a U.S. corn and soybean producer. Following Chiu & Karni (1998), an endogenous adverse selection equilibrium requires by definition that the analyst posit the source of the adverse selection. Such a source in the present paper is taken to be the producer's use of a forward contract.[7] A producer that has a forward contract commitment will choose an effort level that differs from one who does not have this commitment. The resulting heterogeneity in effort levels determines risk heterogeneity. In this sense, adverse selection is *endogenous*.

Evidence from this test of adverse selection delivers, we argue, a sharp policy recommendation: (a) require disclosure of forward contract use, and (b) rate insureds with and without a forward contract as distinct risk classes. In contrast, absence of evidence from the test indicates non-exclusive contracts are not a source of crop insurance inefficiency.

[3] The term quasi-fixed refers to the fact that we consider an optimal decision about this factor, but that the decision is made prior to the resolution of uncertainty over product price.

[4] In the standard case, insurance is for one source of risk and *non-exclusive* contract means purchase of multiple, additive levels of coverage bought from several firms for that one type of insurance. Using a broader notion of non-exclusive contracts, we may consider the price insurance contract to be a supplementary revenue insurance tool, for a multiplicative source of risk.

[5] Pricing under a forward contract, in general, can be a function of quality premiums and discounts, or a basis, for example.

[6] Forward contracts are examples of marketing contracts. These are to be distinguished from production contracts, as considered by Hueth & Ligon (2001) and Ligon (2004), for example.

[7] In our analysis, we assume *use of forward contract* means that the producer makes the delivery stipulated in the contractual commitment.

We have, indeed, evidence that the U.S. corn and soybean yield insurance market exhibits Endogenous Adverse Selection. Insured Indiana producers with a forward contract are found to exhibit a 6 % risk increase. That is, the estimated probability of an indemnity payment is 6% more for these producers than it is for producers with no forward contract. This is substantial, as it exceeds the benchmark risk increase associated with producers who employ high-risk crop rotation practices. Our policy recommendation is to adjust the rate, increasing coverage and lowering price, for the low risk producers (here, insured Indiana producers without a forward contract). Given our model, the expected utility of these low risk producers would rise from the second-best level. Rates for producers in other states need not be adjusted because there is no evidence that forward contract use is a source of adverse selection in those states.

Insurers, including those in the U.S. crop insurance market, examine insureds' loss histories and characteristics and use such information to place insureds into risk classes. Given the classes they identify, insurers' offer of a menu of contracts implies that some degree of adverse selection remains within each risk class. Within a yield-risk class, crop insurance policies are offered at a base price specific to the yield-risk class and to the particular level of coverage. Hence, the offer of a menu of coverage levels may be optimal in a market with adverse selection.[8]

Our empirical analysis considers corn and soybean producers in Illinois, Indianai, and Iowa in 1996. These states are among the largest producers of such crops in the United States, producing a little over 40 % of the U.S. total in each crop. Total U.S. corn and soybean production accounted for 40% and 52%, respectively, of world production in 1996. We use the Agricultural and Resources Management Survey (ARMS) data set, which is a nationally representative cross-sectional sample of individual farm producers' production, cost, income, and assets. We use the 1996 version of ARMS because this was the year just prior to the introduction of many new insurance products. Using the 1996 sample enables us to more accurately conduct a test of the model of Endogenous Adverse Selection. In

[8] A society may be concerned, in addition, with equity (or redistributive) effects of risk classification, as in Hoy (2005).

1996, the primary insurance tool available on the market was crop yield insurance. Because alternatives were largely unavailable, it is reasonable to assume that a grain producer's key decisions were whether to buy yield insurance and whether the coverage should be high or low.

The test we propose requires a different type of data set. It is not enough for us to know only the information insurers knew. Tests of correlation between risk and coverage, conditional on observables, typically have been based on data containing all the information known to the insurer and only that information. Two examples are studies by Chiappori & Salanie (2000) and Finkelstein & Poterba (2004). Our model, test, and data set requires in addition a variable that the insurer presently does not observe but potentially could observe. Forward contract use is such a variable.[9]

Some recent studies, it is true, have attempted in other ways to go beyond tests in the tradition of Chiappori & Salanie (2000). Finkelstein & Poterba (2004) find evidence of adverse selection in the U.K. voluntary annuity market by expanding the set of relevant contract choice variables beyond coverage. Makki & Somwaru (2001) conduct tests of adverse selection in selected U.S. crop insurance programs, contrasting policies that set premiums based on individual yield with those employing regional yields. Yet these studies do not go further than the conclusion that unobservables are the source of the adverse selection problem. An exception is the study by Cardon & Hendel (2001) which, in a structural model framework, checks for variables that might determine the estimated unobservable error variance.

For our empirical analysis, we isolate two key predictions from a zero-subsidy EAS equilibrium informing the research questions addressed by our two hypothesis tests. Our modification of traditional adverse selection tests relies on distinguishing between the role of unobserved and observed variables and on framing one of the tests as a causal test. We begin by providing a selected review of the litera-

[9] The practical import of our methodology is that it can be done ex-ante or prospectively. Indeed, insurers make adjustments ex-post via experience rating. With experience rating, an insurer periodically updates the premium of an individual insured based on historical loss experience. An EAS test is an ex-ante exercise which enables an insurer to establish a finer risk class prior to contract offer. And, in contrast to risk adjustment methods actuaries use for rate setting, our method is based on a model of economic behavior.

ture on the theory and evidence of adverse selection and offer a summary of what we argue are the gaps in the present literature. Second, we discuss institutional details about a crop-yield insurance market in which the seller is the government crop insurer and the buyers are U.S. corn and soybean producers. Third, we give a brief overview of the Endogenous Adverse Selection (EAS) analytical framework, which delivers predictions that motivate the two hypothesis tests.[10] Fourth, we present our empirical strategy, including a discussion of the role of observed and unobserved variables. Finally, in a brief conclusion which follows the presentation of results, we suggest how our proposed methodology can inform legislative proposals for insurance market regulation. The appendix contains further details about notation used for the EAS analytical model, the ARMS data, constructed biological and marketing risk variables, full model results, and the method used to calculate a probability benchmark for interpreting results of one of the hypothesis tests.

1 Literature: Theory and Evidence

Two central predictions emerge from early models of insurance markets. Unobserved buyer heterogeneity causes adverse selection and reduces welfare. Then, welfare will increase if the buyer discloses information about his type. Second, an equilibrium with unobserved effort choice by the buyer, which is the source of the moral hazard problem, requires the insurer to enforce an exclusive contract. Real-world insurance markets, however, are characterized by limited disclosure and disclosure bans (such as recent state level regulation that limits credit-score-based insurance underwriting), and by non-exclusive contracts (such as in life insurance markets).

[10] We have provided the more detailed discussion of crop insurance and of the EAS analytical framework for the interested reader. Otherwise, it is possible to skip these sections and proceed to the section on empirical strategy.

1.1 Theory

Standard adverse selection happens when the insurer cannot observe the type of a particular insured. The insurer does not know if the particular insured is a high or low risk type, but he does know what the two possible loss distributions are. The loss distributions are exogenous. The insured's type is the unobservable characteristic. In this situation, the insurer offers high-price-high-coverage and low-price-low-coverage contracts. The high risk choose the high-price-high-coverage contracts. Rothschild & Stiglitz (1976) found that the offer of such a menu of contracts is optimal when unobserved heterogeneity results in residual risk within a risk class, with one contract designed for low risk types and another for high risk types. Then, if the source of risk heterogeneity becomes observable to the seller, the aggregate welfare increases.

Standard ex-ante moral hazard happens when the insurer cannot contract on effort (that is, cannot observe the insured's effort choice). The insurance makes the insured reduce effort, increasing loss probability. The loss distribution is endogenous. In this situation, the insurer gives partial coverage, making the insured bear some risk. Effort level is less than if effort were contractible (observable by the insurer). Arnott & Stiglitz (1988) found that partial insurance, that is, less than full insurance, tightens incentives and is therefore optimal when there is an unobserved action that results in reduced incentives under moral hazard. Furthermore, since the insured wants to buy more insurance in the equilibrium, an exclusive contract policy, which allows only one insurance policy for the insured risk, would enable such risk sharing between the insured and insurer. Otherwise, the insured could buy several policies, obtain full insurance, and have weak incentives, which in turn would cause market failure.

1.2 Gap in Analytical Framework

We suggest that an analytical framework of an Endogenous Adverse Seleciton equilibrium is a requisite first step towards a modification of tests in the tradition of Chiappori-Salanie. Testing for evidence of an Endogenous Adverse Selection

(EAS) equilibrium requires the analyst to posit the source of the adverse selection. Necessarily, that source must be *presently unobserved* by the insurer. Then, if evidence of this information problem is found to exist, it is useful if this source is *potentially observable*. The policy recommendation, given such evidence, is to require that the insured disclose information regarding this source. In this way, when disclosure is mandatory, the insurer can rate the insureds as members of distinct risk classes that are defined by the, now observed, necessary information.

In our model of the yield insurance market, unobservable heterogeneity in price risk management practice causes heterogeneity in effort choice. This heterogeneity effect then results in heterogeneous yield risk. It is in this way that adverse selection due to heterogeneity of buyer risk is considered *endogenous*. Following Chiu & Karni (1998) we label this structure of information imperfection as Endogenous Adverse Selection. Chiu & Karni (1998) consider conditions which support an equilibrium with no unemployment insurance (either public or private). They consider the possibility of a zero-subsidy separating equilibrium. We consider a framework that, in general, allows for a zero-subsidy or cross-subsidized equilibrium in which insurance can be provided by many private firms or a single government provider.[11] However, under a zero-subsidy EAS equilibrium, there is an unambiguous welfare gain from removing the information problem.

Since we focus on regulatory implications, we abstract away from complications that arise due to non-existence of equilibrium. For this reason, we consider only exclusive quantity contracts, as defined by Arnott & Stiglitz (1991). And, we consider the set of information-constrained contracts that are solutions to the optimal subsidy problem presented by Rothschild & Stiglitz (1976) and Dionne, Doherty & Fombaron (2000). Further, we assume that firms in the market behave with Miyazaki-Wilson foresight (Miyazaki 1977). In this way we can ensure that either a zero subsidy, or a cross-subsidized, separating equilibrium can exist in a competitive market.

[11] For a complete discussion of an Endogenous Adverse Selection equilibrium, see Chapter 2 of my disseration *Risk in US Agriculture: Crop Insurance, Forward Contracts, Adverse Selection and Moral Hazard*. There, a primary goal is to derive a necessary condition under which information disclosure *bans* may be optimal.

1.3 Evidence on Insurance Market Efficiency

Chiappori & Salanie (2000) find no adverse selection in the French auto insurance market. They conduct their tests separately for the class of beginning and of senior drivers. This accounts for experience rating of policies by insurers and enables the researchers to search for adverse selection within the most finely defined homogeneous risk class. They estimate the correlation coefficient between error terms (unobservables) in probit equations for contract choice and accident occurence and conduct a non-parametric test of independence between accident occurence and contract choice. The authors conclude that the French auto insurance market does not face inefficiencies due to adverse selection. This is consistent with the view of French auto insurers who believe that they are able to classify insureds by the observable information available to them. For risk classification, they use driver's sex, age, extent of city driving and make, age and size of car.

Finkelstein & Poterba (2004) suggest that adverse selection may be found along some dimensions of a contract and not others. In their case, annuity contracts are defined by payment amount, temporal payment profile and guarantee of payment to beneficiary in the event of early death. At the time a contract is signed, an insurer does not have any information on whether the particular insured will be long-lived or short-lived. A high risk insured is one who lives too long. They find that unobservables cannot explain the positive relation between contract choice and risk, if contract is defined by payment amount. However, if choice is defined by payment profile or guarantee, then choice and risk are related and unobservables explain this relationship. They estimate a duration model for mortality risk conditional on contract choice characteristics. These researchers conclude that the extent of the adverse selection is substantial since the difference in mortality risk between high and low risk groups is bigger than a benchmark. Their benchmark is the male-female differential in mortality risk.

Makki & Somwaru (2001) find evidence of adverse selection in U.S. crop insurance. They estimate a three-stage-least-squares model for each of the four main types of insurance products. Premium and coverage level are specified as simultaneously determined, conditional on variables observable to the analyst. These variables are probability of loss, yield span (risk class as defined by historical average yield), irrigation use, land tenure arrangement and income. Risk is found to be positively related to coverage. Non-parametric tests of independence between coverage choice and accident occurrence are rejected for all insurance products except for the Group Risk Plan. This latter plan bases insurance indemnity payments upon a regional yield trigger. The researchers conclude that it is the absence of an accurately measured regional trigger that causes the adverse selection for the other policies that are based on individual yield triggers.

1.4 Gap in Empirical Method

Our EAS framework implies we have a model with both adverse selection and moral hazard. Our test is regarding the existence of information problems due to adverse selection. Later in our discussion of model specification, we shall refer to studies of moral hazard in crop insurance by Horowitz & Lichtenberg (1993) and Smith & Goodwin (1996). For now, we note that the source of adverse selection is forward contract use. First, under the EAS separating equilibrium considered here, forward contract use causes the heterogeneity in riskiness. This gives the first hypothesis. Second, given this hypothesis, forward contract use determines choice of coverage. This is the second hypothesis. If both tests reveal evidence in favor of these predictions, the insurer can eliminate this adverse selection problem by requiring disclosure of forward contract usage, and thereby offer each type of insured a contract at a price appropriate for his type, increasing the low risk insured's welfare.

2 Yield Insurance Buyers and Sellers

While the conclusions of the present essay may reasonably apply to other types of insurance markets, we discuss the model of endogenous adverse selection and associated test when the insurance is for crop yield risk, the seller is the crop yield insurer, and the buyer is a U.S. corn and soybean producer.

2.1 Buyers: U.S. Corn and Soybean Producers

For a typical producer, located in a state in the midwest of the U.S. , important decision points occur around planting time, during the growing season and at harvest. Planting time for corn and soybeans is made in April and May for most states and harvest is in October for most states (USDA 1997). The growing season is during the summer months. Figure 1 summarizes key decisions made at these times.

The state of the world is revealed to both the individual farmer and the insurer at

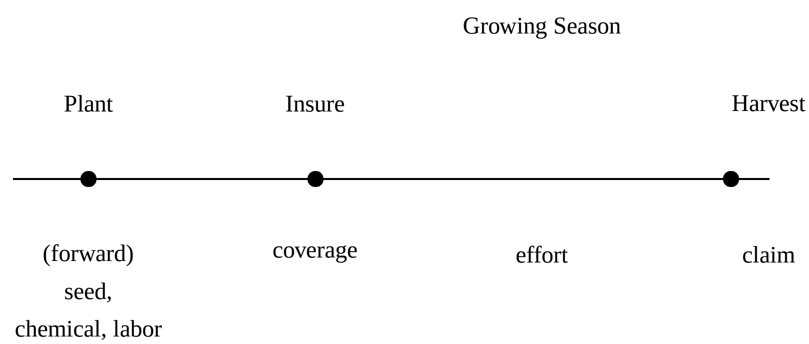

Figure 1: Production and Risk Management Decisions

harvest time, when yield is known and when a claim may be made. A long growing season for wheat permits farmers to wait until well after planting time, no sooner than 180 days before harvest, to make forward contracting decisions. Townsend & Brorsen (2000) suggest, however, that this delay may or may not characterize forward contracting decisions for corn and soybeans. Hence, we conjecture, for our studies of corn and soybean production, that forward contract decisions might be made at the time crop acreage decisions are made around, or just after, planting. Following Smith & Goodwin (1996), we conjecture that farmers wait until the region-specific crop insurance enrollment deadline (which is near the end of the peak planting season), after planting is completed, to decide whether to buy insurance. This enables them to acquire as much information as possible on the possible state of the world which will not be completely revealed until harvest. When a farmer decides to buy crop insurance, he chooses some coverage level.

At planting time, seed, fertilizer and pesticide applications are made using planting time (roughly, first and second quarter) labor. Fertilizer is used mainly at planting time and mainly to enhance yields for corn, since soybeans fix soil nitrogen and thereby replenish soil nutrients (USDA 2003). After this time, during the growing season, a producer may utilize more inputs, such as labor to inspect for pest infestation, post-emergence (after crops are visible on the soil surface) chemicals, or an additional loan for input purchases. All of these inputs utilized during the growing season, after the insurance contract is signed, constitute effort. It is the effect of insurance coverage on incentives for effort that are considered in analytical models with moral hazard. As is usual, in the model considered here, insurance coverage reduces incentives to use effort whether a producer uses a forward contract or not. And, the use of a forward contract determines heterogeneity in effort level. As all models are abstractions from and approximations of the real-world, we consider the effort decision to refer to an input decision made roughly after or contemporaneously with the insurance purchase decision.

Corn and Soybean Risk Management

A producer can use numerous methods to manage price and production (yield) risk. They diversify their crop mix, spread sales across the production cycle, use forward contracts, use derivative hedging instruments, and buy crop insurance. About 80% of producers in the 1996 ARMS data sample reported that they rely on availability of liquid assets (including cash), spreading sales over the year, and having an open line of credit to manage risk. And, no more that 30% to 40% report using futures and options, or even diversification. Insured producers have a greater tendency to spread sales, to buy inputs and to sell outputs using forward contracts and to hedge price risk. The 1996 version of the survey asked specific questions about use of 8 key risk management tools because the 1995 Farm Bill took effect in 1996. We know from responses to these questions that no more than 20% to 40% changed their use of any of the 8 tools in 1996 and no more than 4% to 10% of producers in the data sample changed risk management behavior *due to* the Farm Bill. We can assume then that our study of risk management behavior using 1996 data is not contaminated by Farm Bill policy-induced effects.

For 1996, it is reasonable to assume that an insured grain producer's key decision was level of coverage (high or low) for yield insurance. Insureds chose among basic (low) and buy-up (high) coverage. The yield trigger for basic coverage was 50% of an individual's historical mean yield, and for buy-up it usually was 65%. The 50% trigger is more stringent, so the implied coverage level is less.[12]

A forward contract is a simple, sometimes informal, contract between a farm producer and either a grain elevator, brokerage firm, cooperative or processor, that stipulates the time, quantity, quality and price for crop delivery. In our analytical model we assume a producer cannot default on this contract. If harvest falls short, the producer buys the difference on the spot market and makes the delivery under conditions stipulated in the forward contract. Among the insured producers in our sample of Illinois, Indiana and Iowa farms, about 40% use a forward contract to sell their crop.

2.2 Sellers: U.S. Crop Insurance

The U.S. Risk Management Agency (RMA) and Federal Crop Insurance Corporation (FCIC) together administer the crop insurance program and re-insurance. Rate are set by the RMA and private insurers are responsible for actual sales to producers, and for bearing part of the re-insurance burden. 1996 was the last year before alternatives to yield insurance were introduced by the RMA. Therefore, the

[12] Since we focus on corn and soybean operations we make the assumption that an insured farm with both crops likely insures both crops. Sherrick, Barry, Ellinger & Schnitskey (2004) find in their study of Illinois, Iowa and Indiana farms that 70% of farms buy coverage for both crops, 7% insure only one crop, while 15% do not insure. We require this fact since the survey question on crop insurance asks only whether a producer bought some insurance and does not ask the producer to specify the type of crop.

relevant set of risk management choices in that year included the decision to insure with yield insurance and the choice of high or low coverage level. It is therefore a year that is suitable for the empirical analysis we conduct here. It is important to recognize however that despite the recent increase in popularity of the newer alternatives to yield insurance, the exposure of insurers to yield insurance liability remains substantial. And, for all programs except for the fully subsidized yield insurance, loss ratios are high.

Just et al. (1999) found that crop insurance program suffered welfare loss due to adverse selection, whereby low risk insureds were priced out of the market. Recent reforms under the 2002 Farm Bill have focused on trying to overcome the welfare loss due to adverse selection, by offering subsidies that serve to increase coverage at given premiums (Babcock, Hart & Hayes 2004).

Between 1995 and 2003, yield insurance (Actual Production History, APH) use has fallen and revenue insurance (Crop Revenue Coverage, CRC, and Revenue Assurance, RA) use has increased. Fully subsidized APH insurance accounted for about 30m insured acres each for corn and soybeans in 1995. This dropped to 7m for corn and 8m for soybeans by 2003. Partially subsidized yield insurance acreage has returned to the level of the 1990s for soybeans, about 20m acres. In 2003, there was just under 14m acres for corn and 10m acres for soybeans under the CRC program, and 26m acres for corn and 16m acres for soybeans under the RA program.

Actuarially fair insurance means premiums received equal expected indemnities. That is, receipts must cover expenditures. Loss ratio, typically reported by private and public insurers, is a measure of the efficiency of an insurance policy. Loss ratio equals the ratio of indemnity paid out to premiums collected for a given year. In other words, a loss ratio equal to 1 implies zero profits for an insurance firm. A ratio less than 1 implies the insurer makes positive profits, while a ratio bigger than one implies losses. Figure 2 shows that loss ratios are well in excess of 1 for CRC and RA soybean insurance, and have risen in 2003 above 1996 levels for partially subsidized APH yield soybean insurance. Based on this criterion at the very least we may conclude that lack of efficiency is a problem for soybean

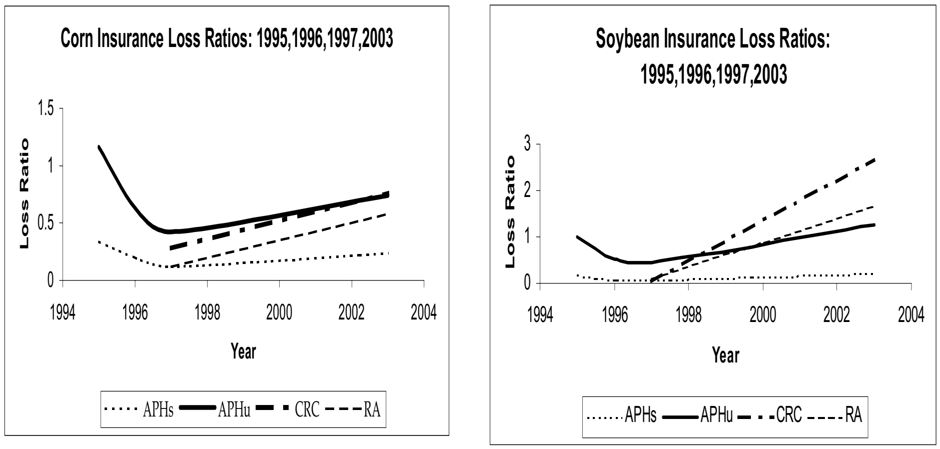

Figure 2: Corn and Soybean Insurance Loss Ratios

insurance. The RMA reports loss ratios that count subsidized premiums as bona fide premiums, so that in fact the reported loss ratio *under-estimates* insurer firm losses in the usual sense. While reported loss ratios in 1996 are less than 1, they are biased downward since they do not account for the premium subsidy given to the insured. It may be useful to emphasize at this point that the inefficiency problem due to adverse selection can exist *even if* an insurer earns zero profits in the usual sense. Second, the notion of *cross-subsidization* or *zero-subsidy* equilibrium we consider refers to the possible incorrect rating for policies offered at different coverage levels within the yield insurance program. [13]

Premiums for yield insurance are set by the RMA at the county level and for the yield risk class associated with the individual farm's historical average yield. Further adjustments are made for production practices, such as crop rotation, irrigation and tillage methods. These adjustments require that a practice factor be applied to the base premium rate to adjust it upward for a user of a high risk practice and downward for a user of a low risk practice. For soybeans a higher risk practice

[13] Recent legislation has stipulated that the Actual Production History yield insurance may operate at loss ratio up to 1.075. That is, the premium can be as low as 93% of expected indemnities, less than actuarially fair. Zero profits are not a requirement for this policy. See (Babcock et al. 2004).

which significantly differentiates producers' yields and loss experiences is the crop rotation practice in which a new crop is planted following planting of another crop which had reached maturity (and thereby depleted the soil of nutrients) within the same production year. For corn, a lower risk practice is the use of irrigated farming which supplements natural rainfall with water available in underground wells and in reservoirs.

3 Theory & Predictions (EAS Equilibrium)

An Endogenous Adverse Selection (EAS) equilibrium can involve no subsidy or can be cross-subsidized. Under a zero-subsidy equilibrium, there is an unambiguous welfare gain from removing the adverse selection. The zero-subsidy endogenous adverse selection separating equilibrium is the contract and subsidy policy that solves the optimal subsidy problem presented by Rothschild & Stiglitz (1976), given that each type of insured chooses an effort level that maximizes his individual expected utility.

There are two types k of risk-averse producers who have identical preferences given by utility function U. Type F buys insurance and has a forward contract. Type N buys insurance and has no forward contract. Each faces high p_H and low p_L price and high y_H and low y_L yield outcomes. Price is p_L with probability π and p_H with probability $1-\pi$. Yield is y_L with probability probability $q_k = q(e_k)$, which is a decreasing function q of type k's effort choice e_k, so $q' < 0$. Definitions for all notation for the model of Endogenous Adverse Selection, presented in Chapter 2 of the dissertation, is reproduced in the appendix here.

A type F producer sells forward $y_m > 0$ of his yield at price p_m per unit. A type N producer has $y_m = 0$. State-contingent wealth for a type F producer includes payment $f_L^F > 0$ in low price states and cost $f_H^F < 0$ in high price states. For type N, $f_L^N = f_H^N = 0$.

The insurer offers contracts distinguished by coverage net of premium α that increases income in the low yield state and premium β that decreases income in high yield state.. Optimal effort $e_k(\alpha, \beta)$ is a function of insurance contract (α, β). The insurer sets contract terms so that the premium is actuarily fair: expected profits are zero. The zero-profit locus (ZPL) for type k is

$$(1) \qquad (1 - \mathrm{q_k})\beta_\mathrm{k} - \mathrm{q_k}\alpha_\mathrm{k} = 0.$$

The ZPL then gives all (α, β) pairs that earn zero expected profit. The ZPL is defined in terms of $q_k = q(e_k)$ and thus assumes that the insured chooses optimal effort.

For each level j of yield, I_j^k gives the yield-specific expected utility for type k, and is equal to

$$\pi U(W_{Lj}^k) + (1 - \pi)U(W_{Hj}^k).$$

For example, for k=L and j=N, $I_L^N = \pi U(W_{LL}^N) + (1 - \pi)U(W_{HL}^N)$, where state-contingent wealth expressions are $W_{LL}^N = R_{LL} + \alpha_N$ and $W_{HL}^N = R_{HL} + \alpha_N$. We call I_j^k a derived utility since it is price-probability weighted utility. It is heterogeneity in derived utility I_j^k that determines the difference in risk aversion between the two types. And, again, this difference is driven entirely by difference in forward net gain $f_i^k, \quad i \in \{L, H\}$. State-contingent wealth and derived utility for each price and yield combination that is considered in Chapter 2 of the dissertation is reproduced in the Appendix included here.

Producers (the insureds) choose effort to maximize expected utility net of effort cost. Expected utility for type k insureds is

$$(2) \qquad \mathrm{EU^k} = \mathrm{q_k I_L^k} + (1 - \mathrm{q_k})\mathrm{I_H^k} - \mathrm{e_k}.$$

Following Jullien, Salanie & Salanie (1999), we use the first order condition for effort e_k, from (2) to establish an ordering for optimal effort across types. Optimal effort for type F is less than that of type N $e_F < e_N$, given any insurance contract. This leads to the first prediction,

$$\text{(EAS-1)} \qquad q_F > q_N,$$

that loss probability for type F exceeds loss probability for type N, given any insurance contract.[14]

The value function associated with equation (2) is the effort-optimized expected utility V. Together with EAS-1, additional conditions establish that $V_F' > V_N'$, so that type F is the high risk type. The slope of effort-optimized expected utility, that gives an expected marginal utility, is higher for type F, given any contract, so he is willing to pay more per unit increase in coverage. This is easy to see in the typical diagram in (β, α) space given below.

With heterogenous yield probability, and type unobserved by the insurer, there is adverse selection. In this case, insureds self-select (find it optimal to chose) the contract designed for their type (α_k, β_k). The insurer does not observe effort e_k chosen by insured type k. This induces moral hazard. The insurer designs the contract assuming that the insured chooses the optimal level of effort given contract pair (α, β).

The equilibrium contract menu $[(\alpha_F, \beta_F), (\alpha_N, \beta_N)]$ is the optimal contract pair that solves the problem:

$$
\text{(EAS)} \qquad \begin{aligned} \max_{\alpha,\, s} \quad & V_N(\alpha, s) \\ \text{s.t.} \quad & V_F(\alpha_F, s) \geq V_F(\alpha, s) \\ & s \geq 0 \\ & \eta \geq 0, \quad \delta \geq 0. \end{aligned}
$$

Endogenous variable α gives the contract that maximizes effort-optimized expected utility for type N. Given that the contract (α, β) satisfies the zero profit constraint (discussed above), β is determined once α is determined. The self-selection constraint says type F prefers the contract that was designed for his type (α_F, β_F) over any other contract α. This problem can have a zero-subsidy $(s = 0)$ or cross-

[14] $q_k = q(e_k) = q(e(\alpha_k, \beta_k))$ is an ex-ante probability of loss, known by the insurer since it is a function of obervable information that includes the contract (α, β). For one-period or static adverse selection equilibrium to hold period after period, the ex-post (or, realized) probability of loss (equivalently, indemnity payment) must equal this ex-ante probability.

subsidized ($s > 0$) solution. δ is the Lagrange multipler for the self-selection constraint and η is the multiplier for the non-negativity constraint. Further details are contained in Chapter 2 of the dissertation.[15]

For a zero-subsidy endogenous adverse selection equilibrium, when $s = 0$, $\eta \geq 0$ and $\delta \geq 0$, the relative proportion γ of high risk type F must be sufficiently large,

$$\gamma > \gamma^*,$$

so that the proportion exceeds some lower bound γ^*. The optimal contract that solves the optimization problem in this case gives the second prediction,

$$\text{(EAS-2)} \qquad \alpha_{\mathrm{F}} > \alpha_{\mathrm{N}} \quad \text{and} \quad \beta_{\mathrm{F}} > \beta_{\mathrm{N}}.$$

This states that the high risk type (which we have determined to be type F) buys more coverage. And, given a zero profit contract (one that is on the ZPL), he pays more for it. It may help to warn the reader that the empirical test will be framed in terms of a parameter α. That will have a separate definition.

An example of a zero-subsidy endogenous adverse selection separating equilibrium is shown in Figure 3 as the contract pair (F,N).[16] At point F, the marginal rate of substitution equals the relative cost at the margin for the zero-profit insurer. The effort optimized expected utility V_F for type F is tangent to the zero-profit locus ZPL_F for type F. This is the outcome we refer to as first best.[17] A similar situation prevails for type N, for the outcome at N'. When the market faces a hidden information problem, there is adverse selection, since now the insurer offers a menu in the equilibrium and type F chooses the policy at point F. Type N chooses the policy at point N: the term 'adverse selection' refers to the fact that the lower risk type N

[15] Although not shown explicitly here, contracts satisfy pre-and post-subsidy zero-profit constraints. Under a cross-subsidized contract, the type N policy is taxed and type F policy is subsidized. Balanced-budget tax t (t = $\gamma \cdot$ s) ensures aggregate expected profits are zero. $\gamma = \frac{\lambda_F}{\lambda_N}$, where λ_k is the fraction of type k in the population of insureds.

[16] The shape of V as shown requires that the risk-aversion effect of increased coverage dominates the effort-incentive effect. This restriction on the shape of V and the shape of ZPL both depend also on the curvature of the probability function q. These are discussed in Arnott & Stiglitz (1988). EAS-1 and risk ordering implies $e_F(\alpha_F, \beta_F)$ at the optimal contract (α_F, β_F) is lower for type F. In order to avoid problems of existence in a competitive market, we assume exclusive quantity contracts (Arnott & Stiglitz 1991) and Miyazaki-Wilson foresight (Miyazaki 1977). Further details are in Chapter 2 of the dissertation.

[17] Some may object to out use of the term 'first best' when we have acknowledged that this is an optimum under hidden action.

is choosing to buy the menu option with less coverage offered at a lower premium, at point N. Type N's first best choice is N'. Utility insreases as coverage α is higher and premium β is lower.

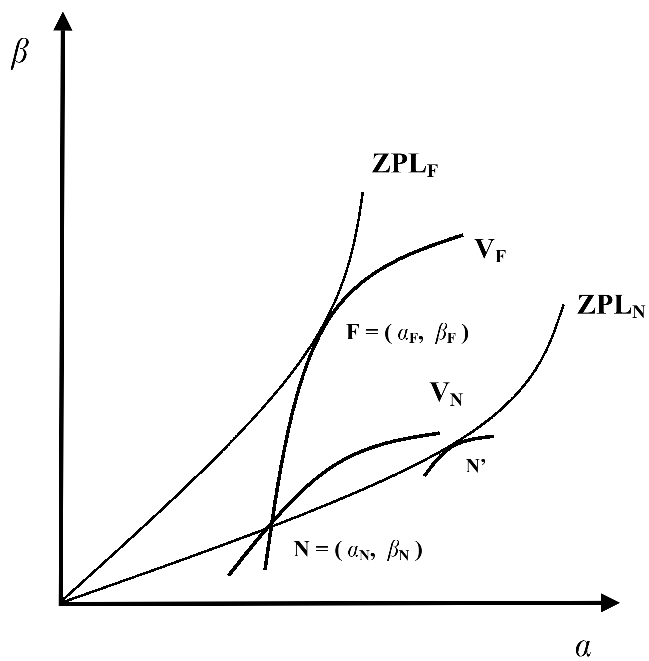

Figure 3: Zero-Subsidy EAS Separating Equilibrium

4 Predictions and Hypothesis Tests

A model of an insurance market with an endogenous adverse selection separating equilibrium delivers two central predictions, EAS-1 and EAS-2. These are the basis for the two hypothesis tests conducted here. If tests conclude there is evidence of such an equilibrium, there not only is an unambiguous welfare gain from its removal, but in addition, the analytical framework and associated tests reveal exactly how to achieve that gain.

Here, the source of adverse selection is forward contract usage. First, under the Endogenous Adverse Selection separating equilibrium considered here, forward contract use causes the heterogeneity in risk of low yield (equivalently, risk of loss). Prediction EAS-1 gives the first hypothesis. And, second, given hypothesis EAS-1 and prediction EAS-2, forward contract determines choice of coverage. This is the second hypothesis. If both tests reveal evidence in favor of these predictions, the insurer can eliminate this adverse selection problem by requiring disclosure of forward contract use, and thereby offer each type of insured a contract at a price appropriate for his type, increasing the low risk insured's welfare. In Figure 3, an example of such a policy with disclosure is given as contract at the point labeled F for type F and contract at the point N' for type N. In this way, there is an unambiguous welfare gain due to higher welfare for low risk type N, from removal of adverse selection, since the contract at point N' is at a higher level of utility than the one at point N.

5 Empirical Strategy

The analytical model delivers two central predictions, EAS-1 and EAS-2. Since the need to modify traditional tests of adverse selection is motivated by the need to make the policy recommendation more precise, we begin by explaining the role of unobserved and observed variables in past tests and in our own. We then describe the variables needed, data sources (with some details relegated to the appendix) and methods. We report only selected results in the main text and refer the reader to the appendix for a complete set of model results.

5.1 Observed and Unobserved Variables

Empirically, the distinguishing feature of the tests performed is due to the role of observed and unobserved variables. There are three important classes. There are variables observed by the analyst, but not observed by the insurer. A variable that is *presently unobserved but potentially observable* by the insurer comes from this set. There are variables observed by the insurer and by the analyst: these are the *observables* in our analysis. And, there are variables not observed by the insurer or analyst; these are the *unobservables* in our analysis. We assume that any variables observed by the insurer and not the analyst play, if anything, a minor role.

In order to understand the distinction between past tests and the present one, we show how our tests and the traditional test are special cases of the following two-equation system. The outcome variable is y, covariates are denoted by the x (including a constant), t is dummy (or, indicator) for the variable whose effect we wish to measure.[18] For the EAS-1 hypothesis test, t is a bona-fide causal variable. t is an indicator for the 'treatment'. If participation in the treatment is endogenous, we may need a variable z to explain variation in t. A model for probability of loss requires that we specify a model for a latent variable

$$y^* = \alpha t + \beta x_1 + u. \tag{3}$$

Observed loss $y = 1$ if $y^* > 0$ and vice versa if $y^* < 0$. t potentially is endogenous. That is, treatment potentially is determined by a choice. The model for choice probability requires that we specify a model for a latent variable

$$t^* = \gamma z + \delta x_2 + v, \tag{4}$$

and observed choice $t = 1$ if $t^* > 0$ and vice versa if $t^* < 0$.

In the traditional test of adverse selection due to Chiappori & Salanie (2000), y is an indicator for occurrence of loss or claim and t is an indicator for coverage choice, and $\gamma = \alpha = 0$. This traditional test is *not* a test of effects. Rather, it is a test of conditional independence. The covariates x contain all relevant variables observed by the insurer. The test involves a test of significance of the correlation between errors u and v, which measure all variables unobserved by the analyst, and importantly, unobserved by the insurer. If the specification uses a bivariate

[18] For the EAS-2 test it is not quite correct to call y an *outcome* variable, since it is variable that is potentially jointly determined with t.

probit, a model for bivariate normal random variables, zero correlation implies independence. For this test, evidence in favor of the alternative hypothesis permits the analyst to conclude that there is adverse selection, and that unobserved variables are the source of the inefficiency problem. One is thus left to conjecture what that source might be.

In contrast, our test of an EAS equilibrium requires first that we posit a source of adverse selection and then conduct tests for hypotheses EAS-1 and EAS-2. In the analytical model, a forward contract (that specifies an insured's type) is taken as an exogenous parameter: one insured is type N and another is type F. Empirically, whether and how forward contract use is exogenous must be considered separately for each hypothesis test. Sufficiently strong evidence in favor of the alternative hypotheses leads to a precise policy recommendation: require disclosure of forward contract use. In order to be clear about the difference between the traditional test and our own, we can put it in the framework presented above.

The first hypothesis test EAS-1 is a test of the effect of forward contract use on loss probability. For this test, y is an indicator for loss probability and t is an indicator for forward contract use, the causal variable. The set of covariates x may contain some variables observed by the insurer and some not observed by the insurer. Their only purpose is to ensure that econometric methods used to determine the effect of forward are valid. For this test α potentially is non-zero and is the parameter of interest. And, γ potentially is non-zero if tests reveal that the causal variable is endogenous in the outcome equation. This test is exactly the same as the test considered by Evans & Schwab (1995), Altonji, Elder & Taber (2005) and model 6 (page 122) by Maddala (1983). In this case, the variable z, that determines outcome only through its effect on t, is local grain elevator capacity. Accounting for this variable z that measures marketing risk helps to solve the problem of omitted variables that may bias the measure of the effect of t on y. The definition and construction of elevator capacity is presented in the appendix. Since the outcome variable is a dichotomous variable the parameter of interest must be converted to a marginal effect: a change in probability. Following Finkelstein & Poterba (2004) we compare that to a relevant benchmark. For crop insurance, a relevant benchmark is the difference in loss probability associated with groups that already are differentially rated by crop insurers. For instance irrigated (low risk) and non-irrigated farms (high risk) are differentially rated in selected counties. Another classification variable already in use is crop rotation practice. Construction of the benchmark is described in the appendix.

The second hypothesis test EAS-2 is a test of the effect of forward on coverage choice. The interpretation of the econometric model differs for this test. Now, we may consider that choices over y and t result from expected utility comparisons. The specification can be shown to be an outcome of a random utility model, where the decision maker (the insured) has to consider two discrete choices. This is model 3 (page 119) in Maddala (1983). The endogeneity that could bias estimates of the effect of t now is due to the simultaneous determination of y and t. z performs the role of a traditional instrumental variable. The difference in specification for this test is that the variable y now is an indicator for high coverage. For this test, it is enough that we consider the significance and sign of parameter α, without converting to a marginal effect.

Econometric methods and tests used by Makki & Somwaru (2001), Finkelstein & Poterba (2004) and Cardon & Hendel (2001) differ from those of Chiappori & Salanie (2000), but they share in common a test for a role of variables unobserved by both the insurer and the analyst. In their tests, evidence in favor of the alternative hypothesis of adverse selection simply is deemed to be due to unknown variables. Again, here, for a test of endogenous adverse selection we require a relevant variable unobserved by the insurer but observed by the analyst. And, to have practical value, this variable should be potentially observable by the insurer. Here, in our analysis, we consider that variable to be forward contract use.

5.2 Specification

Identification here means the parameter of interest is measured with precision and without bias. Variables needed for analysis include outcome, causal, jointly determined, and covariate variables. Here, covariate variables measure riskiness, risk aversion and insurance contract terms.

For the causal test in EAS-1, a relevant, familiar example for many would be a study of the effect of education on earnings. There, relevant covariates try to measure ability. Measures of risk and risk aversion here are the counterpart to ability measures needed for those studies. By accounting for these covariates, the program evaluation can accurately assess if the effect of treatment on outcome is due to the causal variable alone.

Risk aversion depends upon the form of utility function as well as the argument of the utility function. In the analytical model, the argument of the utility function is per acre net income, or per acre revenue less per acre cost, and may in turn be affected by institutional and financial constraints. The relevant form of the utility function may depend upon non-pecuniary variables, such as age and education. For our choice of covariates, we have used theory and past empirical practice as our criteria for inclusion.

The analytical model of endogenous adverse selection does not distinguish between the probability of an accident (occurrence of yield below trigger yield) and the probability of a claim (when indemnity payment is made). They are identical and equal to q_k. Indeed, if all insureds who experience low yield make claims, this poses no problem. Here, for the outcome variable in EAS-1, we use a variable that measures occurence of a claim, so it likely gives a conservative estimate of the extent of accidents (low yield occurences).

The set of covariates that are conditioning variables for a test of causality for EAS-1 must be carefully chosen to ensure identification of the parameter of interest. Such identification is achieved if the chosen covariates satisfy required assumptions for each type of empirical method. We use matching methods, Probit and Bivariate Probit. Matching and Probit require a key exogeneity assumption, called variously, unconfoundedness, selection on observables or conditional independence. Essentially the outcome and causal variable must be independent after conditioning on covariates: $y \perp t | x$.

For EAS-1, covariates cannot contain intermediate outcomes. Intermediate outcomes are outcomes, not of interest for the analyst, that are determined by the causal variable, and in turn determine the outcome of interest to the analyst. So, while we have said risk aversion is determined by arguments of the utility function such as post-planting season off-farm income, we cannot include that as a covariate since it is an intermediate outcome. If off-farm work activity diverts labor resources from the farm, this may result in poor yields, which in turn determines the incidence of a claim for the EAS-1 test.

For both the EAS-1 and EAS-2 tests, covariates can include pre-determined variables. These are variables whose values the analyst knows were determined before the treatment was received, that is, before the forward decision was made. An example relevant for comparison to EAS-1 is the evaluation of government-sponsored work programs. For such evaluations, income in the year prior to introduction of the program is a pre-determined variable. Examples for our analysis are pre-planting season rainfall, soil quality and crop acreage decisions. Pre-determined variables are covariates that may determine the outcome, or even the causal variable. By including variables that may determine the causal variable for the EAS-1 test, it is possible to account for possible non-random assignment to treatment group. We do have non-random assignment here. Some farms choose forward and some do not. There, potentially is bias that could result from the self-selection into the forward contract user group.

For the EAS-1 test the marketing risk variable (local elevator capacity) determines variation in forward contract use, but likely does not determine the incidence of loss. It is a variable we use to specify a bivariate probit model that accounts for and is used to test for the possible selection bias. For the EAS-2 test, forward contract potentially is jointly determined, so that the marketing risk variable serves the role of traditional instrumental variable to correct for the endogeneity of forward contract use.[19]

[19] One complication is that forward in fact determines effort choice and effort choice in turn determines yield or risk of loss. Forward and effort are in fact sequential treatments. We do not account for such multiple treatments, and consider only the initial treatment: forward. The time line in Figure 1 showed that effort consists of input choices made during the growing-season.

5.3 Data

A variable that contains information that presently is unobserved by the insurer but potentially observable is central to our test of a zero-subsidy EAS separating equilibrium. Forward contract use is such a variable. The Agricultural and Resource Management Survey (ARMS) data set is unique in that it contains a nationally representative sample of individual farm level data on production, income, and risk management practices (including insurance and forward contracts). This is our primary data source.

1996 Cross-Section Data

We use the 1996 version of the ARMS data set because this was the year just prior to the introduction of many more insurance products. This enables us to more accurately conduct a test of the simple model of endogenous adverse selection. The primary insurance tool was crop yield insurance (called APH). Furthermore, the primary insurance purchase decision of a corn and soybean producer was whether to buy the insurance, and then whether to buy high or low coverage.

Claim and Coverage

The ARMS survey asks if the individual farm bought any basic (low coverage) or additional buy-up (high coverage) crop yield insurance (APH yield insurance). An indicator for purchase of some high coverage is a jointly determined variable for the EAS-2 test. An ARMS question asks if the farm received any indemnity payments. This indicates if a claim was made and is an outcome variable used for the EAS-1 test.

Forward Contract

The ARMS survey also contains detailed information on forward contracts. We know all terms of the contract: type of seller, price, quantity and crop. An indicator for forward contract is the causal variable for the EAS-1 test. This same indicator serves as a potentially jointly determined variable for the EAS-2 test.

Insurance Premium Base Rates

The ARMS survey does not report unit premiums paid by each individual. We infer this information from Risk Management Agency Actuarial tables. From these tables we know the county level premium schedule each producer would have faced for both levels of coverage. From these tables we construct a variable which measures the premium each producer would have paid. Details are given in the appendix.[20]

Risk

Soil quality (NRCS data) , precipitation (NCDC data) , and local elevator capacity (BNSF data) variables are constructed from other data sources. Soil quality and precipitation measure biological risk. Soil quality is an index of soil organic matter contained within the top 12 inches of the soil surface for an average column of soil measured in a county. Precipitation is total inches recorded in the county from January to April. Elevator capacity (reported in the appendix) is in 1000s of bushels in the county.[21] Elevator capacity measures marketing risk and potentially determines participation in forward contracts. This variable is used as an instrumental variable for our robustness checks to determined whether endogeneity of forward is a problem. The appendix describes how soil quality, precipitation and elevator capacity variables were constructed.

[20] Makki & Somwaru (2001) utilize premium and tenure for their insurance study, and importantly are able to account for RMA's differential rating of producers by yield risk class (called yield span). While we cannot do this precisely, since we do not know the producer's historical average yield, we do use RMA's County Acturial tables that give the premium schedule by coverage level and yield span to construct our measure of premium.

[21] Elevator access is elevator capacity divided by total county level bushels of corn and soybean. The county level total bushels are those reported in the census and not the sample.

Covariates

The moral hazard studies of Smith & Goodwin (1996) and Horowitz & Lichtenberg (1993) that assess the relationship between input use and insurance decisions, and the adverse selection study by Makki & Somwaru (2001), within the agricultural economics empirical crop insurance literature, are our main guides for included covariates. Horowitz & Lichtenberg (1993) utilize their own constructed measure of insurance premium, January-to-March rainfall, total operated acres and percent of owned acreas in their insurance demand equation. Smith & Goodwin (1996) utilize premium and debt-asset ratio in their specification for insurance demand. We use percent of acreage planted with soybeans as a measure of diversification (which also is used for input use specifications by these authors), and is used by Katchova & Miranda (2004) in their study of forward contracts use and by Sherrick et al. (2004) in their study of crop insurance demand. Our covariates also include age and education which are human capital variables that measure risk aversion, and are utilized in these and many other studies. Age is in years and education is measured as a 5-level indicator variable (2=high school and 3 =some college).[22] Effort (called 'Family Labor' in Table 1) is measured as total family (operator, spouse, and other family) hours reported as a share of total operated acres (land devoted to crops as well as other uses).[23] Fertilizer expense is total expenditure (in 1996 dollars) divided by total crop acres. Yield is bushels per acre.

Table 1 in the text gives the mean and t-test for a difference in means among insured producers in Illinois, Indiana and Iowa. t-statistics in bold are significant at the 10% level or higher. A simple difference of means for loss probability and coverage choice suggests no effect of forward contract. Some evidence to the contrary will be presented below where we discuss model results. Insured producers who use a forward contract tend to be younger and more educated, operate more total acreas, but own a smaller share of these operated acres. They utilize less family labor in all quarters. Labor in the 3rd and 4th quarters constitutes

[22] 1=less than high school; 2=high school, 3=some college, 4=college degree, and 5=graduate school.
[23] Total operated acres is appropriate as a measure of farm size.

	No Forward	Forward	t-statistic
Outcome			
Claim	0.19	0.20	0.19
High Coverage	0.47	0.52	0.94
Covariates			
Total Operated Acres	793	1415	**3.77**
Fertilizer Expense	30.50	31.77	0.60
Corn Acreage (Share)	0.47	0.48	0.54
Soybean Acreage (Share)	0.42	0.41	-1.20
January-March Family Labor	0.07	0.05	**-2.90**
April-May Family Labor	0.16	0.11	**-3.86**
Age	50.04	47.28	**-2.03**
Education	2.56	2.87	**3.00**
Debt-Asset Ratio	0.43	0.60	1.17
Owned Acres (Share)	0.42	0.32	**-1.87**
January-April Rainfall	7.90	8.93	**1.83**
Corn Premium Base Rate	0.0256	0.0256	-0.09
Soybean Premium Base Rate	0.0205	0.0217	1.30
Yields			
Corn	128.84	136.10	**2.36**
Soybean	42.58	42.88	0.29
Effort			
July-August Family Labor	0.13	0.08	**-4.08**
September-October Family Labor	0.15	0.10	**-3.90**

For t-stat in bold, p-value < 10%

Table 1: Difference of Means for Analysis Variables

effort, with respect to the insurance coverage decision. They also tend to operate in regions with higher pre-planting season rainfall that prepares the soil nutrient levels for optimal planting conditions.[24]

Additional data in Tables 6, 7 and 8 in the appendix give data means for all Illinois, Indiana, and Iowa farms in the ARMS sample, broken down by insured and uninsured groups, and for insured producers by forward contract use. The difference of means tests discussed for analysis variables above is for the same subset of data contained in the last two columns of the appendix tables. We give

[24]It is not clear whether this indicates they can commit to forward delivery due to reduced biological risk or use a forward since their income is more uncertain. It is the relatively drier areas of Illinois (the North and Central region) at least that tend to be associated with the highest corn yields.

these additional tables so that those familiar with Corn and Soybean production have a point of reference. From these data we see that un-insured producers tend to be older, own more of the acreage they operate, have lower income, equity and debt, and use more inputs. Their average yields are somewhat lower than producers who do use insurance. In addition, the data indicate that they have substantially greater access to grain elevator capacity. To the extent that elevator capacity measures marketing (price) risk, they face less such risk.

5.4 Tests and Methods

The parameter of interest is α in equation 3. α_1 is the effect when (3) is specified with outcome y equal to an indicator for indemnity payment, for test EAS-1. α_2 is the effect when (3) is specified with the jointly determined variable y equal to an indicator for high coverage, for test EAS-2. The magnitude of the marginal effect (abbreviated ME) also is important to give a truly policy-relevant answer to the question addressed by test EAS-1. Only the sign of the effect is required for test EAS-2. In terms of these policy-relevant quantities then, the null and alternative hypotheses for test EAS-1 is

$$\begin{aligned} &H0: \alpha_1 \leq 0 \quad \& \quad HA: \alpha_1 > 0, \quad \textit{or} \\ H0: ME_1 \leq \Delta q^{benchmark} \quad &\& \quad H1: ME_1 > \Delta q^{benchmark}. \end{aligned} \tag{5}$$

Derivation of a benchmark change in probability is discussed in the appendix. This benchmark is about 1% to 2%. For test EAS-2, the null and alternative hypotheses are

$$H0: \alpha_2 \leq 0 \quad \& \quad H1: \alpha_2 > 0. \tag{6}$$

For test EAS-1, we estimate α using a simple difference of means, a matching estimator and a probit equation for the model (3), while for test EAS-2, we use a simple difference of means and a probit equation.

Suspecting lack of overlap, for the test EAS-1, from an inspection of distributions of the estimated propensity score (see Figure 4) and lack of unconfoundedness, we must suspect the validity of a matching estimator. Unconfoundedness, or selection on observables, is a key assumption needed also for a probit model specified with treatment t as an exogenous variable. We therefore test for violation of exogeneity with a test of significance for the correlation coefficient ρ between variables y and t in a Bivariate probit model with equations (3) and (4), separately for EAS-1 and for EAS-2. For the EAS-1 test, such a bivariate probit model is what Heckman called the multivariate probit model with structural shift. See Model 6 in chapter 5 of Maddala (1983) and Vytlacil & Yildiz (2004).[25]

The final set of covariates x in equations (3) and (4) for each test, for each of the models estimated, includes a constant, total acres, soybean acreage, second quarter family labor, age, education, debt-asset ratio, owned acres, rainfall and premium. These are the covariates used for the Matching estimator as well. Probit models include indicators for Iowa and Indiana and interactions of the forward contract indicator with Iowa and Indiana indicators. The base model is for Illinois. For endogeneity tests, we use the set of Illinois and Iowa farms. For these tests, soil quality is included as a covariate and elevator capacity is used as an instrument for potentially endogenous forward (t) in equation (3). Presumably elevator capacity may determine forward contract participation, but is unlikely to be a separate determinant of either occurrence of claim (loss) or coverage choice. We have chosen to present models with the full set of covariates (x) only. As noted above, the specification is informed by economic theory and the empirical literature. It is also parsimonious. So we have no strong grounds to meaningfully omit any of these covariates. We may now state results and offer policy recommendations and suggestions for further work.

[25]We also test for violation of exogeneity using the Rivers-Vuong test .The procedure for the Rivers-Vuong follows pages 474 and 478 of Wooldridge (2003) and validity of the the bivariate probit model in the absence of unconfoundedness follows Wooldridge (2003) and Evans & Schwab (1995).

6 Results

We have conducted two hypothesis tests from a model of a zero-subsidy Endogenous Adverse Selection (EAS) equilibrium, using data on corn and soybean producers in Illinois, Iowa and Indiana.

The two hypothesis tests answer two questions. First, is the probability of an indemnity payment higher for insured producers with forward contracts than without them? This is the question addressed by hypothesis test EAS-1. Second, is the probability of choosing a high coverage policy greater for insured producers with forward contracts than without them? This question is addressed by hypothesis test EAS-2.

For each test, we must assess whether forward contract choice is endogenous. With a proper accounting for the endogeneity[26] of forward contract choice, we can identify the parameter of interest.[27] For test EAS-1, endogeneity is assessed using a Bivariate Probit model for the joint probability of the indemnity payment outcome and of forward contract choice. For test EAS-2, endogeneity is assessed using a Bivariate Probit model for the joint probability of high coverage choice and forward contract choice. In each case, a test of significance of the estimated correlation coefficient $\hat{\rho}$ in the model for two jointly distributed Normal random variables is a test for endogeneity of forward contract choice.

For the EAS-1 model, $\hat{\rho} = -0.004$ with a standard error of 1.46. For the EAS-1 test, we cannot reject the null hypothesis that forward contract choice is exogenous.[28] We instead use results from the univariate probit model, given in Table 10.

Appendix Table 11 provides results for the Bivariate probit model for coverage and forward contract choice, using elevator capacity as an instrumental variable for forward contract choice. Endogeneity of forward contract use however may be a problem for the EAS-2 model and test, since $\hat{\rho} = -0.74$ with a standard error of 0.54. Hence, we report results from both the Bivariate Probit model (in Table 11)

[26] There is relevant endogeneity in the econometric model and in the analytical model. They are related, but not equivalent.
[27] The parameter of interest, labeled α in equation 3.
[28] Alternatively, we can state that the assumption of selection on observables, or of unconfoundedness, holds.

and univariate probit model (in Table 10).

Results for hypothesis test EAS-1 in Table 2 and those for test EAS-2 in Table 3 are given as marginal effects. In Table 2, the marginal effect is the estimated increase (if positive) in probability that an insured producer with a forward contract receives an indemnity payment. The marginal effect in Table 3 is the estimated increase (if positive) in probability that an insured producer with forward contracts would choose high coverage.

In the EAS-1 test, we see from Table 2 that insured Indiana producers with forward contracts exhibit an estimated 6% (SE = 3%) increase in risk. Insured Illinois and Iowa producers with forward contracts exhibit decreases in risk, but these effects are not significant. Estimated effects from a difference of means test using a T-statistic and using a Matching estimator are provided for completeness.

In the EAS-2 test, we see from the Bivariate Probit model result given in Table 3 that the estimated chance of choosing high coverage increases by about $\frac{1}{3}$ (ME=0.37, SE=0.21) for insured producers with a forward contract above those insured and without forward contracts. [29] From the univariate probit model of coverage choice, reported in Table 10, we see from the positive sign of the coefficient of Forward (estimated to be 0.54) that insured Illinois producers tend to choose high coverage. From the estimated coefficient of the Indiana x Forward interaction (estimated to be -0.77) we see that insured producers with forward contracts in Indiana have a reduced propensity to opt for high coverage (since the effect for Indiana is calculated as 0.54 - 0.77 < 0).

Interpretation

Insured Indiana producers' risk increase is estimated to be 6% when they use at least one forward contract. This is substantial when compared to a benchmark risk increase of 1% to 2%. The benchmark we use comes from our calculation of the risk increase implied by published actuarial premium rates for soybean producers with

[29] These results are for the sample that includes Illinois and Iowa because the elevator capacity variable used as an instrument for the Bivariate Probit model is not available for Indiana.

	Difference of Means	Matching	Probit	Bivariate Probit
Forward (overall)	0.01	0.01		na
	[0.05]	[0.05]		na
Illinois Effect			-0.07	
			[0.07]	
Iowa Effect			-0.01	
			[0.13]	
Indiana Effect			**0.06**	
			[0.03]	

SE in square brackets.

Table 2: Effect of Forward on Loss Probability: EAS-1 Test

	Difference of Means	Matching	Probit	Bivariate Probit
Forward (overall)	0.05	na		**0.37**
	[0.06]	na		**[0.21]**
Forward (Illinois)			0.21	
			[0.11]	
Iowa Effect			0.07	
			[0.18]	
Indiana Effect			-0.10	
			[0.16]	

Table 3: Effect of Forward on Coverage Choice: EAS-2 Test

and without a high risk crop rotation practice. A high risk crop rotation practice is one in which there is continuous sequential cropping so that the soil is not allowed to remain fallow and replenish its nutrient content. Since the ARMS data do not ask very detailed questions about the crop insurance policy, we determined from separate analysis that the indemnity payments actually were made for corn yield losses. Hence, our estimated risk increase pertains to an increase in the probability of a corn yield insurance indemnity payment.

Considering all producers as a group, insured producers with a forward contract are more likely to use high coverage than are producers who sell their crop only on the harvest-time spot market. Together with the result for the relatively greater riskiness of insured Indiana producers with forward contracts, we have evidence in favor of endogenous adverse selection. We also conducted a Chiappori-Salanie test (not reported here) and found the correlation between error terms (that measure unobservable variables) was positive and significant. This merely says there is evidence of adverse selection, and that variables the insurer does not observe cause the problem. Our framework and results supplement this finding with a more precise policy prescription, and suggest that Indiana producers without a forward contract who buy corn yield insurance should be placed in a lower risk class.

If we consider results only from univariate probit models, we find that insured Indiana producers with forward contracts pose higher risks to the insurers and are less likely to use high coverage. This is consistent with so-called propitious selection. Propitious selection (due to Hemenway) occurs when lower-risk agents tend to buy higher coverage and be more cautious (and, vice versa for higher-risk agents). Adverse selection, in contrast, predicts that the lower-risk agents buy policies with sub-optimally low coverage levels. In his comprehensive review of econometric studies that test for adverse selection, Siegelman (2004) concludes that evidence has accumulated suggesting propitious selection (not adverse selection) might best characterize the behavior of insureds.

We may find stronger evidence that forward contract use differentiates producers into distinct risk classes when other major corn and soybean producing states are included. If we consider states accounting for most corn and soybean production and for the highest exposure levels (as measured by crop insurance program liability) in 1996, we would include Illinois, Iowa, North Dakota, South Dakota, Kansas, Nebraska and Missouri.

Our tests and analysis have been stated in terms of loss probability, as measured by indicator variables for occurence of indemnity payments. Alternatively, they might have been stated in terms of differences in yield (outcome) distributions, facilitating possible yield distributions heterogeneity. In a simple graphical analysis and test of difference between means, we have insured producers using forward contracts to have corn yield distributions different from insured producers without forward contracts.

7 Conclusion

We have claimed that a weakness of adverse selection tests which follow the seminal work of Chiappori and Salanie (2000) is that they are designed merely to detect inefficiency due to hidden information and cannot provide a precise policy recommendation on how to remove the inefficiency. We argue that a revised question, analytical framework, and tests are necessary to render the analysis informative for a firm or policy maker. The appropriate analytical framework must endogenize the adverse selection. Empirical analysis involves hypothesis tests motivated by two predictions from a model of Endogenous Adverse Selection. As analysts, we must posit the *presently unobserved, but potentially observable* source of hidden information which might cause the inefficiency. Indeed, we find evidence that insured producers' use of forward contracts is a source of inefficiency. The precise policy recommendation is to design a new insurance contract at a higher coverage level at given or lower premiums for the low-risk producers who do not have forward contracts.

Legislative proposals aiming to protect consumers in vulnerable groups frequently advocate information disclosure restrictions and bans. Such legislation usually is opposed by the insurance industry on grounds that the information which would be banned predicts loss experience and thus enables actuarially fair rating. Examples are recent state-level restrictions and bans on homeowner insurance firms' practice of credit-score-based underwriting and rating. A consumer's credit-score is used to make policy offers (underwriting) and as a rating variable to set premium rates. Our proposed analytical method can inform a policy maker whether society has benefited from permitting insurance firms to use the credit-score as a rating variable. The analyst must check if the homeowners' insurance market was characterized by a zero-subsidy endogenous adverse selection equilibrium prior to the insurers' use of credit-scores as a rating variable. If there is evidence that a zero-subsidy endogenous adverse selection equilibrium prevailed historically, prior to the insurers' use of the rating variable, then society's welfare would fall with regulation that imposes a credit-score disclosure ban.

Consumers may, on the other hand, object to an insurer's requirement that new information be disclosed for rating the insurance contract. Policy makers can determine whether the market presently is characterized by a zero-subsidy endogenous adverse selection equilibrium, induced by the rating variable in question. If there is evidence of such an equilibrium, the consumer's disclosure to the insurance company would boost social welfare. Debate recently has been raised about the efficiency and equity effects of informing health and life insurers about genetic test results.

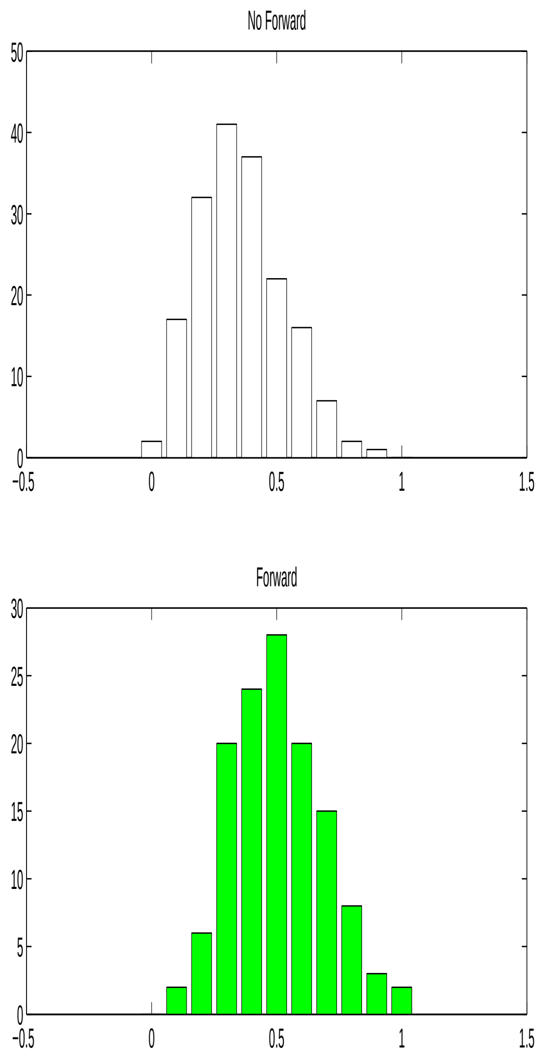

Figure 4: Propensity Score: Treatment and Control

Appendix

The appendix presents details for discussion on the EAS equilibrium, data and model results, referred to in the main text.

EAS Equilibrium Notation

Definitions for basic notation and expressions for state-contingent wealth used for the model of an Endogenous Adverse Selection equiibrium are given in Tables 4 and 5. Complete details for this model is found in Chapter 2 of the dissertation.

Notation	Explanation
U	utility function common to all
p_i	price, low or high, $i \in L, H$
y_j	yield, low or high, $j \in L, H$
π	probability of low price
k	type of insured $k \in F, N$
	forward (k=F), no forward (k=N)
y_m^k	quantity sold forward for type k
f_i^k	net gain with forward, $(p_m - p_i)y_m^k$
	p_m is forward price per unit
α_k	coverage net of premium for type k
β_k	premium for type k
e_k	effort $e_k = e(\alpha, \beta, f^k)$ for type k
q_k	probability of low yield for type k, $q_k = q(e_k)$
R_{ij}	revenue with price i and yield j, $p_i y_j$
W_{ij}^k	wealth for type k, with price i and yield j
I_j^k	indirect utility for type k, given yield j
	$I_j^k = \pi U(W_{Lj}^k) + (1-\pi)U(W_{Hj}^k)$
λ_k	proportion of type k in market

Table 4: Basic Notation

Supplementary Summary Statistics

Tables 6, 7 and 8 give data on income, production and risk measures. Construction of risk measures is considerd in the next section. Yield insurance premium base rates implicitly are measures of yield risk, soil quality and rainfall are measures of biological sources of yield (supply) risk, and local grain elevator capacity and degree of access are measures of marketing or price (demand) risk. These data are referenced in the main text, and some of these data are explicitly used for the estimated models.

Wealth (price & yield state-contingent)

$$W^N_{LL} = R_{LL} + \alpha_N$$
$$W^N_{LH} = R_{LH} - \beta_N$$
$$W^N_{HL} = R_{HL} + \alpha_N$$
$$W^N_{HH} = R_{HH} - \beta_N$$
$$W^F_{LL} = R_{LL} + \alpha_F + f_L$$
$$W^F_{LH} = R_{LH} - \beta_F + f_L$$
$$W^F_{HL} = R_{HL} + \alpha_F + f_H$$
$$W^F_{HH} = R_{HH} - \beta_F + f_H$$

Derived Utility (yield state-contingent)

$$I^N_L = \pi U(W^N_{LL}) + (1-\pi)U(W^N_{HL})$$
$$I^N_H = \pi U(W^N_{LH}) + (1-\pi)U(W^N_{HH})$$
$$I^F_L = \pi U(W^F_{LL}) + (1-\pi)U(W^F_{HL})$$
$$I^F_H = \pi U(W^F_{LH}) + (1-\pi)U(W^F_{HH})$$

Table 5: State-contingent Wealth

Construction of Risk Variables

An insurance premium is the price an insured producer pays for insurance coverage. Also, since an insurer rates a policy for a particular insured based on all the information available to him, it also is a summary measure of that information set. January to April rainfall is a measure of short-run biological risk for corn and soybean production, since planting occurs in April and May. In contrast, soil quality is better thought of as a measure of long-run biological risk. Access to local grain elevator capacity may be thought of as a measure of marketing (price) risk.

Insurance Premium Base Rate

Since the insurer observes enough about the insured to assign a premium, we include premium as a covariate in our specification. In the results we see that coefficients are significant for premium in the indemnity equation and not in the coverage equation. Premium in the indemnity equation serves as a measure of all information available to the insurer. In the coverage equation it serves as a measure of price the insured faces.

Four price bounds are constructed for each county. They are the minimum base rate for low (50%) coverage and the maximum base rate for high (65%) coverage, for corn and for soybeans. The minimum base rate at a given coverage level applies to the highest yield class. Makki & Somwaru (2001) use the technical term yield span to refer to what we call yield class. Likewise, the maximum base rate applies to the lowest yield class.

After several initial model runs, we found that maximum base rates did not improve model fit measures, and so the final set of covariates uses only the minimum base rate. Such a minimum gives the a threshold measure: a producer must be willing to pay at least this minimum to be willing to buy insurance.

The base rates are extracted from USDA Risk Management Agency County Actuarial Table rate files for APH for 1996 (USDA 1996).

	All	Uninsured	Insured	No Forward & Insured	Forward & Insured
Unweighted Obs.	425	120	305	177	128
Age	49.83	52.23	48.88	50.04	47.28
Education	2.65	2.54	2.69	2.56	2.87
Income ($1000)					
Contract Corn	29.32	11.00	36.53	0.00	87.05
Contract Soybean	19.41	17.27	20.26	0.00	48.28
Cash Corn	86.30	66.21	94.20	96.03	91.66
Cash Soybean	61.12	45.07	67.43	71.60	61.68
Livestock	22.73	19.22	24.11	19.70	30.21
Crop	203.39	143.31	227.03	174.61	299.51
Off-farm	3.21	1.08	4.05	4.56	3.33
Gross	252.62	176.71	282.49	217.89	371.83
Income Per Acre	258.21	240.77	265.07	264.50	265.87
Equity ($ 1000)	251.42	182.57	278.50	237.78	334.81
Debt ($1000)	166.48	100.74	192.34	136.71	269.27
Debt-Asset Ratio	0.46	0.36	0.50	0.43	0.60
Total Operated Acres	941.72	656.73	1053.85	792.84	1414.77
Total Crop Acres	857.33	592.48	961.53	726.36	1286.72
Ownership (Percent of Total)	0.44	0.59	0.38	0.42	0.32

Table 6: Illinois, Iowa and Indiana Grain Farms: Income & Asset Data Means

	All	Uninsured	Insured	No Forward & Insured	Forward & Insured
Fertilizer ($ Per Crop Acre)	31.84	33.88	31.03	30.5	31.77
Family Labor					
First Quarter	0.07	0.11	0.06	0.07	0.05
Second Quarter	0.17	0.23	0.14	0.16	0.11
Third Quarter	0.13	0.18	0.11	0.13	0.08
Fourth Quarter	0.15	0.20	0.13	0.15	0.10
Yield (Bushels Per Acre)					
Corn	130.70	127.32	131.91	128.84	136.10
Soybean	41.84	39.58	42.71	42.58	42.88

Table 7: Illinois, Iowa and Indiana Grain Farms: Production Data Means

	All	Uninsured	Insured	No Forward & Insured	Forward & Insured
Premium Base Rate					
Corn Minimum	0.03	–	0.03	0.03	0.03
Corn Maximim	0.11	–	0.11	0.11	0.11
Soybean Minimum	0.02	–	0.02	0.02	0.02
Soybean Maximum	0.11	–	0.11	0.10	0.11
Biological Risk					
Soil Quality	3.78	3.56	3.84	3.98	3.64
January-April Rainfall	9.02	10.72	8.33	7.90	8.93
Marketing Risk					
Elevator Capacity	1180.42	804.84	1296.60	1287.33	1310.86
Access	0.07	0.05	0.07	0.07	0.07

Table 8: Illinois, Iowa and Indiana Grain Farms: Risk Data Means

Rainfall

January through April rainfall is calculated as a county level sum of daily rainfall readings for these months for 1996. Data were obtained from the National Climatic Data Center (NCDC) Cooperative Summary of the Day data files (NCDC 1996). When a county has more than one weather station, we use the average across stations for that county. When a county contains no weather station, we use the average across all surrounding counties. Rainfall is recorded in 100ths of an inch. The data are converted to inches.

Soil Quality

Many measures of soil quality have been conceived. We use soil organic matter data collected at soil sample stations by the National Resources Conservation Survey (NRCS), and recorded in their Map Unit Interpretation Record (MUIR) data base (USDA 1995). As of 2002, these were the most complete set of spatially referenced soil organic matter data. They contain samples collected up to about 1990. This information is suitable for an analysis of 1996 production.

The NRCS collect soil information at multiple sample sites within homogeneous ecological regions called soil sample areas. Within each area one or more locations is sampled; we use a simple average across stations as a measure of quality for each soil sample area. At each location, the sample is obtained at multiple depths. We use the information recorded at the shallowest depth, which covers up to 12 inches below ground.

Soil sample area boundaries do not coincide with county boundaries. We use ARC Info software to overlay soil sample boundaries with county boundaries. Then we aportion soil quality values from soil sample area to each county in proportion to the degree of area overlap. (If area A covers $\frac{1}{3}$ of a county and area B covers the remaining $\frac{2}{3}$, the county has soil quality equal to $\frac{1}{3}$ of soil quality measure for A plus $\frac{2}{3}$ that of B.

Local Elevator Capacity

We obtained elevator longitude and latitude and capacity in bushels for each elevator in the Burlington Northern Santa Fe (BNSF) elevator directory (*Burlington Northern Santa Fe Elevator Directory* 2002). Using ARC Info software, we then constructed a Voronoi tesselation with this set of elevator points. In this way, regions get cut into polygons created from this set of points. Polygons are created by taking perpendicular bisectors between every pair of points. All those bisectors intersect into polygons, some very small in areas with a heavy concentration of points. Each polygon has the feature that it contains one unique elevator point and every other possible geographical point in that polygon, where farms and houses would be located, is closest to the elevator point in that polygon.

We then overlay the voronoi elevator polygons onto county boundaries. Similar to the method for soil quality, each county then shares some amount of area with one or more polygons. The county elevator capacity is a polygon area share weighted average of polygon-specific elevator capacity.

The BNSF elevator directory unfortunately contains no elevators in Indiana since that area is served by other rail lines and elevators. This is why Indiana is excluded from the sample to conduct exogeneity tests that require an instrumental variable. Soil quality simply was not collected for Indiana; this is why specifications that do include Indiana do not include this important measure of biological risk. We can see from visual inspection of maps that the incidence of claims is higher in regions of low soil quality. See the map supplement.

Benchmark Loss Probability

In this appendix we describe how we derive the benchmark change in probability used to assess the economic significance of estimated effects of forward on loss probability. These estimated effects give the difference in probability of loss between producers with and without a forward.

It is not enough merely to test whether the estimated effect is significantly different from zero. Rather, we compare the estimated magnitude to a difference in loss probability between groups which differ by crop rotation practice and irrigation practice. Crop rotation and irrigation have been identified by crop insurers to establish distinct risk class for soybean insurance in certain counties in each state in our sample. Very few counties have soybean insurance premiums that differ by irrigation practice. Actuarial tables give base premium rates for each yield class for each level of coverage, separately for each type of practice. From this information we must infer the implied difference in loss probability.

First, for an actuarily fair premium, premium equals expected indemnities. Premium is premium per unit liability π times liability L. Probability of loss is q and indemnity payment in the event of loss is I. Zero expected profit implies $q(\pi L - I) + (1-q)\pi L = 0$. So, $\pi = \frac{qI}{L}$, which equals expected indemnity divided by liability.

For crop insurance, premium is base rate (π) times insured acres (A) times coverage share (β) times mean yield (μ), or $\pi \cdot A\beta\mu$. Expected Indemnity per acre is the expected value of the difference between yield trigger ($\beta\mu$) and actual yield y. This equals $P(y < \beta\mu) \cdot [\beta\mu - E[y|y < \beta\mu]$. Since insured acres cancel out of the expression, per unit premium π is given by

$$\pi = P(y < \beta\mu) \cdot \left[1 - \frac{E[y|y < \beta\mu]}{\beta\mu}\right] = q \cdot \delta. \tag{7}$$

δ is the term in square brackets. q is the probability that yield falls below trigger, when an indemnity payment would be made, assuming a loss implies a claim is made by the insured. So, the benchmark change in probability $\Delta q = q_1 - q_2$ is derived from the data by calculating

$$\Delta q = \frac{\pi_1}{\delta_1} - \frac{\pi_2}{\delta_2} = \frac{1}{\delta} \cdot \Delta\pi, \tag{8}$$

under the assumption that $\delta_1 = \delta_2 = \delta$. Subscripts 1 and 2 indicate distinct practices. Here we consider crop rotation practice (within-year sequential cropping equals 1 and no within-year sequential cropping equals 2) and irrigation (non-irrigated crop equals 1 and irrigated crop equals 2).

For our benchmark we assume $\beta = 0.65$, since the 65% share is a popular coverage level. By assuming various values for $E[y|y < \beta\mu]$, we assume various values for δ. Since mean yield when yield is below trigger level must be less than the trigger level, $E[y|y < \beta\mu] < 0.65\mu$. We assume two values for this lower tail mean yield: 0.25μ and 0.50μ. Hence, δ equals 0.615 and 0.231, respectively. $\Delta\pi$ is calculated by finding the premium base rate from the Risk Management Agency county level actuarial table which applies to the 1996 NASS county average yield (the value for μ we use for each county) at coverage level $\beta = 0.65$. There is a separate county level table for each type of practice, 1 and 2. Hence we get π_1 and π_2 for each county in this way. We then calculate the mean, median, minimum and maximum difference for the set of $\Delta\pi$. Benchmark Δq, estimated from $\frac{\Delta\pi}{\delta}$, is compared to our estimated effects. The calculated benchmarks are given in Table 9. A reasonable range of benchmarks are given in bold. The range is 1% to 2%.

Model Results

Full results for models estimated for EAS-1 and EAS-2 hypothesis tests are in Table 10 for univariate probit analysis when the source of adverse selection is use of a forward contract, and Table 11 for the bivariate probit model when source of adverse selection is use of a forward contract.

	$E[y\|y<\beta\mu]$	δ	Δq			
			Mean	Median	Min	Max
Crop Rotation						
$\Delta\pi$:			0.005	0.004	-0.018	0.034
	0.25μ	0.615	**0.008**	**0.006**	-0.028	0.052
	0.50μ	0.231	**0.022**	**0.017**	-0.078	0.147
Irrigation						
$\Delta\pi$:					-0.002	0.009
	0.25μ	0.615			-0.003	0.015
	0.50μ	0.231			-0.009	0.039

Table 9: Benchmark Probability Difference: Δq

	Indemnity		**Coverage**	
	Estimate	[T-Ratio]	Estimate	[T-Ratio]
Forward	-0.29	[-0.97]	0.54	*[2.00]
Forward x Iowa	0.23	[0.50]	-0.35	[-0.96]
Forward x Indiana	0.55	[1.25]	-0.77	*[-1.88]
Constant	-1.39	*[-1.82]	0.53	[0.81]
Iowa	-0.49	*[-1.67]	0.53	*[2.07]
Indiana	0.04	[0.14]	0.67	*[2.23]
Total Operated Acres	0.001	[0.14]	-0.02	*[-1.83]
Soybean Acreage (Share)	1.20	*[2.05]	0.61	*[1.22]
April-May Family Labor	-0.39	[-0.41]	-1.36	*[-1.70]
Age	-0.01	[-1.23]	-0.01	[-1.51]
Education	-0.16	[-1.40]	-0.02	[-0.18]
Debt-Asset Ratio	0.03	[1.45]	0.15	[1.00]
Owned Acres (Share)	-0.02	[-0.13]	-0.31	[-1.62]
January-April Rainfall	0.02	[0.70]	-0.07	*[-2.48]
Corn Premium	71.99	*[2.51]	-8.83	[-0.36]
Soybean Premium	-48.12	*[-1.80]	24.98	[1.05]
χ^2	35.52	< 0.01	44.43	<0.01
Correct Predictions	81%		67%	

Table 10: Univariate Probit Models Used for EAS-1 & EAS-2 Tests

	Coverage		Forward Participation	
Forward (t)	1.34	*[1.82]	na	na
Constant	0.40	[1.15]	-0.02	[-0.02]
Iowa	0.54	*[2.02]	-0.56	*[-1.90]
Total Operated Acres	-0.03	*[-2.36]	0.02	[1.43]
Soybean Acreage (Share)	0.70	[1.10]	-0.53	[-0.76]
April-May Family Labor	-1.26	[-1.02]	-1.46	[-0.89]
Age	-0.01	[-1.00]	0.001	[0.07]
Education	-0.02	[-0.13]	0.20	*[1.69]
Debt-Asset Ratio	0.12	0.64	[-0.09]	[0.33]
Owned Acres (Share)	-0.21	[-0.82]	-0.14	[-0.41]
January-April Rainfall	-0.07	[-1.22]	-0.02	[-0.32]
Corn Premium	-6.66	[-0.21]	-22.88	[-0.61]
Soybean Premium	24.67	[0.77]	0.99	[0.27]
Soil Quality	-0.07	[-0.67]	-0.04	[-0.34]
Elevator Capacity (z)	na	na	0.01	[1.44]
ρ			-0.74	
SE($\hat{\rho}$)			0.54	

Table 11: Bivariate Probit Model Used for EAS-2 Test

References

Altonji, J., Elder, T. & Taber, C. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of Political Economy* **113**(1), 151–184.

Arnott, R. & Stiglitz, J. (1988), 'Basic analytics of moral hazard', *Scandinavian Journal of Economics* **90**(3), 383–413.

Arnott, R. & Stiglitz, J. (1991), 'Equilibrium in competitive insurance markets with moral hazard', *NBER Working Paper Series* . WP 3588.

Babcock, B., Hart, C. & Hayes, D. (2004), 'Actuarial fairness of crop insurance rates with constant rate relatives', *American Journal of Agricultural Economics* **86**(3), 563–575.

Burlington Northern Santa Fe Elevator Directory (2002), www.bnsf.com.

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

Chiappori, P.-A. & Salanie, B. (2000), 'Testing for asymmetric information in insurance markets', *Journal of Political Economy* **108**(1), 56–78.

Chiu, W. H. & Karni, E. (1998), 'Endogenous adverse selection and unemployment insurance', *Journal of Political Economy* **106**(4), 806–827.

de Garidel-Thoron, T. (2005), 'Welfare improving asymmetric information in dynamic insurance markets', *Journal of Political Economy* **113**(1), 121–150.

Demeza, D. & Webb, D. C. (2001), 'Advantageous selection in insurance markets', *RAND Journal of Economics* **32**(2), 249–262.

Dionne, G., Doherty, N. & Fombaron, N. (2000), Adverse selection in insurance markets, Kluwer Academic Publishers, Boston. Chapter 7.

Evans, W. & Schwab, R. (1995), 'Finishing high school and starting college: Do catholic schools make a difference?', *Quarterly Journal of Economics* **110**, 941–974.

Finkelstein, A. & Poterba, J. (2004), 'Adverse selection in insurance markets: Policyholder evidence from the uk annuity market', *Journal of Political Economy* **112**(1), 183–208.

Horowitz, J. K. & Lichtenberg, E. (1993), 'Insurance, moral hazard, and chemical use in agriculture', *American Journal of Agricultural Economics* pp. 926–935.

Hoy, M. (2005), 'Risk classification and social welfare', *Department of Economics Discussion Paper* . University of Guelph, 2005-8.

Hueth, B. & Ligon, E. (2001), 'Agricultural markets as relative performance evaluation', *American Journal of Agricultural Economics* **83**(2), 318–328.

Jullien, B., Salanie, B. & Salanie, F. (1999), 'Should more risk averse agents exert more effort', *Geneva Papers on Risk and Insurance Theory* **24**, 19–28.

Just, R., Calvin, L. & Quiggin, J. (1999), 'Adverse selection in crop insurance: Actuarial and asymmetric information incentives', *American Journal of Agricultural Economics* .

Katchova, A. & Miranda, M. (2004), 'Two-step econometric estimation of farm characteristics affecting marketing contract decisions', *American Journal of Agricultural Economics* **86**(1), 88–102.

Ligon, E. (2004), 'Production data to design efficient contracts', *American Journal of Agricultural Economics* **86**(3), 848–853.

Maddala, G. S. (1983), *Limited Dependent Variables and Qualititative Variables in Econometrics*, Cambridge University Press, New York.

Makki, S. S. & Somwaru, A. (2001), 'Evidence of adverse selection in crop insurance markets', *The Journal of Risk and Uncertainty* **68**(4), 685–708.

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

NCDC (1996), 'National climatic data center (ncdc) cooperative summary of the day', CD ROM.

Polborn, M., Hoy, M. & Sadanand, A. (2005), 'Advantageous effects of regulatory adverse selection in the life insurance market', *Economic Journal* . Forthcoming.

Rothschild, M. & Stiglitz, J. (1976), 'Equilibrium in competitive insurance markets: An essay on the economics of imperfect information', *Quarterly Journal of Economics* pp. 629–649.

Sherrick, B., Barry, P., Ellinger, P. & Schnitskey, G. (2004), 'Factors influencing farmers' crop insurance decisions', *American Journal of Agricultural Economics* **86**(1), 103–114.

Siegelman, P. (2004), 'Adverse selection in insurance markets: An exaggerated threat', *Yale Law Review* pp. 1223–1277.

Smith, V. H. & Goodwin, B. K. (1996), 'Crop insurance, moral hazard and agricultural chemical use', *American Journal of Agricultural Economics* **78**, 428–438.

Townsend, J. & Brorsen, B. W. (2000), 'Cost of forward contracting hard red winter wheat', *Journal of Agricultural and Applied Economics* **32**, 89–94.

USDA (1995), *NRCS Map Unit Insterpretation Records (MUIR) Data Base*. CD ROM Version 5.

USDA (1996), 'Risk management agency (rma) county actuarial tables', FTP Site. FCI-35 Rates.

USDA (1997), 'Usual planting and harvesting dates for u.s field crops'.

USDA (2003), *Agricultural Resources and Environmental Indicators*. ERS.

Vercammen, J. (2002), 'Welfare-improving adverse selection', *International Economic Review* **43**(4), 1017–1033.

Vytlacil, E. & Yildiz, N. (2004), Dummy endogenous variables in weakly separable models. Working Paper, Department of Economics, Stanford University.

Wooldridge, J. (2003), *Econometric Analysis of Cross-section and Panel Data*, The MIT Press, Cambridge. Chapters 15 and 18.