TESTING FOR ADVERSE SELECTION
WITH "UNUSED OBSERVABLES"

Amy Finkelstein
James Poterba

Testing for Adverse Selection with "Unused Observables"
Amy Finkelstein and James Poterba
NBER Working Paper No. 12112
March 2006
JEL No. D82, G22

ABSTRACT

This paper proposes a new test for adverse selection in insurance markets based on observable characteristics of insurance buyers that are not used in setting insurance prices. The test rejects the null hypothesis of symmetric information when it is possible to find one or more such "unused observables" that are correlated both with the claims experience of the insured and with the quantity of insurance purchased. Unlike previous tests for asymmetric information, this test is not confounded by heterogeneity in individual preference parameters, such as risk aversion, that affect insurance demand. Moreover, it can potentially identify the presence of adverse selection, while most alternative tests cannot distinguish adverse selection from moral hazard. We apply this test to a new data set on annuity purchases in the United Kingdom, focusing on the annuitant's place of residence as an "unused observable." We show that the socio-economic status of the annuitant's place of residence is correlated both with annuity purchases and with the annuitant's prospective mortality. Annuity buyers in different communities therefore face different effective insurance prices, and they make different choices accordingly. This is consistent with the presence of adverse selection. Our findings also raise questions about how insurance companies select the set of buyer attributes that they use in setting policy prices. We suggest that political economy concerns may figure prominently in decisions to forego the use of some information that could improve the risk classification of insurance buyers.

Amy Finkelstein
Department of Economics
MIT E52 262F
50 Memorial Drive
Cambridge MA 02142
and NBER
afink@mit.edu

James Poterba
Department of Economics
MIT E52 350
50 Memorial Drive
Cambridge MA 02142
and NBER
poterba@mit.edu

Economic theory demonstrates that asymmetric information can hinder the efficient operation of insurance markets. Whether asymmetric information exists in specific markets, however, remains a subject of considerable controversy. Numerous studies have tested for the presence of asymmetric information, typically using what Chiappori et al. (2006) label the "positive correlation" test. This test rejects the null hypothesis of symmetric information when there is a positive correlation between insurance purchases and risk occurrence, conditional on the buyer characteristics that are used to set insurance prices. However, the positive correlation test breaks down when individuals may have private information about characteristics other than risk type, such as risk preferences, and when these other individual attributes may affect insurance demand. This is likely to be a common situation.

This paper therefore develops and implements a new test for asymmetric information that is robust to the existence of preference heterogeneity in insurance demand. When there are insurance buyer characteristics that are not used to price insurance policies, but that are correlated with insurance demand and with subsequent risk experience, then insurance buyers have private information about their risk type. Researchers can identify such cases of "asymmetrically used information" by finding individual characteristics that satisfy these conditions. This situation may occur even when insurance companies observe, or could observe, the relevant individual characteristics, but choose not to use the characteristics in pricing. Asymmetrically used information that results from insurer choices has the same implications for market equilibrium and market efficiency as asymmetric information that results when features of the contracting environment make it impossible for the insurer to observe characteristics of the insured.

Regulation is one potential source of asymmetrically used information. When insurance companies are prevented from using some individual characteristics in pricing insurance policies, insurance buyers who know these characteristics and their relationship to risk type can exploit this information in purchasing insurance. In many insurance markets, however, asymmetrically used information occurs because insurance companies choose not to use risk-related buyer information that they collect, or could collect, to set prices. We explore this ostensible puzzle in more detail below, and suggest that political

economy issues involving regulatory response and consumer backlash may contribute to this behavior.

An attractive feature of the unused observables test is that in some cases it can identify the presence of adverse selection, rather than just asymmetric information that may take the form of either adverse selection or moral hazard. When external evidence suggests that an unused observable characteristic is correlated with risk type even when there are no differences in insurance status across individuals, then finding that individuals with certain values of the unused observable select more insurance suggests the presence of adverse selection. In contrast, neither the positive correlation test nor other recently developed tests for asymmetric information in insurance markets can distinguish between adverse selection and moral hazard. This distinction is of considerable interest because adverse selection is based on ex-ante private information about risk type, while moral hazard relies on ex-post private information about behavior. The two have very different implications for public policy. By mandating insurance coverage, the government can potentially ameliorate inefficiencies produced by adverse selection, but it is unlikely to have any comparative advantage relative to the private sector in redressing inefficiencies caused by moral hazard.

We illustrate the unused observables test by applying it to a new data set containing information on the annuity policies sold by a large U.K. insurance company. The company collects information on the annuitant's place of residence, but this information is not used in setting prices. Conditional on the insurance company's risk classification, which is based on the annuitant's age and gender, place of residence contains information about an annuitant's socio-economic status that helps to predict future mortality experience. Moreover, annuitants in higher socio-economic status residential locations purchase larger annuities. Taken together, these empirical findings reject the null hypothesis of symmetric information. Moreover, since external evidence suggests that socioeconomic status and mortality are related even among individuals with identical annuity policies, the empirical findings support the presence of adverse selection in particular in the U.K. annuity market.

This paper is divided into five sections. The first describes the unused observables test and places it in the context of previous research on asymmetric information. Section two summarizes our data set,

2

which was provided by a large U.K. annuity seller. The third section presents results from the unused observables test. Section four suggests several explanations for why insurance companies often choose not to condition policy prices on all of the observable factors that are related to the insurance buyer's risk type. A brief conclusion considers the implications of our findings for equilibrium in other insurance markets.

1. Testing for Asymmetric Information in Insurance Markets

This section summarizes the standard positive correlation test for asymmetric information in insurance markets, describes our unused observables test, and discusses how this test may be used to distinguish adverse selection from moral hazard.

1.1 The Positive Correlation Test

Several authors, notably Cawley and Philipson (1999) and Chiappori and Salanie (2000), have observed that many models of equilibrium with either adverse selection or moral hazard predict that those with more insurance should be more likely to experience the insured risk. With moral hazard, insurance coverage lowers the cost of the insured outcome and thus increases the expected loss. With adverse selection, the insured knows more about risk type *ex-ante* than the insurance company does. Since the marginal utility of insurance at a given price is increasing in the risk of the insured event, those who know that they are high risk will select contracts with more insurance than those who know that they are low risk.

The positive correlation test estimates the correlation between the amount of insurance an individual buys and his ex-post risk experience, conditional on the observable characteristics that are used in pricing insurance policies. It is essential to condition on all the information that is used to set insurance prices. Finding, for example, that smokers demand more life insurance than non-smokers, and that they also have higher mortality risk, is uninformative, since the price of insurance for smokers is adjusted to reflect this differential. Results from the positive correlation test or from the unused observables test are always conditional on the risk classification that the insurance company assigns to the individual.

A canonical positive correlation test involves estimating two reduced-form econometric models: one

for insurance coverage (C) and the other for risk of loss (L). For simplicity we present linear versions of both models. The explanatory variables (X) in both equations are the set of variables that the insurance company uses to place the buyer into a risk class. The estimating equations are:

(1a) $\quad C_i = X_i * \beta + \varepsilon_i$

and

(1b) $\quad L_i = X_i * \gamma + \mu_i.$

Under the null hypothesis of symmetric information, the residuals in these two equations, ε_i and μ_i, should be uncorrelated. A statistically significant positive correlation between the two implies rejection of the null hypothesis.

The findings from the positive correlation test vary from market to market. In the health insurance market, Cutler and Zeckhauser (2000) review many studies that find a positive correlation, although there are exceptions such as Cardon and Hendel (2001) and Fang, Keane, and Silverman (2006). In the automobile insurance market, Chiappori and Salanie (2000), Dionne et al. (2001), and Chiappori et al. (2006) fail to reject the null hypothesis of symmetric information, while Cohen (2005) finds a positive correlation between risk type and insurance demand. In the life insurance market, Cawley and Philipson (1999) and McCarthy and Mitchell (forthcoming) are unable to reject the null hypothesis that insurance coverage and risk occurrence are uncorrelated. In the annuity market, Finkelstein and Poterba (2002, 2004) and McCarthy and Mitchell (forthcoming) find a positive correlation between annuity demand and longevity.

Life insurance markets and annuity markets insure opposite mortality risks. Annuities insure against the prospect of an unexpectedly long life, while life insurance insures against the prospect of an unexpectedly short one. It is therefore surprising that the positive correlation test for asymmetric information yields different results in these two markets. One possible explanation for this pattern is that insurance demand is determined not only by private information about risk type but also by heterogeneity in risk tolerance. All else equal, more risk averse individuals are likely to demand more annuity coverage and more life insurance. Wealthier individuals might also demand more insurance of both types.

4

However, risk aversion and wealth are likely to be negatively correlated with the risk of dying early, and positively correlated with the risk of living a long time, since more risk averse and wealthier individuals may invest more in life-extending activities.

When individuals have different tastes for insurance, the correlation between ε_i and μ_i can no longer be attributed only to unobserved differences in risk of loss. Imagine that individuals have private information about their risk type (Z_1). They also differ in their risk aversion (Z_2). We can then write:

(2a) $\varepsilon_i = Z_{1,i}*\pi_1 + Z_{2,i}*\pi_2 + \acute{\eta}_i$

and

(2b) $\mu_i = Z_{1,i}*\rho_1 + Z_{2,i}*\rho_2 + v_i.$

In these equations, ε_i and μ_i are the residuals from the coverage equation (1a) and the risk of loss equation (1b), respectively. The positive correlation test is based on the idea that risk type (Z_1) will be positively correlated with both insurance coverage and the risk of loss if there is private information about risk type (i.e. $\pi_1 > 0$ and $\rho_1 > 0$). Assume, however, that risk aversion (Z_2) is also positively correlated with coverage but that it is negatively correlated with risk of loss; thus $\pi_2 > 0$ while $\rho_2 < 0$. The correlation between ε_i and μ_i may therefore be negative or zero. In this case, the positive correlation test would fail to reject the null hypothesis of symmetric information even in the presence of asymmetric information about risk type.

The foregoing example illustrates how unobserved heterogeneity in individual preferences can lead to Type II errors in applications of the positive correlation test. If individuals with stronger preferences for insurance are also lower risk, then preference-based selection may offset risk-based selection, and insurance coverage and risk occurrence may be uncorrelated or even negatively correlated. De Meza and Webb (2001), Jullien, Salanie, and Salanie (2003), and Chiappori et al. (2006), among others, develop formal equilibrium models that illustrate how, when individuals have private information about risk preferences, the positive correlation test can fail to reject the null hypothesis of symmetric information, even though private information about risk type is also present and may impair market efficiency.

Several recent studies also illustrate the empirical importance of heterogeneous preferences for

insurance. In the Israeli automobile insurance market, Cohen and Einav (2005) estimate that private information about preference heterogeneity is quantitatively more important than private information about risk type. They find that private information about risk aversion is positively correlated with private information about risk type. This reinforces the positive correlation between insurance coverage and risk occurrence created by private information about risk type. Several studies in other insurance markets demonstrate that unobserved preferences for insurance can also be negatively correlated with unobserved risk type. Examples include Davidoff and Welke's (2005) analysis of the reverse annuity mortgage market, Fang, Keane and Silverman's (2006) study of the Medigap market, and Finkelstein and McGarry's (forthcoming) study of the long-term care insurance market. The last of these studies finds that equilibrium insurance coverage and risk occurrence are uncorrelated, despite the presence of private information about risk type.

Heterogeneity in individual preferences can also lead to Type I errors in applications of the positive correlation test. Whenever insurance is actuarially unfair, as it is likely to be whenever there are administrative costs or taxes on insurance companies, or whenever the insurance market is imperfectly competitive, more risk-averse individuals and those with a higher variability of outcomes in the absence of insurance will be more likely to purchase insurance. If such individuals also have a higher risk of loss, this can generate a positive relationship between insurance coverage and risk occurrence and lead to spurious rejection of the null hypothesis of symmetric information.

Private information about risk type may impair market efficiency, regardless of the sign of the correlation between the residuals in the insurance demand and risk of loss equations. The buyer's private information affects the supplier's cost of providing insurance. Consider for example an equilibrium in which two types of individuals choose to purchase the same insurance contract. One group has private information that they face an above average risk of loss, while the other group is very risk-averse but knows that it faces a below-average risk of loss. In this setting risk occurrence may be uncorrelated with insurance coverage. However, the two groups of buyers have different expected costs of insurance, yet they pay the same premium. They cannot both face their own-type actuarially fair price on the margin, so

the insurance purchases of at least one group, and possibly both, will be distorted away from the first best.

1.2 Testing for Asymmetric Information Using Unused Observables

When there is symmetric information, conditional on the risk class in which the insurance company places a buyer, there should not be any buyer characteristics that can be observed by the econometrician and that are correlated with both insurance coverage and risk of loss. Finding a characteristic that is known to the insurance buyer, unknown to or unused by the insurer, and correlated with both insurance coverage and ex-post risk of loss constitutes a rejection of symmetric information. The search for such characteristics forms the basis for the unused observables test.

The test can be formalized using the foregoing notation in which X denotes the attributes that are used to assign a potential insurance buyer to a risk class, C denotes insurance coverage and L denotes risk of loss. Let W denote a candidate unused observable variable. It could be an element of either Z_1 (risk type) or Z_2 (risk preference). The estimating equations for the unused observable test are:

(3a) $C_i = X_i * \beta + W_i * \alpha + \varepsilon_i{'}$

and

(3b) $L_i = X_i * \gamma + W_i * \delta + \mu_i{'}.$

Rejecting the joint null hypothesis that $\alpha = 0$ and $\delta = 0$ is tantamount to rejecting the null hypothesis of symmetric information, regardless of the signs of α and δ. By investigating several candidate variables for W, we can also learn which individual characteristics are related to private information in the insurance market.

Implementation of the unused observables test requires individual data on: (i) insurance coverage, (ii) ex-post risk experience, (iii) the characteristics used by insurance companies in pricing insurance, and (iv) at least one individual characteristic that is <u>not</u> used in setting prices. The positive correlation test requires the first three types of data, and the settings in which it has been applied often provide opportunities for collecting the fourth type of data. Household surveys, for example, have been used to implement the positive correlation test for asymmetric information in various insurance markets including life insurance, health insurance, and long term care insurance. Such data often include information on

unused (i.e. unpriced) observables such as – depending on the insurance market studied – wealth, parental

health history, seat belt use, and occupational risk. The other type of data that has been widely used to

implement the positive correlation test is proprietary data sets provided by insurance companies.

Proprietary insurance company data have been used in implementing the positive correlation test in, for

example, both automobile insurance and annuity markets, among others. Such data may be supplemented

with survey data that contain unused observables. For example, Hemenway (1990) conducted an in-

person survey of seat belt use and insurance purchases among rental car drivers, and Ivaldi (1996)

supplemented a French data set on automobile insurance with a survey of the insured's smoking behavior;

neither seat belt use nor smoking is used in pricing the respective insurance products. In some cases,

insurance company records may also already include variables that are collected but not used in setting

prices. For example, a policyholder's address is almost always collected and kept in the data for billing

purposes, but it is not always used in determining prices. Nonetheless, geographic information is often

correlated with individual attributes that may affect both demand for insurance and risk type.

Our unused observables test complements the tests for asymmetric information recently developed by

Chiappori et al. (2006). Their tests, like ours, were designed to counter the limitations of the positive

correlation test when there is unobserved preference heterogeneity. Their tests use data on the claims

distributions associated with two distinct insurance policies to test whether insurance buyers who select

less-generous insurance contracts when more generous ones are available receive higher expected

income, net of losses and premiums, with the less generous than the more generous policy. These tests

require making an assumption about the insurance company's cost structure, while the unused

observables test does not. In addition, these tests only apply to pairs of insurance contracts, while the

unused observables test can be applied to any number of insurance products. On the other hand, a

drawback of the unused observables test is that failure to find individual characteristics that are not used

in pricing, but that are correlated with risk of loss and demand for insurance, may simply reflect a lack of

sufficiently rich data to detect asymmetric information, rather than its absence.

1.3 Distinguishing Adverse Selection from Moral Hazard

The unused observables test, unlike alternative tests, can also provide specific evidence on the presence of adverse selection rather than just asymmetric information. Distinguishing adverse selection from moral hazard hinges on the availability of information on how risk type and unused observable characteristics are related among individuals with the same insurance coverage. Simply observing that a characteristic that insurance companies do not use in pricing is positively correlated with insurance quantity purchased and with ex post risk occurrence is consistent both with individuals selecting more insurance because they face an ex ante lower price (selection) and with the purchase of more insurance leading individuals to incur more risk (moral hazard). When there is external information that certain characteristics are correlated with risk occurrence for reasons other than insurance coverage, then the unused observables test can identify the presence of adverse selection and rule out moral hazard as the exclusive source of the observed correlation between individual attributes, insurance quantity, and risk of loss. It is possible, of course, for both moral hazard and adverse selection to be present.

The unused observables test's capacity to indicate the presence of adverse selection complements recent tests for moral hazard that exploit dynamic panel data with exogenous changes in insurance prices, such as Abbring, Chiappori, and Pinquet's (2003) study of automobile insurance in France and Israel's (2004) investigation of automobile insurance in the United States. It also complements Karlan and Zinman's (2005) field experimental design to test separately for adverse selection and moral hazard in credit card markets.

While the distinction between adverse selection and moral hazard is of some interest in the annuity market, it is likely to be of greater interest in other insurance markets. Despite the theoretical potential for moral hazard in annuity markets, as suggested by Philipson and Becker (1998), its empirical relevance is arguably weaker than in other insurance markets such as health insurance. Yet many public policies regarding health insurance markets are predicated on the assumption of at least some adverse selection. In such a setting, a direct test for adverse selection should be of substantial value.

2. Place of Residence as an Unused Observable in the U.K. Annuity Market

We apply the unused observables test to the compulsory annuity market in the United Kingdom.

Annuities are insurance contracts that pay a pre-specified payment stream to the beneficiary, the annuitant, for as long as he is alive. They provide a way of spreading an accumulated stock of resources over a remaining lifetime of uncertain length and thus provide insurance against the risk of outliving one's resources. From the perspective of an insurance company, a high risk annuitant is one who has a substantial chance of living a long time.

2.1 Insurance Company Data and Descriptive Statistics

Retirees who have accumulated savings in tax-preferred retirement saving accounts in the United Kingdom are required to annuitize a large portion of their accumulated balance. They can, however, choose among a number of annuity options that offer different amounts of insurance. There are currently no restrictions on the characteristics that U.K. insurance companies may use in pricing annuities in this market. During our sample period (1988 – 1998), Ainslie (2000) reports that the vast majority of annuities, including all of the ones sold by the company that provided data for this study, were priced solely on the annuitant's gender and age at the time of purchase. To apply the unused observables test, we need to observe not just these characteristics, but also another characteristic that is related to both survival prospects and annuity demand.

We obtained data from one of the largest U.K. annuity sellers. The data set includes information on all of the company's compulsory annuities that were in force in 1998 and that were sold between January 1, 1988 and December 31, 1998. We observe the annuitant's date of death if he died over the six-year period between January 1, 1998 and February 29, 2004. We also observe detailed information on the type of annuity purchased, and the three characteristics of the annuitant that are used in pricing the annuity: the date of purchase, the annuitant's date of birth, and the annuitant's gender. Finally, we observe a characteristic not used in pricing: the individual's post code, which indicates his place of residence.

For analytical tractability, we restrict our sample in several ways. We limit the sample to the approximately sixty percent of the sample firm's annuities that insure a single life. The mortality experience of the single life annuitant provides a convenient ex-post measure of risk type; measuring the risk type of a joint life policy which insures multiple lives is less straightforward. We also restrict the

sample to the approximately eighty percent of annuitants who hold only one annuity policy, thereby avoiding the complexity of modeling the total annuity stream for individuals who hold multiple policies. We restrict attention to the approximately ninety percent of policies sold in England or Wales because we cannot map postcodes in Scotland into the same type of geographic unit that we can for England and Wales. Finally, we exclude annuitants who purchased annuities before age 50, and those who purchased annuities with guarantee periods not equal to five or ten years; these exclusions affect less than one percent of our sample. Our final sample consists of 52,824 annuitants.

Table 1 presents summary information on our data sample. The average age at annuity purchase is 62, and 59 percent of the purchasers are male. Our sample characteristics appear to match the characteristics of the broader market, described by Murthi et al. (1999), and of other individual firms in the market, such as the one studied in Finkelstein and Poterba (2004).

2.2 Residential Location as an Unused Observable

Each postcode lies wholly within a ward. Wards consist of about 9,000 individuals, while there are about forty individuals per postcode. Our sample includes annuitants from 49,123 unique postcodes and 8,941 unique wards, out of a possible 1.24 million postcodes and 9,527 wards in England and Wales. We link the annuitant's ward to ward-level data on socio-economic characteristics from the 1991 U.K. Census. The public use version of the U.K. Census does not contain postcode-level data.

There are two measures of ward-level socio-economic status available in the U.K census: educational attainment and occupation. Educational attainment is reported as the percent of the ward population aged 18 and over that is "qualified". To be "qualified", an individual has to have an educational credential above the level of the GCE A-level standard, the equivalent of a good high school degree in the United States. Table 2 provides summary statistics on educational attainment. We report two sets of summary statistics, one weighting each ward by its population, and the other weighting each ward by the number of policies from that ward in our sample. Table 2 indicates that the average person in England and Wales comes from a ward in which about 13 percent of individuals are qualified. The average annuitant in our sample, however, comes from a ward in which about 16 percent of individuals are qualified.

The ward-level census data also report the percent of employed people in each ward in different occupational classes, which are called "social classes." We compare three groups: professional and managerial (social classes I and II), skilled manual or non manual (social class III), and partly skilled or unskilled (social classes IV and V). Table 2 shows that the average person in England and Wales comes from a ward in which about one-third of the employed individuals are in professional and managerial occupations, two-fifths are in skilled manual or non-manual occupations, and one-fifth are in partly skilled or unskilled occupations. The average annuitant in our sample is drawn from a higher social class ward than the average individual in the population. This is consistent with Banks and Emmerson's (1999) findings from the U.K. Family Resources Survey that annuitants are of higher SES than non annuitants.

Ward-level census data also provide a measure of health status: the percent of persons in the ward having a "long-term illness, health problem, or handicap which limits his/her daily activities or the work he/she can do? Include problems which are due to old age." The average person in England and Wales comes from a ward in which about 12 percent of the population reports having a long-term illness; the average annuitant in our sample comes from one in which about 11 percent of the population reports such illness. We investigate whether this ward characteristic helps to predict annuitant survival, since it represents a variable that is not directly related to socio-economic status but that may provide annuitants with private information about their mortality prospects.

Characteristics of the ward population convey some predictive information about the characteristics of a randomly drawn individual within the ward, but substantially less information than knowing the individual's own characteristics. This disparity can be calibrated by comparing the variance in a characteristic across wards to the variance of the characteristic across individuals. This ratio can range from 0, in which case the average ward characteristic contains no predictive information about the individual characteristic, to 1, in which case the ward characteristic is perfectly predictive of an individual's characteristic. The value of the ratio is 0.11 for long-term illness, 0.23 for education qualification, 0.26 for an indicator variable for membership in social class I or II, 0.14 for an indicator for social class III, and 0.21 for an indicator for Social Class IV or V. Our estimates of asymmetric

information using ward-level rather than annuitant-level SES measures therefore underestimate the annuitants' private information. Our estimates of the informational value of the characteristics of an annuitant's ward are also likely to understate the information potentially available to insurers, who observe the much finer geographic unit of postcode.

A substantial body of evidence suggests a relationship between socioeconomic status and survival rates independent of any relationship between socioeconomic status and annuity holdings. This evidence includes Attanasio and Hoynes (2000) and Menchik (1993) in the U.S. and Attanasio and Emmerson (2001) in the U.K. Cutler, Deaton, and Lleras-Muney (forthcoming) provide a recent survey of the literature. There are many potential explanations for the positive correlation between socio-economic status and survival rates. Some suggest that causality runs from attributes of socio-economic status, such as education, to mortality, while others suggest that unobserved health attributes that predict mortality are also correlated with the ability to earn income and accumulate wealth. Moral hazard, which would require differential prevalence of annuities across different SES classes and changes in survival-relevant behavior conditional on owning an annuity, is not one of the standard explanations.

Two distinct strands of empirical evidence suggest that the positive relationship between socioeconomic status and longevity exists in the absence of differential annuity coverage. First, there is evidence of a positive correlation between SES and longevity in the United States, where, as Brown et al. (2001) observe, private annuity markets are virtually non-existent. A potential caveat is that employer-provided defined benefit plans in the United States provide annuity-like payouts, and these plans are more prevalent among households with higher incomes. It is possible that behavioral responses to these plans could explain some of the observed correlation, although since publicly-provided annuities from the Social Security system are a much larger proportion of the wealth of lower income households than those with higher lifetime income, the predominant effect is likely to run in the opposite direction.

Second, the positive relationship between SES and health and longevity also exists among pre-retirement individuals who are not receiving any form of annuitized payments. Figure 1 plots cumulative survival probabilities in the U.K. for men below age 55. It indicates that survival probabilities are higher

for men in higher social classes at ages before annuitization typically occurs. Figure 2 shows that this pattern persists among individuals of retirement ages as well.

Given the correlation between socio-economic status and survival prospects among individuals with the same annuity coverage, we can interpret evidence of a link between a ward's socio-economic characteristics and annuitant product choice as support for the presence of at least some adverse selection. Such a finding does not rule out the potential presence of moral hazard as well.

3. Results of the Unused Observables Test in the U.K. Annuity Market

To apply the unused observables test to the annuity data described above, we begin by testing whether the socio-economic characteristics of the annuitant's ward contain any predictive information about the annuitant's survival probability, conditional on the other characteristics that are used in annuity pricing. We then explore the analogous conditional relationship between these socio-economic characteristics and the demand for annuities.

3.1 Geographic Location and Annuitant Survival Rates

To analyze how unused unobservables are related to survival patterns, we estimate a modified version of equation (3b), which imposed a linear relationship between risk of loss and the unused observable. In the annuity context, the "risk of loss" is the risk of survival; this is more appropriately estimated by a proportional hazard model of the length of time the annuitant lives after purchasing an annuity:

(4) $$\lambda(t, x_i, \beta, \lambda_0) = \exp(x_i'\beta)\lambda_0(t)$$

$\lambda(t, x_i, \beta, \lambda_0)$ denotes a hazard function for the probability that an annuitant with characteristics x_i dies t periods after 1998, conditional on living until t. Following Cox (1972, 1975), we estimate a continuous-time, semi-parametric, partial likelihood proportional hazard model. This allows us to estimate the β coefficients without imposing any parametric assumptions about the form of the baseline hazard function $\lambda_0(t)$. The Cox model readily handles the left truncation and right censoring in our data. In our earlier study of another company's annuitant data, Finkelstein and Poterba (2004), we obtained very similar results using the Cox model and alternative models that allow for a discrete rather than continuous

non-parametric baseline hazard as in Han and Hausman (1990). The main covariates of interest are socio-economic status measures of the annuitant's ward and the annuitant characteristics that are used in pricing.

Table 3 presents our findings. The first column only includes covariates for the annuitant characteristics used in pricing. The only coefficient shown is for the indicator variable identifying male annuitants; not surprisingly, mortality hazards are higher for males. The other covariates, single year- and age-specific indicator variables, are not reported to conserve space. The second and third columns add ward-level SES measures to this basic specification. They show that, conditional on the characteristics that are used in pricing, the socioeconomic status of the annuitant's ward is statistically significantly and positively correlated with annuitant survival. Column (2) indicates that annuitants from wards with a greater proportion of the population who are educationally qualified have a statistically significantly lower mortality hazard. Column (3) indicates that those from wards with a greater proportion in managerial and professional occupations (social class I and II) have a statistically significantly lower mortality hazard than both those in wards with a greater proportion in skilled occupations (social class III) and those in our reference category, which are wards with a greater proportion in partly skilled or unskilled occupations (social class IV and V). Finally, column (4) indicates that annuitants from wards in which a greater proportion of the population suffer from long-term illness have a statistically significant higher mortality hazard.

To illustrate the substantive importance of the coefficient estimates in Table 3, we use the estimate of the baseline hazard and the hazard model coefficients to compute the implied impact of a change in ward characteristics on the 5-year annuitant mortality rate. Table 4 shows the results. We estimate, for example, that a 65 year old male annuitant who purchases a policy in 1994 in a ward with the average proportion of qualified individuals and survives until 1998 has a 10.7 percent chance of dying within the next five years. The same individual from a ward one standard deviation above the average in the proportion educationally qualified has only a 9.7 percent chance of dying. Similarly, the same 65 year old male has

only a 9.3 percent chance of dying if he is from a ward in which the fraction of the population from managerial and professional occupations is one standard deviation above average.

Survival differences of this magnitude can affect the expected present discounted value of an annuity payout stream. We illustrate this by computing how much annual annuity payments would change if insurance companies adjusted prices in an actuarially fair way to account for the relationship we find between ward-level socio-economic status and annuitant mortality. This illustrative calculation of course ignores any demand response to such price changes. The actuarially fair annual payment from an annuity depends on the characteristics of the annuity, the annuitant mortality table used, and the interest rate. We focus on a nominal annuity with no guaranteed payments. Since we can only estimate mortality over a six year span using our data, for this illustrative calculation we use the annuitant mortality tables for the compulsory annuity market described in Finkelstein and Poterba (2002) for our baseline mortality hazard. We consider a 65 year old who purchases an annuity on January 1, 1998, and discount future annuity payments using the zero-coupon yield curve of nominal U.K. Treasury securities.

The mortality differences we estimated in Table 3 imply that if annuity payments were adjusted in an actuarially fair manner based on the proportion of the ward that is educationally qualified, eleven percent of male 65 year old annuitants would experience a payout change of at least five percent. For 65 year old women, the analogous estimate is four percent. If payments were adjusted based on the proportion of the ward in the managerial and professional class, about 17 percent of men and 8 percent of women would experience a change in annuity payments of 5 percent of more. Presumably the changes in payments associated with using annuitant SES – rather than ward-level SES – would be even higher.

3.2 Place of Residence as Predictor of Product Selection

The second component of the unused observables test requires examining whether annuitant ward characteristics are correlated with the amount of insurance purchased, conditional on the annuitant characteristics used in pricing. Since we have external evidence that SES and survival are positively related even among individuals who face the same annuity structure, evidence of a relationship between the annuitant's ward-level SES, which is not priced, and the quantity of insurance purchased provides

evidence that there is adverse selection in this market.

In the spirit of equation (3a), we relate insurance purchases and ward characteristics as follows:

(5) $C_{iw} = \alpha * X_i + \beta * WARD_w + \varepsilon_{iw.}$

In this equation, C_{iw} denotes the quantity of insurance purchased by annuitant i in ward w, and X_i denotes the annuitant characteristics that are used in pricing. As before, X_i consists of indicator variables for annuitant's gender, age at time of purchase, and year of annuity purchase. The main coefficient of interest is β, which describes the conditional correlation between a ward-level characteristic and the amount of insurance coverage purchased.

Three different aspects of the annuity contract affect the effective quantity of insurance (C_{iw}). First, the insurance amount is increasing in the initial annual annuity payment. Second, the amount of insurance depends on the tilt of the annuity payment stream. Table 1 indicates that 90 percent of the annuities pay a constant nominal payment stream; the rest provide a payment stream that increases in nominal terms over time. For a given initial annuity payment, these more "backloaded" annuities provide more insurance than annuities which pay a constant nominal amount each period. Third, the amount of insurance is decreasing in the length of the guarantee period of an annuity. During the guarantee period, the insurance company will continue to make payments to the annuitant's estate, even if the annuitant has died. Guarantees therefore decrease the effective amount of insurance in the contract, since they turn mortality-contingent payments into certain payouts during the guarantee period. Annuitants are allowed to select guarantee periods of up to 10 years. Table 1 indicates that about 82 percent of the annuitants choose guarantees, of which 89 percent are five year guarantees, and the other 11 percent are ten year guarantees.

We follow two different strategies for measuring C_{iw}, the quantity of insurance in a particular annuity contract. First, we stratify our sample of annuity contracts into sub-samples that vary on only one contract dimension, such as the amount of initial payout. We can then look at the relationship between ward-level SES and the amount of insurance purchased along the one variable dimension of insurance quantity. Specifically, we stratify the sample into constant nominal annuities with no guarantee, constant nominal annuities with 5-year guarantees and constant nominal annuities with 10-year guarantees. Within

17

each of these three sub samples, we examine the relationship between ward-level SES and the quantity of insurance as measured by the log of the initial annual annuity payment. The logarithmic transformation corrects for the skewness in the distribution of initial payments. This approach can only examine selection on one dimension of the contract at a time, while stratifying on other, potentially endogenous, dimensions.

Our second approach addresses this difficulty. It measures the quantity of insurance by combining the different features of each annuity product into a single measure of insurance quantity. A constant nominal annuity policy with a guarantee has both a bond component and an insurance component. We measure the insurance component by subtracting the present value of the bond component from the EPDV of the entire payment stream.

$$(6) \qquad \text{Quantity} = \sum_{t=1}^{T} \frac{AS_t}{\Pi^t_{j=1}(1+i_j)} - \sum_{G} \frac{A}{\Pi^t_{j=1}(1+i_j)}$$

In this equation, A denotes the annual nominal annuity payment, S_t denotes the probability that the annuitant survives until period t, G denotes the number of years in the guarantee period, and i_j denotes the expected nominal short-term interest rate at time period j. This insurance quantity measure is increasing in the amount of the initial payment and decreasing in the length of the guarantee.

Evaluating (6) requires both a table of survival probabilities (S_t) and a term structure for discounting future payments (i_j). Following Finkelstein and Poterba (2004), we use the U.K. population cohort mortality table provided by the Government Actuaries' Department to measure survival rates. This mortality table, which we condition on year of purchase, provides current and projected future mortality rates by age and sex. For the term structure of interest rates, we use data from the Bank of England on the zero-coupon yield curve of nominal U.K. Treasury securities on the first day of the month in which the annuity was purchased.

Table 5 reports results from estimating equation (5) using both the stratified-sample approach as well as the insurance quantity metric given by equation (6) to measure the quantity of insurance C_{iw}.. Because neither of the two approaches is well suited to accounting for differences in the degree of backloading of

the annuity, we restrict the sample to the 90 percent of policies that provide constant nominal payments. Our other results are robust to this sample restriction.

The four columns in Table 5 report results for different approaches to defining the dependent variable, as indicated in the column headings. The three different panels report results using different ward-level characteristics as right hand side variables. The table thus presents results from twelve separate regressions. Across all dependent variables (columns) and all ward-level measures (panels), the results suggest that individuals in wards of higher socio-economic status or better health are likely to purchase a greater quantity of annuity insurance. The only exception is the coefficient on the percentage of the ward in skilled occupations when the dependent variable is defined using the EPDV insurance quantity metric (Panel B, column 4). All the findings are statistically significant.

One concern with these results is that our sample of policies is left-truncated, since the annuitant must survive from the date of policy purchase until 1998. While such left-truncation is easily handled in the hazard model analysis in Table 3, it may bias the linear regression analysis in Table 5. We therefore verified that our results are robust to limiting the sample to the subsample of policies, about 13 percent, sold in 1998. The left truncation problem does not apply to those policies, and the basic findings for this subsample are similar to those for the full sample.

While statistically significant, the magnitude of the relationship between ward-level characteristics and annuity quantity is modest. For example, a one-standard deviation, or 8.1 percentage point, increase in the proportion of the annuitant's ward that is educationally qualified is associated with only a 0.13 to 0.22 percent increase in amount of insurance purchased. Results using the other ward-level measures are similarly small in magnitude.

Even if the substantive magnitude of the coefficients is modest, the qualitative finding that ward-level SES attributes are correlated with insurance demand, taken in conjunction with our earlier finding of a link between these characteristics and survival rates, constitutes a rejection of the null hypothesis of symmetric information. Since an external literature suggests that the positive relationship between socio-economic status and survival is not simply due to differences in insurance coverage by SES, we further

19

interpret the relationship between the SES of the annuitant's ward and the quantity of insurance purchased as supporting the presence of at least some adverse selection in this market. Individuals who face lower effective prices, because their high SES indicates they have a greater chance of living a long life than the insurance company's gender- and age-based prices assume, select annuity contracts with more insurance.

3.3 The Sources of Asymmetric Information

Our findings also provide some information about the *form* of the private information in annuity markets. The correlation between ward-level socio-economic status and annuity demand suggests that some of the selection in annuity markets may be based on socio-economic status. This may reflect "active" selection as prospective annuity buyers recognize that their socio-economic status may affect their survival prospects. It could also reflect "passive" or "preference-based" selection if socio-economic status affects demand for insurance for reasons other than its effect on longevity risk, for example because it is correlated with risk aversion. Regardless of which type of selection is operating, there are still potentially adverse efficiency consequences. Our finding that the share of the annuitant's ward reporting long-term illness is also related to the amount of insurance purchased seems to offer some support for traditional "active" selection, since long-term illness is less likely to be a marker for preferences for insurance than for risk type. However, ward-level health and socio-economic characteristics are highly correlated, which makes it difficult to determine the relative importance of these various selection factors.

A related question is whether the positive correlation between annuity quantity and annuitant survival found in earlier studies can be explained by the unused observables we have identified, or whether there are likely to be other unobservable factors underlying selection. To investigate this, we begin by replicating the previous positive correlation finding in our new annuity data. We follow the approach of Finkelstein and Poterba (2004) and estimate a proportional hazard model of length of time lived after purchasing an annuity, as in equation (4). The covariates of interest are the three annuity product characteristics that affect the quantity of insurance in the annuity contract: initial annual annuity payment, length of guarantee period, and degree of backloading. We also control for the annuitant characteristics used to price the annuity as well as the frequency of annuity payments.

The results of this replication exercise are shown in the first column of Table 6 and corroborate the presence of the positive correlation property in our annuity data. Annuitants who purchase guaranteed policies, which offer less effective insurance than policies without guarantees, display higher mortality rates (i.e. are lower risk) than annuitants who do not purchase guarantees. Those who choose larger initial annuity policies have a lower mortality risk (i.e. are higher risk). Finally, people who purchase constant nominal annuities exhibit higher mortality rates than individuals who purchase more back-loaded annuities, which provide more insurance; however, this relationship is not statistically significant.

The remaining columns of Table 6 add controls for characteristics of the annuitant's ward to the analysis in the first column. The results in columns (2) through (4) indicate that the addition of ward-level characteristics does not fully attenuate the positive correlation between dimensions of the insurance contract that provide additional coverage and ex-post risk type. This suggests that additional unobserved annuitant characteristics that we have not measured affect selection in this market.

4. The Determinants of the Attributes Used in Pricing Insurance Policies

Our empirical results suggest that U.K. insurance companies do not use all of the publicly available information that is related to mortality risk and annuity demand in pricing annuities. This raises the question of why insurers forgo pricing on observable, relevant information. A similar question arises in many other insurance markets. For example, automobile insurance companies often choose not to price on occupation or geographic location, even though both have been shown to be related to claims. In addition, in the French automobile insurance market, Ivaldi (1996) finds that automobile accident rates are higher for smokers than for non-smokers in France, by as much as the difference in accident rates between men and women, yet insurance is not priced on the basis of smoking status. In the U.S. automobile insurance market, Carter (2005) reports that most insurers use simple pricing formulae based on a driver's age and place of residence, although some insurers are now moving toward more sophisticated pricing structures that condition premiums on driving history and demographics as well as other characteristics, such as credit history and the driver's occupation. In the long-term care insurance market, Brown and Finkelstein (2004) find that premiums are constant across place and gender, even

though there are substantial differences in expected nursing home utilization related to these characteristics.

In general, one would expect insurers to use a characteristic of the insured in pricing whenever the costs of collecting the information is less than the incremental profitability of using the variable in differentiating prices. This simple rule may fail if insurance regulation prohibits the use of a given attribute. For example, many states restrict the characteristics that may be used in pricing automobile insurance. In such cases, it is relatively uninteresting to test for asymmetric information, which is likely to exist, but it is important to measure the extent of asymmetric information created by regulation and to quantify its efficiency effects. The unused observables test can play an important role in this endeavor.

The more interesting cases involve information on individual characteristics that insurance companies could collect and use in setting prices, but that they choose not to use. The examples from the U.K. annuity market, the U.S. long-term care insurance market, and the French automobile insurance market discussed above are but a few illustrations of such cases. The puzzle of unexploited information is particularly acute for variables such as gender and geography that are collected by default. Costs of information acquisition cannot explain the limited use of such data.

We consider four potential explanations for why insurance companies choose not to use information that they collect, or could collect at low cost, in pricing insurance. While a definitive explanation is beyond the scope of this paper, political economy concerns regarding potential regulation appear to be an important element. Before considering this possibility, we discuss three other potential explanations which, in our reading of the available evidence, do not appear to be of primary importance.

First, insurance companies may choose not to use easily available, relevant information in pricing if such information is not quantitatively important in improving the prediction of loss outcomes. While this may explain why some buyer characteristics are not used in pricing, our estimates suggest that the association between ward-level SES and annuitant mortality is large enough to translate into non-trivial changes in payouts for a substantial fraction of annuity buyers. Presumably the relationship between annuitant SES, a variable which is not currently measured but which could be collected, and annuitant

mortality is even larger. There are large disparities based on non-priced attributes in other markets, too. For example, Brown and Finkelstein (2004) document that the unisex pricing of long-term care insurance generates gender-based effective load differentials valued at nearly half of the policy cost.

Second, the predictive content of characteristics such as place of residence may be limited by the extent to which such characteristics are subject to change in response to characteristic-based pricing differentials. For a sufficiently large difference in expected annuity payments based on geographic location, would-be annuitants might try to modify their mailing address or even to move. While this may explain why place of residence is not used in pricing various types of insurance, it seems unlikely to be a general answer to the lack of characteristic-based pricing. There are a number of difficult to change characteristics, such as educational attainment in the annuity market and gender in the long-term care insurance market, that are strongly correlated with the risk of loss but are not used in pricing.

Third, using a previously-unexploited buyer characteristic for pricing may involve considerable up-front investment to determine the appropriate pricing structure, and competitors may copy the pricing rule without incurring the initial costs. We discussed this possibility with a number of executives at U.K. insurance companies, who addressed our questions by reference to the emerging "impaired life" annuity market that has developed in the U.K. since the end of our sample period. Firms in this market offer substantial discounts to smokers and other individuals who are likely to be in poor health. The initial pricing of these impaired life products involved both considerable investment in actuarial analysis and product development. One of the developers of impaired life products analyzed a database of medical records from life insurance sales around the world to try to predict the relationship between various medical conditions and annuitant mortality. Another impaired life annuity seller contracted with one of the U.K. health authorities for their data on the mortality of individuals in nursing homes and hospitals and then devoted considerable resources to analyzing these data to derive relationships between mortality and health conditions. Even with these efforts, there was substantial uncertainty surrounding initial estimates of the prices at which impaired life annuities would break even.

While the firms that introduced impaired life annuities were concerned about pricing errors, they do

not appear to have worried about other firms free-riding on their pricing decisions without paying the costs of determining the appropriate pricing structure. Executives at firms that created the impaired life annuity business explained to us that one of the incentives to enter this market early is to build up statistical experience that can be used to refine subsequent pricing. Early entrants thus apparently gain an informational advantage relative to later-entering competitors. More importantly, simply observing the new policy's pricing structure will not enable a competitor to mimic the innovator's sales practices. Potential imitators will not know the innovator's underwriting rule, and the criteria that it uses to deny coverage to some applicants. Underwriting rules can play an important role in markets such as this. Firms that emulate the innovator by introducing policies with similar pricing would likely suffer from a form of the winner's curse in which individuals who were denied policies by the innovator obtain insurance from the emulator, making the emulator's risk pool less profitable than that of the innovator.

These discussions led us to rule out the risk of emulation as a primary factor discouraging the use of additional information in insurance pricing, at least in the annuity market. Nevertheless, changing practices at other insurers is likely to reduce the profitability of any innovation in policy pricing. If the innovator's rivals ultimately adopt pricing rules that condition on the newly-exploited individual characteristics, the result may simply be an equilibrium in which all firms incur more up-front costs in pricing insurance policies. The gain in profitability in such a setting may be much smaller than the gain to a monopoly insurer using new information in pricing.

Finally, we consider the political economy issues surrounding the use of new individual characteristics for insurance pricing. It appears that introducing more refined pricing distinctions can have large public relations costs for individual firms and for the insurance industry. It may also trigger regulatory changes to ban the use of such information in pricing. Surprisingly, this appears to be the case even in a market in which it is the wealthy who would face the higher prices. Several U.K. insurers recently considered adjusting annuity prices based on the annuitant's postcode. There was a sharply negative public reaction to such proposals, illustrated by newspaper stories on "Postcode Prejudice" (*Sunday Times*, July 13 2003), and "Postcode Peril" (*Manchester Evening News* July 7, 2003). Insurance

firms contemplating more refined pricing may be concerned about the direct costs of negative publicity, as well as by the prospect of triggering new regulatory initiatives in the largely unregulated annuity market. Mohl (2005) illustrates the same adverse public reaction to proposals to use more variables describing a driver's lifestyle in pricing automobile insurance in U.S. states.

The behavior of large and small firms provides some support for these political economy concerns. Adverse publicity and fear of future regulation should have less impact on small firms or new entrants who do not internalize the costs of increased regulation or lost good will to the same extent that large existing firms do. Consistent with this, Ainslie (2000) reports that impaired life annuities were introduced to the U.K. market by new, start-up companies formed expressly for the purpose of offering the impaired annuity products to individuals in observably poor health. Incumbent firms did not follow suit, until, about five years after the introduction of these products, the impaired life market had grown to the point where the cream skimming of good risks by the impaired life companies created pressure on the existing companies to expand their pricing system. By that point, the political economy costs of offering impaired life annuities had presumably declined as the public had become accustomed to such products.

5. Conclusion

This paper applies a new test for asymmetric information in insurance markets that we label the unused observables test. The test is based on the identification of individual characteristics that are not used in policy pricing but that are correlated with both insurance coverage and risk occurrence. It offers a more robust approach to testing for asymmetric information than the standard positive correlation test that has been used in much of the previous literature. In some settings, it can also identify the presence of adverse selection directly, rather than merely providing a joint test for the presence of either adverse selection or moral hazard.

We implement this new test using information on the place of residence of U.K. annuitants as an unused observable that is not used in pricing. We show that the socio-economic characteristics of an annuitant's geographic location are correlated with both his survival probability and the amount of insurance he purchases on average. This provides evidence of asymmetric information. External

evidence of a positive relationship between socio-economic status and longevity among individuals with similar annuity coverage allows us to further interpret our findings as supporting the presence of adverse selection in this market. Part of the private information available to annuity buyers consists of information about their socio-economic status. Insurance companies could reduce the extent of such private information by conditioning policy premiums on where annuity buyers live. They choose not to do so for reasons that we suspect involves potential public backlash and regulatory risk.

Our findings have implications beyond the operation of annuity markets. There is no a priori reason to expect socio-economic selection to operate in the same direction in all insurance markets, and empirical evidence suggests that it does not. In the annuity market, our findings suggest that socio-economic selection draws longer-lived, and therefore higher risk, individuals into the market. It therefore reinforces any selection based directly on private information about risk type. In the life insurance market, Banks and Tanner (1999) find that selection based on socio-economic status also appears to draw longer-lived individuals into the market. Such individuals, however, are low-risk life insurance buyers. These patterns of socio-economic selection may help explain why individuals in the life insurance market do not appear to be higher risk than the population at large, while annuitants do appear to be higher risk than the general population. Socio-economic selection may also help more generally in explaining the observed disparities across different insurance markets in the correlations between risk of loss and the quantity of insurance purchased.

A complete understanding of the limited use in pricing of available or collectible risk-related information on insurance buyers remains an open issue. Our reading of the available evidence suggests that the political economy of insurance regulation may play an important role in determining the pricing function. Studying the history of characteristic-based pricing of insurance policies, and the evolution of such pricing in various markets, may offer further insights into how insurance companies decide which variables to use in setting prices. Comparing insuring prices in states with elected and appointed insurance commissioners, for example, might offer insights on the role of endogenous regulation in affecting pricing behavior. These issues are left for future study.

References

Abbring, Jaap, Pierre-Andre Chiappori, and Jean Pinquet. 2003. "Moral hazard and dynamic insurance data." *Journal of the European Economic Association* 1: 767-820.

Ainslie, Ross. 2000. "Annuity and Insurance Products for Impaired Lives." Presented to the Staple Inn Actuarial Society, May 9.

Attanasio, Orazio and Carl Emmerson. 2001. *Differential Mortality in the UK*. Institute for Fiscal Studies working paper W01/16. http://www.ifs.org.uk/wps/wp0116.pdf

Attanasio, Orazio and Hilary Hoynes. 2000. "Differential mortality and wealth accumulation," *Journal of Human Resources* 35(1): 1-29.

Association of British Insurers. 1999. Insurance Statistics Year Book. London: ABI.

Banks, James, and Carl Emmerson, 1999. "U.K. Annuitants." *IFS Briefing Note.* London: Institute for Fiscal Studies.

Banks, James and Sarah Tanner. 1999. "Household Savings in the UK." *IFS Reports*, R62. London: Institute for Fiscal Studies.

Brown, Jeffrey and Amy Finkelstein. 2004. "Supply or Demand: Why is the Market for Long-Term Care Insurance So Small." NBER Working Paper 10782.

Brown, Jeffrey, Olivia Mitchell, James Poterba and Mark Warshawsky. 2001. The Role of Annuity Markets in Financing Retirement. MIT Press: Cambridge, MA.

Cardon, James and Igal Hendel. 2001. "Asymmetric information in Health Insurance: Evidence from the National Medical Expenditure Survey." *Rand Journal of Economics* 32: 408-27.

Carter, Adrienne. 2005. "Telling the Risky from the Reliable." *Business Week* August 1, 57.

Cawley, John and Tomas Philipson. 1999. "An empirical examination of information barriers to trade in insurance." *American Economic Review* 89: 827-846.

Chiappori, Pierre-Andre. 2000. Econometric Models of Insurance under Asymmetric Information. In Georges Dionne, ed. Handbook of Insurance Economics. London: Kluwer, 365-394.

Chiappori, Pierre-Andre and Bernard Salanie. 2000. "Testing for Asymmetric Information in Insurance

Markets." *Journal of Political Economy*, 108, 56-78.

Chiappori, Pierre-Andre, Bruno Jullien, Bernard Salanie, and Francois Salanie. 2006. "Asymmetric

Information in Insurance: General Testable Implications." *Rand Journal of Economics.*

Cohen, Alma. 2005. "Asymmetric Information and Learning in the Automobile Insurance Market."

Review of Economics and Statistics 87: 197-207.

Cohen, Alma and Liran Einav. 2005. "Estimating Preferences from Deductible Choice." Working Paper,

Stanford University.

Cox, David. 1972. "Regression models and life-tables (with discussion)." *Journal of the Royal Statistical*

Society, Series B 34: 1987-220.

Cox, David. 1975. "Partial Likelihood." *Biometrika* 62: 269-276.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. forthcoming. "The Determinants of

Mortality." *Journal of Economic Perspectives.*

Cutler, David and Richard Zeckhauser. 2000. The Anatomy of Health Insurance, in The Handbook of

Health Economics, A. J. Culyer and J. P. Newhouse, editors.

Davidoff, Thomas and Gerd Welke. 2004. "Selection and moral hazard in the reverse mortgage market."

Unpublished mimeo, University of California – Berkeley.

de Meza, David and David Webb. 2001. "Advantageous selection in insurance markets." *Rand Journal*

of Economics. Summer, 249-262.

Dionne, Georges, Christian Gourieroux and Charles Vanasse. 2001. "Testing for evidence of adverse

selection in the automobile insurance market: a comment." *Journal of Political Economy* 109(2):

444-451.

Fang, Hanming, Michael Keane, and Dan Silverman. 2006. "Sources of Advantageous Selection:

Evidence from the Medigap Insurance Market," mimeo, Yale University Department of

Economics.

Finkelstein, Amy and Kathleen McGarry. forthcoming. "Multiple Dimensions of Private Information:

Evidence from the Long-Term Care Insurance Market." *American Economic Review*.

Finkelstein, Amy and James Poterba. 2002. "Selection Effects in the Market for Individual Annuities: New Evidence from the United Kingdom." *Economic Journal* 112, 28-50.

Finkelstein, Amy and James Poterba. 2004. "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market." *Journal of Political Economy* 112, 183-208.

Han, Aaron and Jerry Hausman. 1990. "Flexible Parametric Estimation of Duration and Competing Risk Models." *Journal of Applied Econometrics* 5:1-28.

Hemenway, David. 1990. "Propitious Selection." *Quarterly Journal of Economics* 105 (4): 1063-1069.

Israel, Mark, 2004. "Do We Drive More Safely When Accidents are More Expensive? Identifying Moral Hazard from Experience Rating Schemes." Mimeo, Kellogg School, Northwestern University.

Ivaldi, Marc. 1996. "Sinistralité des jeunes conducteurs novices, incitations a la sécurité routière et tarifications de l'assurance automobile." [Accident rates of young drivers, incentives to road safety, and auto insurance pricing.] Toulouse, IDEI Report.

Jullien, Bruno, Bernard Salanie, and Francois Salanie. 2003. "Screening Risk-Averse Agents Under Moral Hazard: Single Crossing and the CARA Case," mimeo, University of Toulouse. .

Karlan, Dean and Jonathan Zinman. 2005. "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment." Working Paper, Yale University.

McCarthy, David and Olivia Mitchell, forthcoming, "International Adverse Selection in Life Insurance and Annuities." In S. Tuljapurkar, N. Ogawa, and A. Gautheir, eds., Riding the Age Wave: Responses to Aging in Industrial Societies. Amsterdam: Elsevier.

Menchik, Paul. 1993. "Economic Status as a Determinant of Mortality Among Black and White Older Men: Does Poverty Kill?," *Population Studies* 47, 427-436.

Mohl, Bruce. 2005. "What Makes a Good Driver? Insurers Often Say Lifestyle, Job." *Boston Globe* July 24, F1-F5.

Murthi, Mamta, J. Michael Orszag, and Peter Orszag. 1999. The Charge Ratio on Individual Accounts: Lessons for the U.K. Experience. Birkbeck College, University of London Working Paper 99-2.

Office for National Statistics. 1997. *Health Inequalities.* London: The Stationary Office Limited.

Philipson, Tomas and Gary Becker. 1998. "Old-Age Longevity and Mortality-Contingent Claims."

    Journal of Political Economy 106: 551-573.

Rothschild, Michael and Joseph Stiglitz. 1976. "An Essay on the Economics of Imperfect Information."

    Quarterly Journal of Economics 90: 629-649.

Smith, James. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation between Health and

    Economic Status," *Journal of Economic Perspectives* 13: 145-166.

Table 1: Summary Statistics on Annuitant Population at Sample Firm

Number of policies	52,824
Number (%) of annuitants who die within sample period	5,592 (10.6%)
Number (%) of annuitants who are male	31,329 (59.3%)
Average age at purchase	62.2
Number (%) of policies that are constant nominal payout	47,370 (89.7%)
Number (%) of policies that have guarantees	43,259 (81.9%)
Mean initial payment (£)	1,819
Median initial payment	901
Standard deviation of initial payment (£)	3,682
Average premium	19,550

Note: The sample consists of single life compulsory annuities sold between 1988 and 1998 that were still in force in 1998. The text describes further sample restrictions. Mortality experience covers the period January 1 1998 through February 29, 2004. Policies that do not have constant nominal payouts have payouts that increase over time in nominal terms. Policies with guarantees continue to make payments to annuitant estate if the annuitant dies during the guarantee period. Premium and initial payment are converted to £1998 using annual values of the Retail Prices Index (RPI).

Table 2: Summary Statistics on Ward-Level Socio-Economic Status and Health Status

Social Economic Status Measure	Population-weighted		Annuitant-weighted	
	Average	Std. Dev	Average	Std. Dev.
Qualified	13.4%	8.00	15.9%	8.15
Social Class: Professional and Managerial (I & II)	31.6	12.13	36.1	12.13
Social Class: Skilled (III)	43.6	6.95	41.7	7.48
Social Class: Partly Skilled or Unskilled (IV & V)	21.6	8.03	19.4	2.47
Presence of Long-term illness	12.1	3.44	11.4	3.12

Note: Based on ward-level statistics from 1991 UK census. Population-weighted estimates are constructed weighting each ward by its population; annuitant-weighted estimates are constructed weighting each ward by the number of policies the sample firm has in that ward. The omitted social class, which consists of those in the armed forces, receiving annuity payments through government schemes, and "unknown," accounts for 3 percent (2.8 percent) of the population-weighted (annuitant-weighted) sample.

Table 3: Hazard Models Relating Annuitant Mortality Experience to Annuitant Gender and Ward-Level SES Characteristics

	(1)	Education (2)	Occupation (3)	Illness (4)
Male	0.638*** (0.0349)	0.629*** (0.0347)	0.628*** (0.0348)	0.636*** (0.0347)
Percentage of Ward that is Educationally Qualified		-0.0150*** (0.0017)		
Percentage of Ward in Professional or Managerial Occupations (Social Class I & II)			-0.0118*** (0.0017)	
Percentage of Ward in Skilled Occupations (Social Class III)			-0.0029 (0.0027)	
Percentage of Ward with Long Term Illness				0.0248*** (0.0043)

Note: Coefficients are from Cox Proportional Hazard Model of time lived since 1998 (see equation 4). N = 52,824. In addition to the covariates shown in the table, all regressions contain indicator variables for age at purchase and year of purchase. Heteroscedasticity-robust standard errors clustered at the ward level are in parentheses. In column 3, the omitted category is percentage of ward in partly skilled or unskilled occupations (Social Class IV or V). ***, **, * denotes statistical significance at the 1 percent, 5 percent and 10 percent level respectively.

Table 4: The Effect of Varying Ward Characteristics on Implied Five-Year Mortality Rates for Annuitants

	Fraction of Ward Qualified		Fraction of Ward in Social Class I or II		Fraction of Ward with Long-Term Illness	
	Average	One Std Dev Above Average	Average	One Std Dev Above Average	Average	One Std Dev Below Average
Male	10.7	9.7	10.7	9.3	10.9	10.2
Female	4.3	3.7	4.3	3.7	4.2	3.8

Notes: Table reports the post-1998 5-year cumulative mortality probability of an individual who purchased an annuity at age 65 in 1994, conditional on having survived until 1998. Cumulative mortality probabilities are derived from the coefficient estimates in Table 3 and the corresponding estimate of the baseline hazard (not reported). For the change in the proportion of the ward in Social Class I or II, the individuals are moved to Social Class IV or V.

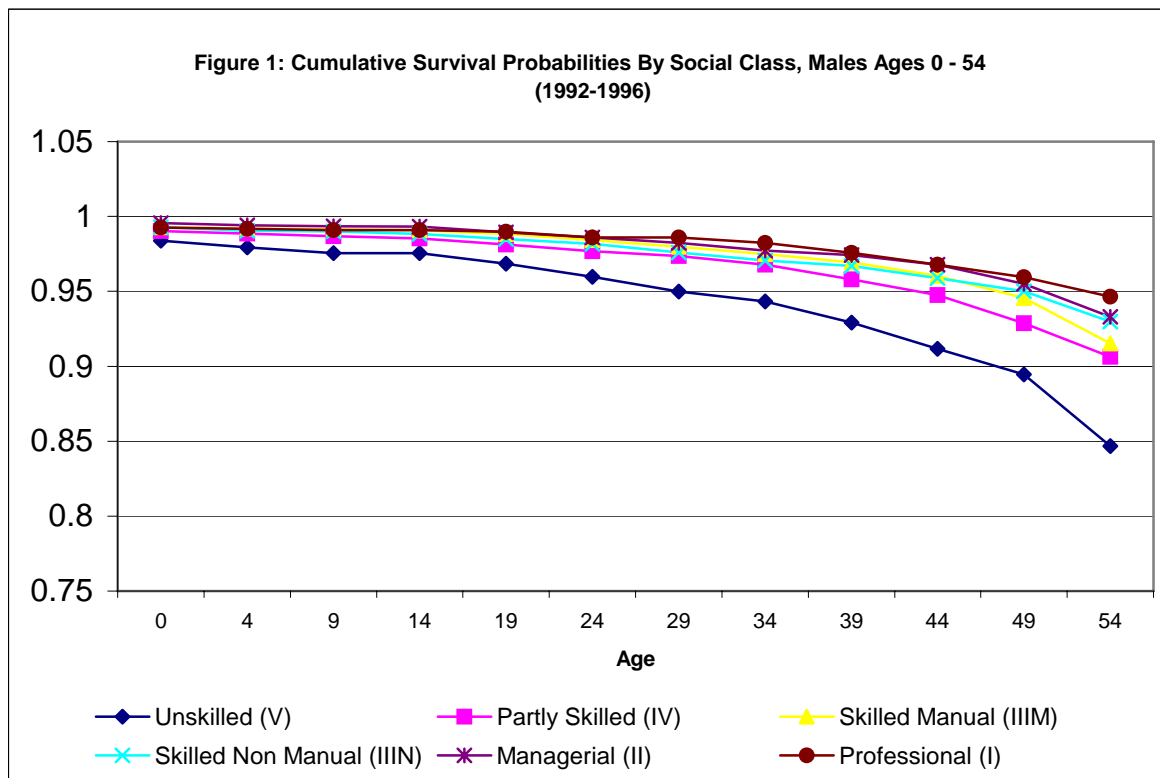Table 5: Ward SES Characteristics and Quantity of Insurance Purchased

	Dependent Variable: Log Initial Payment			Dependent Variable: EPDV of Annuity Component of Policy
	Policies with No Guarantee	Policies with 5-Year Guarantee	Policies with 10-Year Guarantee	
	[N=7,964]	[N=35,042]	[N=4,366]	[N=47,370]
Panel A: Education measure of SES				
Percentage of Ward that is Educationally Qualified	0.0223*** (0.0017)	0.0271*** (0.0011)	0.0160*** (0.0022)	223.4*** (12.6)
Panel B: Occupational Measure of SES				
Percentage of Ward in Professional or Managerial Occupation (Social Class I & II)	0.0154*** (0.0018)	0.0201*** (0.0013)	0.0103*** (0.0022)	136*** (9.6)
Percentage of Ward in Skilled Occupations (Social Class III)	-0.0012 (0.0029)	-0.0010 (0.0020)	-0.0054 (0.0035)	-49.2*** (16.4)
Panel C: Health Measure				
Percentage of Ward with Long-Term Illness	-0.0373*** (0.0046)	-0.0438*** (0.0029)	-0.0284*** (0.0052)	-330.1*** (25.3)
Mean Dependent Variable				
(£1998)	6.63	6.30	7.23	8,842

Note: The table reports OLS estimates of equation (5) on the sample of policies with constant nominal payouts. Different columns report results using different dependent variables; they are all measured in constant, 1998 £'s. Different panels report results using different ward characteristics on the right hand side. Each cell (defined by a column and a panel) reports a coefficient from a different regression. In addition to the covariates shown in the table, all regressions include indicator variables for age and year of purchase and for gender of annuitant. In panel B, omitted category is partly or unskilled social class (Social Class IV or V). Standard errors, heteroscedasticity-robust and clustered at the ward level to allow for within-ward correlation in the error term, are shown in parentheses. ***, **, * denotes statistical significance at the 1 percent, 5 percent, and 10 percent levels respectively.
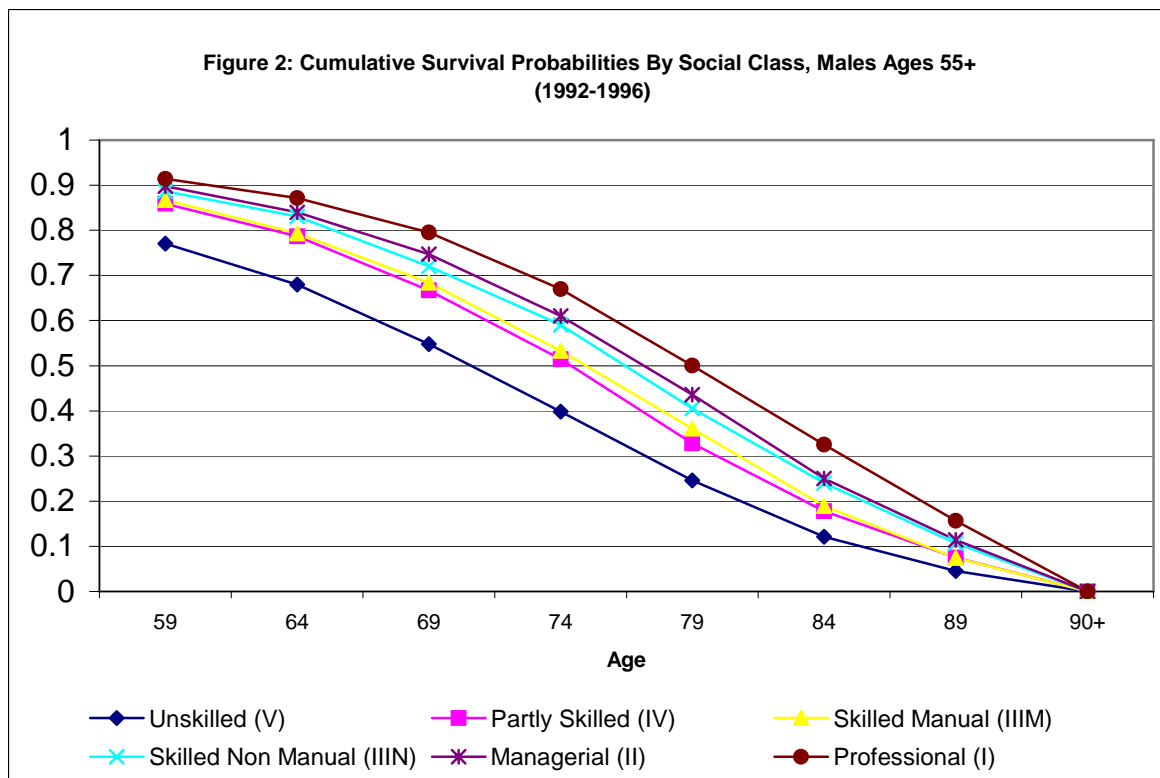
TABLE 6 – Hazard Model Related Mortality to Annuity Policy Characteristics and Ward Characteristics

	(1)	(2)	(3)	(4)
Male	0.630***	0.621***	0.628***	0.620***
	(0.0355)	(0.0354)	(0.0354)	(0.0355)
Constant Nominal Indicator	0.047	0.048	0.045	0.050
	(0.049)	(0.049)	(0.049)	(0.049)
Guarantee Indicator	0.083**	0.076*	0.079**	0.076*
	(0.0391)	(0.0400)	(0.0400)	(0.0400)
Initial Payment (£1,000)	-0.013***	-0.009	-0.012**	-0.009
	(0.0040)	(0.0058)	(0.0059)	(0.0058)
Percentage of Ward that is Educationally Qualified		-0.014***		
		(0.0018)		
Percentage of Ward with Long Term Illness			0.024***	
			(0.0043)	
Social Class (Omitted Category = % of Ward in Partly or Unskilled Occupations)				
Percentage of Ward in Professional or Managerial Occupation (Social Class I & II)				-0.011***
				(0.0017)
Percentage of Ward in Skilled Occupations (Social Class III)				-0.003
				(0.0027)

Note: Coefficients are from Cox Proportional Hazard Model of time lived since 1998 (see equation 4). N = 52,824. In addition to the covariates shown in the table, all regressions contain individual dummies for age at purchase and year of purchase (1988-1998) and frequency of annuity payments. Standard errors are in parentheses. They are heteroscedasticity-robust standard errors and are clustered at the ward level to allow for within-ward correlation in the error term. ***, **, * denotes statistical significance at the 1 percent, 5 percent and 10 percent level respectively. Reference category for "constant nominal indicator" is a more backloaded annuity. In column 4, the omitted category is Social Class IV & V (partially skilled or unskilled occupation).

Figure 1: Cumulative Survival Probabilities By Social Class, Males Ages 0 - 54 (1992-1996)

Source: ONS (1997).

Figure 2: Cumulative Survival Probabilities By Social Class, Males Ages 55+ (1992-1996)

Source: ONS (1997).