

# Effect of Insurance on Mortality in an HIV-Positive Population in Care

Dana P. GOLDMAN, Jayanta BHATTACHARYA, Daniel F. McCAFFREY, Naihua DUAN,  
Arleen A. LEIBOWITZ, Geoffrey F. JOYCE, and Sally C. MORTON

---

As policymakers consider expanding insurance coverage for individuals infected with human immunodeficiency virus (HIV), it is useful to ask if insurance has any effect on health outcomes and, if so, whether its magnitude has changed with recent efficacious but expensive treatments. By using data from a nationally representative cohort of HIV-infected (HIV+) persons receiving regular medical care, we estimate the impact of insurance on mortality in this population. A naïve single-equation model confirms the perverse result found by others in the literature—that insurance increases the probability of death for HIV+ patients. We attribute this finding to a correlation between unobserved health status and insurance status in the mortality equation for two reasons. First, the eligibility rules for Medicaid and Medicare require HIV+ patients to demonstrate a disability, almost always defined as advanced disease, to qualify. Second, if unobserved health status is the cause of the positive correlation, then including measures of HIV+ disease as controls should mitigate the effect. Including measures of immune function (CD4 lymphocyte counts) reduces the effect size by approximately 50%, although it does not change sign. To deal with this correlation, we develop a two-equation parametric model of both insurance and mortality. The effect of insurance on mortality is identified through the judicious use of state policy variables as instruments (variables related to insurance status but not mortality, except through insurance). The results from this model indicate that insurance does have a beneficial effect on outcomes, lowering the probability of 6-month mortality by 71% at baseline and 85% at follow-up. The larger effect at follow-up can be attributed to the recent introduction of effective therapies for HIV infection, which have magnified the returns to insurance for HIV+ patients (as measured by mortality rates).

KEY WORDS: AIDS; Bivariate normal; Instrumental variable; Potential outcome; Probit; Switching regression.

---

## 1. INTRODUCTION

Treatment of human immunodeficiency virus (HIV) infection poses tremendous challenges to our public and private health care financing system. With the advent of efficacious—but expensive—new therapies, it is possible that health outcomes for HIV-infected (HIV+) persons will be very responsive to the availability of insurance. The HIV Cost and Services Utilization Study (HCSUS) assembled the first nationally representative cohort of HIV-infected persons receiving regular medical care (Bozzette et al. 1998; Shapiro et al. 1999a). This dataset provides a unique opportunity to produce national estimates of the impact of insurance on health outcomes, particularly mortality.

Measuring the efficacy of insurance in deferring mortality among HIV+ patients is complicated by the association between disease severity and insurance availability. For example, in the early stages of the epidemic, HIV affected a relatively affluent homosexual male population. Many were

employed and had private insurance at the time of infection. However, as their health deteriorated, they were forced to leave their jobs and they lost their insurance. Eventually they qualified for public insurance through Medicaid or Medicare, but eligibility rules for these programs require HIV+ patients to demonstrate a disability—almost always associated with advanced disease.

These patterns of insurance coverage pose a challenge for analysts. Indeed, in Lancaster and Intrator's (1998) study of the hospitalization rates for HIV+ patients using longitudinal data from a large multisite study, they found the perverse result that health insurance increases the risk of death for HIV+ individuals. Lancaster and Intrator were appropriately suspicious of this result, calling it "striking," and identified it as an area for further investigation. In this article, we attempt to estimate the true effect of insurance on mortality by modeling both simultaneously.

Our article has applicability to a broader set of issues in medicine. Much clinical research is devoted to analyzing the effects of various treatments on health outcomes. Typically, this is done with controlled clinical trials that randomize patients to treatment group. However, for various reasons, including cost and ethical concerns, randomized trials often are infeasible or too narrowly defined to be useful to policymakers. Researchers can circumvent these problems by using observational data in studies that evaluate outcomes. However, as McClellan, McNeil, and Newhouse (1994) and others have argued, such studies may be subject to bias—differences in outcomes between treatment groups may reflect underlying differences in characteristics that are not measured and cannot be controlled for via modeling. In this quasi-experimental setting, analysts typically view these unobservable characteristics

---

Dana P. Goldman is Senior Economist, RAND, Santa Monica, CA 90407 (E-mail: [Dgoldman@rand.org](mailto:Dgoldman@rand.org)). Jayanta Bhattacharya is Associate Economist, RAND, Santa Monica, CA 90407 (E-mail: [Jay@rand.org](mailto:Jay@rand.org)). Daniel F. McCaffrey is Statistician, RAND, Santa Monica, CA 90407 (E-mail: [Danielm@rand.org](mailto:Danielm@rand.org)). Naihua Duan is Professor in Residence, School of Medicine, UCLA, Los Angeles, CA 90095. Arleen A. Leibowitz is Professor, School of Public Policy and Social Research, UCLA, Los Angeles, CA 90095 (E-mail: [Arleen@ucla.edu](mailto:Arleen@ucla.edu)). Geoffrey F. Joyce is Associate Economist, RAND, Santa Monica, CA 90407 (E-mail: [Gjoyce@rand.org](mailto:Gjoyce@rand.org)). Sally C. Morton is Senior Statistician, RAND, Santa Monica, CA 90407 (E-mail: [Smorton@rand.org](mailto:Smorton@rand.org)). The HIV Cost and Services Utilization Study is being conducted under cooperative agreement U01HS08578 between RAND and the Agency for Health Care Policy and Research. Substantial additional funding for this cooperative agreement was provided by the Health Services Resources Administration, the National Institute for Mental Health, the National Institute on Drug Abuse, and the National Institutes of Health Office of Research on Minority Health through the National Institute for Dental Research. Additional support was provided by the Robert Wood Johnson Foundation, Merck and Company, Glaxo-Wellcome, Inc., the National Institute on Aging, and the Office of the Assistant Secretary for Planning and Evaluation in the U.S. Department of Health and Human Services. The authors thank Afshin Rastegar for research assistance and three anonymous reviewers and an associate editor for their useful comments.

as reflecting health status differences between the treatment and control groups.

Few observational studies of health care interventions deal effectively with unobserved disease severity bias. An exception is McClellan et al., who apply instrumental variables (IV) techniques to Medicare claims data to analyze the effect of more intensive treatment on the mortality of elderly acute myocardial infarction patients. In IV estimation, analysts use variables that influence treatment choice, but not health outcomes, to purge the estimates of bias on the margin. Essentially, this amounts to quasi-randomization of patients to groups that differ in their likelihood of receiving treatment.

We develop a similar but distinct approach that may be used in a nonlinear setting where both the outcome and the treatment variable are binary, namely, a joint model of insurance and mortality. In constructing our model, we use state policy variables that affect the ease with which patients can obtain public insurance as instruments. These variables are clearly related to patient insurance status but should not directly affect death rates (except through insurance). By allowing arbitrary correlation between insurance and mortality in a parametric setting, and choosing our instrumental variables judiciously, we account for bias that arises because the sickest patients get public insurance.

Why might insurance status affect mortality? Several antiretroviral drugs were developed to treat HIV disease. The earliest effective drugs were the nucleoside reverse transcriptase inhibitors, such as zidovudine (AZT). More recently, newer drugs, such as protease inhibitors and nonnucleoside reverse transcriptase inhibitors, have come into clinical practice. Certain combinations of these drugs—commonly referred to as highly active antiretroviral therapies (HAART)—have been shown to reduce mortality in HIV+ patients (Palella et al., 1998). A panel of clinicians recommended the initiation of HAART even before the development of AIDS symptoms for many HIV+ patients (Carpenter et al., 1998). However, HAART costs roughly \$10,000 per year of treatment (United States General Accounting Office, 1998), so it not surprising that patients with insurance are much more likely to initiate therapy (Shapiro et al., 1999b). Therefore, with these drugs, it is reasonable to expect that patients with insurance are less likely to die.

In addition, one may ask why estimating the true effect of insurance on mortality is interesting. Clinicians are probably more concerned with the direct effect of new treatments on mortality. Such questions, although important, are best addressed via randomized trials. Policymakers, however, cannot affect treatment directly without interceding into the patient-provider relationship. This is something they are probably reluctant to do. In the context of HIV, then, most policy efforts have focused on extending Medicaid coverage to all HIV+ patients. For example, the Work Incentives Improvement Act of 1999 authorizes demonstration programs to provide Medicaid coverage to nondisabled workers. Our strategy allows us to predict the effects of novel policies, such as extending insurance to all HIV+ patients, that cannot be addressed with reduced form estimates alone.

The next section describes the innovative HCSUS dataset used to address this issue. Following that, we provide some

context by presenting a naïve model that treats insurance as a covariate, not as a joint outcome. We then present our joint model of insurance and mortality and provide the details of its estimation and fit diagnostics. The final sections discuss the limitations of our analysis and conclusions. Appendixes A and B detail our computation of standard errors given the clustered sample design and detail our sensitivity analyses.

## 2. DATA

Our goal is to measure the effect of health insurance on the mortality of the HIV+ population in care. As the only nationally representative dataset of HIV+ patients, the HCSUS data are ideally suited to address this question. (A public-use data file from HCSUS is available from the Agency for Healthcare Research and Quality.)

HCSUS employs a multistage national probability sample design described in detail elsewhere (Frankel et al. 1999). The reference population is HIV+ patients over 18 years old who made at least one visit for regular care in the contiguous United States in early 1996. HCSUS samples Metropolitan Statistical Areas (MSAs) and clusters of rural counties (Lam and Liu 1998) in the first stage of sampling, health care providers in the second stage, and patients in the third stage. Women and patients of private, staff-model HMOs are oversampled. The HCSUS sample does not include HIV+ patients whose only contact with the health care system was through military, prison, or emergency department facilities.

HCSUS collected three rounds of interviews: baseline, a first follow-up, and a second follow-up. Initially, 2,864 baseline interviews were conducted between January 1996 and April 1997, with a response rate of 68% (Duan et al. 1999). The first follow-up interview was conducted with 2,466 subjects between December 1996 and July 1997, and the second follow-up was conducted with 2,267 subjects between August 1997 and January 1998. Much of the attrition between the waves is due to mortality.

Toward the end of data collection for the baseline survey, new efficacious therapies were introduced into clinical practice. These HAART's disseminated widely as the first follow-up survey was in the field. To account for this change, we estimate two sets of models. In the first set, we use the baseline sample of 2,864 respondents with covariates measured at the time of the baseline interview (except mortality, which for obvious reasons is measured at later time points). In the second set, we use the first follow-up sample of 2,466 subjects with covariates measured at first follow-up.

Mortality status is unknown for all respondents, but we impute missing values for essential covariates by using a standard hot-deck strategy described in detail elsewhere (Duan et al. 1999). The theoretically troublesome assumption needed for the validity of this hot-deck procedure is ignorability of nonresponse. In practice, this is not a concern, given the small number of missing respondents, 4.9% for CD4 cell counts and less than .5% for all other covariates including insurance. The results of our model are not sensitive to excluding respondents with imputed values.

We construct analytic weights to adjust the sample to the reference population. A respondent's analytic weight, which

may be interpreted as the number of people in the population represented by that respondent, is the product of three patient-specific quantities: the sampling weight, the multiplicity weight, and the nonresponse weight. The sampling weight adjusts for oversampling (of women, for example); the multiplicity weight adjusts for patients who could potentially enter the sample via multiple providers; and the nonresponse weight adjusts for differential cooperation (Duan et al. 1999). All analyses presented in this article use these weights. Adjustment of standard error estimates for the complex sample design is discussed later.

We use mortality 6 months after baseline and follow-up interviews. Mortality is based on data collected by our interviewers in the field and from Equifax, Inc., a company that tracks deaths in the United States. The main explanatory variable we are interested in is insurance status, derived from the HCSUS data.

Descriptive means for all model variables are given in Table 1. Most of the variables are self-explanatory. In some models, we include measures of the lowest ever CD4+ T-lymphocyte cell count, a critical measure of the function of a patient's immune system. A depletion in these cells correlates strongly with the worsening of HIV disease and the risk of developing an acquired immunodeficiency syndrome (AIDS) defining opportunistic infection (Harrison et al. 1997). In this article, we categorize CD4+ counts, following Bozzette et al. (1998), into four categories, as shown in Table 1. Patients with CD4+ lymphocyte counts below 50 have a very poor prognosis in general; those with counts above 500 are considered much healthier.

We use state policy regarding Medicaid as our instrumental variables. In our model, we would like these instruments to predict whether an individual is covered by insurance but not to affect mortality (except via insurance). Medicaid is the most common form of insurance for the HIV+ population in care, covering 46% of the population. States vary in the generosity of Medicaid, both in terms of eligibility restrictions and in the generosity of coverage. For each patient, we define the following two variables based on generosity of Medicaid in the state in which the patient is sampled:

1. Medically needy program: A dummy variable indicating whether the state has a "medically needy" provision that allows it to extend Medicaid eligibility to qualified persons, especially the disabled, who may have too much income to qualify under statutory requirements. This provision allows the disabled to spend down to Medicaid eligibility by incurring medical or remedial care expenses to offset their excess income.
2. AFDC threshold. The state's income eligibility threshold for Aid to Families with Dependent Children (AFDC) eligibility in 1997. AFDC eligibility makes a person immediately eligible for Medicaid, and so it is a potentially important source of Medicaid coverage for HIV+ women.

Clearly, these state policies affect Medicaid participation, but the question arises whether they affect other types of insurance, especially private insurance. In other contexts, economists have found that public insurance availability can

encourage people to reduce private insurance—a crowd-out effect (Cutler and Gruber 1996). This is especially a concern for the HIV+ population, for which anecdotal evidence suggests that patients try to keep working to maintain private coverage until they are sick enough to qualify for Medicaid or Medicare.

These incentives are stronger when public insurance is more generous. Thus, we also define a set of instrumental variables based on the generosity in a patient's state. In addition to the generosity of Medicaid programs, we include measures of the generosity of Aids Drug Assistance Programs (ADAP's). ADAP's are state-administered drug reimbursement programs that pay for HIV-related medication for low-income HIV-infected with no insurance or inadequate insurance. These programs are funded through a combination of federal and state resources and provide an important source of funds for the HIV+ population (Doyle, Jefferys, and Kelly 1998). The generosity measures include

1. Per-prescription limits. A dummy variable indicating whether the state Medicaid has a restriction on the number of drugs per prescription.
2. Less than three covered prescriptions per month: A dummy variable indicating whether the state Medicaid program limits the number of covered prescriptions per month to three, a particularly severe restriction for HIV+ patients. Either all states have no limits on monthly prescriptions, or they set a maximum of three, five, or six prescriptions per month.
3. Five or six covered prescriptions per month: A dummy variable indicating whether the state Medicaid program limits the number of prescriptions per month to five or six, a moderately severe restriction for HIV+ patients.
4. Generous ADAP program: Whether the state has a generous ADAP program, as determined by a grade of A or B from the National ADAP Monitoring Project (Doyle, Jefferys, and Kelly 1998).

We prefer to fit models that include variables that control for Medicare eligibility because they are most likely to be independent of unobserved health status. Variables reflecting the generosity of Medicaid or ADAP within states seem more likely to be correlated with unobserved variables that predict mortality separately from insurance status. However, we estimate models with these variables to test the sensitivity of our results to the choice of the instrumental variable.

### 3. A NAÏVE MODEL

To illustrate the consequences of selection bias, we estimate a naïve probit model, where insurance is treated as just another covariate. Because mortality risk depends on factors other than insurance coverage, we include in the model other covariates that are arguably uncorrelated with the error, which we interpret as unobserved determinants of mortality risk. This model treats insurance status as an exogenous variable.

Because we will contrast the resulting probit estimate of the association between mortality and insurance status with the results from our structural model, we discuss this naïve model in some formal detail here. Let  $m_i^*$  represent an index function that measures the mortality propensity for HIV+ patient  $i$

Table 1. Weighted Descriptive Statistics

Variable	Baseline (n = 2,864) proportion or mean	Follow-up (n = 2,466) proportion or mean
<b>Demographics</b>		
Insured	79%	82%
Age	38 years (8.5)	39 years (8.4)
Nonwhite	51%	49%
Female	30%	29%
<b>Presumed exposure route</b>		
Men who have sex with men	52%	54%
Intravenous drug user	20%	20%
Heterosexual or other	28%	26%
<b>Education</b>		
Less than high school degree	25%	25%
High school degree	56%	56%
College degree or more	18%	19%
<b>Years since diagnosis</b>		
Less than 1 year	10%	10%
1–2 years	12%	12%
3–4 years	22%	22%
5+ years	56%	56%
<b>State policies</b>		
Medically needy program	82%	83%
AFDC threshold	\$5,489 (\$1,887)	\$5,486 (\$1,800)
Per-prescription limits	35%	34%
Less than 3 covered prescriptions/month	6%	6%
5–6 covered prescriptions/month	46%	47%
Generous ADAP program	68%	69%
<b>Lowest ever CD4 lymphocyte count (cells per mm<sup>3</sup>)</b>		
500+	9%	7%
200–499	38%	38%
50–199	30%	32%
0–49	23%	24%
<b>Mortality (postinterview)</b>		
Died by 12 months	7%	5%

NOTE: Number in parentheses indicate the population SD.

1 year after interview. The indicator variable  $m_i \in \{0, 1\}$  represents whether patient  $i$  died 1 year after interview. If insured <sub>$i$</sub>  represents whether patient  $i$  is insured, then the naïve model is

$$m_i^* = c_1 + \gamma \times \text{insured} + \beta_1' X_i - \varepsilon_i, \quad (1)$$

where  $X_i$  represents exogenous covariates that determine mortality propensity, and  $c_1, \beta_1$ , and  $\gamma$  are parameters to be estimated. The term  $\varepsilon_i$  is an identically distributed error term representing unobserved determinants of  $m_i^*$  that we assume to be independent of both  $X_i$  and insured <sub>$i$</sub>  in this specification. We define  $m_i$  as

$$m_i = \begin{cases} 1 & \text{if } m_i^* > 0, \\ 0 & \text{if } m_i^* \leq 0. \end{cases} \quad (2)$$

To complete the model, we specify a distribution for  $\varepsilon_i$  as standard normal, with zero mean and unit variance. This distributional assumption implies a probit model for  $m_i$ , where the probability of death, conditional on insured <sub>$i$</sub>  and  $X_i$  is

$$P[m_i = 1 \mid \text{insured}_i, X_i] = P[m_i^* > 0 \mid \text{insured}_i, X_i] \\ = \Phi(c + \text{insured}_i \times \gamma + X_i \beta), \quad (3)$$

where  $\Phi(\cdot)$  is the cumulative distribution function for the standard normal distribution. With (3), we obtain naïve estimates of  $c, \beta$ , and  $\gamma$  by using a (weighted) maximum likelihood

estimator. This estimator is naïve because of the assumption that insured <sub>$i$</sub>  is independent of  $\varepsilon_i$ . The impact of insurance on mortality propensity is given by  $\gamma$ .

#### 4. RESULTS FROM THE NAÏVE MODEL

We estimate the naïve model twice, once with baseline data and once with follow-up data. We do so because one of our objectives is to see whether the insurance effect on mortality ( $\gamma$ ) has changed with time. There are some strong reasons to believe it has. In particular, between the time of the baseline and the first follow-up survey, new and powerful drugs (such as protease inhibitors) became available. These HAART's are remarkably effective in combating the progression of AIDS (Collier et al. 1996).

Comparing the two values of  $\gamma$  that we obtain from the two runs of our model, we have an estimate of the change between baseline and follow-up time periods in the effectiveness of insurance in reducing mortality among AIDS patients. Table 2 summarizes the results from the naïve models. To demonstrate the importance of including information on health status, we include two sets of estimates for each sample—one set with controls for disease progression, particularly CD4 cell count, and one without.

To provide a better metric for the magnitude of the results, we also predict the probability of death for every-one in our sample both with and without insurance. We then

Table 2. Results for Naïve Weighted Probit Model, 6-Month Mortality

Coefficient <sup>a</sup>	Baseline		Follow-up	
	No severity controls	With severity controls	No severity controls	With severity controls
<b>Demographics</b>				
γ (insurance)	.3918*	.1458	.3462	.0891
Age	-.0317	-.0619	-.0191	-.0462
Age <sup>2</sup>	.0003	.0007**	.0002	.0006
Nonwhite	.0012	.0233	.1126	.1245
Female	-.2760**	-.1148	-.4340**	.4022**
<b>Presumed exposure route<sup>b</sup></b>				
Men who have sex with men	-.1508	-.0579	-.2373*	-.2421*
Intravenous drug user	-.2936**	-.2737**	-.0792**	-.0977**
<b>Education<sup>b</sup></b>				
High school degree	-.1448	-.2120	-.1131	-.1427
College degree or more	-.3293	-.4064*	-.1195	-.1469
<b>Years since diagnosis<sup>b</sup></b>				
Less than 1 year	-.4036**	-.4263**	-.1628	-.3567*
1–2 years	-.4152**	-.3698*	-.2512	-.1976
3–4 years	-.3410**	-.4297**	-.3599**	-.3697**
<b>Lowest ever CD4 lymphocyte count (cells per mm<sup>3</sup>)<sup>b</sup></b>				
500+	NA	-1.1235**	NA	-.6511*
200–499	NA	-1.4900**	NA	-1.3730**
50–199	NA	-.9844**	NA	-.7204**
Constant	-.9945	.3809	-1.5653	-.3157
Prob(died   ins = 0)	.016	.027	.011	.019
Prob(died   ins = 1)	.039	.035	.025	.032
Change in relative mortality	144%	31%	132%	21%

<sup>a</sup>All model coefficients are multiplied by 100.

<sup>b</sup>The reference groups are heterosexual or other for presumed exposure route; less than high school for education; 5 years or more for years since diagnosis; and 0–49 for lowest ever CD4 lymphocyte count.

NOTE: \* .05 < p < .10, \*\* p ≤ .05.

compute the relative impact of insurance on average mortality (Δ; labeled “Increase in relative mortality” in the model results in Tables 2–4). More precisely, we compute

$$\begin{aligned} \Delta &\equiv \frac{\text{Avg mortality insured}}{\text{Avg mortality not insured}} - 1 \\ &= \frac{\sum_i w_i \Phi(\hat{c} + \hat{\beta} \cdot X_i + \hat{\gamma})}{\sum_i w_i \Phi(\hat{c} + \hat{\beta} \cdot X_i)} - 1, \end{aligned} \tag{4}$$

where  $w_i$  is the sample weight for observation  $i$ , and the parameters with carats (^) represent estimated values.

In the baseline sample, insurance is associated with an increased probability of death by first follow-up, and the effect is statistically and substantively significant without severity controls. This finding is analogous to that of Lancaster and Intrator (1998). This positive finding persists in the mortality analysis of the first follow-up data without severity controls, although the effect is no longer statistically significant.

In both samples, the use of severity controls reduces the magnitude of the insurance effect substantially. For instance, in the baseline sample, the results without severity controls indicate that insurance increases relative mortality by 144%. Including health status measures like CD4 cell count reduces this effect to 31%, but it does not disappear. Similar results are obtained in the follow-up sample. We attribute these findings to a spurious correlation between severity of illness and insurance status for HIV+ patients, i.e., insured<sub>*i*</sub> is correlated with  $\varepsilon_i$ . Unobserved elements of severity of illness and disease progression are among the most important elements in  $\varepsilon_i$ ,

the unobserved determinants of mortality propensity. If we had enough information about disease severity, we might even reverse the sign of this effect. Unfortunately, such data are not available in a patient survey. In the next section, we present a structural model to allow for correlation between insurance status and mortality from unobservable health status.

### 5. A STRUCTURAL MODEL

In the structural version of our model, we retain Equations (1) and (2) but drop the assumption that  $\varepsilon_i$  is independent of insured<sub>*i*</sub>. Instead, we specify an additional equation for insured<sub>*i*</sub> that embodies the mechanism that we believe is driving the naïve results. Let insured<sub>*i*</sub><sup>\*</sup> represent the propensity of patient  $i$  to have health insurance at the baseline interview. We specify

$$\text{insured}_i = c_2 + \beta_2' Z_i - \eta_i, \tag{5}$$

where  $Z_i$  represents variables that determine insurance status, including our set of instrumental variables (that is, variables that belong in the insurance equation but not in the mortality equation);  $c_2$  and  $\beta_2$  are additional parameters to be estimated; and  $\eta_i$  represents the unobserved determinants of insurance status. We assume

$$\text{insured}_i = \begin{cases} 1 & \text{if insured}_i^* > 0, \\ 0 & \text{if insured}_i^* \leq 0. \end{cases} \tag{6}$$

Table 3. Results for Structural Model Mortality Equation, 6-Month Mortality

Coefficient <sup>a</sup>	Baseline		Follow-up	
	No severity controls	With severity controls	No severity controls	With severity controls
<b>Demographics</b>				
$\gamma$ (insurance)	.0142	-.7804	-.0801	-1.1301
Age	-.0289	-.0565*	-.0163	-.0399
Age <sup>2</sup>	.0003	.0007**	.0002	.0006
Nonwhite	-.0140	-.0024	.0997	.0913
Female	.2726**	-.1003	-.4255**	-.3593
<b>Presumed exposure route<sup>b</sup></b>				
Men who have sex with men	-.1635	-.0915	-.2478*	-.2776*
Intravenous drug user	.2776**	-.2604	-.0627	-.0705
<b>Education<sup>b</sup></b>				
High school degree	-.1352	-.1853	-.1050	-.1453
College degree or more	-.3084	-.3765	-.1020	-.1094
<b>Years since diagnosis<sup>b</sup></b>				
Less than 1 year	-.4658*	-.5700**	-.2286	-.4953**
1-2 years	-.4419*	-.4327	-.2729	-.2293
3-4 years	.3521**	-.4540**	-.3734	-.4247**
<b>Lowest ever CD4 lymphocyte count (cells per mm<sup>3</sup>)<sup>b</sup></b>				
500+	NA	-1.1764**	NA	-.7637**
200-499	NA	-1.5410**	NA	-1.5295**
50-199	NA	-1.0330**	NA	-.7671**
Constant	-.7412	1.1330	-1.2639	.7303
Correlation <sup>c</sup>	.211	.281	.233	.583**
Prob(died   ins = 0)	.033	.111	.026	.141
Prob(died   ins = 1)	.034	.032	.022	.021
Change in relative mortality	3%	-71%	-17%	-85%

<sup>a</sup>All model coefficients are multiplied by 100.

<sup>b</sup>The reference groups are heterosexual or other for presumed exposure route; less than high school for education; 5 years or more for years since diagnosis; and 0-49 for lowest ever CD4 lymphocyte count.

NOTE: \*.05 <  $p$  < .10. \*\* $p$  ≤ .05.

To complete the model, we need to assume a joint distribution for  $\varepsilon_i$  and  $\eta_i$ . We posit a bivariate normal distribution with arbitrary correlation ( $\rho$ ) across the errors (and unit variances to reflect a necessary scaling assumption):

$$\begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix} \sim BVN \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right). \quad (7)$$

We assume a bivariate normal distribution of the errors because it allows considerable flexibility in the correlation structure across the errors and because it is a natural generalization of the probit assumption in the naïve model. Still, one might question whether such strong distributional assumptions are necessary to consistently measure the structural parameter ( $\gamma$ ) that we are most interested in. Unfortunately, in our case with a discrete dependent variable ( $m_i^*$ ) and a probit naïve model, we do not have available the easiest method with which to avoid such distributional assumptions, an IV estimator. In our structural model, we can identify the causal effect of insurance on mortality rates because we have plausible instruments. We use variables measuring the generosity and eligibility thresholds of the ADAP and the Medicaid program in the patient's state as variables that predict whether a patient has insurance but do not directly belong in the mortality equation. Because our hypothesis is that the naïve finding is caused by late-stage AIDS patients enrolling in Medicaid, these arguably exogenous state policy variables seem to be just the right instrument set.

The difficulty with an instrumental variables estimator for  $c_1$ ,  $\beta_1$ , and  $\gamma$  is that there is no easy way to recover the structural error  $\varepsilon$  from a transformation of the data, because  $m^*$  is never observed. The usual IV approach requires isolating  $\varepsilon_i = m_i^* - E(m_i^*)$  from Equation (1) and using the orthogonality of the instrument set  $Z_i$  with the structural error to define a consistent estimator. This kind of estimator does not require an explicit unverifiable assumption about the form of the full distribution for  $\varepsilon_i$ . It relies entirely on the assumption of orthogonality of the instrument set with the structural error (which at some level is also unverifiable). Of course, because  $m_i$  is observed, we can construct  $\mu_i = m_i - E(m_i^*)$ , but the relationship between  $\varepsilon_i$  and  $\mu_i$  is not obvious. In particular, we cannot tell whether the orthogonality of  $Z_i$  and  $\varepsilon_i$  implies the orthogonality of  $Z_i$  and  $\mu_i$  without more explicit assumptions (such as a full specification of the  $\varepsilon_i$  distribution). Thus, we cannot define an IV estimator by using only observed information.

Another popular procedure for estimating binary dependent variable models in the presence of endogenous right side variables is a two-step estimator. The endogenous variable is regressed on the instrumental variables (and other exogenous variables), and predicted values from this first step regression are inserted into the structural equation. This procedure yields consistent estimates of the structural parameters when the second stage involves a continuous dependent variable. In fact, two-stage least squares (2SLS) estimates are equivalent to the IV estimator in the presented case with a continuous depen-

Table 4. Results for Structural Model Insurance Equation, 6-Month Mortality

Coefficient <sup>a</sup>	Baseline		Follow-up	
	No severity controls	With severity controls	No severity controls	With severity controls
<b>Demographics</b>				
Age	.0187	.0048	.0123	.0045
Age <sup>2</sup>	-.0000	.0001	.0000	.0003
Nonwhite	-.1849*	-.1918*	-.1630**	-.1755**
Female	-.0068	.0624	.0422	.1094
<b>Presumed exposure route<sup>b</sup></b>				
Men who have sex with men	-.2024	-.1799	-.1746	-.1496
Intravenous drug user	.0947	.0960	.1539	.1532
<b>Education<sup>b</sup></b>				
High school degree	.0755	.0608	.0268	.0181
College degree or more	.1847	.1546	.0966	.0641
<b>Years since diagnosis<sup>b</sup></b>				
Less than 1 year	-.6631**	-.6789**	-.6762**	-.7222**
1–2 years	-.3462**	-.3015**	-.2643**	-.1848**
3–4 years	-.1797**	-.1791**	-.2047**	-.1881**
<b>State policies</b>				
Medically needy program	-.3940**	-35.62**	-.4761**	-.4627**
AFDC threshold	.0001**	.01**	.0001**	.0001**
<b>Lowest ever CD4 lymphocyte count (cells per mm<sup>3</sup>)<sup>b</sup></b>				
500+	NA	-.7158**	NA	-.8454**
200–499	NA	-.6880**	NA	-.8199**
50–199	NA	-.5226**	NA	-.5832**
Constant	8.78	.7771	.3556	1.1729**
Wald test	14.01	17.10	13.25	16.41
Degrees of freedom	2	2	2	2
p-value	<.01	<.01	<.01	<.01

<sup>a</sup>All model coefficients are multiplied by 100.

<sup>b</sup>The reference groups are heterosexual or other for presumed exposure route; less than high school for education; 5 years or more for years since diagnosis;

NOTE: \*.05 < p < .10, \*\* ≤ .05.

dent variable. However, in the case of a binary probit, this procedure will not yield consistent estimates of the structural parameters. Bhattacharya, McCaffrey, and Goldman (1999) reviewed the research related to such estimators and concluded that the literature offers no simple consistent two-step estimator.

Angrist (1991), Evans and Schwab (1995), Angrist and Evans (1998), and Evans, Farrelly and Montgomery (1999) noted that the 2SLS estimation procedure applied to a linear probability model for the dichotomous outcome can sometimes recover an acceptable approximation to the treatment effect. However, for our data, this linear approximation provided results that differed substantially from the results of the bivariate probit model; in particular, they resulted in estimated effects that were implausibly large. By using 2SLS with the same instruments and covariates of the bivariate probit model, we find that the predicted probability of death for those with insurance is .5%, and it is 10.1% for those without insurance. These results confirm that using an instrumental variables approach, even in a linear probability model context, leads one to predict that insurance is protective. However, the underlying mortality risk for these individuals is too low for an HIV+ patient population, even with HAART. Further, in other specifications (not shown), the 2SLS procedure yielded negative mortality probabilities for the insured. Thus, we develop a joint structural model of mortality and insurance.

We use maximum likelihood to estimate the parameters of our model. Details on the bivariate probit likelihood function can be found in Greene (2000). In Appendix A, we discuss how we calculate standard errors that account for the multistage sampling design of HCSUS. In Appendix B, we show how our structural framework is directly related to the Neyman–Fisher–Cox–Rubin causal model, which may be more familiar to statisticians than the structural framework we use.

## 6. IDENTIFICATION OF STRUCTURAL EFFECTS

The distributional assumptions aside, as in all IV-based studies, the credibility of our study rests on the believability of our instruments. Our state policy instruments could fail in at least two ways. First, the estimators perform poorly if the instruments are only weakly correlated with the treatment variable, i.e., insurance status (Nelson and Startz 1990; Bound, Jaeger, and Baker 1995; Staiger and Stock 1997). Thus, we report Wald statistics for the joint significance of our instruments in predicting insurance status.

Second, our instruments might be correlated with unobserved determinants of mortality (like unmeasured health status variables). The assumption that an instrumental variable is uncorrelated with the outcome measure cannot be directly tested. For these reasons, some researchers have argued that IV estimates in this context should be viewed with caution (Bound et al.).

However, in our application, it seems clear that patients have little direct influence at an individual level on state policies, so our state policy instruments are *prima facie* exogenous. This argument is not enough to establish exogeneity, however, if there are unobserved state-level variables that determine both health and insurance status. In that case, state policies would be endogenous in our model despite the lack of control by patients over these policies. To address this issue, we develop some indirect evidence that our instruments are not simply picking up differences in unobserved health across states. If the latter hypothesis were true, then one would expect to find that our state-level policy instruments predict the mortality of patients even in a non-HIV, non-Medicaid population.

To check this assumption, we estimate a logit model of 1-year mortality by using data from the Medicare Current Beneficiary Survey (MCBS). On the right side, this mortality model includes a sparse set of health status indicators, such as measures of activities of daily living (ADL's) and a general health index, and our state policy instruments. Because this elderly or disabled population is by definition insured by Medicare, our instruments should not predict their mortality unless they proxy for unobserved state-level effects. In fact, we find that our instruments are not statistically significant in the model, with odds ratios near 1 in response to a one-standard deviation movement in their values. Complete results of this logistic regression are available on request from the authors. Of course, these results do not prove that unobserved state effects are unimportant in the HCSUS population, but they are certainly suggestive. There is no good reason to expect that such effects should be present for HIV+ patients when they are not present for the elderly or disabled.

### 7. RESULTS FROM THE STRUCTURAL MODEL

The parameter estimates for the structural model using 6-month postinterview mortality as the outcome variable are shown in Tables 3 and 4. Table 3 shows the results from the mortality equation (1) and the predicted probabilities (4); Table 4 gives the parameter estimates for the insurance equation (5). We also report the joint significance of the excluded instruments from (4) by using a Wald statistic because these are considered a test of the small sample bias associated with the IV estimator (Staiger and Stock 1997).

Table 3 shows several important differences from the naive results. First, in our preferred specifications, *i.e.*, those with severity controls, we see that insurance decreases the likelihood of death, as one would expect. Second, the magnitude of the effect is greater in the follow-up sample. For instance, the presence of insurance decreases 6-month mortality by 71% in the baseline sample and 85% in the follow-up sample.

We attribute these findings to the rapid proliferation of new efficacious treatments in the HCSUS sample (Bozzette et al. 1998) and the strong correlation between the use of these drugs and insurance (Shapiro et al. 1999b). Our main result, that the efficacy of insurance in preventing death increased through time, is consistent with the fact that highly effective treatments for AIDS (such as protease inhibitors) did not exist at the time of the baseline survey. Thus, those patients who had insurance at baseline were not helped as much by it as those who had insurance at the time of the follow-up survey

(when protease inhibitors did exist). We fully expect the efficacy of insurance to grow with the diffusion of these new and effective pharmaceutical therapies for AIDS.

### 8. MODEL DIAGNOSTICS

We conduct several tests of our structural model to examine its robustness. First, we implement a goodness-of-fit test to see how well the model predictions fit the data. Second, we estimate the model under a variety of different assumptions to determine whether our results are only a special case. We find that our main conclusions do not alter much under different assumptions

#### 8.1 Goodness-of-Fit Tests

To test the goodness of fit for our bivariate model, we follow Hosmer and Lemeshow's (1989, sect. 8.1.4) approach to testing the goodness of fit of models for polytomous outcomes data. Our bivariate 0-1 data can be treated as a polytomous outcome with four levels (a)  $m_i = 0$  and  $insured_i = 0$ ; (b)  $m_i = 0$  and  $insured_i = 1$ ; (c)  $m_i = 1$  and  $insured_i = 0$ ; and (d)  $m_i = 1$  and  $insured_i = 1$ . For a four-level polytomous outcome, Hosmer and Lemeshow suggest testing goodness of fit by using their test statistic for logistic regression on three dichotomous submodels. For our model, this approach means that we first create a pseudo-outcome for each observation,  $Y_{11}$ , where  $Y_{11} = 1$  if  $m_i = 1$  and  $insured_i = 1$ ,  $Y_{11} = 0$  if  $m_i = 0$  and  $insured_i = 0$ , and  $Y_{11}$  is set to missing otherwise. We then use our model to predict the probability that  $Y_{11}$  equals 1 for each observation where  $Y_{11}$  is not missing. We create 10 bins based on the deciles of the estimated probabilities and calculate the statistic

$$\widehat{C}_{11} = \sum_{k=1}^{10} \frac{(o_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)},$$

where  $o_k$  is the weighted sum of  $Y_{11}$  in the  $k$ th bin,  $\bar{\pi}_k$  is the weighted average of the estimated probabilities for observations in the  $k$ th bin, and  $n_k$  is the sum of the weights for the bin. We then use the analogous procedure to calculate  $\widehat{C}_{10}$  (for testing the fit by using the submodel for predicting  $m_i = 1$  and  $insured_i = 0$  versus  $m_i = 0$  and  $insured_i = 0$ ) and  $\widehat{C}_{01}$  (for testing the fit by using the submodel for predicting  $m_i = 0$  and  $insured_i = 1$  versus  $m_i = 0$  and  $insured_i = 0$ ).

We compare the statistics to quantiles of  $\chi^2_8$  distribution. Because of the complex nature of the HCSUS sample design, we expect that each of the  $\widehat{C}$  statistics will tend to be more variable than will a  $\chi^2_8$  distribution (Rao and Scott 1981). Hence, using  $\chi^2_8$  as the reference distribution is conservative in that the probability that a test finds lack of fit when the model is adequate is greater than the nominal value of the test.

Table 5 provides the results of the tests for lack of fit. For both the baseline and the follow-up data, we find no evidence of a significant lack of fit for the models that do not include disease severity. Because the  $p$ -values are too liberal, we do not interpret the  $p$ -value of .03 for  $\widehat{C}_{11}$  for the baseline data as an indication of significant lack of fit. Although the lack-of-fit statistics are large for models with disease severity measures fitted to the follow-up data, closer inspection of the data indicate that the large values of these statistics are due to the

Table 5. Goodness-of-Fit Tests for 6-month Mortality Models

Model	$\hat{C}_{01}$	<i>p</i> -value	$\hat{C}_{10}$	<i>p</i> -value	$\hat{C}_{11}$	<i>p</i> -value
Baseline						
Without disease severity	3.81	.87	13.92	.08	3.79	.89
With disease severity	5.60	.69	10.62	.22	16.77	.03
Follow-up						
Without disease severity	3.20	.92	7.07	.53	3.30	.92
With disease severity	17.50	.025	22.69	<.01	37.10	<.01

scarcity of deaths rather than model lack of fit. The statistics are highly sensitive to a few deaths that occur among people who are predicted to have a low probability of death. In addition, we explored interactions in these models and found that the interactions were not statistically significant and did not reduce the values of the lack-of-fit statistics. We believe the model accurately represents the data, and the results presented in Tables 3 and 4 are not biased by lack of fit.

## 8.2 Robustness of Results

The qualitative results from these models are robust to different model specifications. We try several other specifications, including changing the dependent variable to 12-month mortality, expanding the list of instruments to include insurance generosity, and expanding the list of other covariates. The expanded list includes other, more detailed indicators for clinical disease stage than CD4.

In the main results section, we show results when 6-month mortality is used as the dependent variable. For reasons related to timing of the baseline and follow-up HCSUS surveys, we prefer our estimates using 6-month mortality, as shown in Tables 3 and 4. However, we also estimated the model by using 12-month rather than 6-month mortality, as the key dependent variable, with little substantive change in our results. In addition, we estimated other specifications, with the instrument set augmented to include other state policy variables affecting insurance coverage (like prescription limits and the adequacy of the ADAP programs in the state). The results from these specification checks are presented in an appendix that is available from the authors on request.

All of these variations yield essentially the same conclusions:

1. Single-equation models that do not model insurance lead to the finding that insurance has either a positive effect or no effect on mortality.
2. The bivariate model finds that insurance has a beneficial effect on mortality.
3. Severity adjustments yield a stronger insurance effect (in the sense that the point estimate for  $\gamma$  is reduced).
4. Insurance has a stronger protective effect in the post-HAART era.

## 9. DISCUSSION

Our main findings are that (a) ignoring observed and unobserved health status in our structural mortality equation leads one to conclude that insurance may not be protective for HIV patients; (b) once observed and unobserved health status are accounted for, one recovers the more intuitively pleasing

causal inference that insurance does not kill HIV patients (and may actually protect HIV patients); and (c) the effectiveness of insurance in protecting HIV patients improved after the development and introduction of HAART therapy.

Our work has three main limitations. First, because our strategy for identifying the causal effect of insurance requires the use of state policy variables, we must exclude state fixed effects that may directly enter the structural mortality equation. The presence of such unobserved state fixed effects may bias our measurement of the causal effect of insurance. Of course, if these unobserved state fixed effects are independent of  $\varepsilon$ , then there is no bias. Recall that in Section 6 we derived indirect evidence from mortality in the Medicare population that our instruments are not simply proxies for state-level fixed effects. However, even if these random variables are correlated, this limitation should not substantially alter any of our three main findings. State fixed effects cannot explain the change that we observe in the causal effect of insurance when richer health status information is included in the model. Furthermore, they cannot explain the increase in the protective power of insurance coincident with the development of effective therapy for HIV disease. Even if our point estimates are biased by the presence of state fixed effects in the mortality equation, such effects (which are, by definition, fixed across time and across individuals of differing health status) cannot explain the intuitively plausible changes in the causal effects that we estimate.

Second, we lump patients with different types of insurance into a single category. In our sample, many patients are insured via private insurance, whereas others have Medicaid or other public insurance. Because these types of insurance differ in quality and in the extent of coverage, we would expect that they differ in their efficacy in preventing premature AIDS deaths. An alternative approach, which we briefly examined, involves estimating a three-equation system, one equation for mortality (as before) and one for each type of insurance. Unfortunately, estimating this model is computationally challenging with a maintained hypothesis of joint normality across the errors in these three equations. Our results can be seen as a composite estimate of the effect of these different insurance types on mortality.

The third limitation arises out of the HCSUS sampling scheme. In particular, the HIV+ patients in the follow-up survey are the same people in the baseline survey, with one exception—they survived to take the follow-up survey. To the extent that the follow-up sample likely has harder HIV+ patients than the patients in the full baseline sample, our results are not representative of the effect of insurance on the general HIV+ population in care as a whole. Instead, our follow-up results should be seen as estimates of the mortality hazard rates of the people sampled in the baseline survey, conditional on their survival to the follow-up survey.

## 10. CONCLUSIONS

We find that insurance increases the probability of death in HIV+ patients when we use a single-equation model of mortality. We attribute this finding to the correlation between unobserved health status and insurance status in the mortality equation, a finding that is buttressed by the observation that

including severity controls reduces the effect size by approximately 50%. To deal with this correlation, we developed a bivariate probit model to address the unobserved heterogeneity. The results from the structural model indicate that insurance does have a beneficial effect on outcomes, lowering the probability of 6-month mortality by 71% at baseline and 85% at follow-up. Our results also indicate that the development of more effective therapies for HIV infection has magnified the returns to insurance for AIDS patients, at least as measured by mortality rates. The results suggest that policies to expand insurance coverage to the uninsured HIV+ population could save many lives.

### APPENDIX A: STANDARD ERRORS

The goal of the HCSUS survey was to accurately represent a national sample of HIV+ patients in the United States. Rather than focusing mainly on academic medical centers, as did other studies of the same population, the HCSUS sampling scheme was designed to place a positive probability of sampling nearly every HIV+ patient that seeks care. One consequence of this complex sampling design is that each patient has an associated sampling weight,  $w_i$ , which we incorporate into our model as weights on the likelihood functions. These sample weights account for differential sample selection probabilities, but they do not account for correlation in the model errors across patients in the HCSUS sample. There is some reason to believe that these correlations do not equal 0 (as they would in simple random sample).

In the HCSUS sampling design, clinics that serve HIV+ patients are first drawn at random from the universe of all such clinics in the United States. Then, from each clinic, a random sample of HIV+ patients is drawn. Patients from the same clinic are more likely to be similar on unobservable dimensions than patients from different clinics. Unless the clustering of patients in clinics is taken into account in the standard errors of our parameter estimates, there is the possibility of mistaken inferences.

To allow for the complex multistage sample design used to obtain the HCSUS cohort, we follow the pseudomaximum likelihood estimation approach as described by Skinner (1989) and Binder (1983). In this approach, the target parameter of interest,  $\theta^*$ , is the solution to  $T(\theta) = Eu(\theta) = 0$ , where  $\theta$  is the vector of parameters in our model,  $u(\theta) = \partial \ell / \partial \theta$ , and  $\log L = \sum_i \ell(m_i, \text{insured}_i, X_i, Z_i, \theta)$ . The pseudomaximum likelihood estimate  $\hat{\theta}_{PML}$  of  $\theta$  is the value of  $\theta$  that solves  $\hat{T}(\theta) = 0$ , where  $\hat{T}(\theta) = \sum_{i=1}^n w_i u_i(\theta)$ . Note that if our likelihood model is correct for the entire population, then the value of  $\theta$  that solves  $\sum_{i=1}^n w_i u_i(\theta) = 0$  is the maximum likelihood estimate (ML) for the entire population. Therefore, we can think of  $\hat{T}(\theta)$  as an unbiased estimate of the population score.  $\hat{\theta}_{PML}$  is a consistent estimate of  $\theta^*$ .

We estimate the variance of  $\hat{\theta}_{PML}$  by the linearization method described by Skinner (1989). By using a first order Taylor's expansion of  $\hat{T}(\theta)$ , we find that variance of  $\hat{\theta}_{PML}$  is approximately  $I(\hat{\theta}_{PML})^{-1} V[\hat{T}(\hat{\theta}_{PML})] I(\hat{\theta}_{PML})^{-1}$ , where  $I(\hat{\theta}_{PML})$  is the information matrix, the matrix of the expected value of the second partial derivatives and  $V[\hat{T}(\hat{\theta}_{PML})]$  denotes the variance of  $\hat{T}(\hat{\theta}_{PML})$ . In our estimate of  $V(\hat{\theta}_{PML})$ ,  $VL(\hat{\theta}_{PML})$ , we use the observed information matrix to estimate  $I(\hat{\theta}_{PML})$  and estimate  $V[\hat{T}(\hat{\theta}_{PML})]$  by

$$V_L[\hat{T}(\hat{\theta}_{PML})] = \sum_{h=1}^H \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (z_{hi} - \bar{z})(z_{hi} - \bar{z})'$$

where  $z_{hi}$  is the sum of the  $w_{hij}u(m_{hij}, \text{insured}_{hij}, X_{hij}, Z_{hij}, \hat{\theta}_{PML})$  across units  $h_{ij}$  in analysis PSU  $h_i, i = 1, \dots, n_h$  of analysis

stratum  $h$ . For analysis purposes, PSU's are grouped in strata  $h$ , according to order of sample selection (Duan et al., 1999). The estimator  $V_L[\hat{T}(\hat{\theta}_{PML})]$  obtains by noting that  $\hat{T}(\hat{\theta}_{PML})$  is the weighted sample mean of the  $u_{hij}$  and  $V_L[\hat{T}(\hat{\theta}_{PML})]$  is an approximate variance estimator for this mean given the complex HCSUS design.

The estimator  $\hat{\theta}_{PML}$  and the variance estimator  $V_L(\hat{\theta}_{PML})$  can be motivated as robust estimators that allow for possible misspecification in our model. The weighted PML will converge to the finite population ML estimates, and the variance estimator is consistent even if the likelihood model is incorrect. The variance estimator is analogous to variance estimator suggested by Huber (1967) for misspecified models.

The accuracy of both  $\hat{\theta}_{PML}$  and the variance estimator  $V_L(\hat{\theta}_{PML})$  requires correct specification of the likelihood. In particular, our models assume that the relationship between mortality and insurance status does not vary across states. If such interactions do exist, then our parameter estimates could be biased, and our standard errors might estimate the true variability of these parameter estimates.

### APPENDIX B: RELATIONSHIP BETWEEN OUR STRUCTURAL MODEL AND A POTENTIAL OUTCOMES FRAMEWORK

In this article, we fitted a parametric structural equation model to model the relationship between mortality and insurance status, adjusting for observed covariates and the possible selection bias that results because the eligibility requirements of Medicaid increase the likelihood that the sickest patients will be insured. Such structural models are widely used in economic analyses and are presented in introductory econometrics texts [for example, Greene (2000)].

Recent literature (Angrist and Imbens 1994; Angrist, Imbens, and Rubin 1996; Imbens and Rubin 1997; Heckman and Vytlacil 1999) focused on an average treatment effect, and in the statistics literature this discussion has been motivated by the Neyman-Fisher-Cox-Rubin potential outcomes (PO) model, sometimes referred to as the Rubin causal model (Holland 1986). Briefly, in the PO model,  $m_i$  summarizes the information on mortality that is observed in the sample for any patient  $i$ , but there are two potential outcomes of interest— $m_{1i}$ , which represents mortality if the patient is insured, and  $m_{2i}$ , which is mortality if the patient is uninsured. The quantity of interest is the treatment effect or the expected value of the difference  $m_{1i} - m_{2i}$ . The difference is not observable because either  $m_{1i}$  or  $m_{2i}$  is observed but never both. Therefore, additional assumptions are required to estimate the treatment effect. Angrist et al. (1996) provided full details on the assumptions necessary for the consistent estimation of a local average treatment effect using instrumental variables in the context of the PO model. In the context of our study, the local average treatment effect is the protective effect of insurance for the subset of the HIV+ population who would change their insurance status in response to differences in state policy on Medicaid eligibility.

Heckman and Vytlacil (1999), among others, point out that there is a natural mapping from the potential outcomes approach to a switching regression model (Quandt 1972; Roy 1951; Heckman and Honore 1990). A switching regression representation of the mortality potential outcomes is given by

$$m_i = m_{1i} \text{insured}_i + m_{2i} (1 - \text{insured}_i).$$

Our structural model is a special case of either the PO or the switching regression model in which we make explicit assumptions about the joint distribution of  $m_{1i}$  and  $m_{2i}$ . In particular, we make the following assumptions.

*Assumption 1 (linearity and latent variable model assumption).* The models generating  $m_1$ ,  $m_2$ , and  $\text{insured}_i$  are given by

$$\begin{aligned} m_{1i}^* &= \beta_1 X_i + \varepsilon_{1i}; & m_{1i} &= 1 \text{ if} \\ & m_{1i}^* \geq 0, = 0 & \text{otherwise;} \\ m_{2i}^* &= \beta_2 X_i + \varepsilon_{2i}; & m_{2i} &= 1 \text{ if} \\ & m_{2i}^* \geq 0, = 0 & \text{otherwise;} \\ \text{insured}_i^* &= \alpha Z_i + \eta_i; & \text{insured}_i &= 1 \text{ if} \\ & \text{insured}_i^* \geq 0, = 0 & \text{otherwise,} \end{aligned}$$

where  $\alpha$ ,  $\beta_1$ , and  $\beta_2$  are parameter vectors.

*Assumption 2 (additivity).*  $\varepsilon_{1i} - \varepsilon_{2i} = \gamma \text{Insured}_i + \varepsilon_i$ .

*Assumption 3 (no interactions).*  $\beta_1 = \beta_2$ .

Rubin (1991) notes that Assumption 2 is common in the treatment-effects literature. It implies that insurance status shifts the location of the distribution of unobserved errors but does not alter the shape. Given the latent variable structure of our model, this is not likely to be controversial. The important empirical fact that we want to model is the change in probability of death caused by insurance status. Whether this probability shift is caused by a change in the shape or just the location of the underlying latent variable representation is not of particular interest to us. Assumption 3 is perhaps more controversial; it implies that any interactions will be subsumed in the estimate of  $\gamma$ , which provides an estimate of the average of variable effects. Because our data are from a probability sample of the entire HIV+ population in care, and we are not interested in the context of this analysis on the differential effect of insurance on particular subgroups, Assumption 3 seems appropriate.

In addition, we assume bivariate normality for the joint distribution of  $\varepsilon_i$  and  $\eta_i$ . This assumption specifies the link between the expected value of mortality and the covariates and effects estimates of the covariance parameters. However, unlike as in Assumptions 1–3, we tested the assumption about the distribution empirically and found no evidence of obvious violations of this assumption.

Under either the PO or the switching regression framework, our model requires additional assumptions for the estimate of  $\gamma$  to estimate the average treatment effect.

*Assumption 4 (stable unit treatment value assumption).* For each person  $i$  in the population, we assume that each potential outcome  $m_{1i}$ ,  $m_{2i}$ ,  $\text{insured}_i$ , and  $Z_i$  is well defined in the sense that they are a complete representation of all possible outcomes for observation  $i$ . This implies that for any patient  $i$ , the values of  $m_{1i}$ ,  $m_{2i}$ ,  $\text{insured}_i$ , and  $Z_i$  are independent of the values for any other patient  $j$ .

We follow Heckman and Vytlačil (1999) to define our Assumption 5, although alternative specifications for the continuity and non-degeneracy assumptions exist.

*Assumption 5 (nondegeneracy, continuity, and exogeneity).*

5.1.  $\alpha Z$  is a nondegenerate random variable conditional on  $X = x \forall x$ .

5.2.  $(\varepsilon_1, \eta)$  and  $(\varepsilon_2, \eta)$  are independent of  $(Z, X)$ ; this implies that the model for  $Z$  is ignorable.

5.3.  $(\varepsilon_1, \eta)$  and  $(\varepsilon_2, \eta)$  are absolutely continuous with respect to Lebesgue measure on  $\mathfrak{R}^2$ .

5.4.  $P(\text{insured}_i = 1) > 0 \forall i$ .

Assumptions 5.1 and 5.2 are the most critical and cannot be verified through empirical tests. Assumption 5.1 implies the necessity of an exclusion restriction; otherwise, the independence of the errors

and the  $X$ 's would not be possible. In our case, the exclusion restriction assumes that conditional for insurance status, mortality does not depend on state policy. Assumption 5.2 ensures the exogeneity of all right-side variables. Assumption 5.3 is necessary to guarantee full support for  $P(\text{insured}_i = 1 | Z_i = z)$  (the propensity score), which in turn is required to guarantee that all interesting functions of the treatment effect are identified (Heckman and Vytlačil, 1999). A different assumption is required in the PO framework to identify the local average treatment effect (Angrist, Imbens, and Rubin, 1996). Loosely, this assumption is nonexistence of defiers—people who would be uninsured in states with generous public policies but insured in states with less generous policies. Such an assumption seems less appropriate in this setting than does Assumption 5.3. Assumption 5.4 seems uncontroversial.

Implicit in Assumption 5.2 is the assumption that other characteristics associated with the state do not affect mortality after conditioning on the observable patient level covariates and insurance status. If such characteristics affect mortality, then  $\varepsilon_1$  and  $\varepsilon_2$  would not be independent of  $Z$ . Unfortunately, this assumption cannot be directly tested, because the independence assumption is necessary in our model to derive the estimates in the first place. Nevertheless, we do not believe that such fixed unobserved state characteristics can explain our results, which show an increasing mortality protective effect of insurance after the introduction of HAART. In addition, in Section 6, we provided indirect evidence from the Medicare population that our instruments are not simply serving as proxies for these unobserved state characteristics.

[Received September 1999. Revised December 2000.]

## REFERENCES

- Angrist, J. D. (1991), "Instrumental Variables Estimation of Average Treatment Effects in Econometrics and Epidemiology," Technical Paper 115, National Bureau of Economic Research.
- Angrist, J. D., and Evans, W. N. (1998), "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 88, 450–477.
- Angrist, J. D., and Imbens, G. W. (1994), "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.
- Angrist, J. D., Imbens, G., and Rubin, D. B. (1996), "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455; 468–472.
- Bhattacharya, J., McCaffrey, D., and Goldman, D. (1999), "Estimating Probit Models with Endogenous Covariates," unpublished mimeo, RAND Corporation.
- Binder, D. A. (1983), "On the Variances of Asymptotically Normal Estimators From Complex Surveys," *International Statistical Review*, 51, 279–292.
- Bound, J., Jaeger, D., and Baker, R. (1995), "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443–450.
- Bozzette, S. A., Berry, S. H., Duan, N., Frankel, M. R., Leibowitz, A. A., Lefkowitz, D., Emmons, C., Senterfitt, J. W., Berk, M. L., Morton, S. C., and Shapiro, M. F. (1998), "The Care of HIV-Infected Adults in the United States," *New England Journal of Medicine*, 339, 1897–1904.
- Carpenter, C. C. J., Fischl, M. A., Hammer, S. M., Hirsch, M. S., Jacobsen, D. M., Katzenstein, D. A., Montaner, J. S. G., Richman, D. D., Saag, M. S., Schooley, R. T., Thompson, M. A., Vella, S., Yeni, P. G., and Volberding, P. A. (1998), "Antiretroviral Therapy for HIV Infection in 1998: Updated Recommendations of the Internal AIDS Society-USA Panel," *Journal of the American Medical Association*, 280, 78–86.
- Collier, A. C., Coombs, R. W., Schoenfeld, D. A., Bassett, R. L., Timpone, J., Baruch, A., Ones, M., Facey, K., Whitacre, C., McAuliffe, V. J., Friedman, H. M., Merigan, T. C., Reichman, R. C., Hooper, C., and Corey, L. (1996), "Treatment of Human Immunodeficiency Virus Infection with Saquinavir, Zidovudine, and Zalcitabine," *New England Journal of Medicine*, 334, 1011–1017.

- Cox, D. R. (1958), *Planning of Experiments*, New York: Wiley.
- Cutler, D. M., and Gruber, J. (1996), "The Effect of Medicaid Expansions on Public Insurance, Private Insurance, and Redistribution," *The American Economic Review*, 86, 378-383.
- Doyle, A., Jefferys, R., and Kelly, J. (1998), "National ADAP Monitoring Project: Interim Technical Report," National Alliance of State and Territorial AIDS Directors and the AIDS Treatment Data Network. Available at <http://www.aidsinfonyc.org/adap>.
- Duan, N., McCaffrey, D. F., Frankel, M. R., St. Clair, P. A., Beckman, R., Keesey, J. W., Chien, C., Goldman, D. P., Smith, S. M., and Morton, S. C. (1999), "HCSUS Baseline Methods Technical Report: Weighting, Imputation and Variance Estimation," Report MR-1060-AHCPR, RAND Corporation.
- Evans, W. N., and Schwab, R. M. (1995), "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *Quarterly Journal of Economics*, 110, 941-974.
- Evans, W. N., Farrelly, M. C., and Montgomery, E. (1999), "Do Workplace Smoking Bans Reduce Smoking?" *American Economic Review*, 89, 728-747.
- Frankel, M., Shapiro, M., Duan, N., Morton, S., Berry, S., Brown, J., Burnam, A., Cohn, S., Goldman, D., McCaffrey, D., Smith, S., St. Clair, P., Tebow, J., and Bozzette, S. (1999), "National Probability Samples in Studies of Low-Prevalence Diseases. Part II: Designing and Implementing the HIV Cost and Services Utilization Study Sample," *Health Services Research*, 34, 969-992.
- Greene, W. H. (2000), *Econometric Analysis*, (4th ed.), Upper Saddle River, NJ: Prentice-Hall.
- Harrison, T. R., Fauci, A. S., Hauser, S. L., and Braunwald, E. (eds.) (1997), *Harrison's Principles of Internal Medicine* (14th ed.), New York: McGraw-Hill.
- Heckman, J. J., and Honoré, B. (1990), "The Empirical Content of the Roy Model," *Econometrica*, 58, 1121-1149.
- Heckman, J. J., and Vytlacil, E. J. (1999), "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences*, 96, 4730-4734.
- Holland, P. (1986), "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81, 945-970.
- Hosmer, D. W., Jr., and Lemeshow, S. (1989), "Special Topics. Assessment of Fit and Diagnostics for Polytomous Logistic Regression Model," in *Applied Logistic Regression*, New York: Wiley.
- Huber, P. J. (1967), "The Behavior of Maximum Likelihood Estimates Under Non-Standard Conditions," *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*, 1, 221-233.
- Imbens, G. W., and Rubin, D. B. (1997), "Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance," *Annals of Statistics*, 25, 305-327.
- Kaiser Family Foundation (2000), "The State of the HIV/AIDS Epidemic in America," *Capitol Hill Briefing Series on HIV/AIDS*.
- Lam, N. S. N., and Liu, K. B. (1998), "Use of Space-Filling Curves in Generating a National Rural Sampling Frame for HIV/AIDS Research," *Professional Geographer*, 48, 321-332.
- Lancaster, T., and Intrator, O. (1998), "Panel Data with Survival: Hospitalization of HIV-Positive Patients," *Journal of the American Statistical Association*, 93, 46-53.
- McClellan, M., McNeil, B., and Newhouse, J. P. (1994), "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables," *Journal of the American Medical Association*, 272, 859-866.
- Nelson, C. R., and Startz, R. (1990), "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator," *Econometrica*, 58, 967-976.
- Palella, F. J., Delaney, K. M., Moorman, A. C., Loveless, J. F., Satten, G. A., Aschman, D. J., and Holmberg, S. D. (1998), "Declining Morbidity and Mortality Among Patients With Advanced Human Immunodeficiency Virus Infection: HIV Outpatient Study Investigators," *New England Journal of Medicine* 338, 853-860.
- Quandt, R. (1972), "A New Approach to Estimating Switching Regressions," *Journal of the American Statistical Association*, 67, 306-310.
- Rao, J. N. K., and Scott, A. J. (1981), "The Analysis of Categorical Data from Complex Sample Surveys: Chi-Squared Tests for Goodness-of-Fit and Independence in Two-Way Tables," *Journal of the American Medical Association*, 76, 221-230.
- Roy, A. D. (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers* (New Series), 3, 135-146.
- Rubin, D. B. (1978), "Bayesian Inference for Causal Effects: The Role of Randomization," *Annals of Statistics*, 7, 34-58.
- Rubin, D. B. (1991), "Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism," *Biometrics*, 47, 1213-1234.
- Shapiro, M. F., Berk, M. L., Berry, S. H., Emmons, C., Athey, L. A., Hsia, D. C., Leibowitz, A. A., Maida, C. A., Marcus, M., Perlman, J. F., Schur, C. L., Schuster, M. A., Senterfitt, W. A., and Bozzette, S. A. (1999a), "National Probability Samples in Studies of Low-Prevalence Diseases, Part 1: Perspectives and Lessons Learned from the HIV Cost and Services Utilization Study," *Health Services Research*, 34, 969-992.
- Shapiro, M. F., Morton, S. C., McCaffrey, D. F., Andersen, R. M., Berk, M. L., Bing, E. G., Cleary, P. D., Cunningham, W. E., Fleishman, J. A., Kanouse, D. E., Senterfitt, J. W., and Bozzette, S. A. (1999b), "Variations in the Care of HIV-Infected Adults in the United States. Results From the HIV Cost and Services Utilization Study," *Journal of the American Medical Association*, 281, 2305-2315.
- Skinner, C. J. (1989), "Introduction to Part A" in *Analysis of Complex Surveys*, C. J. Skinner, D. Holt, and T. M. F. Smith (eds.), New York: Wiley.
- Staiger, D., and Stock, J. H. (1997), "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557-586.
- United States General Accounting Office (1998), "HIV/AIDS Drugs: Funding Implications of New Combination Therapies for Federal and State Programs," Report GAO-HEHS-99-2.