Using Trivariate Copulas to Model Sample Selection and Treatment Effects: Application to Family Health Care Demand

David M. Zimmer
Department of Economics
Wylie Hall
Indiana University
Bloomington, IN 47405, U.S.A
(damzimme@indiana.edu)

Pravin K. Trivedi†
Department of Economics
Wylie Hall
Indiana University
Bloomington, IN 47405, U.S.A
(trivedi@indiana.edu)

July 08, 2005

*We thank Douglas Miller, James Prieger, Casey Quinn, and especially Tong Li for helpful comments on earlier versions of this paper. We also received excellent suggestions for improvements and changes from an Associate Editor and two anonymous referees of this journal. Zimmer also acknowledges research support provided by the dissertation fellowship award R03 HS013870-01 by the Agency for Health Care Research and Quality. We alone are responsible for the research and conclusions reported in this paper. Acknowledgments: We thank Douglas Miller, James Prieger, and especially Tong Li for helpful comments on an earlier version. Zimmer also acknowledges research support provided by the dissertation fellowship award R03 HS013870-01 by the Agency for Health Care Research and Quality. We alone are responsible for the research and conclusions reported in this paper.

†Corresponding author
ABSTRACT

Simultaneous nonlinear econometric models with discrete outcomes are often difficult to implement. This paper considers the use of the copula approach for a model with three jointly determined outcomes. It also deals with the discrete case in which outcomes include a mixture of dichotomous choices and discrete count data. We apply this technique to study self-selection and interdependence between health insurance and health care demand among married couples. The full model consists of a dichotomous choice equation for family insurance, and a separate negative binomial equation for each spouse’s health care use.

Keywords: Simultaneous equations; Interdependence; Medical Expenditure Panel Survey; Count data.

JEL Classification Codes: C35, I11
1. Introduction

Many issues in econometrics involve joint modeling of interdependent events in a nonlinear setting. Popular alternative modeling strategies are to either use a fully parametric model or to use a limited information semiparametric framework that requires partial specification of the interdependent equations. A fully parametric specification is in some cases difficult because the joint distribution of variables of interest may be unknown, or because it is difficult to simultaneously respect important features of the data to be modeled. One approach for generating joint distributions is based on correlated unobserved heterogeneity (Munkin and Trivedi, 1999; 2003). Typically this requires distributional assumptions on the latent heterogeneity components followed by the application of computationally intensive simulation-based estimation methods to integrate out the latent variables. In the parametric versions of this approach, the identification strategy exploits functional form restrictions. The computational difficulties discourage attempts at fully parametric modeling. The copula-based approach of this paper is also parametric but it avoids having to specify parametrically the distribution of latent heterogeneity, does not require the simulation machinery, and hence is computationally considerably more tractable. Joint distributions derived using copulas provide an approach to parametric modeling with other attractive features. This motivates this paper’s exploration of the copula approach to joint modeling of discrete outcomes.

A copula is a functional form of a multivariate joint distribution derived using information solely on the marginal distributions of jointly dependent random variables. When the
marginal distributions are members of well known parametric families, and they are mathematically tractable, then their properties can be analyzed and studied in depth. By contrast, joint distributions are less tractable. By beginning with marginal distributions, copula-based joint modeling offers a promising alternative. However, in econometrics only a handful of published articles have used copulas, and the literature focuses almost exclusively on bivariate analyses.

This paper extends the copula approach in two respects. First, it considers trivariate models. Second, it addresses a discrete case in which outcomes include both dichotomous choices and discrete count data. We apply our model to study health insurance and health care demand among married couples, a case in which self-selection is expected a priori. We use trivariate copulas to estimate family insurance arrangement decisions and health care demands of cohabiting spouses. The model consists of three equations, including a dichotomous choice equation for the selection of insurance status and one equation each for the use of health services by each spouse.

While copulas have been long-established in the statistics literature, applications in econometrics are more recent. The contributions of Hoeffding (1940) and Fréchet (1951) were followed by a key result of Sklar (1959) who showed that all continuous joint distributions possess unique copula representations. More recently Clayton (1978) has studied a parametric family of copulas with application to chronic disease incidence. Frank (1979) and Cook and Johnson (1981) introduce their own classes of copulas. Oakes (1982) makes use of the family of copulas studied by Clayton to model bivariate survival data. Genest (1987) provides a detailed study of
Frank’s family of copulas. Genest and Rivest (1993) study the properties of a special class of copula functions known as Archimedean copulas. For a more detailed survey of copula research in statistics, see Joe (1997) and Nelson (1999).


The remainder of the paper is as follows. Section 2 reviews some key results on bivariate copulas and discusses the concept of interdependence. Section 3 shows how copulas are extended to the trivariate case. Section 4 discusses an empirical application to family insurance
and health care demand. Section 5 presents and discusses the results of estimation, and the final section concludes.

2. The Copula Approach

This section summarizes several results on bivariate copulas, ahead of the treatment of the trivariate case which is the main focus of this paper.

2.1. Theory of Copulas

Every continuous multivariate cumulative distribution function (cdf) can be decomposed into univariate marginal cdfs that are connected by a copula function, which can be parameterized to account for dependence between the marginals. A copula is a multivariate cdf that possesses uniform univariate marginal cdfs.

Consider a bivariate cdf \( F(y_1, y_2) \) with corresponding marginal cdfs \( F_1(y_1) \) and \( F_2(y_2) \) for random variables. Following Joe (1997), the copula \( C \) associated with \( F \) is a cdf with mapping \( C : [0, 1]^i \to [0, 1] \) for \( i = 1, 2 \) such that

\[
F(y_1, y_2) = C(F_1(y_1), F_2(y_2)).
\]  

(1)

If \( F \) is continuous, and if the univariate marginal cdfs \( F_i \) have corresponding inverses \( F_i^{-1} \), then the copula in equation (1) is unique. (Uniqueness is due to two properties. First, if \( H \) is a univariate cdf with inverse \( H^{-1} \), and if \( u \) is a uniformly distributed random variate, then \( H^{-1}(u) \sim H \). Second, if \( H \) is continuous, and \( x \sim H \), then \( H(x) \) is uniformly distributed.)
The mapping $C : [0, 1]^i \to [0, 1]$ indicates that copulas are expressed in terms of cdfs, $F_1(y_1)$ and $F_2(y_2)$, that are uniform with support on the interval $[0, 1]$.

An important feature of a copula is that the functional form of a copula does not depend on the functional form of its component univariate marginal cdfs. This is useful for empirical applications because if the bivariate cdf $F(y_1, y_2)$ is unknown but the univariate marginal cdfs are known, then an appropriate choice of copula function $C$ in expression (1) provides a representation of the unknown bivariate distribution. Each marginal cdf can be specified to be conditioned on a vector of covariates. In addition, the copula can be specified to include a vector of parameters $\theta$ that captures the dependence structure between the univariate margins. It is possible for the vector to be multidimensional, but for bivariate copulas it is customary to parameterize it as a scalar measure of dependence. It is possible to derive a unique copula representation for every continuous multivariate distribution, but the same is not true for discrete random variables. The argument above does not hold for discrete measures because if $H$ is not continuous and $x \sim H$, then $H(x)$ does not have a $U(0, 1)$ distribution. For a more detailed proof, see Joe (1997, page 14).

The lack of uniqueness of a copula representation for discrete distributions is a theoretical issue which needs to be confronted in analytical proofs but does not inhibit empirical applications. Finding a unique copula for a joint distribution requires one to know the form of the joint distribution. Researchers use copulas because they do not know the form of the joint distribution, so whether working with continuous or discrete data, a pivotal modeling problem is to
choose a copula that adequately captures dependence structures of the data without sacrificing attractive properties of the marginals. We take up this issue again in Section 3.

Applying copulas to discrete data is not as straightforward as in the case of continuous data. Interpreting the dependence parameter of a copula for a discrete distribution can be difficult. Dependence parameters for most copulas are not bound on the region \([-1, 1]\), so they do not provide easy interpretations of dependence. For continuous copulas, it is customary to convert the dependence parameter to a measure such as Kendall’s tau or Spearman’s rho which are both bounded on the interval \([-1, 1]\), and they do not depend on the functional forms of the marginal distributions. For discrete data, however, Marshall (1996) and Tajar et al. (2001) explain that the usefulness of both measures is problematic because they depend on the choice of marginal distributions. Tajar et al. (2001) provides proofs for the case of binary outcome variables.

One final theoretical issue is important for interpreting copulas. Consider any \(m\)-variate joint cdf \(F(y_1, \cdots, y_m)\) with univariate marginal cdfs \(F_1, \cdots, F_m\). By definition, each marginal distribution can take any value in the range \([0, 1]\). The joint cdf is bounded below and above by the Fréchet lower and upper bounds, \(F^-\) and \(F^+\), defined as

\[
F^-(y_1, \cdots, y_m) = \max \left[ \sum_{j=1}^{m} F_j - m + 1, 0 \right]
\]

\[
F^+(y_1, \cdots, y_m) = \min[F_1, \cdots, F_m].
\]

Copulas are also restricted by the Fréchet bounds. Knowledge of Fréchet bounds is important in selecting an appropriate copula. Every copula places bounds on permissible values for its
dependence parameter $\theta$. A desirable feature of a copula is that as $\theta$ approaches the lower (upper) bound of its permissible range, the copula approaches the Fréchet lower (upper) bound. However, the parametric form of a copula may impose restrictions such that one or both Fréchet bounds are not included in the permissible range. A copula is said to be comprehensive when it has full coverage. If a researcher wishes to extend a bivariate copula to include a third marginal, it might not be desirable for the copula to allow the lower Fréchet bound in its range (see Joe, 1993). The implication is that one should choose a family of copulas that best fits the empirical application. The appendix include several examples of bivariate copulas that are popular in statistics.

2.2. Estimation Using Copulas

Estimation proceeds by first selecting an appropriate copula $C$ and marginal distributions $F_1(y_1 | X_1, \beta_1)$ and $F_2(y_2 | X_2, \beta_2)$ where $X_1$ and $X_2$ are covariates, and $\beta_1$ and $\beta_2$ are unknown parameters. ($X_1$ and $X_2$ need not be different sets of covariates. For notational convenience, we will suppress conditioning on $X$ and express the marginal distributions as $F_1(y_1)$ and $F_2(y_2).$) Then standard maximum likelihood techniques are applied to the joint distribution $C (F_1(y_1), F_2(y_2))$. If $y_1$ and $y_2$ are continuous variables and $F_1$ and $F_2$ are continuous cdfs, then the corresponding joint pdf for each observation $i = 1, \cdots, n$ is calculated as

$$
c (F_1(y_{1i}), F_2(y_{2i})) = \frac{d}{dy_2dy_1} C (F_1(y_{1i}), F_2(y_{2i})) = C_{12} (F_1(y_{1i}), F_2(y_{2i})) \cdot f_1(y_{1i}) \cdot f_2(y_{2i})
$$

(2)

where $f_1$ and $f_2$ are univariate pdfs. The log-likelihood function is formed by taking the log-
arithm of expression (2) and summing over all observations. When $y_1$ and $y_2$ are discrete variables and $F_1$ and $F_2$ are discrete cdfs, the joint probability mass function (pmf) is formed by taking differences. For the bivariate case, the pmf is

$$
c((F_1(y_{1i}), F_2(y_{2i})) = C(F_1(y_{1i}), F_2(y_{2i})) - C(F_1(y_{1i} - 1), F_2(y_{2i})) - C(F_1(y_{1i}), F_2(y_{2i} - 1)) + C(F_1(y_{1i} - 1), F_2(y_{2i} - 1)).
$$

(3)

For the trivariate case the pmf is

$$
c((F_1(y_{1i}), F_2(y_{2i}), F_3(y_{3i})) = C(F_1(y_{1i}), F_2(y_{2i}), F_3(y_{3i})) - C(F_1(y_{1i} - 1), F_2(y_{2i}), F_3(y_{3i})) - C(F_1(y_{1i}), F_2(y_{2i} - 1), F_3(y_{3i})) - C(F_1(y_{1i} - 1), F_2(y_{2i} - 1), F_3(y_{3i})) + C(F_1(y_{1i} - 1), F_2(y_{2i} - 1), F_3(y_{3i}) - 1)) + C(F_1(y_{1i} - 1), F_2(y_{2i} - 1), F_3(y_{3i}).
$$

(4)

The log-likelihood function is maximized using a quasi-Newton iterative algorithm requiring only first derivatives. After estimating the parameters simultaneously, variances of the estimates are obtained using the robust sandwich formula.

3. Trivariate Extensions

The Clayton, Frank, and Gumbel copulas are members of a special class of “Archimedean” copulas that take the general form

$$
C(u, v) = \varphi (\varphi^{-1}(u) + \varphi^{-1}(v))
$$

(5)

where $\varphi(t)$ is a generator function. For example, $\varphi(t) = (1 + t)^{-1/\theta}$ generates the Clayton copula. (See Frees and Valdez (1999) for a summary of Archimedean copula generators.)
virtue of their construction, Archimedean copulas have an attractive feature – they can be extended to easily include a third marginal distribution. We exploit this feature in our application. Among the handful of Archimedean copulas that have been used in other empirical applications, Frank’s copula is one of the few that is comprehensive, and in the bivariate case it also permits negative dependence. These properties make Frank’s copula the most attractive choice in our application.

A simple trivariate extension for Archimedean copulas is

$$ C(u, v, w) = \varphi \left( \varphi^{-1}(u) + \varphi^{-1}(v) + \varphi^{-1}(w) \right). \quad (6) $$

This construction is restrictive in empirical applications because it only has one dependence parameter. It is not possible to model separately the dependence between the three pairs $(u, v), (v, w), \text{and} (u, w)$. Next, we show how trivariate copulas with a more flexible dependence structure can be constructed from bivariate Archimedean copulas through the use of “mixtures of powers.”

**Mixtures of Powers**

The mixtures of powers approach produces two dependence parameters for a trivariate copula, so there is not a distinct parameter for each bivariate marginal distribution. However, the nature of the dependence is appropriate for many empirical applications.

Extending Archimedean copulas to include a third marginal requires the use of Laplace transformations. If $M$ is a univariate cdf, then for $s \geq 0$, the Laplace transformation of $M$ is
defined as
\[ \phi(s) = \int_{0}^{\infty} e^{-s\alpha} dM(\alpha). \]  

The cdf \( M \) is the “mixing function.” To illustrate mixtures of powers, consider the univariate case. (For a detailed discussion of mixtures of powers, see Joe (1993, 1997)). Following Joe (1993), for any univariate cdf \( F(x) \) and mixing function \( M \), there exists a unique cdf \( G(x) \) satisfying
\[ F(x) = \int_{0}^{\infty} G(x|\alpha) dM(\alpha) \equiv \phi(-\log G(x)). \]  

The term \( \alpha, \alpha > 0 \), can be thought of as unobserved heterogeneity that affects \( x \). By specifying the distribution of \( \alpha \) and integrating over its domain, the calculation in equation (8) yields a cdf \( F \) in which the heterogeneity term has been integrated out (averaged). The choice of \( M \) determines the functional form of the Laplace transformation \( \phi \), which, in turn, determines the functional form of \( F \). Solving this equation for \( G \) gives \( G(x) = \exp(-\phi^{-1}(F(x))) \). (Note that \( \alpha \) must be positive, or else one cannot solve for \( G \) (see Joe, 1997 page 86).) Bivariate distributions can also be expressed as mixtures of powers. Since copulas are distribution functions, bivariate copulas can be expressed in terms of mixtures of powers as
\[ C(u, v; \theta) = \int_{0}^{\infty} G(u|\alpha)G(v|\alpha) dM(\alpha) \equiv \phi(-\log G(u) - \log G(v)) \]
\[ = \phi(\phi^{-1}(F(u)) + \phi^{-1}(F(v))) \]

where \( M \) is a univariate mixing function, and \( \alpha \) is a heterogeneity term that enters the distributions of both \( u \) and \( v \). (Any functional forms for the \( G \)s can be used, in principle, but if they are not in terms of powers with respect to \( \alpha \), then the integral in expression (9) usually does
not have a closed form expression.) Similar to the univariate case, $G(u) = \exp(-\phi^{-1}(u))$, and $G(v) = \exp(-\phi^{-1}(v))$. If $F(u)$ and $F(v)$ are $U(0,1)$, then (9) takes the Archimedean form given in expression (5). Therefore, Archimedean copulas such as the Clayton, Frank, and Gumbel can be expressed in terms of mixtures of powers.

We exploit the mixtures of powers representation to extend the Frank copula to include a third marginal. Since the Clayton and Gumbel copulas are also members of the Archimedean family, they can also be extended to include a third marginal. For the empirical application in this paper, we use the Frank copula because of its popularity in the statistics literature, because it provides the best fit to the data used in our empirical application as well as other applications in which we have been involved, and because it appears to be more stable in dealing with large count values. (In separate research in which we have been involved, the three copulas have produced nearly identical results. Therefore, we do not believe inference in this paper is the result of a particular choice of copula.)

The trivariate mixtures of power representation is

$$C(u, v, w) = \int_0^\infty \int_0^\infty G(u|\beta)G(v|\beta)dM_2(\beta; \alpha)G(w|\alpha)dM_1(\alpha)$$

where $G(u) = \exp(-\phi^{-1}(u))$, $G(v) = \exp(-\phi^{-1}(v))$, $G(w) = \exp(-\psi^{-1}(w))$, and $\psi$ is a Laplace transformation. In this formulation, the heterogeneity term $\alpha$ affects $u$, $v$, and $w$, and a second heterogeneity term $\beta$ affects $u$ and $v$. The distribution $M_1$ has Laplace transformation $\psi(\cdot)$, and $M_2$ has Laplace transformation $((\psi^{-1} \circ \phi)^{-1}(\alpha^{-1} \log(\cdot)))^{-1}$. Thus, the copula approach accommodates unobserved heterogeneity in the three outcomes. When $\phi = \psi$
expression (10) simplifies to expression (6). (The mathematical notation $f \circ g$ denotes the functional operation $f(g(x))$.) When $\phi \neq \psi$, the trivariate extension of (5) corresponding to (10) is

$$C(u,v,w) = \psi \left( \psi^{-1} \circ \phi[\phi^{-1}(u) + \phi^{-1}(v)] + \psi^{-1}(w) \right).$$

(11)

Therefore, different functional forms of the Laplace transformation produce different families of trivariate copulas.

Expression (6) has symmetric dependence in the sense that it produces one dependence parameter $\theta = \theta_{uv} = \theta_{uw} = \theta_{vw}$. But the dependence properties of three different marginals are rarely symmetric in empirical applications. The trivariate representation of expression (11) is symmetric with respect to $(u,v)$ but not with respect to $w$. Therefore, (11) is less restrictive than (6).

Ideally, an $m$-variate copula would have $m(m-1)/2$ dependence parameters corresponding to each pair of marginals. For a trivariate copula, this means there would be three dependence parameters: $\theta_{uv}, \theta_{uw},$ and $\theta_{vw}$. However, the partially symmetric formulation of expression (11) yields two dependence parameters, $\theta_1$ and $\theta_2$, such that $\theta_1 \leq \theta_2$. The parameter $\theta_2 = \theta_{uv}$ measures dependence between $u$ and $v$. The parameter $\theta_1 = \theta_{uw} = \theta_{vw}$ measures dependence between $u$ and $w$ as well as between $v$ and $w$, and the two must be equal. (Distributions greater than three dimensions also have a mixtures of powers representations. But this technique yields only $m - 1$ dependence parameters for an $m$-variate distribution function. Therefore, the mixtures of powers approach is more restrictive for higher dimensions.) While this restric-
tion constitutes a potential weakness of the approach, it is less restrictive than formulation (6) which yields only one dependence parameter. Moreover, (11) allows a researcher to explore several dependence patterns by changing the ordering of the marginals. For example, instead of \((u, v, w)\), one could order the marginals \((w, v, u)\), which provides a different interpretation for the two dependence parameters.

As will be discussed in more detail below, the empirical application in this paper orders the marginals as follows: (1) wife’s health care utilization; (2) husband’s health care utilization; (3) family health insurance choice. This ordering provides a convenient interpretation for the dependence parameters: \(\theta_2\) measures correlation between spouses’ utilizations, and \(\theta_1\) is a general indicator of the presence of family self-selection into an insurance arrangement. This setup of trivariate dependence is consistent with the so-called “shared” unobserved heterogeneity models with two pairs of correlated components, one pair for the utilization equations, and another that is common to each utilization equation and the insurance equation.

There is no known general multivariate extension of a bivariate family of copulas that has a dependence parameter for each pair of marginals. It is important to note that \(m\)-variate copulas with \(m(m - 1)/2\) dependence parameters do exist that are not extensions of bivariate copulas, but they are not easy to implement. The most obvious is the normal copula discussed by Lee (1983), which can be extended to include a third marginal. The covariance structure of the multivariate normal distribution has three dependence parameters. However, implementing the trivariate normal copula requires calculation of a triple integral without a closed form solution,
which must be approximated numerically. Husler and Reiss (1989) provide a multivariate copula with $m(m - 1)/2$ dependence parameters, but their specification relies on the multivariate normal distribution and therefore also requires numerical integration. Joe (1990) provides multivariate copulas with flexible dependence, but they fail to produce consistent results. In a later paper (Joe, 1994), he develops multivariate copulas with $m(m - 1)/2$ dependence parameters, but these models also require numerical integration; some of his models required approximately four weeks to achieve convergence at the time of his publication. Although advances in computational power have rendered less serious the practical difficulties associated with numerical integration, a simulation-based model, presented in Section 5, that mimics properties of the trivariate copula still may require substantial computational resources as well as knowledge of simulation acceleration techniques. Given these complications, the mixtures of powers approach appears to be appropriate for most trivariate applications.

The Trivariate Frank Copula

Using the mixtures of powers approach, the Frank copula is extended to the trivariate case.

Frank Copula:

If $\phi(s) = \exp(-s^{1/\theta})$ and $(\psi^{-1} \circ \phi)(s) = s^{\theta_1/\theta_2}$, then expression (11) becomes

$$C(u, v, w; \theta_1, \theta_2) = -\theta_1 \log \left\{ 1 - c_1^{-1} \left( 1 - [1 - c_2^{-1}(1 - e^{-\theta_2 w})(1 - e^{-\theta_2 w})]^{\theta_1/\theta_2} \right) \left( 1 - e^{-\theta_1 w} \right) \right\}$$

(12)

where $\theta_1 \leq \theta_2$, $c_1 = 1 - e^{-\theta_1}$, and $c_2 = 1 - e^{-\theta_2}$. (The proof is complicated; see Joe (1993).)

Both dependence parameters $\theta_1$ and $\theta_2$ must be greater than zero.
This extension to include the third marginal has one limitation. We sacrifice the property of being able to model negative dependence. That is, our formulation requires $\theta_1, \theta_2 \geq 0$. In the context of our application this restriction is quite plausible. Further, since the trivariate Frank imbeds a bivariate Frank marginal, our bivariate marginal will be modelled with greater flexibility, which is desirable. Notwithstanding these arguments, AIC or BIC model selection criteria are always available for formal (non-nested) model comparison and selection.

4. Application to Family Health Care Demand and Health Insurance

In this section, we apply the trivariate copula methodology to jointly model family health care demands and health insurance arrangement decisions. The goal is to assess to what extent family insurance arrangements and health care demand are interrelated.

It is well known that the majority of private health insurance coverage in the United States is financed through employers (Cutler and Zeckhauser, 2000). Families with two working spouses might be confronted with more insurance choices than if only one spouse is employed. Even when one or both spouses are not employed, insurance might still be available from past employers through continuing coverage as part of a severance package. The implication is that spouses must decide whether to enroll together in the same insurance plan, or they might acquire separate plans. Health care demands might be related to a couple’s choice of a particular insurance arrangement.

The question of empirical interest in this application is whether enrollment in separate
plans and spousal health care use are jointly determined. If a husband’s and wife’s utilizations and their insurance choice are simultaneously determined, then estimating the three outcomes jointly provides efficiency gains compared to estimating the three marginals separately. Estimation is complicated because the joint distribution of spouses’ health care utilizations and the family’s insurance choice might not assume an analytically tractable form if the outcomes are discrete and/or nonlinear. Therefore, we employ the trivariate copula approach. The copula’s facility to accommodate unobserved heterogeneity is especially important in an application of health care demand which necessarily includes unobserved individual specific health-related variables because a claim that one has controlled for all factors that simultaneously affect insurance choice and health care consumption is not credible.

The copula used here respects the dependence structure between the three marginals by permitting two types of interdependence. First, the formulation of the copula approach allows for a measure of family selection into insurance status. Selection on unobservables is present if unobserved traits that lead a family to choose a certain insurance arrangement also cause the family to alter its health care demands. For example, if a family anticipates particular health care needs in the future, it might choose an insurance arrangement that provides generous coverage of appropriate services. Second, the copula approach allows for a measure of interdependence between spouses’ health care demands based on unobserved family heterogeneity. Individuals tend to marry partners who are similar in income, race, age, education, life style, and health traits. Consequently, spouses share environments and behaviors that affect medical needs. In
view of marital matching and shared behaviors of spouses, it is important that a model of family health care utilization allow for interdependence in utilization between spouses.

The marginal distributions are ordered as follows: (1) the wife’s health care demand; (2) the husband’s health care demand; and (3) the family’s insurance arrangement choice. This ordering provides a convenient interpretation for the dependence parameters: $\theta_2$ measures correlation between spouses’ utilizations, and $\theta_1$ is a general measure of a family’s selection into a particular insurance arrangement based on unobservables. To check whether the degree of dependence varies by type of use, we estimate separate empirical models for three different measures of health care use.

4.1. Specification of the Marginal Distributions

The wife and husband in family $i$ have latent propensities to use health care services denoted as $y_{i,w}^*$ and $y_{i,h}^*$, and the couple’s latent propensity to enroll in separate plans is denoted as $\text{DIFF}_i^*$. The equations underlying these latent variables are assumed to be linear,

\begin{align}
y_{i,w}^* &= x_{i,w}' \beta_w + u_{i,w} + \lambda_{v,w} v_i \\
y_{i,h}^* &= x_{i,h}' \beta_h + u_{i,h} + \lambda_{v,h} v_i \\
\text{DIFF}_i^* &= z_i' \alpha + \lambda_u (u_{i,w} + u_{i,h}) + \epsilon_i
\end{align}

where $x_{i,w}$ and $x_{i,h}$ are vectors of explanatory variables that affect the wife’s and husband’s utilizations, respectively. The vector $z_i$ consists of variables that affect insurance choice. The terms $\beta_w$, $\beta_h$, and $\alpha$ are coefficients to be estimated, and the $\epsilon_i$ are independently distributed
error terms, assumed to be uncorrelated with \((u_{i,w}, u_{i,h}, v_i)\). Dependence between spouses’ utilizations and insurance choice arises from the shared terms \(u_i\) and \(v_i\), which are not observed by the researcher. The terms \(u_{i,w}\) and \(u_{i,h}\) represents unobserved characteristics that affect insurance choice as well as each spouse’s utilization. For example, health is likely to affect both insurance choice and utilization, but health status is only partially observed in household surveys. In a similar manner, the term \(v_i\) represents unobserved family traits, independent from those contained in \(u_i\), that cause spouses’ utilizations to be correlated. For example, wives and husbands might share similar diets and lifestyles that affect spousal health care consumption.

In the above exposition the outcome variables \(y_{i,w}^*, y_{i,h}^*,\) and \(\text{DIFF}_i^*\) are unobserved latent tendencies with corresponding observable variables that can be studied. The observable utilization variables \(y_{i,w}\) and \(y_{i,h}\) are recorded as discrete counts of the number of visits to certain health care providers. The wife’s utilization probability for a particular medical service is assumed to follow a negative binomial 2 specification (NB2),

\[
f(y_{i,w}|\mu_{i,w}) = \frac{\Gamma(y_{i,w} + \psi)}{\Gamma(\psi)\Gamma(y_{i,w} + 1)} \left( \frac{\psi}{\lambda_{i,w} + \psi} \right)^\psi \left( \frac{\lambda_{i,w}}{\lambda_{i,w} + \psi} \right)^{y_{i,w}},
\]

where \(\mu_{i,w} = \exp(x_{i,w}'\beta)\) is the conditional mean, and \(1/\psi (\psi > 0)\) is an overdispersion parameter in the conditional variance \(\mu_{i,w}(1 + \psi \mu_{i,w})\). This functional form provides substantial modeling flexibility (Cameron and Trivedi, chapter 4, 1998). There is a corresponding probability function for the husband. We consider three measures of utilization: physician visits, nonphysician visits, and emergency room visits. Examples of nonphysicians include nurse practitioners and physiotherapists.
The variable $\text{DIFF}_i$ is a dichotomous variable indicating whether the husband and wife in couple $i$ are enrolled in separate insurance plans. Couple $i$’s decision to enroll in separate plans follows a probit specification for which the contribution to the unlogged likelihood function is $L_{\text{DIFF}} = [\Phi(z'_i \alpha)]^{\text{DIFF}_i} [1 - \Phi(z'_i \alpha)]^{(1 - \text{DIFF}_i)}$, where $\Phi$ represents the cumulative standard normal distribution.

Expressions (16) and $L_{\text{DIFF}}$ are pmf representations while copulas use cdfs as arguments. The cdf of the NB2 specification for the wife is obtained by summing from zero to $r_w$. If the wife in family $i$ uses a medical service $r_w$ times, then her cdf is calculated as

$$\Pr[y_{i,w} \leq r_w | \mu_{i,w}] = \sum_{k=0}^{r_w} f(y_{i,w} = k | \mu_{i,w}). \quad (17)$$

The husband’s cdf is calculated analogously. For the insurance arrangement decision a cdf representation is also needed. When $\text{DIFF}_i = 1$, the cdf is $L_{\text{DIFF}=0} + L_{\text{DIFF}=1}$, and when $\text{DIFF}_i = 0$, the cdf is $L_{\text{DIFF}=0}$.

Spouses’ utilizations and family insurance choice depend, in part, on the error terms in equations (13), (14), and (15). Joint estimation is preferable to single equation estimation because variation in utilization and insurance choice due to heterogeneity in $u_{i,w}$, $u_{i,h}$, and $v_i$ should be captured by the dependence parameters of the trivariate copula. The first two marginals in the copula are the wife’s NB2 cdf, and the husband’s NB2 cdf, respectively, for health care use. The third marginal is the probit model. The three marginals are substituted into the Frank copula specified in the previous section, and the joint probability mass is obtained through the difference equation given by (4).
An important practical issue in our application is that there are many observations where the wife and/or husband have zero utilization. And, of course, when spouses are enrolled in the same plan, the outcome of the third marginal is zero, as well. As is evident from an inspection of the terms in the equation (4), this will lead to the presence of some terms defined at negative support points. However, the marginals have by definition zero probability mass at these points, and hence the corresponding terms in the likelihood are also zero. Note that this problem has no connection with the well-known issue of “excess zeros” that has been extensively discussed in both count data and medical expenditure literature.

4.2. Data and Summary Statistics

Data come from the 1996, 1997, 1998, and 1999 waves of the Medical Expenditure Panel Survey (MEPS) conducted by the Agency for Healthcare Research and Quality (AHRQ). MEPS data are nationally representative and contain information on socioeconomic variables, health care utilization, health status, employment, and income. Our choice of these covariates is fairly standard in the health economics literature.

MEPS data refer to individuals, but household identifiers are available so that family members can be matched. By matching individuals with their reported spouses, a sample was created consisting of husband/wife pairs. The sample is restricted to married couples residing in the same household. Since this paper focuses on private insurance arrangements, the sample includes only couples reporting some form of private insurance and no form of public insurance. It is necessary to restrict the age range of the sample. Individuals younger than 18 are often
dependents in parents’ insurance policies and have minimal input in family insurance decisions. At the other end of the spectrum, most individuals avail themselves of Medicare benefits at age 65. Many people purchase private insurance to supplement Medicare coverage, but those decisions are likely affected by the presence of Medicare coverage. Consequently, the sample is limited to persons between 18 and 64 years of age. The final sample size is 6,636 married couples.

Table 1 presents variable definitions and summary statistics. Socioeconomic variables include age, education, family size, income, race, veteran status, and employment characteristics. Health indicators include self-reported health and number of chronic health conditions. Each person reports whether he is the principal holder of the insurance policy under which he is covered. When each spouse reports to be the principal holder, it is assumed that they both hold separate insurance plans, which is indicated by the dummy variable DIFF. (The validity of this assumption has been checked with MEPS data experts at the AHRQ.) It can be argued that employment and health insurance access are jointly determined (Gruber, 2000), so we eliminate from the sample individuals who are not privately insured.

Wives tend to be more intensive users of healthcare than husbands, and spouses in separate plans appear to have slightly more physician visits and a similar number of nonphysician visits and emergency room visits compared to couples enrolled in the same plan. Couples enrolled in separate plans have a higher proportion of employed wives and lower rates of self-employment for both spouses. Moreover, couples enrolled in separate plans appear to have slightly higher
incomes, fewer children, and a higher proportion of blacks. For men, a dummy variable indicates whether the husband served in any twentieth century military conflict. Approximately 19 percent of the population served in the military. There have been several studies of post-war experiences of military veterans (Berger and Hirsch 1983, Angrist, 1990, Angrist and Krueger, 1994), but they are primarily concerned with earnings and not health care utilization. Regardless, anecdotal evidence that military veterans experience injuries and health conditions that require increased medical attention suggests the veteran status dummy as a variable to include in a health utilization equation. Sample averages of other variables are consistent with national averages for the population of privately insured married couples.

5. Results

This section presents estimation results. Insurance choice results are discussed first, and utilization results follow.

5.1. Insurance Choice Results

The empirical model considers three measures of utilization: physician visits, nonphysician visits, and emergency room visits. For each measure, the Frank copula is estimated with three marginals: wife’s utilization, husband’s utilization, and family insurance choice. Since insurance choice is modeled as a family decision, this part of the copula contains family specific covariates such as family size and income. But including the full set of covariates for each spouse causes potential collinearity problems because spouses share similar characteristics. For
example, interracial couples are rare, so including race dummies for each spouse produces a data matrix that is nearly colinear. Accordingly, spouse specific variables measuring age and race of the husband are included in the insurance choice equation, but the wife’s age and race are excluded. On the other hand, education and employment characteristics of each spouse appear in the insurance choice specification since they frequently differ between spouses. To emphasize this feature, variables that correspond to the husband have prefix “H_”, and those that correspond to the wife have the prefix, “W_”; e.g., W_EMPLOYED indicates whether the wife is employed, and H_EMPLOYED indicates whether the husband is employed. In preliminary analyses, self-reported health and chronic conditions were not found to affect insurance arrangement decisions, so they are not included in the insurance choice specification.

Employment variables and family income are included in the insurance choice equations because of the prevalence of employer-provided insurance in U.S. private health insurance markets. Employment variables include employment status of each spouse and whether each spouse is self-employed. Children variables are also included such as whether the couple has a child between ages 0 and 6, whether the couple has a child between ages 7 and 17, and whether the couple has a child who is less healthy than other children of the same age. In preliminary analysis, employment, income, and children variables were found to have a negligible effect on utilization, so they are not included in the usage equations.

It is important to note that employment and children variables are not exclusion restrictions in the sense of identification in simultaneous equations models because each marginal involves
only one endogenous variable. Specifically, insurance choice does not appear as an explanatory variable in the utilization equations. This reflects the fact that copulas are expressed in terms of marginal distributions rather than conditional distributions. Therefore, there is no issue of identification to be considered in this application. Nevertheless, as discussed below, it is possible post estimation to back out the direct effect of insurance on utilization. (However, the model involves exclusion restrictions in the sense of the seemingly unrelated regressions (SUR) model. It is well known that in the linear case simultaneous estimation of all equations leads to efficiency gains only if the explanatory variables in different equations are not identical. In such cases the term “exclusion restrictions” refers to the presence of nonoverlapping variables in different equations.)

Insurance choice results appear to be robust with respect to the three measures of utilization, but there is some variation in the precision of estimates. Results are shown in Table 2. Larger families (FAMSIZE) tend to enroll in common coverage in one plan. When the husband is black or Hispanic (H_BLACK or H_HISPANIC), the family is more likely to enroll in separate plans. Higher income couples are more likely to enroll in separate plans. Several employment variables are significant. When either spouse is employed (H_EMPLOYED and W_EMPLOYED), the family is more likely to enroll in separate plans, but when either spouse is self-employed (H_SELFEMP and W_SELFEMP), the couple is less likely to enroll in separate plans. This is probably because employed individuals are more likely to have insurance options offered at discounted group rates, but self-employed workers are less likely to have such options available.
If a couple has a child younger than 17 (KID0-6 and KID7-17), then spouses are more likely to enroll in the same plan.

5.2. Utilization Results

During estimation of the physician visits and nonphysician visits models, largely for reasons of computational stability we eliminate from the sample individuals reporting more than 30 visits in a given year. Computationally, large count integers cause convergence instability. The complication stems from calculation of negative binomial cdfs in equation (17). Since the negative binomial pmf contains gamma functions that approach infinity for large count values, the summation of many such pmfs causes instability. This sample restriction reduces the sample size of each model by approximately 1 percent, perhaps mainly pertaining to patients who may require chemotherapy or radiology, often at weekly or biweekly frequency. Although such patients are not representative of the “average user”, they provide information about the behavior in the upper tail of the user distribution. Whether such truncation also affects the estimates of the key dependence parameters of interest, especially $\theta_1$, is an unresolved issue. In preliminary analysis under the assumption of insurance exogeneity, where the computation of cdfs is not involved, the inclusion of these observations did not substantially alter the empirical results. Hence we do not suspect that this sample truncation leads to distorted inference. Further, the results from a simulation-based model discussed below also support the finding that reported results are robust.

Estimates of the utilization parts of the models are presented in Table 3. Copula estima-
tion yields interesting results for the dependence parameters $\theta_1$ and $\theta_2$. A positive value of $\theta_1$ indicates that the family’s insurance decision is related to each spouse’s health care utilization even after controlling for covariates. The parameter $\theta_1$ is positive and significant with respect to emergency room visits and marginally significant with respect to physician visits. The interpretation is that unobserved family characteristics that increase the probability of spouses enrolling in separate plans also induce more frequent use of these particular services. If unobserved family characteristics are related to self-perceived health problems not known to insurers, then this result can be interpreted as a form of adverse selection in which relatively unhealthy families choose separate plans instead of common coverage. This would be especially true if separate plans tend to provide each spouse with less expensive access to particular medical services. However, as discussed in Section 3, it is not possible to determine whether selection is due to wives, husbands, or both; the dependence must be symmetric between the two spouses. It is also impossible to account for favorable selection, because dependence parameters must be positive. (This is a potential weakness of the trivariate copula approach, but results of a simulation-based model, discussed later, suggest that the selection parameter is positive for two of the three measures of utilization.) The lack of significance of $\theta_1$ in the nonphysician visits equation is intuitive because many nonphysician services are not covered by typical private insurance plans. Therefore, it is not surprising that insurance decisions and the consumption of nonphysician services do not appear to be jointly determined.

A positive value for $\theta_2$ indicates that spouses’ utilizations are positively correlated after
conditioning on observed factors. Results indicate that positive association is present for all three measures of utilization. This result is intuitive in light of the fact that individuals tend to marry partners with similar life styles and interests. Such similarities put spouses at risk to acquire similar illnesses and incur related injuries. Moreover, since married people live in close proximity, spouses are at risk of passing contagious ailments to each other. It is important to note that the trivariate Frank copula cannot account for negative correlation between spouses’ utilizations because $\theta_2$ cannot be negative. (Results from a simulation-based model, discussed below, suggest that correlation between spouses’ utilizations is positive, so the restriction is data consistent. It might be argued that families are subject to a budget constraint, and, thus, utilizations might be negatively correlated. However, our sample includes only privately insured individuals. After premium payments, health care costs are close to zero for much of the sample. Therefore, the budget constraint effects will be minimal.)

Table 3 presents the full sets of results for the utilization equations. Overall, blacks and Hispanics have fewer physician visits and nonphysician visits than their nonblack and nonHispanic counterparts. Age is positively associated with husbands’ usage of physician services, but age is negatively associated with wives’ utilization of physician and nonphysician services. This is probably because we are unable to adequately control for pregnancies, which tend to occur when spouses are relatively young. Age is negatively associated with emergency room visits. More education is associated with greater use of physician and nonphysician services for both spouses but negatively related to emergency room utilization. Spouses from large families have
fewer physician and nonphysician visits. The most important variables both in terms of magnitude and significance are health status variables. Not surprisingly, individuals in poor health and those afflicted with chronic conditions are more intense users of health care services.

5.3. Measures of Fit

Likelihood values for each copula model are presented at the bottom of Table 3. The ratio of $\hat{\theta}_1$ and $\hat{\theta}_2$ to their estimated standard errors are large except in the case of emergency room visits. These estimates indicate positive association. Because hypotheses tests about the association parameters here involve boundary hypotheses, the likelihood ratio (LR) test statistic has a non-standard distribution, whereas under the standard set-up it would follow a $\chi^2(2)$ distribution because the copula models include two dependence parameters. The LR statistics for physician visits, nonphysician visits, and emergency room visits are 175.68, 157.01, and 33.03. If the $\chi^2(2)$ distribution provided a reasonable approximation to the correct asymptotic distribution of the test, then these values would easily reject the null hypothesis of zero interdependence and would indicate that the copula provides a superior fit compared to separate marginal models. Although caution is necessary, the precision with which one or both dependence parameters are estimated in the models suggests that there is evidence of simultaneity.

Table 4 shows actual and fitted utilization frequencies calculated by substituting coefficient estimates in the marginal distributions. For each individual in the sample, the calculated probability corresponds to the use of a particular medical service $r$ times for $r = 1, \ldots, 9$; the values are then averaged across all observations. For example, among the sample of wives, 22 percent
have no doctor visits while 18 percent have one doctor visit. The Frank copula predicts that 22 percent and 17 percent of wives have zero and one doctor visit, respectively. Both models appear to provide satisfactory fits with respect to all three measures of utilization.

Despite the fact that employment variables have been successfully used as explanatory variables for insurance choices in similar applications, it may be argued that individual heterogeneity across different statuses of employment are large enough to warrant separate models for subgroups. To explore this possibility, the models were reestimated for two subsamples: couples in which both spouses are employed, and couples in which both spouses are employed but not self-employed. Results from the subsamples are quantitatively similar to results from the full sample, but there is some variation in the precision of the estimates.

5.4. Treatment Effects

The dependence parameter $\theta_1$ measures the degree to which a family’s insurance arrangement decision is related to its health care utilization. This relation is composed of two separate effects. First, as discussed above, $\theta_1$ includes the indirect selection effect of being enrolled in different plans. This is the effect of unobserved family traits that effect insurance choice and each spouse’s health care utilization. In addition, $\theta_1$ also measures the direct causal effect on utilization of being enrolled in different plans. Despite the fact that DIFF is not included as an explanatory variable in the utilization equations, its direct effect on utilization can be extracted from estimates of $\theta_1$.

The joint density $f(y_w, y_h, \text{DIFF}|x_w, x_h, z, u_w, u_h, v)$ can be decomposed by Bayes’ Rule
\[ f(y_w, y_h, \text{DIFF}|x_w, x_h, z, u_w, u_h, v) = f(y_w, y_h|x_w, x_h, \text{DIFF}, v) \cdot f_{\text{DIFF}}(\text{DIFF}|z) \quad (18) \]

where \(u_w\) and \(u_h\) no longer appears on the right hand because utilization is conditional on DIFF. Therefore, the distribution of utilization conditional on insurance choice is expressed as

\[ f(y_w, y_h|x_w, x_h, \text{DIFF}, v) = \frac{f(y_w, y_h, \text{DIFF}|x_w, x_h, z, u_w, u_h, v)}{f_{\text{DIFF}}(\text{DIFF}|z)}. \quad (19) \]

The numerator in expression (19) is estimated by the Frank copula, and the denominator is estimated by maximum likelihood estimation of a probit model. By substituting estimated coefficients and dividing these two quantities, a representation of the conditional distribution is obtained from which the causal effect of DIFF on utilization is determined.

The effect of DIFF on \(y_w\) is calculated by taking the expectation with respect to \(y_w\) and summing over all possible values of \(y_h\),

\[ E(y_w|\text{DIFF}, x_w, x_h, \Omega) = \sum_{y_h=0}^{\infty} \sum_{y_w=0}^{\infty} \frac{f(y_w, y_h, \text{DIFF}|x_w, x_h, z|\Omega)}{f_{\text{DIFF}}(\text{DIFF}|z|\Omega)} y_w \quad (20) \]

where \(\Omega\) represents estimation coefficients. The average treatment effect (ATE) of being enrolled in different plans is estimated as

\[ \hat{\text{ATE}} = E(y_w|\text{DIFF} = 1, x_w, x_h, z|\hat{\Omega}) - E(y_w|\text{DIFF} = 0, x_w, x_h, z|\hat{\Omega}) \quad (21) \]

where the covariates are set to their sample means. The husband’s treatment effect is calculated analogously. To calculate the standard deviation of \(\hat{\text{ATE}}\), a Monte Carlo procedure is employed.
in which we draw $\Omega$ from a normal distribution with mean $\hat{\Omega}$ and variance equal to the estimated covariance matrix of coefficients and recalculate ATE. This procedure is replicated 500 times, and the standard deviation of these 500 draws is also reported along with ATE in Table 5, panel A.

It is not surprising that treatment effects are small and relatively widely dispersed. The only link between DIFF and utilization is through the dependence parameter $\theta_1$. Estimates of $\theta_1$ are quantitatively small compared to estimates of $\theta_2$ for the three measures of utilization. The interpretation is that the extent to which a family’s insurance arrangement is related to its health care utilization is small relative to the extent to which spouses’ utilisations are related to each other. Moreover, most of the relationship between insurance arrangements and utilization is not explained by a direct causal effect of insurance arrangements, but rather it is explained by spouses selecting themselves into particular insurance arrangements based on unobserved family characteristics that also affect utilization. In a separate empirical paper Zimmer (2005) analyzes in greater detail alternative explanations of observed health insurance arrangements among the married couples.

5.5. Comparison with Simulation-Based Estimates

There are a number of simulation-based methods for estimating nonlinear simultaneous equations models with correlated unobserved heterogeneity. However, maximum simulated likelihood (MSL) seems appropriate in the present context. We construct a latent factor model in the MSL framework that is comparable to the trivariate copula model that we have developed.
We recognize the disadvantage of an MSL method that the number of simulation draws must increase with the sample size.

Recall from equations (13), (14), and (15) that the composite error terms corresponding to the outcome variables are

\[
\eta_{i,w} = u_{i,w} + \lambda_{v,w} v_i \\
\eta_{i,h} = u_{i,h} + \lambda_{v,h} v_i \\
\eta_{\text{DIFF},i} = \lambda_u (u_{i,w} + u_{i,h}) + \epsilon_i.
\]

If \( u_w, u_h, v, \) and \( \epsilon_i \) are independent and standard normally distributed random variables, then the covariances of the composite error terms are

\[
cov(\eta_w, \eta_h) = \lambda_{v,w} \cdot \lambda_{v,h} \\
cov(\eta_w, \eta_{\text{DIFF}}) = \lambda_u \\
cov(\eta_h, \eta_{\text{DIFF}}) = \lambda_u.
\]

If the \( \lambda \) coefficients could be estimated, they would provide measures of dependence similar to \( \theta_1 \) and \( \theta_2 \) in the copula approach: \( \lambda_u \) would measure how insurance choice is related to family health care consumption, and \( \lambda_{v,w} \lambda_{v,h} \) would indicate the degree of spousal correlation in utilization. However, latent variables \( u_w, u_h, \) and \( v \) are not observed. Therefore, random values of the three latent variables are drawn from independent standard normal distributions, and the simulated values, \( (u_w, u_h, v) \) are treated as observed variables in the joint density function of
the three outcomes. This process is repeated $S$ times, and the average of the $S$ draws provides a simulated density,

$$
\tilde{f}(y_w, y_h, DIFF|x_w, x_h, z) = \frac{1}{S} \sum_{s=1}^{S} f(y_w, y_h, DIFF|x_w, x_h, z, \tilde{u}_{ws}, \tilde{u}_{hs}, \tilde{v}_s) \tag{22}
$$

Maximization of the likelihood function based on the simulated density is called Maximum Simulated Likelihood (MSL.) (For a more detailed exposition MSL, see Train (2003).) Because variances of latent factors cannot be identified, a normalization on either $\lambda_{v,w}$ or $\lambda_{v,h}$ is required. In results discussed below, $\lambda_{v,w}$ is normalized to unity so that the estimate $\hat{\lambda}_{v,h}$ measures correlation between spouses’ utilizations.

The top half of Table 6 reproduces results from Table 3, and the bottom half shows results from the MSL model. Some coefficients between the two models cannot be directly compared, because the functional forms of the copula model and the MSL specification are different. Nevertheless, a qualitative comparison of the two models reveals several similarities. Both models indicate significant correlation between spouses’ utilizations (measured by $\theta_2$ in the copula model and $\lambda_{v,h}$ in the MSL model), and both reveal that correlation is largest in magnitude for nonphysician utilization followed by ER and physician usage. Regarding selection (measured by $\theta_1$ in the copula model and $\lambda_u$ in the MSL model), the MSL model does not reveal significant selection with respect to any of the measures of utilization. Although the copula model finds selection with respect to ER utilization and modest selection with respect to physician use, both estimation approaches reveal that selection effects are quantitatively small, especially in comparison to correlation between spouses’ utilizations.
Table 5, panel B, shows MSL treatment effects and their estimated standard errors. Similar to the copula model, treatment effects are small and dispersed. Coefficients of explanatory variables in the conditional mean function, available upon request from the authors, are virtually the same between the two estimation techniques.

Because both models contain the same number of parameters, differences in information criteria are due to differences in log-likelihood values. The log-likelihood for the copula model is larger (less negative) for physician visits but smaller for nonphysician visits and ER utilization. The implication is that, a priori, it is not clear which of the two nonnested models is significantly superior; model performance depends on the measure of utilization. One advantage of the MSL approach is that individuals with more than 30 visits can be included in estimation, because equation (22) does not require calculation of negative binomial cdfs as does the copula approach. Table 7 shows that including these observations does not change either the quantitative or qualitative conclusions.

In our example, the ability of the MSL model to accommodate large counts comes at a substantial computational cost. The joint density in equation (22) is based on three simulated latent factors, which requires many more simulation draws than a comparable model with only one latent factor. Results presented in Table 6 require \( S = 3000 \), which is much larger than the number of simulations used in other MSL studies (e.g., Munkin and Trivedi, 1999). For the MSL model to achieve convergence in a reasonable amount of time, we employ a variety of simulation acceleration techniques, including antithetic sampling and quasi-random Halton
draws. (See Train (2003) for a discussion of simulation acceleration techniques for simulation-based models of health care demand.) Without such acceleration methods, the MSL model requires approximately twenty to thirty times longer to achieve convergence. By comparison, the trivariate copula model in equation (4) is the basis for a conventional likelihood function, albeit a complicated function, and convergence is usually achieved in less than 60 minutes. Furthermore, the simulation-based model requires approximately three times more memory than the copula model, a significant disadvantage when working with large samples. Thus, a major advantage of the copula approach is that complicated nonlinear simultaneous equations models that were previously beyond the scope of conventional likelihood estimation can be consistently and efficiently estimated with relative ease.

6. Conclusion

Econometricians are often interested in modeling outcomes that are jointly determined, but existing methods of estimating joint models present challenges, particularly for nonlinear or discrete outcomes. The joint distribution can be approximated by simulation-based techniques, usually under strong functional form restrictions, but these procedures can be computationally burdensome. This paper presents the alternative copula-based approach, which produces a closed form expression for the joint distribution, estimable by standard maximum likelihood techniques, and without the intermediate step of specifying the explicit distribution of unobserved factors that induce correlation. The copula approach produces dependence parameters
that provide estimates of association between dependent variables. Thus some degree of model
parsimony is achieved via the application of Occam’s Razor.

We apply the copula approach to jointly estimate insurance decisions and health care de-
mands among married couples. The purpose is to determine the extent to which being enrolled
in different plans and spouses’ utilizations of health care services are interrelated. The results
have implications for policy makers. Health economists are concerned with potential overuti-
lization of health care services where individuals consume health care more than they would in
equilibrium. Copula results suggest two possible sources of overutilization. First, spouses’ uti-
lizations are positively correlated. Health care reforms aimed at increasing health care access of
females might have the unintended consequence of increasing utilization by their husbands as
well. Moreover, when spouses select themselves into generous insurance arrangements based
on unobserved family characteristics, health care utilization is higher that it would be if insur-
ance arrangements were randomly assigned.

The copula approach and the empirical model presented in this paper have many potential
empirical applications. Any instance in which two jointly determined outcomes are both directly
or indirectly affected by a third variable could potentially be modeled with a similar formulation.
For example, the voting decisions by two senators from the same state on a crime bill might be
related to a crime characteristic of the state’s constituency, and, therefore, jointly determined
with the senators’ voting decisions. Or the emission of two toxic pollutants from a factory might
be related to profit expectations. With the use of copulas, researchers can explore many issues
that were previously beyond the scope of existing estimation techniques.
Appendix – Examples of Bivariate Copulas

The following families of bivariate copulas are potential choices for empirical work.

\textit{Product Copula:} The simplest copula is the product copula,

\[ C(u, v) = uv \]

where \( u \) and \( v \) are uniformly distributed over \((0, 1)\). The product copula is important as a frame of reference because it corresponds to independence.

\textit{Farlie-Gumbel-Morgenstern Copula:} The Farlie-Gumbel-Morgenstern (FGM) copula takes the form

\[ C(u, v; \theta) = uv (1 + \theta (1 - u)(1 - v)) . \]

The FGM copula was first proposed by Morgenstern (1956). The FGM copula is a perturbation of the product copula; if the dependence parameter \( \theta \) equals zero, then the FGM copula collapses to independence. It is attractive due to its simplicity, and Prieger (2002) advocates its use in modeling selection into health insurance plans. However, it is restrictive because the dependence parameter is bounded on the interval \([-1, 1]\), and this interval does not correspond to either Fréchet bound. Therefore, this copula is only useful when dependence between the two marginals is modest in magnitude.

\textit{Normal Copula:} The normal copula takes the form

\[ C(u, v; \theta) = \Phi_B\left( \Phi^{-1}(u), \Phi^{-1}(v), \theta \right) \]
where $\Phi$ is the cdf of the standard normal distribution, and $\Phi_B(u,v)$ is the standard bivariate normal distribution with correlation parameter $\theta$ restricted to the interval $(-1, 1)$. This is the copula function proposed by Lee (1983) for modeling selectivity. As the dependence parameter approaches $-1$ and $1$, the normal copula attains the Fréchet lower and upper bound, respectively. The normal copula is flexible in that it allows for equal degrees of positive and negative dependence and includes both Fréchet bounds in its permissible range.

**Clayton Copula:** The Clayton (1978) copula, originally studied by Kimeldorf and Sampson (1975), takes the form

$$C(u,v; \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

with the dependence parameter $\theta$ restricted on the region $(0, \infty)$. As $\theta$ approaches zero, the marginals become independent. As $\theta$ approaches infinity, the copula attains the Fréchet upper bound, but for no value does it attain the Fréchet lower bound. The Clayton copula cannot account for negative dependence, but it is useful for modeling variables with high degrees of positive dependence.

**Frank Copula:** The Frank (1979) copula takes the form

$$C(u,v; \theta) = -\theta^{-1} \log \left\{ \left( \eta - (1 - e^{-\theta} u)(1 - e^{-\theta} v) \right) / \eta \right\}$$

where $\eta = 1 - e^{-\theta}$. The dependence parameter can assume any real value $(-\infty, \infty)$. Values of $-\infty$, $0$, and $\infty$ correspond to the Fréchet lower bound, independence, and Fréchet upper bound,
respectively. Because each Fréchet bound is included in the permissible range, the Frank copula has been popular in empirical applications (Meester and MacKay, 1994).

**Gumbel Copula:** The Gumbel (1960) copula takes the form

\[ C(u, v; \theta) = \exp \left( -\left( \tilde{u}^{\theta} + \tilde{v}^{\theta} \right)^{1/\theta} \right) \]

where \( \tilde{u} = -\log u \) and \( \tilde{v} = -\log v \). The dependence parameter is restricted to the interval \([1, \infty)\). Values of 1 and \( \infty \) correspond to independence and the Fréchet upper bound, but this copula does not attain the Fréchet lower bound for any value of \( \theta \).
References


<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Sample Means by Insurance Arrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enrolled in same plan, N = 5250</td>
</tr>
<tr>
<td></td>
<td>Wives</td>
</tr>
<tr>
<td>Utilization</td>
<td>DOC &quot;# of physician visits&quot;</td>
</tr>
<tr>
<td></td>
<td>NONDOC &quot;# of nonphysician visits&quot;</td>
</tr>
<tr>
<td></td>
<td>ER &quot;# of emergency room visits&quot;</td>
</tr>
<tr>
<td>Personal Demographic</td>
<td>AGE &quot;age/10&quot;</td>
</tr>
<tr>
<td></td>
<td>EDUC &quot;years of school&quot;</td>
</tr>
<tr>
<td></td>
<td>BLACK &quot;= 1 if black&quot;</td>
</tr>
<tr>
<td></td>
<td>HISPANIC &quot;= 1 if Hispanic&quot;</td>
</tr>
<tr>
<td></td>
<td>VET &quot;= 1 if military veteran&quot;</td>
</tr>
<tr>
<td>Employment</td>
<td>EMPLOYED &quot;= 1 if employed&quot;</td>
</tr>
<tr>
<td></td>
<td>SELFEMP &quot;= 1 if self-employed&quot;</td>
</tr>
<tr>
<td>Self-reported Health</td>
<td>POOR &quot;= 1 if poor&quot;</td>
</tr>
<tr>
<td></td>
<td>TOTCHR &quot;# of chronic conditions&quot;</td>
</tr>
<tr>
<td>Family Demographic</td>
<td>INCOME &quot;income/1000&quot;</td>
</tr>
<tr>
<td></td>
<td>FAMSIZE &quot;family size&quot;</td>
</tr>
<tr>
<td></td>
<td>KID0-6 &quot;= 1 if kid between 0 and 6&quot;</td>
</tr>
<tr>
<td></td>
<td>KID7-17 &quot;= 1 if kid between 7 and 17&quot;</td>
</tr>
<tr>
<td></td>
<td>KIDSICK &quot;= 1 if kid less healthy than others&quot;</td>
</tr>
</tbody>
</table>
TABLE 2
Probit Insurance Choice Results

Dependent Variable = DIFF

<table>
<thead>
<tr>
<th>Variable</th>
<th>DOCVIS</th>
<th></th>
<th>NONDOC</th>
<th></th>
<th>ER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.993**</td>
<td>0.039</td>
<td>-2.085**</td>
<td>0.219</td>
<td>-1.982**</td>
<td>0.298</td>
</tr>
<tr>
<td>H_AGE</td>
<td>0.004**</td>
<td>0.002</td>
<td>0.005</td>
<td>0.005</td>
<td>0.004*</td>
<td>0.002</td>
</tr>
<tr>
<td>FAMSIZE</td>
<td>-0.071**</td>
<td>0.017</td>
<td>-0.069**</td>
<td>0.031</td>
<td>-0.072**</td>
<td>0.035</td>
</tr>
<tr>
<td>INCOME</td>
<td>0.002**</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002**</td>
<td>0.001</td>
</tr>
<tr>
<td>W_EDUC</td>
<td>0.022</td>
<td>0.021</td>
<td>0.021</td>
<td>0.029</td>
<td>0.021</td>
<td>0.014</td>
</tr>
<tr>
<td>H_EDUCYR</td>
<td>-0.010</td>
<td>0.019</td>
<td>-0.009</td>
<td>0.012</td>
<td>-0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>H_BLACK</td>
<td>0.369**</td>
<td>0.068</td>
<td>0.380**</td>
<td>0.168</td>
<td>0.378**</td>
<td>0.061</td>
</tr>
<tr>
<td>H_HISPANIC</td>
<td>0.124**</td>
<td>0.041</td>
<td>0.144**</td>
<td>0.054</td>
<td>0.118**</td>
<td>0.059</td>
</tr>
<tr>
<td>W_EMPLOYED</td>
<td>1.044**</td>
<td>0.032</td>
<td>1.076**</td>
<td>0.189</td>
<td>1.050**</td>
<td>0.102</td>
</tr>
<tr>
<td>W_SELFEMP</td>
<td>-0.556**</td>
<td>0.057</td>
<td>-0.556</td>
<td>0.638</td>
<td>-0.549**</td>
<td>0.071</td>
</tr>
<tr>
<td>H_EMPLOYED</td>
<td>0.229**</td>
<td>0.031</td>
<td>0.255**</td>
<td>0.050</td>
<td>0.210</td>
<td>0.358</td>
</tr>
<tr>
<td>H_SELFEMP</td>
<td>-0.390**</td>
<td>0.051</td>
<td>-0.410</td>
<td>0.948</td>
<td>-0.397**</td>
<td>0.059</td>
</tr>
<tr>
<td>KID0–6</td>
<td>-0.195**</td>
<td>0.019</td>
<td>-0.193**</td>
<td>0.073</td>
<td>-0.194**</td>
<td>0.055</td>
</tr>
<tr>
<td>KID7-17</td>
<td>-0.190**</td>
<td>0.018</td>
<td>-0.197</td>
<td>0.192</td>
<td>-0.193**</td>
<td>0.054</td>
</tr>
<tr>
<td>KIDSICK</td>
<td>-0.040</td>
<td>0.101</td>
<td>0.004</td>
<td>0.232</td>
<td>-0.004</td>
<td>0.085</td>
</tr>
</tbody>
</table>

*: significant at the .10 level

**: significant at the .05 level
### TABLE 3

#### Utilization Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>WIVES</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>1.213**</td>
<td>0.042</td>
<td>−0.553**</td>
<td>0.084</td>
<td>−0.636**</td>
<td>0.171</td>
<td></td>
</tr>
<tr>
<td>W_AGE</td>
<td>−0.009**</td>
<td>0.002</td>
<td>−0.007**</td>
<td>0.002</td>
<td>−0.028**</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>W_EDUC</td>
<td>0.024**</td>
<td>0.006</td>
<td>0.077**</td>
<td>0.007</td>
<td>−0.046**</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td>W_BLACK</td>
<td>−0.299**</td>
<td>0.065</td>
<td>−0.879*</td>
<td>0.471</td>
<td>0.335**</td>
<td>0.169</td>
<td></td>
</tr>
<tr>
<td>W_HISPANIC</td>
<td>−0.029</td>
<td>0.025</td>
<td>−0.373**</td>
<td>0.005</td>
<td>0.108</td>
<td>0.314</td>
<td></td>
</tr>
<tr>
<td>W_POOR</td>
<td>0.434**</td>
<td>0.095</td>
<td>0.452</td>
<td>0.363</td>
<td>0.445*</td>
<td>0.260</td>
<td></td>
</tr>
<tr>
<td>W_TOTCHR</td>
<td>0.412**</td>
<td>0.015</td>
<td>0.559**</td>
<td>0.097</td>
<td>0.423**</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td>FAMSIZE</td>
<td>−0.032**</td>
<td>0.014</td>
<td>−0.096**</td>
<td>0.039</td>
<td>−0.064</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>1/Ψ&lt;sub&gt;w&lt;/sub&gt;</td>
<td>0.960</td>
<td>0.037</td>
<td>4.333</td>
<td>0.034</td>
<td>4.680</td>
<td>0.573</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>HUSBANDS</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>−0.430**</td>
<td>0.041</td>
<td>−1.655**</td>
<td>0.136</td>
<td>−0.729</td>
<td>0.785</td>
<td></td>
</tr>
<tr>
<td>H_AGE</td>
<td>0.016**</td>
<td>0.002</td>
<td>4e-4</td>
<td>0.015</td>
<td>−0.020**</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>H_EDUC</td>
<td>0.016**</td>
<td>0.008</td>
<td>0.084**</td>
<td>0.019</td>
<td>−0.065**</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>H_BLACK</td>
<td>−0.338**</td>
<td>0.084</td>
<td>−0.929**</td>
<td>0.383</td>
<td>0.073</td>
<td>0.175</td>
<td></td>
</tr>
<tr>
<td>H_HISPANIC</td>
<td>−0.126**</td>
<td>0.016</td>
<td>−0.560**</td>
<td>0.081</td>
<td>−0.234</td>
<td>0.189</td>
<td></td>
</tr>
<tr>
<td>H_VET</td>
<td>0.037</td>
<td>0.024</td>
<td>0.198</td>
<td>0.159</td>
<td>0.156</td>
<td>0.152</td>
<td></td>
</tr>
<tr>
<td>H_POOR</td>
<td>0.871**</td>
<td>0.251</td>
<td>1.097**</td>
<td>0.185</td>
<td>1.187**</td>
<td>0.318</td>
<td></td>
</tr>
<tr>
<td>H_TOTCHR</td>
<td>0.578**</td>
<td>0.027</td>
<td>0.666</td>
<td>0.830</td>
<td>0.392**</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>FAMSIZE</td>
<td>−0.036</td>
<td>0.023</td>
<td>−0.071**</td>
<td>0.029</td>
<td>−0.031</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td>1/Ψ&lt;sub&gt;h&lt;/sub&gt;</td>
<td>1.156</td>
<td>0.040</td>
<td>7.356</td>
<td>0.576</td>
<td>4.529</td>
<td>1.322</td>
<td></td>
</tr>
</tbody>
</table>

θ<sub>1</sub> | 0.025 | 0.014 | 0.039 | 0.083 | 0.098 | 0.026 |
θ<sub>2</sub> | 1.127 | 0.037 | 2.273 | 0.199 | 1.844 | 0.252 |

Log Likelihood | −30713.92 | −17364.86 | −7396.02 |

* : significant at the .10 level

** : significant at the .05 level
### TABLE 4

Fitted Frequencies

<table>
<thead>
<tr>
<th>Visits</th>
<th>DOC WIFE Actual</th>
<th>DOC WIFE Predicted</th>
<th>DOC HUSB Actual</th>
<th>DOC HUSB Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.22</td>
<td>0.22</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>1</td>
<td>0.18</td>
<td>0.17</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>2</td>
<td>0.14</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>0.10</td>
<td>0.10</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>0.08</td>
<td>0.08</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>0.06</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>7</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>8</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>9</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Visits</th>
<th>NONDOC WIFE Actual</th>
<th>NONDOC WIFE Predicted</th>
<th>NONDOC HUSB Actual</th>
<th>NONDOC HUSB Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.66</td>
<td>0.67</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>1</td>
<td>0.15</td>
<td>0.12</td>
<td>0.10</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.06</td>
<td>0.06</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>0.03</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>7</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Visits</th>
<th>ER WIFE Actual</th>
<th>ER WIFE Predicted</th>
<th>ER HUSB Actual</th>
<th>ER HUSB Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>1</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
TABLE 5

A: Treatment Effects from the Copula Model

<table>
<thead>
<tr>
<th>Utilization Measure</th>
<th>Wife</th>
<th>Husband</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATE</td>
<td>St. Dev.</td>
</tr>
<tr>
<td>Physician Visits</td>
<td>0.040</td>
<td>0.214</td>
</tr>
<tr>
<td>Nonphysician Visits</td>
<td>0.018</td>
<td>0.072</td>
</tr>
<tr>
<td>ER Visits</td>
<td>0.004</td>
<td>0.008</td>
</tr>
</tbody>
</table>

B: Treatment Effects from the MSL Model

<table>
<thead>
<tr>
<th>Utilization Measure</th>
<th>Wife</th>
<th>Husband</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATE</td>
<td>St. Dev.</td>
</tr>
<tr>
<td>Physician Visits</td>
<td>0.031</td>
<td>0.028</td>
</tr>
<tr>
<td>Nonphysician Visits</td>
<td>−0.001</td>
<td>0.012</td>
</tr>
<tr>
<td>ER Visits</td>
<td>−0.0001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>DOC Coeff.</td>
<td>DOC St. Err.</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>Trivariate Copula Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection Coefficient ($\theta_1$)</td>
<td>0.025</td>
<td>0.014</td>
</tr>
<tr>
<td>Correlation Coefficient ($\theta_2$)</td>
<td>1.127</td>
<td>0.037</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>−30713.92</td>
<td></td>
</tr>
<tr>
<td><strong>Simulation-Based Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection Coefficient ($\lambda_u$)</td>
<td>0.023</td>
<td>0.024</td>
</tr>
<tr>
<td>Correlation Coefficient ($\lambda_{v,h}$)</td>
<td>0.438</td>
<td>0.034</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>−30743.33</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 7

<table>
<thead>
<tr>
<th></th>
<th>DOC</th>
<th></th>
<th>NONDOC</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fewer than 30 visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection Coefficient ($\lambda_u$)</td>
<td>0.023</td>
<td>0.024</td>
<td>-0.052</td>
<td>0.042</td>
</tr>
<tr>
<td>Correlation Coefficient ($\lambda_{v,h}$)</td>
<td>0.438</td>
<td>0.034</td>
<td>1.585</td>
<td>0.082</td>
</tr>
<tr>
<td>All observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection Coefficient ($\lambda_u$)</td>
<td>0.027</td>
<td>0.024</td>
<td>-0.039</td>
<td>0.045</td>
</tr>
<tr>
<td>Correlation Coefficient ($\lambda_{v,h}$)</td>
<td>0.436</td>
<td>0.036</td>
<td>1.811</td>
<td>0.077</td>
</tr>
</tbody>
</table>