Dynamic Cost-offsets of Prescription Drug Expenditures: Panel Data Analysis Using a Copula-based Hurdle Model

Partha Deb
Departments of Economics
Hunter College and the Graduate Center, CUNY, and NBER

Pravin K. Trivedi
Indiana University

David M. Zimmer
Western Kentucky University

March 2010

Abstract

This paper presents a new copula-based approach for analyzing panel data in which (1) the observed multivariate outcomes are a mixture of zeros and continuously-measured positives, with the zeros accounting for a nontrivial proportion of the sample; (2) both the zero and positive outcomes show state dependence and dynamic interdependence (cross-equation lagged dependence); and (3) the zeros and the positives display contemporaneous association. The copula-based approach is especially appealing because the contemporaneous associations may involve asymmetric dependence. We apply our methodology to the analysis of cost-offsets between prescription drug and other non-drug healthcare expenditures using data from the Medical Expenditure Panel Survey. There is evidence of modest cost-offsets of expenditures on prescribed drugs.

JEL codes: C51, C33, I11

Keywords: Medical Expenditure Panel Survey; random effects; Clayton copula; dynamic dependence

1. Introduction

This paper presents a new copula-based approach for analyzing panel data that have the following features: (1) the observed multivariate outcomes are a mixture of zeros and continuously-measured positives, with the zeros accounting for a nontrivial proportion of the sample; (2) both the zero and positive outcomes show persistence (i.e., dependence on their own past values) and dynamic interdependence (cross-equation lagged dependence); and (3) the zeros and the positives display contemporaneous association. The copula-based approach is especially appealing because contemporaneous associations may involve tail dependence, in which extreme values of related variables tend to occur in tandem (Embrechts, McNeil, and Straumann, 2002). We label such a model a dynamic multivariate copula-based hurdle model and focus on the bivariate special case of it. Such nonlinear dynamic models are potentially useful for analyzing expenditure data for goods and services that may be dynamic substitutes or complements. The model is applicable to a variety of bivariate situations in which outcomes exhibit dynamic dependence, and where corner solutions are likely. Examples include expenditures on drug and nondrug healthcare services, values of related portfolio assets, household consumption of generic and name-brand products, and firms' investments in domestic and foreign markets. Variants of this model can be generated also for the case where the positive-valued outcome is discrete (e.g. a count, as in Bien, Nolte, and Pohlmeier, 2007).

We apply our methodology to an analysis of cost-offsets between prescription drug and other non-drug healthcare expenditures in the USA. In essence, this hypothesis states that drug- and non-drug expenditures are potential substitutes such that increased drug expenditures would be off-set by reductions in non-drug expenditures. Our choice of application is motivated by two considerations. First, prescription drug expenditures constitute an important item in the national health budget, accounting for around 10% of the US health budget in 2006. For the elderly this component of healthcare expenditures is even more important, accounting for expenditures exceeding $120 billion in 2005 and amounting to nearly $2,800 per person (Kaiser Family Foundation, 2005). Second, a rigorous test of the cost-offset hypothesis calls for a nonlinear dynamic model capable of capturing patterns of dynamic substitution and complementarity in the use of prescribed drug and non-drug healthcare expenditures. The model is estimated on quarterly panel data from the Med-

ical Expenditure Panel Survey. Because there is substantial heterogeneity of individuals and medical services in their degree of substitutability, so that the relationship between drug therapy and other medical usage varies across a range of medical services, we conduct separate analyses by health status and other conditions.

In the remainder of the paper, section 2 elaborates the statement of the cost-offset hypothesis. Sections 3 and 4 deal with the model specification, including that of its dynamic features. The data and the empirical results are described and discussed in sections 5 and 6. Section 7 concludes.

2. The cost-offset hypothesis

Should policy makers be concerned about the growth of prescription drug expenditures, or do expenditures on prescription drugs pay for themselves through reduced usage of other, possibly more expensive, health services? The strength and extent of substitution between prescription drugs and other medical services is the key aspect of the issue. However, there are plausible arguments that complementarity between prescribed drug and non-drug expenditures might be expected. The topic has timely policy implications, as many states in the US are intensifying efforts to control rising prescription drug costs in their public programs (Cunningham, 2005).

The study of cost-offsets in structural settings has a sound basis but is challenging for reasons already noted. Previous analyses of these questions have followed a variety of different approaches. One strand of the literature analyzes the relationship between Medicare supplemental insurance and the utilization of prescription drugs of the elderly; see, for example, Poisal and Murray (2001) and Goldman and Philipson (2007). These papers examine the effect of a change in the price of prescription drugs, brought about by a change in insurance coverage for drug purchases, on spending on drugs. Relationships between spending on drugs and nondrug care are implied but not directly described. Of greater direct relevance to this paper is the strand of literature which concentrates on the relationship between drug expenditures and cost sharing and on the substitution effects resulting from changes in cost sharing; see Gaynor, Li and Vogt (2007), Goldman et al. (2004), Joyce et al. (2002). Such an approach is "structural" in the sense that it focuses on the mechanism though which changes in prescription drug prices impact usage of

drugs and other types of health care. For example, Gaynor et al. (2007) use individual level data on health insurance claims and benefits; they report evidence of substitution between outpatient care and prescription drug expenditures, with 35 percent of reductions in prescription drug expenditures being offset by increases in other medical expenditures. Another example of a structural approach is Shang and Goldman (2007) who use Medicare Current Beneficiary Survey (MCBS) panel data to examine spending of Medicare beneficiaries with and without supplemental drug coverage. They report that " ... a $1 increase in prescription drug spending is associated with a $2.06 reduction in Medicare spending. Furthermore, the substitution effect decreases as income rises, and thus provides support for the low-income assistance program of Medicare Part D." Stuart and Grana (1995) and a series of coauthored articles by Stuart (2004, 2005, 2007) are other examples of studies that investigate cost-offsets of prescription drugs.

A different reduced form approach to uncovering potential substitution is illustrated by Lichtenberg (1996, 2001) who analyzes the direct impact of (especially newer) prescription drug expenditures on other types of expenditures, especially hospital care. His analysis based on Medical Expenditure Panel Survey (MEPS) data involves direct regression of other expenditures on measures of prescription drug use. His results indicate that " ... persons consuming newer drugs had significantly fewer hospital stays than persons consuming older drugs." Some health policy advocates argue that, on average, use of new prescription drugs reduces total health care costs, but Zhang and Soumerai (2007) show that those results are not robust to changes in specification.

Another strand of literature emphasizes complementarity between drug and nondrug spending. For example, Stuart et al. (2007) argue,

> "Economic theory also posits that when the price of a complementary good falls, both the demand for the good itself and the complement will rise. This leads to a second way in which Part D might affect Medicare Part A (hospitals) and Part B (medical) spending. Because physician services complement to prescription drug fills, we expect that people with prescription drug coverage will be more likely to visit physicians and thereby spend more on Medicare Part B services. Furthermore, increased physician usage could lead to increased rates of diagnostic checks, surgeries, and other expensive procedures."

The extant literature on the direct non-structural approach for testing the cost-offset hypothesis is potentially problematic. Indeed some features of this approach are at odds with the standard static consumer behavior theory. For example, standard static models of consumer demand do not directly introduce current or past expenditures as explanatory variables for explaining other expenditure variables, but such dependence can clearly arise in a dynamic setting. For example, purchases of durable consumer goods at time t will generally affect consumption of nondurables and durables beyond t. Analogously, the longer lasting health effects of prescribed medications, if they exist, may impact the use of other medical services in the future. Thus, it is of interest to test whether expenditures on prescribed medications have predictive value for other future medical expenditures (after controlling for the effects of socioeconomic factors, as well as insurance and health status).

Currently there is not available a rigorous derivation of a cost-offset model from a dynamic model of health care consumption. While our approach uses somewhat ad hoc functional forms and distributional assumptions, it provides a starting point for developing models suitable for empirical study of dependence structures. It addresses several important econometric and modeling issues that will typically arise in such contexts.

We model the cost-offset hypothesis within the statistical framework of the joint bivariate distribution of two types of expenditures, prescribed medications and other nondrug expenditures, denoted y_1 and y_2, respectively. We allow for a potentially long term nonlinear dynamic impact of current medical expenditures on health status and on future health-related expenditures. Within such a framework we attempt to estimate the time profile of the impact of drug expenditures on current and future nondrug expenditures, the key parameters of interest being $\partial y_{2,t}/\partial y_{1,t-\tau}$. Identification of these parameters requires panel data. Within our framework, the cost-offset hypothesis implies negative intertemporal dependence between the two types of expenditures, i.e., $\partial y_{2,t}/\partial y_{1,t-\tau} < 0$. More strictly, a cost-offset that implies a reduction in aggregate medical spending satisfies $\partial y_{2,t}/\partial y_{1,t-\tau} < -1$.[1] We adopt a copula framework which accommodates a flexible formulation of dependence and marks a departure from the usual assumption of linear dependence. Note that, although the effect of interest is the intertemporal relationship that runs from y_1 to y_2, our model has no restrictions of the Granger-causal type.

[1] In our empirical analysis, we focus on the case in which $\tau = 1$.

3. Model specification

The distributions of quarterly drug and nondrug expenditures have substantial numbers of zeros, approximately 60-70 percent for drug expenditures and 30-40 percent for nondrug expenditures. To capture this feature we propose a bivariate hurdle model of expenditures. In the univariate case, the hurdle or two-part model is ubiquitous in the health economics literature (Pohlmeier and Ulrich, 1995). Either logit or probit is the commonly used functional form for the first part, which describes whether spending is positive. For the second part, which models positive spending, much of the older literature used OLS to estimate the parameters of the logarithm of expenditures. More recently, models based on the gamma distribution have been preferred (Manning, Basu, and Mullahy, 2005), in part because they tend to fit the data better, and also because they have the added advantage of not requiring, post estimation, a retransformation to the raw scale. We take this basic setup from the literature on expenditures and extend it in two significant ways. First, it is adapted to the bivariate case. Second, because the cost-offset hypothesis is inherently dynamic, we specify dynamic relationships within the bivariate hurdle framework. We develop a model of the joint distribution of drug and nondrug expenditures because this will lead to a number of parameters relevant to the cost-offsets hypothesis.

Consider two non-negative outcomes y_1 and y_2 each with a significant fraction of zeros. The bivariate hurdle model specifies a statistical process for each of the four configurations of outcomes, $y_1 = 0, y_2 = 0$ (denoted by (y_1^0, y_2^0) in what follows); $y_1 > 0, y_2 = 0$ (y_1^+, y_2^0); $y_1 = 0, y_2 > 0$ (y_1^0, y_2^+) and $y_1 > 0, y_2 > 0$ (y_1^+, y_2^+). Each configuration maps to a data distribution given by a product of a bivariate hurdle probability and a density for the positive outcomes, such that

$$\begin{aligned}
y_1^0, y_2^0 &\longrightarrow F(y_1 = 0, y_2 = 0) \times f_{00}(y_1 = 0, y_2 = 0) \\
y_1^+, y_2^0 &\longrightarrow F(y_1 > 0, y_2 = 0) \times f_{+0}(y_1, y_2 = 0 | y_1 > 0) \\
y_1^0, y_2^+ &\longrightarrow F(y_1 = 0, y_2 > 0) \times f_{0+}(y_1 = 0, y_2 | y_2 > 0) \\
y_1^+, y_2^+ &\longrightarrow F(y_1 > 0, y_2 > 0) \times f_{++}(y_1, y_2 | y_1 > 0, y_2 > 0).
\end{aligned} \tag{1}$$

where F is a bivariate distribution defined over binary outcomes, and f_{00}, f_{+0}, f_{0+} and f_{++} are bivariate densities. The definitions given in equation (1) can be simplified further.

First, note that the density $f(y_1 = 0, y_2 = 0)$ is degenerate and equal to one. Next, note that

$$f_{+0}(y_1, y_2 = 0|y_1 > 0) = f_2(y_2 = 0)f(y_1|y_1 > 0, y_2 = 0).$$

Since $f_2(y_2 = 0)$ is degenerate, the model requires only a specification for $f(y_1|y_1 > 0, y_2 = 0)$. We specify this density using a parametric form for $f_1(y_1|y_1 > 0)$. This specification follows the logic of the two-part or hurdle model (Cragg, 1971; Duan et al., 1983) as being specified "free" of y_2 but this does not mean that $y_1|y_1 > 0$ and y_2 are independent events. Note that, we could specify the joint density $f(y_1, y_2 = 0|y_1 > 0)$ directly, but any dependence parameter specified would be unidentified unless there were other sources of identification.[2] Duan et al. (1983, 1984) provide theoretical and numerical demonstrations of these properties. The density $f(y_1 = 0, y_2|y_2 > 0)$ can be similarly decomposed and simplified. Consequently,

$$\begin{aligned}
y_1^0, y_2^0 &\longrightarrow F(y_1 = 0, y_2 = 0) \\
y_1^+, y_2^0 &\longrightarrow F(y_1 > 0, y_2 = 0) \times f_1(y_1|y_1 > 0) \\
y_1^0, y_2^+ &\longrightarrow F(y_1 = 0, y_2 > 0) \times f_2(y_2|y_2 > 0) \\
y_1^+, y_2^+ &\longrightarrow F(y_1 > 0, y_2 > 0) \times f_{12}(y_1, y_2|y_1 > 0, y_2 > 0),
\end{aligned} \tag{2}$$

Below, we first describe parametric forms for the univariate densities $f_j, j = 1, 2$. Then we describe the joint distribution F and the joint density f_{12}. Note that, for notational convenience, we first describe the setup without conditioning variables. Conditioning on covariates and lagged dependent variables is described later.

3.1. Specification of f_j

Positive expenditures are specified according to the gamma density,

$$f_j(y_j|y_j > 0) = \frac{\exp(-y_j/\mu_j)y_j^{\eta_j - 1}}{\mu_j \Gamma(\eta_j)} \text{ for } j = 1, 2; \mu_j > 0; \eta_j > 0. \tag{3}$$

[2] There are two ways in which a dependence parameter might be identified for $f(y_1, y_2 = 0|y_1 > 0)$. First, dependence parameters might be identified if there were variable exclusion restrictions. Second, one could specify the same dependence parameter for $f(y_1, y_2 = 0|y_1 > 0)$ as is specified for $f(y_1, y_2|y_1 > 0, y_2 > 0)$. However, this adds no informational value or modeling flexibility, but it does add computational complexity.

Note that $\mathrm{E}(y_j|y_j > 0) = \eta_j\mu_j$, $j = 1, 2$ and skewness and kurtosis of the gamma distributions are positively related to $1/\eta_j$. Thus the specification allows the shape parameter to be different for drug and nondrug expenditures.

3.2. Specification of F and f_{12}

It is likely that stochastic dependence between drug and nondrug expenditures is asymmetric, with equally plausible arguments in favor of lower or upper tail dependence. Unlike the typical bivariate probit setup for joint binary outcomes or the seemingly unrelated linear regression setup, both of which emphasize linear correlations, copula-based dependence measures allow for more flexible patterns. Dependence in a copula-based model derives from the functional form of the copula itself, which is specified by the researcher. Some copulas exhibit dependence that is highly nonlinear and asymmetric. Thus, a copula-based model has the potential to more accurately capture the complex, nonlinear relationship between drug and nondrug expenditures. Our statistical framework uses the copula approach to generate the desired joint distributions, F and f_{12}.

3.2.1. Copula basics

The copula approach to multivariate distributions was pioneered by Sklar (1973) and extended to conditional distributions by Patton (2006). Within this framework the copula parameterizes a multivariate distribution in terms of its marginal distributions conditional on information set $\mathcal{I}_{t-1}$. For an m-variate joint distribution function G, the copula satisfies

$$G(y_{1t}, ..., y_{mt}|\mathcal{I}_{t-1}) = \ C(G_1(y_{1t}|\mathcal{I}_{t-1}), ..., G_m(y_{mt}|\mathcal{I}_{t-1}); \theta), \tag{4}$$

where $G_j(y_{jt}|\mathcal{I}_{t-1})$ denotes the marginal distribution function of the j^{th} component and θ is a scalar-valued dependence parameter. Given the marginal distributions, and a copula function $C(\cdot)$, the above equation generates a joint conditional distribution. A fully parametric implementation requires the choice of suitable functional forms of marginal distributions $G_1, ..., G_m$, and the functional form of the copula.

The literature offers a vast array of copula functional forms from which to choose (Nelsen, 2006). Because we have no *a priori* expectations regarding the dependence structure for our data, we have experimented with a variety of copulas: (1) Gaussian; (2)

Clayton; (3) Survival Clayton; (4) Frank. The functional form of each of these copulas appears in Table 3. The Gaussian and Frank copulas both permit positive and negative dependence, and any observed dependence is symmetric with no tail dependence allowed in the Gaussian case and limited tail dependence in the Frank copula. In contrast, the Clayton and Survival Clayton copulas allow for asymmetric tail dependence, but neither allows for negative dependence. Clayton exhibits lower tail dependence, whereas Survival Clayton shows upper tail dependence. By changing the functional form of the copula, many different dependence patterns between marginal distributions can be explored. Properties of these well established functional forms are discussed in the literature (Joe, 1997; Nelsen, 2006; Cherubini, Luciano, and Vecchiato, 2004; Trivedi and Zimmer, 2007).

Anticipating our results, it generally appears that the best fit to the data is obtained using the Clayton copula. The main benefit of the Clayton copula is its ability to capture lower tail dependence, which information criteria measures indicate is omnipresent in health care expenditures data. The bivariate Clayton (1978) copula takes the form

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \; \theta > 0 \tag{5}$$

where $u_j = G_j(y_j|\mathcal{I}_{t-1})$ with the dependence parameter θ restricted to the region $(0, \infty)$. As θ approaches zero, the marginals become independent. The Clayton copula exhibits asymmetric dependence in that dependence in the lower tail is stronger than in the upper tail, but this copula cannot account for negative dependence. It is not always easy to interpret estimates of θ for different copulas. Thus it is helpful to transform θ to more easily interpreted measures of concordance such as Kendall's τ (Nelsen, 2006) which is comparable across copulas. For the Clayton copula the formula for converting θ is $\tau = \theta/(\theta + 2)$.

In using the Clayton copula, contemporaneous dependence between drug and nondrug spending is restricted to be positive, and therefore, not surprisingly, we find that contemporaneous dependence is positive. However, in our formulation the choice of copula does not restrict the direction of *dynamic* dependence, which is our principal concern. Preliminary analysis indicated that other copulas that permit negative contemporaneous dependence also produced positive contemporaneous dependence. Therefore, our findings of positive contemporaneous dependence appear to be robust across different copula specifications.

3.2.2. Specification of F

We use the probit formulation for the marginal distributions for the bivariate hurdle part of the model, i.e., $\Pr(y_j > 0) = \Phi_j(\cdot)$. Let the joint probability distribution of positive drug and nondrug expenditures be

$$F(y_1 > 0, y_2 > 0) = C(\Phi_1(\cdot), \Phi_2(\cdot); \theta_0) \tag{6}$$

where C is one of the copula functions described above, and θ_0 is a dependence parameter. It is easy to derive the following related probabilities:

$$\begin{aligned} F(y_1 = 0, y_2 = 0) &= 1 - \Phi_1(\cdot) - \Phi_2(\cdot) + C(\Phi_1(\cdot), \Phi_2(\cdot); \theta_0); \\ F(y_1 > 0, y_2 = 0) &= \Phi_1(\cdot) - C(\Phi_1(\cdot)\Phi_2(\cdot); \theta_0); \\ F(y_1 = 0, y_2 > 0) &= \Phi_2(\cdot) - C(\Phi_1(\cdot)\Phi_2(\cdot); \theta_0). \end{aligned} \tag{7}$$

3.2.3. Specification of f_{12}

We use the gamma density for the marginal distributions for the copula-based joint distribution of positive drug and nondrug expenditures. That is,

$$f_j^+(y_j|y_1 > 0, y_2 > 0) = y_j^{\eta_j^+ - 1} \frac{\exp(-y_j/\mu_j)}{\mu_j^{\eta_j^+}\Gamma(\eta_j^+)} \text{ for } j = 1, 2; \mu_j > 0; \eta_j^+ > 0 \tag{8}$$

and

$$f_{12}(y_1, y_2|y_1 > 0, y_2 > 0) = c\left(F_1^+(\cdot), F_2^+(\cdot); \theta_+\right) \times f_1^+(\cdot) \times f_2^+(\cdot) \tag{9}$$

where lower case $c(\cdot)$ represents the copula density, and F_j^+ is the cumulative distribution function (cdf) version of f_j^+. Note that, while we have specified μ_j, which we parameterize to be the same as in the specifications of f_j for parsimony, η_j^+ is not necessarily the same as η_j, a proposition we test in our empirical analysis. Allowing η_j^+ to be different from η_j may seem odd at first glance as they seem inconsistent with each other. In fact, this is one of the advantages of the copula approach. The starting point of the analysis consists of specifications of the marginal densities in the most data-consistent manner. Then the use of a copula allows one to combine the marginals and obtain the joint distribution, which, by Sklar's theorem, satisfies all properties of a density / distribution function.

4. Dynamics and estimation

We now introduce the specifications for conditioning on covariates, dynamics via lagged dependent variables and individual-level random effects. We first describe how they are specified for the bivariate hurdle specification and then we describe how they are specified for the models of positive expenditures.

4.1. Specification of conditional means in F

For the marginal distributions $\Phi_1\left(\cdot|\mathcal{I}_{t-1}, \mathbf{x}_{it}\right)$, $\Phi_2\left(\cdot|\mathcal{I}_{t-1}, \mathbf{x}_{it}\right)$ we specify

$$\Pr(y_{1it} > 0) = \Phi\left(h_{01}(\{y_{k,it-j}\}) + \mathbf{x}_{it}'\boldsymbol{\beta}_{01} + \alpha_{01i}\right) \tag{10}$$

$$\Pr(y_{2it} > 0) = \Phi\left(h_{02}(\{y_{k,it-j}\}) + \mathbf{x}_{it}'\boldsymbol{\beta}_{02} + \alpha_{02i}\right), \tag{11}$$

$i = 1, ..., N$; $t = 3, ..., T$; $k = 1, 2$; $j = 1, 2, ..., J$ and $J < T$. The functions h_{0k} are defined over the elements of the set $\{y_{k,it-j}\}$ which includes lagged outcomes and α_{0ji} are random intercepts.

We allow for independent effects of lagged binary indicators of expenditures in addition to lagged continuous expenditure variables. Thus the specifications for h_{0l} are given by

$$h_{0l}\{y_{k,it-j}\} = \sum_{k=1}^{2}\sum_{j=1}^{J}\gamma_{lkj}\mathbf{1}(y_{k,it-j} > 0) + \sum_{k=1}^{2}\sum_{j=1}^{J}\delta_{lkj}\ln\left(\max(y_{k,it-j}, 1)\right) \text{ for } l = 1, 2. \tag{12}$$

That is, the lagged expenditures are entered as their logarithms when they are positive, zero otherwise, along with an indicator for whether the lagged expenditure is greater than zero or not.[3]

In the empirical analysis, we restrict the model to include only first and second quarter lags, i.e., $j = 1, 2$. A model with one lag has many appealing analytical features but is almost certainly too restrictive. The model with two lags allows for considerably more general autocorrelation and cross-correlation structures. In principle, there is no reason to limit the model to have just two lags. Unfortunately, because of the length of the time dimension in our panel data (8 quarters) combined with the large number of control variables, we observed a number of cases of non-convergence when we estimated models

[3] There are no positive expenditures less than \$1.

with 3 or more lags. Therefore, we restrict our attention in this paper to models with two lags.

The random intercepts are further specified as

$$\alpha_{0ki} = \overline{\mathbf{x}}_i'\lambda_{0k} + \sum_{k=1}^{2} \tau_k \mathbf{1}(y_{ki0} > 0) + \sum_{k=1}^{2} \varsigma_k \ln\left(\max(y_{ki0}, 1)\right) + \varepsilon_{0ki};\ k = 1, 2. \tag{13}$$

This extends the standard random effect panel model along two dimensions. Following Mundlak (1978) and Chamberlain (1984), we allow for correlation between α_{0ki} and $\mathbf{x}_{it}$ by including a vector of person-specific time-averaged means, $\overline{\mathbf{x}}_i'$. Following Wooldridge (2005), we allow for the effects of initial conditions by specifying α_{0ki} to be a function of $\mathbf{y}_{ki0}$, a vector of initial values of the outcome variables, allowing for separate effects for the binary indicator and continuous expenditure variables. The term ε_{0ki} may be interpreted as unobserved heterogeneity uncorrelated with $\mathbf{x}_{it}$ and $\mathbf{y}_{ki0}$. For the specification to be valid, we assume that x_{it} are strongly exogenous and not predetermined. To allow for possible dependence between y_{1it} and y_{2it} induced by unobserved heterogeneity, $(\varepsilon_{01i}\ \varepsilon_{02i})$ have a joint bivariate distribution whose functional form is not initially explicitly stated. Given this distribution, the correlated random effects bivariate model integrates out the random effects $(\varepsilon_1, \varepsilon_2)$. Different functional forms of the joint distribution arise from different parametric assumptions about the joint distribution of the random effects. Whereas we do not explicitly carry out this integration, we use several different functional forms of the bivariate joint distribution (i.e. the hurdle copula). Underlying each functional form is some form of dependence. We let the data decide which functional form best fits the data.

The estimation of the univariate dynamic probit model in the presence of initial conditions has been discussed by Heckman (1981) and more recently by Wooldridge (2005); Arumapalam and Stewart (2009) compare the two approaches. We follow Wooldridge in employing a maximum likelihood approach conditional on the observed values of the initial conditions, which is analogous to assuming that the initial conditions are nonrandom. Although we acknowledge the tenuousness of this assumption, fully addressing this concern would require specifying distributions for the initial conditions and integrating them out of (13). This would add a substantial layer of complexity to an already-large model.

4.2. Specification of conditional means in f_1, f_2 and f_{12}

For the marginal distributions $f_j(y_j|y_j > 0, \mathcal{I}_{t-1}, \mathbf{x}_{it})$ and $f_j^+(y_j|y_1 > 0, y_2 > 0, \mathcal{I}_{t-1}, \mathbf{x}_{it})$ we specify

$$\mu_{1it} = \exp\left[h_{+1}(\{y_{k,it-j}\}) + \mathbf{x}'_{it}\boldsymbol{\beta}_{+1} + \alpha_{+1i}\right] \quad (14)$$

$$\mu_{2it} = \exp\left[h_{+2}(\{y_{k,it-j}\}) + \mathbf{x}'_{it}\boldsymbol{\beta}_{+2} + \alpha_{+2i}\right], \quad (15)$$

$i = 1, ..., N$; $t = 3, ...T$; $k = 1, 2$; $j = 1, 2, ..., J$ and $J < T$, in parallel to the specifications for the marginal distributions in the hurdle part of the model. Again, the functions h_{+k} are defined over the elements of the set $\{y_{k,it-j}\}$ which includes lagged outcomes and α_{+ki} are random intercepts. As in the specification for the binary choices, we allow for independent effects of lagged binary indicators of expenditures in addition to lagged continuous expenditure variables via

$$h_{+l}(\{y_{k,it-j}\}) = \sum_{k=1}^{2}\sum_{j=1}^{J}\gamma_{+lkj}\mathbf{1}(y_{k,it-j} > 0) + \sum_{k=1}^{2}\sum_{j=1}^{J}\delta_{+lkj}\ln\left(\max(y_{k,it-j}, 1)\right) \text{ for } l = 1, 2, \quad (16)$$

and

$$\alpha_{+ki} = \overline{\mathbf{x}}'_i\lambda_{+k} + \sum_{k=1}^{2}\tau_{+k}\mathbf{1}(y_{k,i0} > 0) + \sum_{k=1}^{2}\varsigma_{+k}\ln\left(\max(y_{k,i0}, 1)\right) + \varepsilon_{+ki};\ k = 1, 2. \quad (17)$$

As is typical in gamma regressions, the parameters η_1 and η_2 are specified as scalars. Finally, in the empirical analysis, as motivated above we restrict the model to include only first and second quarter lags, i.e., $j = 1, 2$.

4.3. Estimation and inference

As described in equation (2), in our set-up there are four categories of bivariate realizations: (1) $y_1 = y_2 = 0$; (2) $y_1 > 0, y_2 = 0$; (3) $y_1 = 0, y_2 > 0$; (4) $y_1 > 0, y_2 > 0$. The joint likelihood is formed using the probability expression for each realization. Using the marginal and joint expressions described above, the log likelihood function for the bivariate

hurdle model is

$$\begin{aligned}\ln L &= \sum_{0,0}\left[\ln\left(F(y_1=0,y_2=0),\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)\right] \\ &+\sum_{+,0}\left[\ln\left(F(y_1>0,y_2=0),\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)+\ln\left(f_1\left(\cdot|\mathcal{I}_{t-1},\mathbf{x}_{it}\right)\right)\right] \\ &+\sum_{0,+}\left[\ln\left(F(y_1=0,y_2>0),\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)+\ln\left(f_2\left(\cdot|\mathcal{I}_{t-1},\mathbf{x}_{it}\right)\right)\right] \\ &+\sum_{+,+}\left[\ln\left(F(y_1>0,y_2>0),\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)+\ln f_{12}(y_1,y_2|y_1>0,y_2>0,\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_+)\right].\end{aligned} \tag{18}$$

For purposes of estimation, it is convenient to note that the log likelihood decomposes into two parts which can be maximized separately, i.e., $\ln L = \ln L_1 + \ln L_2$ where

$$\begin{aligned}\ln L_1 &= \sum_{0,0}\ln\left(F(y_1=0,y_2=0),\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)+\sum_{+,0}\ln\left(Fy_1>0,y_2=0,\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right) \\ &+\sum_{0,+}\ln\left(F(y_1=0,y_2>0),\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)+\sum_{+,+}\ln\left(Fy_1>0,y_2>0,\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_0\right)\end{aligned} \tag{19}$$

and

$$\begin{aligned}\ln L_2 &= \sum_{+,0}\ln\left(f_1\left(\cdot|\mathcal{I}_{t-1},\mathbf{x}_{it}\right)\right)+\sum_{0,+}\ln\left(f_2\left(\cdot|\mathcal{I}_{t-1},\mathbf{x}_{it}\right)\right) \\ &+\sum_{+,+}\ln f_{12}(y_1,y_2|y_1>0,y_2>0,\mathcal{I}_{t-1},\mathbf{x}_{it};\theta_+).\end{aligned} \tag{20}$$

$\ln L_1$ and $\ln L_2$ are maximized separately using a Newton-Raphson algorithm with numerical derivatives. Upon convergence, robust standard errors that adjust for clustering at the individual level are calculated and used for inference throughout.

The dynamic copula-based hurdle model accommodates dependence between drug and nondrug spending in several unique respects. First, the likelihood function contains two dependence parameters capturing *contemporaneous* dependence: θ_0 measures dependence between the probabilities of having any drug and nondrug expenditures, and θ_+ represents dependence between drug and nondrug expenditures when both are positive. Second, as each marginal distribution specifies spending as a function of one- and two-period lagged indicators of positive spending, one- and two-period lagged spending amounts, and initial conditions, the model also incorporates *dynamic* dependence, which should be of more interest from a policy perspective. As illustrated in more detail below, the models exhibit highly nonlinear structures for both contemporaneous and dynamic dependence, which

contrasts with previous studies on this topic where the main focus was often linear measures of contemporaneous dependence.

The model proposed in this paper shares similarities with the set-up developed by Bien, Nolte, and Pohlmeier (2007). Their count data model also seeks to accommodate large proportions of zero-valued outcomes in a multivariate setting. Although numerous differences separate their model and ours, the most substantive departure is that while Bien et al. do address dynamic issues by way of an error correction mechanism, their main focus remains contemporaneous dependence. In contrast, our model, as well as its attendant policy implications, emphasizes nonlinear dynamic dependence by way of an explicit autoregressive structure introduced directly into the marginal behavior of medical spending.

5. Data

The data for this study come from the 1996-2006 waves of the Medical Expenditure Panel Survey (MEPS) collected by the Agency for Healthcare Research and Quality (AHRQ) from which we construct a number of subsamples of substantive interest. MEPS consists of a series of five interviews over a two-and-a-half year period from which an 8 quarter panel is constructed for each respondent. Person-specific socioeconomic information and monthly health insurance status comes from the Household Component Full Year files. Information on monthly health care spending comes from the Household Component Event files. Spending is accumulated at the quarterly level and includes spending from all sources on the following services: prescription drugs (including refills), office-based visits, outpatient visits, inpatient hospital visits, and emergency room visits. The latter three categories include both facility and separately-billed-doctor expenses. The sample excludes individuals who report quarterly drug or nondrug spending above the 99.5 percentile of all positive spenders. Finally, all spending measures are adjusted for inflation using the medical CPI (http://www.bls.gov/cpi), with fourth quarter 2006 serving as the base period. Although these data permit additional studies in which nondrug spending is delineated into specific subcategories of expenditures, our baseline results emphasize total expenditures due to the policy relevance of aggregate spending and to keep the quantity of results at a managable level.

We construct six subsamples of data for analysis, as MannWhitney U-tests indicated substantial heterogeneity in medical spending in the full sample. Moreover, the six subsamples are relevant for targeted policy considerations. Each sample considers individuals 18 years of age and older. The first two subsamples attempt to introduce homogeneity along age and insurance dimensions. Thus they consist of (1) an elderly sample consisting of individuals ages 65 and older ($N = 78,162$), and (2) well-insured individuals covered by both medical and prescription drug insurance ($N = 289,374$). Four additional subsamples focus on subjects with specific health ailments: (3) diabetes ($N = 42,702$), (4) mental illness ($N = 76,848$), (5) arthritis ($N = 91,230$), (6) heart problems ($N = 120,552$).

Table 1 presents descriptive statistics for quarterly drug and nondrug spending. Not surprisingly, the probability of positive spending appears to vary somewhat with respect to health problems, insurance, and age. The same is true for spending among positive spenders, with the highest spending occurring among the elderly and those with diabetes, arthritis, and heart conditions. The relatively large means of quarterly medical spending, in comparison to the smaller medians, indicate long upper tails. Also, as expected, the quarterly data exhibit substantial serial dependence. In Table 1, we also report the first order serial correlation coefficient for the indicator of positive spending as well as for the logarithm of expenditure (with its value set to zero when expenditure is zero). Two patterns are immediately apparent. First, nondrug expenditures display substantially more serial correlation than drug expenditures. Second, the serial correlation in the indicator variable is uniformly larger than the serial correlation in the corresponding continuous expenditure variable.

All models estimated below contain a common set of explanatory variables, with sample means appearing in Table 2. These are variables that we expect, *a priori*, to correlate with health care demand. Our socioeconomic controls are standard in health care demand models: age, gender, race, marital status, education, region of residence, employed status, and person-level income (both wage and nonwage). We also include measures of firm size and an indicator of government employment, both of which equal zero if the subject is not employed. In addition to indicators for insurance status, we also include self-reported health (on the standard excellent-to-poor scale), indicators of physical limitations, injuries, and number of chronic conditions. The motivation behind including a diverse set of health

variables, as opposed to a simple health index, is that some health indicators are more pertinent to specific populations; for example, certain chronic conditions are especially prevalent among the elderly. Although some of the health measure might be somewhat collinear, our large sample sizes allow precise estimation of their individual impacts on health care spending.

As indicated in Table 2, the elderly sample has lower rates of employment, smaller family sizes, and higher rates of public insurance. The diabetes, arthritis, and heart condition samples are older and have larger numbers of blacks, lower rates of employment, and higher rates of public insurance. The mental illness sample has more females, fewer blacks, and higher divorce rates compared to the other samples. The sample of individuals with prescription drug coverage is younger, whiter, healthier, more educated, and more likely to be employed and married. Differences between the subsamples highlight heterogeneity in health care markets and motivate separate consideration of the different groups.

5.1. Covariates in marginal distributions

All marginal models use a common vector of covariates. Specifically, the lag structure is specified to be the same for all outcomes. This restriction follows from the results of Patton (2006) who developed "conditional" copula modeling by including lagged dependent variables on the right hand side similar to what is proposed here. The model is a nonlinear vector autoregressive system of equations. By including previous-period expenditures variables on the right hand side, the model captures dynamic dependence between drug and nondrug expenditures.

In most previous applications of conditional copulas, usually in models of continuous outcomes, and in the literature on dynamic binary response models, the lag is restricted to one period. For potential flexibility, given that our data periodicity is quarterly, we include two lags on both $1[y_j > 0]$ and y_j^+. Specifically we use four variables at one- and two-period lags to measure past expenditures:

1. one and two-period lagged values of a dichotomous indicator for positive drug expenditures;

2. one and two-period lagged values of a dichotomous indicator for positive nondrug expenditures;

3. one and two-period lagged values of log of drug expenditures with the variable coded as zero when the expenditure is zero.

4. one and two-period lagged values of log of nondrug expenditures with the variable coded as zero when the expenditure is zero.

The vector $\mathbf{x}$ includes all explanatory variables listed in Table 2, with dummies for individual chronic conditions rather than the number of chronic conditions.[4] The vector $\mathbf{x}$ also includes measures of age squared and an interaction between age and female and its square. Counting control variables in $\mathbf{x}$, quarter dummies, lagged spending measures, initial conditions, and the "Mundlak terms", each marginal distribution includes a total of 91 explanatory variables plus an intercept term.[5]

6. Results

We first report on model selection criteria for the choice between different copulas, the results of a number of specification tests of a number of key features of the bivariate hurdle model, and parameter estimates of the dynamic relationships. We then report on the properties of contemporaneous association and tail dependence highlighted by the copula. Finally, because the dynamic relationships inherent in the parameter estimates are quite complicated, we report on calculations of partial effects which illustrate the dynamics much more transparently.

Table 4 reports maximized log likelihood values for several copulas. Because each of the objective functions has the same number of parameters across copulas, picking the model with the largest log likelihood is equivalent to choosing the model with the smallest Akaike or Bayesian Information Criterion. For each subsample, the Clayton copula provides a superior fit in both parts of the model. Examining each part of the model separately,

[4] The chronic condition dummies indicate the presence of cancer, diabetes, arthritis, asthma, hypertension, a mental condition, a urine condition, and a heart condition.

[5] For some of the subsamples, the number of explanatory variables is less because some variables are omitted. For example, the sample consisting of subjects with prescription drug coverage omits the indicator for prescription drug coverage. The time-varying variables used to calculate Mundlak terms are: age, age squared, female*age, female*age squared, married, widow, divorced, family size, education, log of income, employed, firm size, govtjob, private insurance, public insurance, prescription drug coverage, very good health, good health, fair health, poor health, physical limitation, injury, cancer, diabetes, mental illness, arthritis, asthma, urine condition, hypertension, and heart condition.

we note that the Clayton copula provides a substantially better fit for the hurdle part of the model in each case, and also generally for the conditional part, except for three cases for which there is little discrimination across models. The last two columns in Table 4 present estimates of contemporaneous dependence for the different copulas. Despite the different dependence structures implied by the different copulas, estimates reveal a qualitative consistency in dependence magnitudes.

Parameter estimates from the bivariate hurdle model with Clayton copulas are reported in Tables 5-10. The left panel of results corresponds to the hurdle part, and the right panel reports findings for positive expenditures. Only estimates of the autoregressive parameters are shown in the tables along with a number of specification tests and the copula parameters. The models include a rich set of controls, as outlined above, but these are not shown in the tables in the interest of brevity. Tables of results for the full models are available upon request. Although not shown, we note that the estimated coefficients of the control variables are similar in sign to previous studies of medical care access and spending. Not surprisingly, the most important determinants of medical spending, both in terms of magnitude and statistical significance, are health status measures. Individuals with health problems and/or physical limitations are more likely to have positive spending and have higher levels of spending compared to their more healthy counterparts.

The dynamic relationships between the two types of expenditures are captured by the coefficients of functions of lagged expenditures, both as binary indicators of any expenditure and the logarithm of expenditures. There is clear evidence of own and cross lagged effects of spending in both the binary response or hurdle part and the continuous part of the model. Rather that discussing every own and cross effect in Tables 5-10, the discussion that follows focuses on the relationship between lagged drug spending and current period nondrug spending, as this relationship informs on the presence and magnitude of cost-offsets.

In the hurdle component of the model, a consistent pattern emerges across the subsamples: Indicators of lagged positive drug spending are associated with lower probabilities of present-quarter nondrug spending. The 1-quarter lagged indicator of positive drug spending is negative and significant in all six subsamples, while the 2-quarter lagged indicator is negative and significant in the well insured, mental illness, and arthritis samples. In

contrast, the actual amounts of lagged (logged) drug spending are positively related to the probability of present-quarter nondrug spending. (The only lagged logged drug spending measure that is not significant is the 2-quarter lag in the 65 and older sample.) Although negative coefficients of the lagged binary indicators are larger in magnitude than the positive coefficients of the lagged (logged) spending variables, it is difficult to ascertain whether this is evidence of cost-offsets, as the lagged measures correspond to different scales. Furthermore, contemporaneous dependence, discussed in the following subsection, appears to be unambiguously positive. We attempt to quantify these various off-setting effects below.

In the second part of the model, which describes positive spending, none of the lagged measures of drug spending, either binary of logged amounts, appears to be significantly related to nondrug spending. Therefore, we expect that cost-offsets, to the extent that they exist, are largely driven by the hurdle part of the model.

The chi-square test of the null hypothesis that the initial conditions have zero coefficients is reported in Tables 5-10; this refers to the $\boldsymbol{\tau}_k$ and $\boldsymbol{\tau}_{+k}$ terms in (13) and (17). The joint null is rejected in every case, for both parts of the model, at $p < 0.01$. The tables also report a chi-square test that the "Mundlak terms" are jointly insignificant. This refers to the joint significance of the $\boldsymbol{\lambda}_{0k}$ and $\boldsymbol{\lambda}_{+k}$ coefficients in (13) and (17) of the correlated random effects specification. This null hypothesis is also conclusively rejected in every case. Both these tests support the desirability of our more flexible random effects specification. Finally, the tables report tests of the hypothesis that $\eta_j = \eta_j^+$ for $j = 1, 2$, i.e., the shape parameters of the gamma distributions for y_1^+ and y_2^+ are the same in the specifications of the densities in y_1^+, y_2^0, y_1^0, y_2^+ and y_1^+, y_2^+. The null hypothesis of equality is rejected in every case, with the exceptions of drug spending in the diabetes and heart condition samples. In addition, although we do not report test statistics, it is clear that skewness and kurtosis are significantly higher for drug than for nondrug expenditures.

6.1. Contemporaneous and tail dependence

The copula dependence parameters θ_0 and θ_+, reported at the bottom of Tables 5-10, measure contemporaneous dependence between drug and nondrug spending, after controlling for the influence of all explanatory and lagged spending variables. Although less interesting from a policy perspective, contemporaneous dependence represents an important

benchmark, as most previous studies have estimated contemporaneous cost-offsets based on cross sectional data. Our results indicate that the Clayton copula gives the best fit, and this copula supports positive contemporaneous dependence. The results show strong evidence of positive contemporaneous dependence in all subsamples and for both parts of the model. Both θ_0 and θ_+ are estimated with high degrees of precision, so this appears to be a robust finding.

Contemporaneous dependence is larger in magnitude in the hurdle part, with θ_0 between 1.00 and 1.30 (Kendall's tau between 0.33 and 0.39). By comparison, in the second part, θ_+ is between 0.20 and 0.25 (Kendall's tau between 0.09 and 0.11). The interpretation is that an individual's probabilities of positive drug and nondrug spending are more closely related than the amounts of drug and nondrug spending.

To illustrate contemporaneous dependence, post estimation we set explanatory variables equal to their mean values and coefficients equal to their estimated values for each subsample. From the estimated bivariate density, we then draw 2000 Monte Carlo realizations of $(\Pr(y_1 > 0), \Pr(y_2 > 0))$ for the hurdle part and (y_1, y_2) for the second part. These simulated pairs appear graphically in Figures 1 and 2. The figures attempt to depict the nature of dependence implied by the estimated dependence parameter, although skewness in the marginals, themselves, prevents a clean visual interpretation of the dependence. Nevertheless, as the lower tail dependency measure for the Clayton copula is $2^{-1/\theta}$, which indicates that a larger θ value is associated with greater lower tail dependency, the precisely estimated dependence parameters suggest that quarters in which an individual has low probability of incurring drug expenses tend to be the same quarters of low probability of nondrug expenses. Similarly, quarters of low drug spending also exhibit low nondrug spending.

6.2. Dynamic dependence and partial effects

We compute measures of effects that are analogous to the average partial effect proposed by Wooldridge (2005). We define the average partial effect (APE) of $y_{j,t-1}$ on $y_{k,t}$ as

$$APE_k(y_{j,t-1}) = E(y_{k,t}|y^{(1)}_{j,t-1}, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*) - E(y_{k,t}|y^{(0)}_{j,t-1}, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*) \qquad (21)$$

where $j, k = 1, 2$ and $y^{(1)}_{j,t-1}$ and $y^{(0)}_{j,t-1}$denote values of $y_{j,t-1}$ over which the partial effect is desired. All other covariates in the model, including other lagged endogenous regressors

$y_{-k,t-1}, \mathbf{y}_{t-2}$ and exogenous covariates $\mathbf{x}$ are fixed at representative values denoted by "$*$". Different conventions may be used to set $\mathbf{x}^*$; see, for example Stuart et al. (2007). Although this measure is limited because it only captures the one-period impact on $y_{it}^{(2)}$ of the lagged change in binary-valued variable $y_{it-1}^{(1)}$ (the model includes two lags), we estimate one period APE for two reasons. First, the complexity of the model makes it impossible to convey, in an intuitive way, the effects from all coefficients of lagged spending measures. Second, although two-period lagged measures of spending are sometimes significant, particularly in the hurdle part of the model, coefficients on the one-period lags are several times larger in magnitude. Thus, we expect any cost-offsets to derive primarily from one-period effects.

Specifically

$$\begin{aligned} & E(y_{k,t}|y_{j,t-1}^{(m)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \qquad (22) \\ = \; & \Pr(y_{k,t}=1|y_{j,t-1}^{(1)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \times E(y_{k,t}|y_{k,t}>0, y_{j,t-1}^{(1)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \\ + \; & \Pr(y_{k,t}=1|y_{j,t-1}^{(0)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \times E(y_{k,t}|y_{k,t}>0, y_{j,t-1}^{(0)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \end{aligned}$$

where $\Pr(y_{k,t}=1|y_{j,t-1}^{(m)}, y_{-k,t-1}^*, y_{t-2}^*, x^*)$ and $E(y_{k,t}|y_{k,t}>0, y_{j,t-1}^{(m)}, y_{-k,t-1}^*, y_{t-2}^*, x^*)$ for $m=0,1$ are obtained directly from the marginal probit and gamma distributions respectively.

In this paper, we calculate APEs corresponding to the cost offset hypothesis, i.e., we calculate the effects of drug expenditures at time $(t-1)$ on non-drug spending at time t. Specifically, we set $y_{j,t-1}^{(0)}=0$ (no drug expenditure) and $y_{j,t-1}^{(1)}>0$ (positive values of drug expenditure) to calculate APEs. Because there is no convention regarding what value of $y_{j,t-1}^{(1)}$ to use, we calculate APEs repeatedly over a range of values of $y_{j,t-1}^{(1)}$.[6] These are reported in the 6 panels of Figure 3. The gray "45 degree" line demarcates the region in which a dollar in drug spending is offset by more than a dollar in non-drug spending (below and to the left of the line) from the region in which the cost offset is less than one dollar (above and to the right of the line).

Estimated APE is negative in each case throughout most of the range of drug spending, although it does become zero in two cases when the change in drug spending is assumed to be large ($180). For all six samples, the magnitude of APE decreases over the range of

[6] The minimum ($30) and maximum ($180), roughly represent the 25th and 75th percentiles of observed values of drug expenditure, conditional on it being positive.

previous drug spending changes. Taking the 65 and older sample as an example, previous quarter drug spending of \$30 is associated with a current quarter reduction in nondrug spending of approximately \$60. Similarly previous quarter drug spending of \$180 translates to a current quarter reduction in nondrug spending of approximately \$15. Estimates of APE over the range of previous quarter drug spending for the other samples are as follows: continuously insured: −\$25 to \$0, diabetes: −\$100 to −\$40, mental illness: −\$25 to \$0, arthritis: −\$60 to −\$15, heart condition: −\$80 to −\$20. The APE estimates suggest modest cost-offsets in nondrug spending in the quarter following an increase in drug expenditures. Over most of the distribution of drug spending, the magnitudes of cost-offsets are less than dollar-for-dollar, indicating that increases in drug spending translate to increases in aggregate medical spending. In two cases, for the sample of continuously insured and for those with a mental illness, while there is a decrease in non-drug spending, it is less than the amount spent on drugs in the previous quarter.

6.3. An alternative measure of cost-offset

The APEs defined above estimate partial effects that are "marginal" over the distribution of current drug expenditures, which includes a substantial fraction of zeros. But, given that drug expenditures at time t are often predicated on nondrug spending at time t via prescription refill rules and/or physician monitoring behavior, it is important to identify cost offsets conditional on specific values of current drug expenditures, especially as the preferred Clayton-copula formulation suggests positive contemporaneous association along with left tail dependence between the two types of spending. Therefore, we define the conditional average partial effect ($CAPE$) on $y_{k,t}$ given $y_{j,t}$ of the effect of $y_{j,t-1}$ as

$$\begin{aligned} CAPE_k(y_{j,t-1}) &= E(y_{k,t}|y^{(1)}_{j,t-1}, y_{j,t}, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*) \\ &- E(y_{k,t}|y^{(0)}_{j,t-1}, y_{j,t}, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*); \end{aligned} \tag{23}$$

where $j, k = 1, 2$ and $j \neq k$ and $y^{(1)}_{j,t-1}$ and $y^{(0)}_{j,t-1}$denote values of $y_{j,t-1}$ over which the partial effect is desired. The key difference between APE described by equation (21) and $CAPE$ described by equation (23) is the additional conditioning on $y_{j,t}$ in the calculation of $CAPE$. Thus $CAPE_k$ shows how $APE_k(y_{j,t-1})$ changes with $y_{j,t}$. Calculation of the conditional (on $y_{j,t}$) expectations is considerably more complicated than the unconditional

expectations needed for the calculation of the APEs. For the hurdle probabilities,

$$\Pr(y_{k,t} = 1|y^{(1)}_{j,t-1}, y_{j,t} = 0, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*) = \frac{\Pr(y_j = 0, y_k > 0)}{\Pr(y_j = 0)} \tag{24}$$

and

$$\Pr(y_{k,t} = 1|y^{(1)}_{j,t-1}, y_{j,t} > 0, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*) = \frac{\Pr(y_j = 0, y_k > 0)}{\Pr(y_j > 0)} \tag{25}$$

where the terms in the numerator involve the copula formulation and the terms in the denominator are the probit marginals. For the expectations in the "positives" part of the model,

$$E(y_{k,t}|y_{k,t} > 0, y^{(1)}_{j,t-1}, y_{j,t}, y^*_{-k,t-1}, \mathbf{y}^*_{t-2}, \mathbf{x}^*) = \int_{\lim v_k \to 0}^{\infty} v_k \frac{f_{12}(y_j, v_k|y_j > 0, y_k > 0)}{f^+_j(y_j|y_j > 0, y_k > 0)} dv_k, \tag{26}$$

which is computed using numerical integration.

The panels of Figure 4 display values of $CAPE$ for two values of $y_{k,t}$, \$30 and \$180, chosen because they represent the 25th and 75th percentiles of positive drug spending. The figures show that the cost offset is greater the higher is drug spending in the current quarter. Evidence of greater than dollar cost offsets (values below and to the left of the "45 degree" line) are substantially greater here than in the unconditional case of the APE. In other words, real cost offsets are more likely for individuals who have drug expenditures in the current period as compared to those who do not. In contrast to the APE results, cost offsets do not tend towards zero, underlining the distinction once again, between the extensive and intensive margins of spending on drugs.

7. Conclusion

Previous research on the relationship between drug and nondrug spending has produced mixed results. This is due to several empirical complications. First, with high proportions of zeros, health care spending measures cannot be easily described by a single statistical distribution. Second, the bivariate dependence between drug and nondrug spending might exhibit substantial departures from normality. Third, the contemporaneous relationship between drug and nondrug spending might be fundamentally different from the economically more relevant dynamic relationship. Fourth, as medical effects of prescription drugs

might be fast-acting, investigating the dynamic relationship between drug and nondrug spending requires panel data recorded at relatively high frequency.

This paper proposes a dynamic nonlinear multivariate hurdle model and applies it to the case of drug and nondrug spending. More generally, the model also applies to situations in which dynamic dependence between pairs of outcomes are of interest, and where corner solutions are likely, i.e., there are substantial numbers of zeros. Using nationally-representative quarterly data on medical expenditures, the model is estimated for six policy-relevant subsamples. The models produce evidence of positive contemporaneous dependence, somewhat similar to previous studies. However, the models produce negative dynamic dependence across numerous samples and specifications, which we interpret as evidence of cost-offsets. Average partial effects (APE), analogous to those proposed by Wooldridge (2005), suggest that cost-offsets are smaller than dollar-for-dollar. Conditional average partial effects ($CAPE$), calculated similarly to APE but conditioned on specific values for current quarter drug spending, reveal that cost-offsets are larger than dollar-for-dollar for reasonably large current period drug spending.

These results hold important implications for public health insurance policies. If cost-offsets are larger than dollar-for-dollar, then aggregate health care spending might be reduced by encouraging increased spending on prescription drugs. Although our results indicate larger than dollar-for-dollar cost-offsets might exist under certain conditions, those conditions are likely to be too unpredictable to allow formulation of appropriate policies. For example, $CAPE$ estimates suggest that larger than dollar-for-dollar cost-offsets exist between previous quarter drug spending and current quarter nondrug spending when the patient has current quarter drug spending. It seems difficult to implement policies based on these conditions, as spending for certain drugs might be highly unexpected.

Our choice of modeling strategy, especially along the dimension of lag structure, is driven in part by the structure of our panel dataset. While we are able to ascertain whether there are relatively rapid cost-offsets and evaluate the magnitudes of such offsets, for many health conditions and drug classes, effects may not be expected so rapidly. Instead, one might expect no short-term effects but significant long-term effects, even as distant in the future as five years ahead or over the individual's remaining life. Neither is our dataset nor is our model designed to address such issues.

References

Arulampalam, W. and Stewart, M.B. (2009). Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and a Comparison with Alternative Estimators. *Oxford Bullentin of Economics and Statistics*, forthcoming.

Bien, K., Nolte, I. and Pohlmeier, W. (2007). An Inflated Multivariate Integer Count Hurdle Model: An Application to Bid and Ask Quote Dynamics, *CoFE discussion paper 07/04*.

Chamberlain, G. (1984). Panel Data. In *Handbook of Econometrics*, Volume II, ed. by Z. Griliches and M. Intriligator, 1247-1318. Amsterdam: North-Holland.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula Methods in Finance*. New York: John Wiley.

Clayton, D.G. (1978). A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence. *Biometrika*, 65, 141-151.

Cragg, J.G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods, *Econometrica*, 39, 829–844.

Cunningham, P.J. (2005). Medicaid Cost Containment and Access to Prescription Drugs, *Health Affairs*, 24, 780-789.

Duan, N., W.G. Manning, et al. (1983). A Comparison of Alternative Models for the Demand for Medical Care, *Journal of Business and Economics Statistics* 1, 115-126.

Duan, N., W.G. Manning, et al. (1984). Choosing Between the Sample-Selection Model and the Multi-Part Model, *Journal of Business and Economic Statistics* 2, 283-289.

Embrechts, P., McNeil, A., and D. Straumann (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls. In *Risk Management: Value at Risk and Beyone*, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, 176-223.

Gaynor, M., Li, J., and Vogt, W.B. (2007). Substitution, Spending Offsets, and Prescription Drug Benefit Design. *Forum for Health Economics and Policy*, 10, 1-31.

Goldman, D.P. et al. (2004). Pharmacy Benefits and the Use of Drugs by the Chronically Ill. *Journal of the American Medical Association*, 291, 2344-2350.

Goldman, D. and Philipson, T. (2007). Integrated insurance design in the presence of multiple medical technologies. *American Economic Review,* 97, 427-432.

Heckman, J.J. (1981). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. in C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Joyce, G.F., Escarce, J.J., Solomon, M.D., and Goldman, D.P. (2002). Employer Drug Benefit Plans and Spending on Prescription Drugs. *Journal of the American Medical Association*, 288, 1733–1739

Lichtenberg, F.R. (1996). Do (More and Better) Drugs Keep People out of Hospitals? *American Economic Review*, 86, 384-388.

Lichtenberg, F.R. (2001). Are the Benefits of Newer Drugs Worth their Cost? Evidence from the 1996 MEPS. *Health Affairs*, 20, 241-251.

Manning, W., Basu, A., and Mullahy, J. (2005). Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Data. *Journal of Health Economics*, 24, 465-88.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46, 69-85.

Nelsen, R.B. (2006). *An Introduction to Copulas. Second edition*. Springer, New York.

Patton, A. (2006). Modelling Asymmetric Exchange Rate Dependence. *International Economic Review,* 47, 527–556.

Pohlmeier, W. and Ulrich, V. (1995). An Econometric Model of the Two-Part Decision-Making Process in the Demand for Health Care. *Journal of Human Resources*, 30, 339-361.

Poisal, J.A. and Murray, L.A. (2001). Growing differences between Medicare beneficiaries with andwithout drug coverage. *Health Affairs,* 20, 75-85.

Shang, B. and Goldman, D.P. (2007). Prescription Drug Coverage and Elderly Medicare Spending. NBER Working Paper No. 13358.

Sklar, A. (1973). Random Variables, Joint Distributions, and Copulas. *Kybernetica*, 9, 449-460.

Smith, M. (2003). On Dependency in Double-Hurdle Models. *Statistical Papers*, 44, 581-595.

Stuart, B. and Grana, J. (1995). Are Prescribed and Over-the-Counter Drugs Economic Substitutes? The Effects of Health Insurance on Medicine Choices by the Elderly. *Medical Care*, 33, 487-501.

Stuart, B., Doshi, J., Briesacher, B., Shea, D., and Wrobel, M. (2007).Will Part D Produce Savings in Part A and Part B? The Impact of Prescription Drug Coverage on Medicare Program Expenditures. *Inquiry*, 44, 146-156.

Stuart, B., Doshi, J., Briesacher, B., Wrobel, M., Baysac, F. (2004). Impact of Prescription Coverage on Hospital and Physician Costs: A Case Study of Medicare Beneficiaries with Chronic Obstructive Pulmonary Disease. *Clinical Therapeutics*, 26, 1688-1699.

Stuart, B., Simoni-Wastila, L., Chauncey. D. (2005). Assessing the Impact of Coverage Gaps in the Medicare Part D Drug Benefit. *Health Affairs*, Web Exclusive, April.

Trivedi, P.K. and Zimmer, D.M. (2007). Copula Modeling: An Introduction for Practitioners, *Foundations and Trends in Econometrics*, 1, 1-110. (Also published separately as a book.)

Wooldridge, J. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20, 39-54.

Zang, Y. and Soumerai, S. (2007). Do Newer Prescription Drugs Pay for Themselves? A Reassessment of the Evidence. *Health Affairs*, 26, 880-886.

Table 1 : Quarterly medical spending by subsample

	Drug expenditure			Nondrug expenditure		
Statistic	all	1(> 0)	| > 0	all	1(> 0)	| > 0
			65 and older			
Mean	49.18	0.29	168.42	2824.92	0.67	2824.92
Median	0.00		90.68	422.58		422.58
1st order serial corr	0.19	0.18		0.43	0.38	
			Continuously insured - medical and Rx			
Mean	29.20	0.21	139.14	1925.60	0.46	1925.60
Median	0.00		74.15	322.85		322.85
1st order serial corr	0.19	0.18		0.42	0.37	
			Diabetes			
Mean	69.09	0.34	202.31	3258.39	0.70	3258.39
Median	0.00		110.23	470.04		470.04
1st order serial corr	0.17	0.15		0.42	0.35	
			Mental Illness			
Mean	55.79	0.31	178.63	2406.50	0.61	2406.50
Median	0.00		98.09	414.17		414.17
1st order serial corr	0.20	0.19		0.44	0.38	
			Arthritis			
Mean	51.75	0.32	162.53	2608.91	0.66	2608.91
Median	0.00		91.44	442.35		442.35
1st order serial corr	0.18	0.17		0.43	0.37	
			Heart condition			
Mean	58.02	0.32	180.73	2750.15	0.66	2750.15
Median	0.00		99.11	404.76		404.76
1st order serial corr	0.16	0.15		0.40	0.33	

Table 2: Sample means by subsample

	65 and older	Fully insured	diabetes	mental illness	arthritis	heart condition
Socioeconomic						
Age	74.6	44.8	59.2	47.7	57.7	60.7
Female	0.59	0.53	0.56	0.68	0.63	0.58
Black	0.12	0.11	0.19	0.10	0.15	0.18
Hispanic	0.12	0.13	0.24	0.16	0.15	0.14
Never married			omitted			
Married	0.52	0.69	0.58	0.49	0.56	0.58
Divorced	0.10	0.11	0.16	0.21	0.17	0.15
Widow	0.34	0.04	0.17	0.11	0.17	0.18
Family size	1.90	2.98	2.60	2.65	2.41	2.41
Education (total years of schooling)	11.25	13.40	11.14	12.30	11.89	11.81
Northeast residence			omitted			
Midwest residence	0.22	0.24	0.19	0.22	0.22	0.21
West residence	0.38	0.35	0.43	0.36	0.39	0.42
South residence	0.21	0.23	0.23	0.26	0.23	0.20
Metropolitan statistical area	0.74	0.81	0.75	0.77	0.75	0.75
Employed	0.17	0.83	0.43	0.62	0.50	0.47
Log income (person-level, wage and nonwage)	5.15	5.23	5.15	5.16	5.17	5.17
Firm size	1.07	12.89	5.73	7.73	6.19	6.22
Government job	0.02	0.16	0.08	0.11	0.09	0.09
Health						
Excellent health			omitted			
Very good health	0.27	0.35	0.17	0.26	0.25	0.25
Good health	0.32	0.26	0.35	0.31	0.32	0.35
Fair health	0.18	0.07	0.29	0.19	0.21	0.21
Poor health	0.06	0.02	0.14	0.09	0.10	0.09
Physical limitation	0.56	0.19	0.55	0.43	0.55	0.48
Injury	0.21	0.21	0.23	0.30	0.32	0.23
Number of chronic conditions	1.75	0.74	2.52	1.79	2.08	2.13
Insurance						
Private Insurance	0.55	1.00	0.55	0.63	0.64	0.63
Public Insurance	0.44	0.00	0.35	0.24	0.28	0.29
Have Prescription drug insurance	0.34	1.00	0.46	0.56	0.52	0.52

Table 3: Different parameteric copulas

Copula type	Function $C(u_1, u_2)$	$\theta-$domain	Kendall's τ
Clayton	$(u_1^{-\theta}+u_2^{-\theta}-1)^{-1/\theta}$	$\theta \in (0, \infty)$	$\frac{\theta}{\theta+2}$
Survival Clayton	$((1-u_1)^{-\theta}+(1-u_2)^{-\theta}-1)^{-1/\theta}$	$\theta \in (0, \infty)$	$\frac{\theta}{\theta+2}$
Frank	$\frac{1}{\theta}\log\left(1+\frac{(\exp(\theta u_1-1))(\exp(\theta u_2-1))}{\exp(\theta)-1}\right)$	$\theta \in (-\infty, \infty)$	$1+\frac{4}{\theta}\left[\int_0^\theta \frac{t}{\theta(e^t-1)}dt-1\right]$
Gaussian	$\Phi[\Phi^{-1}(u_1)\Phi^{-1}(u_2);\theta]$	$-1<\theta<+1$	$\frac{2}{\pi}\arcsin(\theta)$

Table 4: Maximized log likelihoods for models with alternative copulas

Copula	Subsample	Hurdle lnL	Conditional lnL	Overall lnL	Hurdle $\widehat{\tau}$	Conditional $\widehat{\tau}$
		Models with Clayton copula				
Clayton	65 and older	-81341*	-594359	-675700*	0.333	0.101
	Continuously insured	-280333*	-1425303*	-1705636*	0.366	0.093
	Diabetes	-45777*	-347241	-393018*	0.347	0.098
	Mental illness	-81647*	-539333*	-620981*	0.388	0.110
	Arthritis	-96854*	-689050	-785904*	0.384	0.091
	Heart condition	-131270*	-916224*	-1047494*	0.348	0.106
		Models with Survival Clayton copula				
Survival Clayton	65 and older	-81934	-594359	-676293	0.320	0.021
	Continuously insured	-283909	-1425319	-1709228	0.500	0.013
	Diabetes	-46009	-347249	-393258	0.324	0.015
	Mental illness	-82470	-539352	-621822	0.410	0.019
	Arthritis	-97865	-689041*	-786907	0.373	0.019
	Heart condition	-131991	-916232	-1048223	0.351	0.021
		Models with Frank copula				
Frank	65 and older	-81531	-594357*	-675888	0.338	0.059
	Continuously insured	-281689	-1425343	-1707032	0.446	0.043
	Diabetes	-45822	-347237*	-393059	0.341	0.053
	Mental illness	-81887	-539344	-621231	0.405	0.061
	Arthritis	-97167	-689057	-786224	0.387	0.050
	Heart condition	-131411	-916235	-1047646	0.353	0.059
		Models with Gaussian copula				
Gaussian	65 and older	-81527	-594369	-675896	0.332	0.045
	Continuously insured	-281123	-1425348	-1706471	0.440	0.029
	Diabetes	-45837	-347246	-393084	0.339	0.037
	Mental illness	-81864	-539357	-621221	0.408	0.045
	Arthritis	-97176	-689061	-786237	0.386	0.040
	Heart condition	-131448	-916243	-1047691	0.354	0.047

* denotes model with best fit for given subsample

Table 5: Bivariate two-part model: coefficients of lagged variables
Sample: age 65 and older

	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	0.089**	-0.117**	-0.149**	-0.046
	(0.032)	(0.036)	(0.050)	(0.072)
$1(\text{nondrug}_{t-1} > 0)$	0.002	0.105**	-0.096*	-0.910**
	(0.025)	(0.028)	(0.043)	(0.062)
$ln(\text{drug}_{t-1})$	0.003	0.038**	0.032**	-0.001
	(0.007)	(0.008)	(0.010)	(0.015)
$ln(\text{nondrug}_{t-1})$	0.022**	0.066**	0.022**	0.166**
	(0.004)	(0.004)	(0.006)	(0.008)
$1(\text{drug}_{t-2} > 0)$	0.009	0.040	-0.180**	0.023
	(0.032)	(0.036)	(0.051)	(0.074)
$1(\text{nondrug}_{t-2} > 0)$	0.091**	0.274**	0.022	-0.458**
	(0.026)	(0.028)	(0.044)	(0.064)
$ln(\text{drug}_{t-2})$	0.023**	0.004	0.057**	-0.004
	(0.007)	(0.008)	(0.010)	(0.015)
$ln(\text{nondrug}_{t-2})$	0.004	0.020**	-0.009	0.065**
	(0.004)	(0.004)	(0.006)	(0.008)
χ^2 test for initial conditions = 0	242.6**	781.5**	35.5**	17.0**
χ^2 test for Mundlak terms = 0	195.9**	151.4**	105.5**	88.4**
χ^2 test for $\eta_j = \eta_j^+$	–	–	10.7**	1811**
$\theta_0; \theta_+$	0.997		0.224	
	(0.020)		(0.022)	
Kendall's Tau	0.333		0.101	
	(0.004)		(0.009)	
lnL	-81341		-594359	
N	78162		55848	

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 6: Bivariate two-part model: coefficients of lagged variables
Sample: well insured - medical and Rx

	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	-0.007	-0.171**	-0.228**	-0.020
	(0.020)	(0.020)	(0.034)	(0.053)
$1(\text{nondrug}_{t-1} > 0)$	0.077**	0.015	-0.003	-0.995**
	(0.016)	(0.016)	(0.029)	(0.045)
$ln(\text{drug}_{t-1})$	0.027**	0.053**	0.059**	-0.015
	(0.004)	(0.004)	(0.007)	(0.011)
$ln(\text{nondrug}_{t-1})$	0.013**	0.078**	0.013**	0.213**
	(0.003)	(0.003)	(0.005)	(0.006)
$1(\text{drug}_{t-2} > 0)$	-0.008	-0.082**	-0.293**	0.016
	(0.019)	(0.019)	(0.035)	(0.052)
$1(\text{nondrug}_{t-2} > 0)$	0.122**	0.249**	0.030	-0.395**
	(0.016)	(0.016)	(0.031)	(0.046)
$ln(\text{drug}_{t-2})$	0.026**	0.031**	0.079**	-0.007
	(0.004)	(0.004)	(0.007)	(0.011)
$ln(\text{nondrug}_{t-2})$	-0.002	0.005	-0.000	0.075**
	(0.003)	(0.003)	(0.005)	(0.007)
χ^2 test for initial conditions = 0	804.6**	1543**	57.2**	32.9**
χ^2 test for Mundlak terms = 0	1091**	900.1**	305.2**	210.1**
χ^2 test for $\eta_j = \eta_j^+$	–	–	11.5**	3431**
$\theta_0; \theta_+$	1.155		0.204	
	(0.010)		(0.017)	
Kendall's Tau	0.366		0.093	
	(0.002)		(0.007)	
lnL	-280333		-1425303	
N	289374		139969	

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 7: Bivariate two-part model: coefficients of lagged variables
Sample: diabetes

	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	0.076	-0.161**	-0.215**	-0.084
	(0.041)	(0.047)	(0.062)	(0.093)
$1(\text{nondrug}_{t-1} > 0)$	-0.052	0.030	-0.057	-1.023**
	(0.033)	(0.038)	(0.055)	(0.086)
$ln(\text{drug}_{t-1})$	0.007	0.047**	0.049**	0.004
	(0.008)	(0.010)	(0.012)	(0.019)
$ln(\text{nondrug}_{t-1})$	0.021**	0.074**	0.015	0.173**
	(0.005)	(0.006)	(0.008)	(0.011)
$1(\text{drug}_{t-2} > 0)$	0.006	-0.052	-0.169**	-0.005
	(0.041)	(0.047)	(0.063)	(0.098)
$1(\text{nondrug}_{t-2} > 0)$	0.031	0.169**	-0.091	-0.707**
	(0.034)	(0.038)	(0.055)	(0.084)
$ln(\text{drug}_{t-2})$	0.020*	0.026**	0.042**	-0.005
	(0.008)	(0.009)	(0.012)	(0.019)
$ln(\text{nondrug}_{t-2})$	0.010	0.032**	0.016*	0.100**
	(0.005)	(0.006)	(0.008)	(0.011)
χ^2 test for initial conditions = 0	90.5**	158.2**	21.8**	9.90**
χ^2 test for Mundlak terms = 0	225.5**	115.5**	101.1**	92.5**
χ^2 test for $\eta_j = \eta_j^+$	–	–	1.78	931.3**
$\theta_0; \theta_+$	1.062		0.195	
	(0.028)		(0.031)	
Kendall's Tau	0.347		0.089	
	(0.006)		(0.013)	
lnL	-45777		-347241	
N	42702		31680	

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 8: Bivariate two-part model: coefficients of lagged variables
Sample: mental illness

	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	0.001	-0.229**	-0.145**	0.051
	(0.031)	(0.034)	(0.050)	(0.073)
$1(\text{nondrug}_{t-1} > 0)$	0.004	0.020	-0.174**	-1.106**
	(0.026)	(0.029)	(0.044)	(0.067)
$ln(\text{drug}_{t-1})$	0.029**	0.062**	0.032**	-0.021
	(0.006)	(0.007)	(0.010)	(0.015)
$ln(\text{nondrug}_{t-1})$	0.018**	0.085**	0.037**	0.206**
	(0.004)	(0.005)	(0.007)	(0.009)
$1(\text{drug}_{t-2} > 0)$	-0.012	-0.100**	-0.218**	0.074
	(0.030)	(0.033)	(0.050)	(0.071)
$1(\text{nondrug}_{t-2} > 0)$	0.033	0.176**	0.003	-0.532**
	(0.027)	(0.029)	(0.047)	(0.067)
$ln(\text{drug}_{t-2})$	0.025**	0.029**	0.061**	-0.026
	(0.006)	(0.007)	(0.010)	(0.014)
$ln(\text{nondrug}_{t-2})$	0.010*	0.026**	0.004	0.086**
	(0.004)	(0.005)	(0.007)	(0.009)
χ^2 test for initial conditions = 0	195.3**	371.3**	24.4**	40.8**
χ^2 test for Mundlak terms = 0	317.5**	282.5**	101.8**	87.1**
χ^2 test for $\eta_j = \eta_j^+$	–	–	9.13**	1293**
$\theta_0; \theta_+$	1.267		0.246	
	(0.021)		(0.023)	
Kendall's Tau	0.388		0.110	
	(0.004)		(0.009)	
lnL	-81647		-539333	
N	76848		49601	

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 9: Bivariate two-part model: coefficients of lagged variables
Sample: arthritis

	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	0.016	-0.145**	-0.283**	-0.058
	(0.028)	(0.032)	(0.044)	(0.063)
$1(\text{nondrug}_{t-1} > 0)$	-0.020	0.030	-0.087*	-0.967**
	(0.023)	(0.026)	(0.039)	(0.059)
$ln(\text{drug}_{t-1})$	0.022**	0.046**	0.062**	-0.001
	(0.006)	(0.007)	(0.009)	(0.013)
$ln(\text{nondrug}_{t-1})$	0.023**	0.076**	0.021**	0.187**
	(0.003)	(0.004)	(0.006)	(0.008)
$1(\text{drug}_{t-2} > 0)$	-0.023	-0.105**	-0.175**	-0.029
	(0.028)	(0.031)	(0.045)	(0.066)
$1(\text{nondrug}_{t-2} > 0)$	0.123**	0.184**	0.026	-0.416**
	(0.024)	(0.026)	(0.041)	(0.059)
$ln(\text{drug}_{t-2})$	0.030**	0.036**	0.057**	0.005
	(0.006)	(0.007)	(0.009)	(0.013)
$ln(\text{nondrug}_{t-2})$	-0.003	0.023**	-0.005	0.060**
	(0.004)	(0.004)	(0.006)	(0.008)
χ^2 test for initial conditions = 0	303.3**	560.7**	42.7**	43.7**
χ^2 test for Mundlak terms = 0	322.5**	196.6**	144.6**	88.9**
χ^2 test for $\eta_j = \eta_j^+$	–	–	11.9**	1743**
	–	–		
$\theta_0; \theta_+$	1.245		0.200	
	(0.020)		(0.021)	
Kendall's Tau	0.384		0.091	
	(0.004)		(0.009)	
lnL	-96854		-689050	
N	91230		62983	

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table 10: Bivariate two-part model: coefficients of lagged variables
Sample: heart conditions

	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	0.059*	-0.139**	-0.131**	-0.081
	(0.025)	(0.027)	(0.040)	(0.061)
$1(\text{nondrug}_{t-1} > 0)$	-0.049*	-0.016	-0.098**	-1.112**
	(0.020)	(0.022)	(0.033)	(0.053)
$ln(\text{drug}_{t-1})$	0.009	0.043**	0.031**	0.004
	(0.005)	(0.006)	(0.008)	(0.012)
$ln(\text{nondrug}_{t-1})$	0.023**	0.073**	0.024**	0.192**
	(0.003)	(0.003)	(0.005)	(0.007)
$1(\text{drug}_{t-2} > 0)$	-0.000	-0.042	-0.208**	-0.020
	(0.025)	(0.027)	(0.041)	(0.061)
$1(\text{nondrug}_{t-2} > 0)$	0.057**	0.215**	-0.035	-0.499**
	(0.020)	(0.022)	(0.035)	(0.053)
$ln(\text{drug}_{t-2})$	0.018**	0.021**	0.056**	0.006
	(0.005)	(0.006)	(0.008)	(0.012)
$ln(\text{nondrug}_{t-2})$	0.006*	0.021**	0.004	0.073**
	(0.003)	(0.003)	(0.005)	(0.007)
χ^2 test for initial conditions = 0	246.4**	548.1**	42.4**	22.6**
χ^2 test for Mundlak terms = 0	548.9**	244.5**	139.2**	174.8**
χ^2 test for $\eta_j = \eta_j^+$	–	–	1.56	2520**
$\theta_0; \theta_+$	1.070		0.238	
	(0.016)		(0.018)	
Kendall's Tau	0.348		0.106	
	(0.003)		(0.007)	
lnL	-131270		-916224	
N	120552		84980	

Robust standard errors in parentheses
** p<0.01, * p<0.05

Figure 1: Simulated probabilities from hurdle part (2000 points plotted)

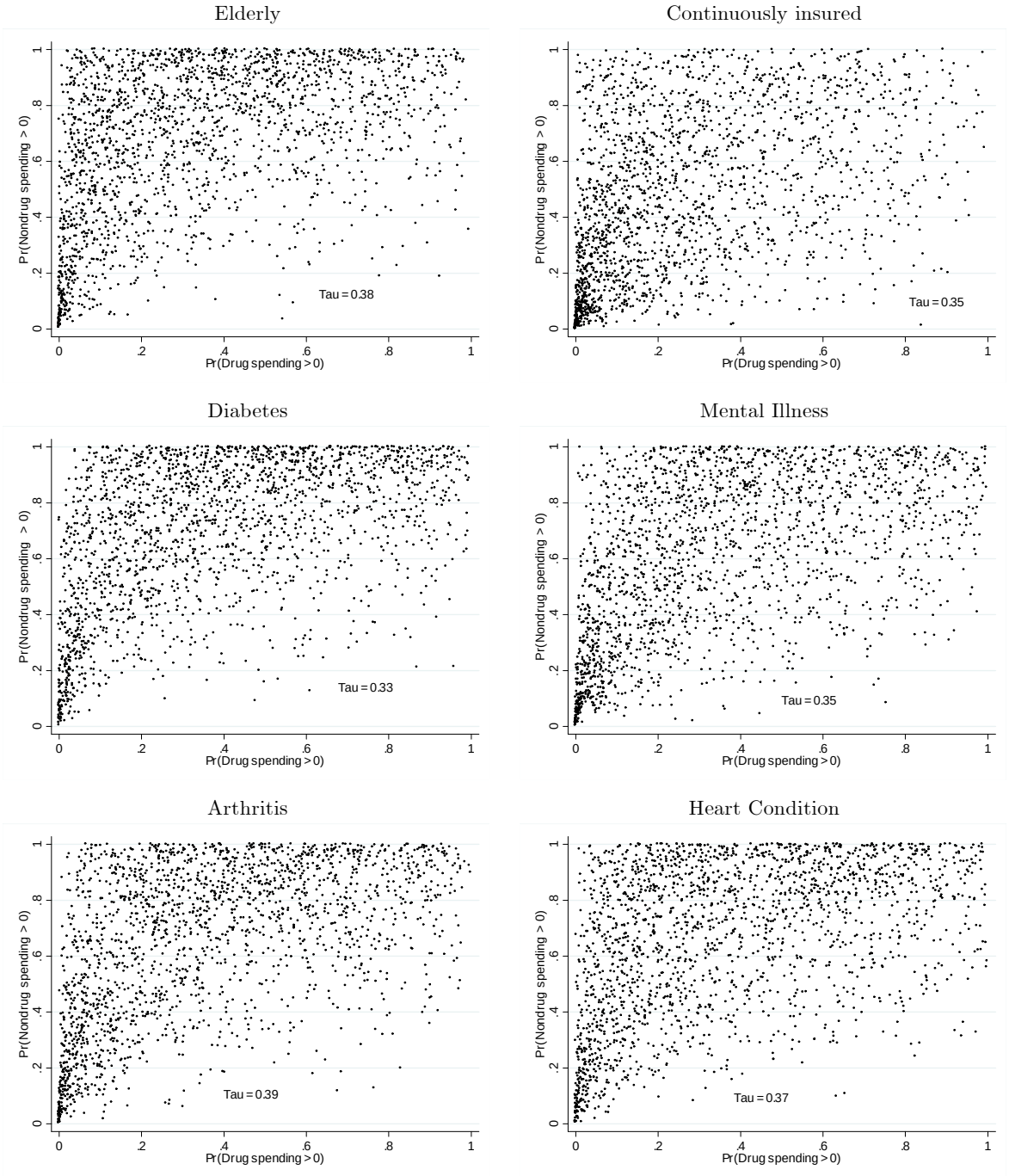

Figure 2: Simulated spending from positive spending part (2000 points plotted)

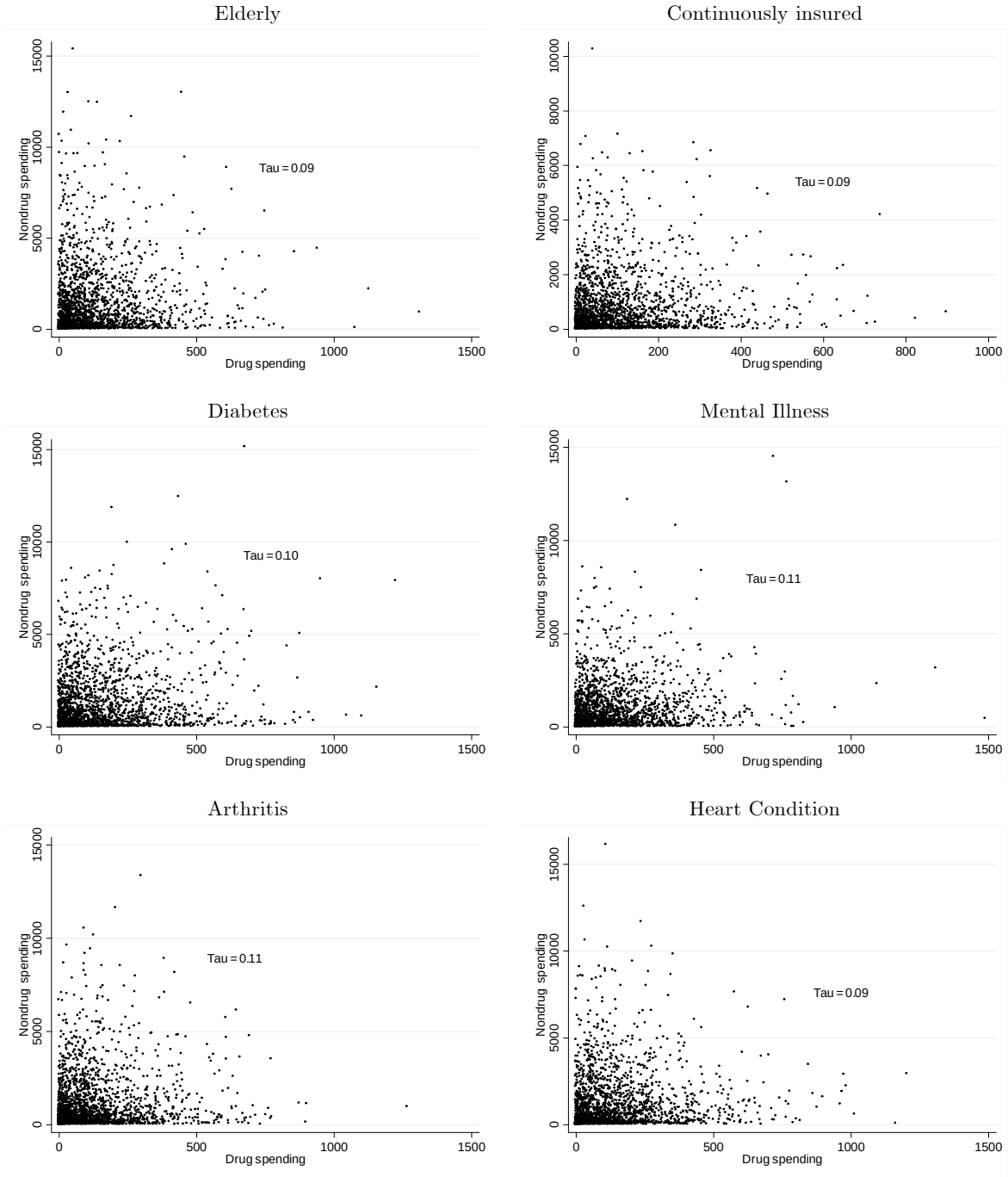

Figure 3: Average Partial Effects

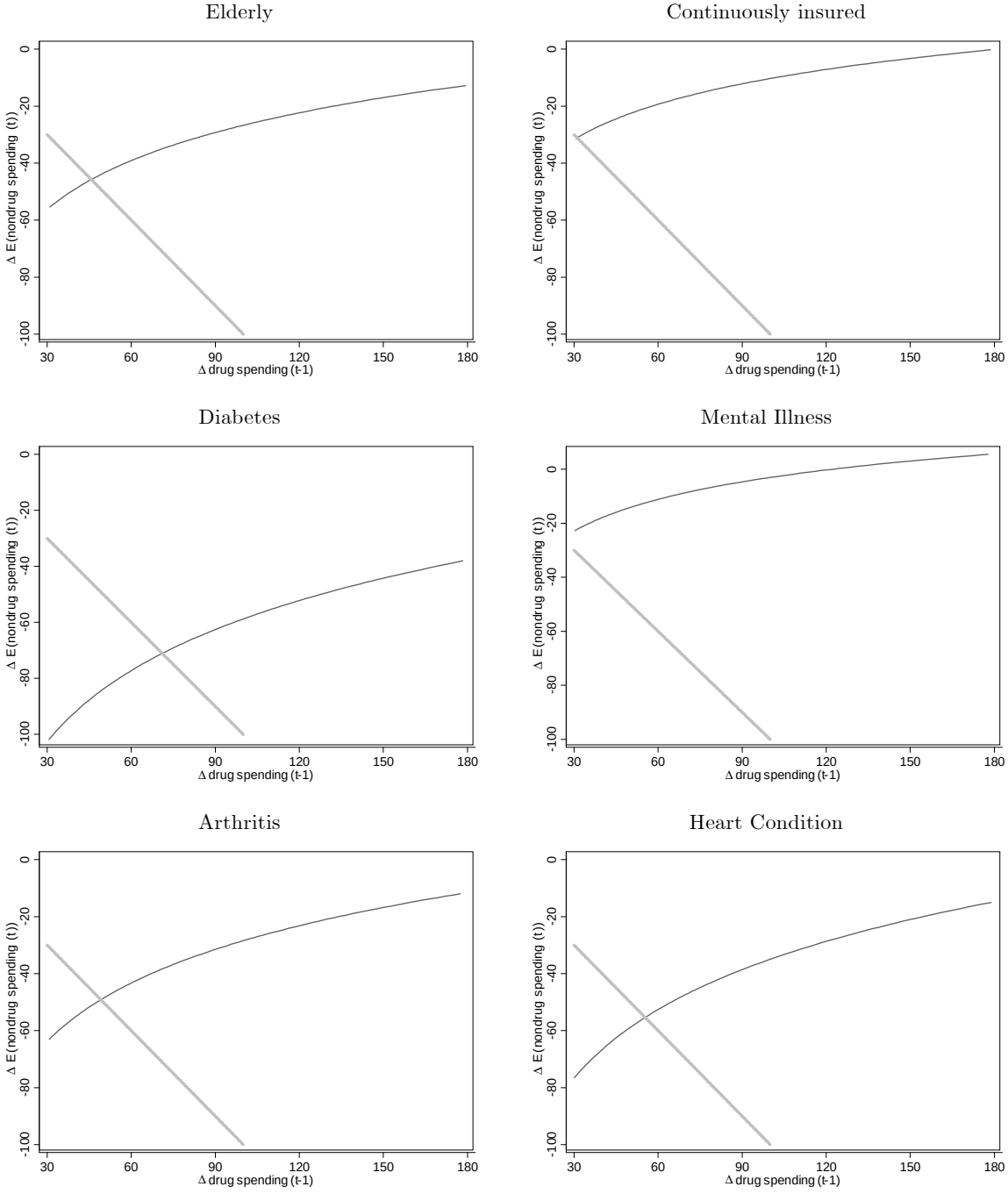

Figure 4: Conditional Average Partial Effects