



THE UNIVERSITY *of York*

HEDG Working Paper 09/15

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

Partha Deb

Pravin K. Trivedi

David M. Zimmer

July 2009

ISSN 1751-1976

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

June 2009

Abstract

This paper presents a new multivariate copula-based modeling approach for analyzing cost-offsets between drug and nondrug expenditures. Estimates are based on panel data from the Medical expenditure panel Survey (MEPS) with quarterly measures of medical expenditures. The approach allows for nonlinear dynamic dependence between drug and non-drug expenditures as well as asymmetric contemporaneous dependence. The specification uses the standard hurdle model with two significant extensions. First, it is adapted to the bivariate case. Second, because the cost-offset hypothesis is inherently dynamic, the bivariate hurdle framework is extended to accommodate dynamic relationships between drug and nondrug spending. The econometric analysis is implemented for six different groups defined by specific health conditions. There is evidence of modest cost-offset impact effect of expenditure on prescribed drugs.

JEL codes: C51, C33, I11

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

1. Introduction

In 2006 prescription drug expenditures accounted for around 10% of the US health budget. For the elderly this component of the healthcare expenditures is even more important, accounting for expenditure exceeding \$120 billion in 2005 and amounting to nearly \$2,800 per person (Kaiser Family Foundation, 2005). The expansion of Medicare Part D through the Medicare Modernization Act of 2005 is expected to increase the growth rate of this expenditure component even further.

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 expenditure might be expected. Overall the empirical evidence is mixed. In the U.S., while some health policy advocates argue that the use of new prescription drugs reduces total health care costs, many states are intensifying efforts to control rising prescription drug costs in their Medicaid programs (Cunningham, 2005). Thus it appears that answers to these questions remain unresolved in both scientific and policy arenas. This paper investigates these issues using a new econometric framework applied to quarterly panel data on prescribed drug and non-drug expenditures.

Econometric measurement of substitution and complementarity is complex. At the very least, a bivariate model is required, and one which can accommodate the presence of a significant proportion of zero-valued outcomes for both categories of expenditures. An additional complication comes from the heterogeneity of individuals and medical services that vary in their degree of substitutability, as the relationship between drug therapy and other medical usage varies across a range of medical services. This motivates disaggregation by health status and condition. A further complication is that both substitution and complementarity have a dynamic dimension, but most econometric studies to date have used static equilibrium frameworks and cross section data. This feature motivates our formulation of a dynamic panel data version of the bivariate two-part model. A leading issue is the difficulty of separating the pure incentive effects of health insurance from those due to adverse or advantageous selection, especially when such selection cannot be fully

controlled through observed characteristics. Thus it is not surprising that econometric results and inferences tend to vary across different studies.

This paper presents a new multivariate copula-based modeling approach for analyzing cost-offsets between drug and nondrug expenditures based on panel rather than the more commonly used cross section data. Such data support a dynamic reduced-form type approach which does not focus on the details of the mechanism through which cost-offsets may arise. The estimation approach, based on copula functions, can potentially explain why existing empirical results are varied. The approach allows for nonlinear dynamic dependence between drug and non-drug expenditures as well as asymmetric contemporaneous dependence. Our richer data set is built up from monthly event files of individual respondents. The resulting sample also permits disaggregation at the level of specific health conditions, thus allowing us to test the main hypotheses using data that are relatively less heterogenous than in many cross section settings.

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

The study of the 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 the Medicare supplemental insurance and the utilization of prescription drugs of the elderly; see, for example, Poisal and Murray, (2001), and Goldman and Philipson (2007). Of greater direct relevance to this paper is the strand in the 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 through which prescription drug usage impacts 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 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 non-linear 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 dependence between the two types of expenditures. We adopt a copula framework which accommodates a flexible formulation of dependence and marks a departure from the usual assumption of linear dependence.

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.¹ 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,

$$\begin{aligned} y_1^0, y_2^0 &\longrightarrow F(y_1 = 0, y_2 = 0) \\ y_1^+, y_2^0 &\longrightarrow F(y_1 > 0, y_2 = 0) \times f_1(y_1 | y_1 > 0) \\ y_1^0, y_2^+ &\longrightarrow F(y_1 = 0, y_2 > 0) \times f_2(y_2 | y_2 > 0) \\ y_1^+, y_2^+ &\longrightarrow F(y_1 > 0, y_2 > 0) \times f_{12}(y_1, y_2 | y_1 > 0, y_2 > 0), \end{aligned} \tag{1}$$

where F is a bivariate distribution defined over binary outcomes, f_1 and f_2 are univariate densities defined over positive, continuously distributed outcomes, and f_{12} is a bivariate density defined over a pair of positive, continuously distributed outcomes. We first describe 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.

¹There are some similarities in the framework of this paper and that of Bien et. al (2008), who use a bivariate hurdle for counts with an application to financial data.

3.1. Specification of f_j

Positive expenditures are specified according to the gamma density,

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

Note that $E(y_j|y_j > 0) = \eta_j\mu_j$, $j = 1, 2$ 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

$$G(y_{1t}, \dots, y_{mt}|\mathcal{I}_{t-1}) = C(G_1(y_{1t}|\mathcal{I}_{t-1}), \dots, G_m(y_{mt}|\mathcal{I}_{t-1}); \theta), \quad (3)$$

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, \dots, G_m , and the functional form of the copula.

The literature offers a vast array of copula functional forms 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) Clayton survival; (4) Frank. By changing the functional form of the copula, many different dependence patterns between marginal distributions can be explored, including both nonlinear and asymmetric tail dependence. Properties of these well established functional forms are discussed in the literature (Joe, 1997; Nelsen, 2006; Cherubini et al., 2004; Trivedi and Zimmer, 2007).²

Anticipating our results, we have found strong evidence that, in general, the best fit to the data is obtained using the Clayton copula. The bivariate Clayton (1978) copula takes the form

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \theta > 0 \quad (4)$$

where $u_j = G^{-1}(y_j | \mathcal{I}_{t-1})$ with the dependence parameter θ restricted to the region $(0, \infty)$. As θ approaches zero, the marginals become independent. As θ approaches infinity, the copula attains the Fréchet upper bound, but for no value does it attain the Fréchet lower bound. 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, we not surprisingly 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

²The Gaussian copula does not permit any tail dependence. The Clayton copula supports only positive dependence and lower tail dependence. The Clayton survival and Gumbel copulas are suitable for modeling positive upper tail dependence. Frank's copula captures both positive and negative dependence.

contemporaneous dependence appear to be robust across different copula specifications. The main benefit of the Clayton copula is its ability to capture lower tail dependence, which, as demonstrated below, is omnipresent in health care expenditures data.

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

$$F(y_1 > 0, y_2 > 0) = C(\Phi_1(\cdot), \Phi_2(\cdot); \theta_0) \quad (5)$$

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:

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

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,

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

and

$$f_{12}(y_1, y_2 | y_1 > 0, y_2 > 0) = c(F_1^+(\cdot), F_2^+(\cdot); \theta_+) \times f_1^+(\cdot) \times f_2^+(\cdot) \quad (6)$$

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.

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 $\Phi_1(\cdot | \mathcal{I}_{t-1}, \mathbf{x}_{it})$, $\Phi_2(\cdot | \mathcal{I}_{t-1}, \mathbf{x}_{it})$ we specify

$$\Pr(y_{1it} > 0) = \Phi(h_{01}(\{y_{k,it-j}\}) + \mathbf{x}'_{it}\boldsymbol{\beta}_{01} + \alpha_{01i}) \quad (7)$$

$$\Pr(y_{2it} > 0) = \Phi(h_{02}(\{y_{k,it-j}\}) + \mathbf{x}'_{it}\boldsymbol{\beta}_{02} + \alpha_{02i}), \quad (8)$$

$i = 1, \dots, N$; $t = 3, \dots, T$; $k = 1, 2$; $j = 1, 2, \dots, J$ 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

$$h_{0l}\{y_{k,it-j}\} = \sum_{k=1}^2 \sum_{j=1}^J \gamma_{lkj} \mathbf{1}(y_{k,it-j} > 0) + \sum_{k=1}^2 \sum_{j=1}^J \delta_{lkj} \ln(\max(y_{k,it-j}, 1)) \text{ for } l = 1, 2. \quad (9)$$

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 As we explain in greater detail below, the dynamics of this bivariate model are characterized by $(\gamma_{ljk}, \delta_{ljk}, j = 1, 2; k = 1, 2; l = 1, 2)$. The random intercepts are further specified as

$$\alpha_{0ji} = \bar{\mathbf{x}}'_i \lambda_{0j} + \sum_{k=1}^2 \tau_k \mathbf{1}(y_{k,i0} > 0) + \sum_{k=1}^2 \varsigma_k \ln(\max(y_{k,i0}, 1)) + \varepsilon_{0ki}; \quad k = 1, 2. \quad (10)$$

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} and, following Wooldridge (2005), we allow for the effects of initial conditions by specifying

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

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} . 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. The estimation of this model requires a further assumption about initial conditions. We follow Wooldridge's conditional maximum likelihood approach under the assumption that the initial conditions are nonrandom.

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

$$\mu_{1it} = \exp [h_{+1}(\{y_{k,it-j}\}) + \mathbf{x}'_{it}\boldsymbol{\beta}_{+1} + \alpha_{+1i}] \quad (11)$$

$$\mu_{2it} = \exp [h_{+2}(\{y_{k,it-j}\}) + \mathbf{x}'_{it}\boldsymbol{\beta}_{+2} + \alpha_{+2i}], \quad (12)$$

$i = 1, \dots, N$; $t = 3, \dots, T$; $k = 1, 2$; $j = 1, 2, \dots, J$ and $J < T$; in parallel to the specifications for the marginal distributions in the hurdle part of the model. Again, the functions h_{0k} are defined over the elements of the set $\{y_{k,it-j}\}$ which includes lagged outcomes and α_{0ki} 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

$$h_{+l}(\{y_{k,it-j}\}) = \sum_{k=1}^2 \sum_{j=1}^J \varphi_{+lkj} \mathbf{1}(y_{k,it-j} > 0) + \sum_{k=1}^2 \sum_{j=1}^J \phi_{+lkj} \ln(\max(y_{k,it-j}, 1)) \text{ for } l = 1, 2, \quad (13)$$

and

$$\alpha_{+ki} = \bar{\mathbf{x}}_i' \lambda_{+k} + \sum_{k=1}^2 \tau_{+k} \mathbf{1}(y_{k,i0} > 0) + \sum_{k=1}^2 \varsigma_{+k} \ln(\max(y_{k,i0}, 1)) + \varepsilon_{+ki}; \quad k = 1, 2. \quad (14)$$

As is typical in gamma regressions, the parameters η_1 and η_2 are specified as scalars.

4.3. Estimation and inference

As described in equation (1), 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

$$\begin{aligned} \ln L = & \sum_{0,0} [\ln(F(y_1 = 0, y_2 = 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0)] \\ & + \sum_{+,0} [\ln(F(y_1 > 0, y_2 = 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) + \ln(f_1(\cdot | \mathcal{I}_{t-1}, \mathbf{x}_{it}))] \\ & + \sum_{0,+} [\ln(F(y_1 = 0, y_2 > 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) + \ln(f_2(\cdot | \mathcal{I}_{t-1}, \mathbf{x}_{it}))] \\ & + \sum_{+,+} [\ln(F(y_1 > 0, y_2 > 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) + \ln f_{12}(y_1, y_2 | y_1 > 0, y_2 > 0, \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_+)]. \end{aligned} \quad (15)$$

Note that the log likelihood function contains two dependence parameters; θ_0 captures dependence between the probabilities of having any drug and nondrug expenditures. Similarly, the term θ_+ represents dependence between drug and nondrug expenditures when both are positive.

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

$$\begin{aligned} \ln L_1 = & \sum_{0,0} \ln(F(y_1 = 0, y_2 = 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) + \sum_{+,0} \ln(F(y_1 > 0, y_2 = 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) \quad (16) \\ & + \sum_{0,+} \ln(F(y_1 = 0, y_2 > 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) + \sum_{+,+} \ln(F(y_1 > 0, y_2 > 0), \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_0) \end{aligned}$$

and

$$\begin{aligned}\ln L_2 = & \sum_{+,0} \ln (f_1 (\cdot | \mathcal{I}_{t-1}, \mathbf{x}_{it})) + \sum_{0,+} \ln (f_2 (\cdot | \mathcal{I}_{t-1}, \mathbf{x}_{it})) \\ & + \sum_{+,+} \ln f_{12}(y_1, y_2 | y_1 > 0, y_2 > 0, \mathcal{I}_{t-1}, \mathbf{x}_{it}; \theta_+).\end{aligned}\quad (17)$$

$\ln L_1$ 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.

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.

We construct six subsamples of data for analysis as the full sample is likely too heterogeneous to be insightful. Each sample considers individuals 18 years of age and older as children are likely to have different health conditions and treatment protocols. 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.

Sample means for explanatory variables appear in Table 2. The samples exhibit differences in socioeconomic and health characteristics. 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, less 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

The specification of the mean of the marginal distributions, controlling for both the initial conditions and correlated random effects, was provided in the preceding section. We now discuss the covariates in greater detail.

First, 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, it captures dynamic dependence between drug and nondrug expenditures. Note that the model is not a simultaneous equations system in the traditional sense.

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 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 X also includes measures of age squared and an interaction between age and female and its square. Counting control variables in 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

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

⁵For some of the subsamples, the number of explanatory variables is less because some variables are

6. Results

We first report model selection criteria for choice between different copulas. Next, we describe the results of a number of specification tests to highlight the importance of a number of key specification 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 3 reports Bayes Information Criteria (BIC) statistics for several combinations of copulas. For each subsample, the Clayton copula provides a superior fit in both parts of the model, except for three models for continuous expenditures for which there is little discrimination across models. Therefore, all results presented and discussed below are based a version of equation (15) in which all copulas are specified as Clayton.

Parameter estimates from the bivariate hurdle model with Clayton copulas are reported in Tables 4-9. 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

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.

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 than discussing every own and cross effect in Tables 4-9, 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 elderly, arthritis, and heart condition 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 diabetes 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 or 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 4-9; this refers to the τ_j term in (10). 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 λ_j coefficients in (10) 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. 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 4-9, 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 are reported graphically in Figures 1 and 2. The figures illustrate the degree of lower tail dependence, formally defined as $\lim_{v \rightarrow 0^+} \frac{C(v,v)}{v}$. Informally, lower tail dependence is evident when events that occur with lower cumulative probabilities tend to occur together. Lower tail dependence is visually summarized by the extent of clustering in the lower left corners of Figures 1 and 2, and is stronger for the positive part than the hurdle part. Note that 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

implication from Figure 1 is 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, Figure 2 indicates that quarters of low drug spending also exhibit low nondrug spending.

6.2. Dynamic dependence and partial effects

In principle, the dynamic dependence and partial effects are functions of each of the lagged expenditure coefficients. However, the complexity of the model makes it impossible to fully understand these effects directly from coefficients. Therefore, we compute measures of effects that are analogous to the average partial effect proposed by Wooldridge (2005). We define the average partial effect (*APE*) on $y_{k,t}$ of the effect of $y_{j,t-1}$ as

$$APE_k(y_{j,t-1}) = E(y_{k,t}|y_{j,t-1}^{(1)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) - E(y_{k,t}|y_{j,t-1}^{(0)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \quad (18)$$

where $j, k = 1, 2$ and $y_{j,t-1}^{(1)}$ and $y_{j,t-1}^{(0)}$ 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). This measure is limited because it only captures the one-period impact on $y_{it}^{(2)}$ of the lagged change in binary-valued variable $y_{it-1}^{(1)}$.

In the context of the bivariate hurdle model, it is also insightful to examine the decomposition of *APE* into the effects on the probability or hurdle part of the model and the continuous outcome, conditional on it being positive. Thus we define

$$APE_k^0(y_{j,t-1}) = \Pr(y_{k,t} = 1|y_{j,t-1}^{(1)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) - \Pr(y_{k,t} = 1|y_{j,t-1}^{(0)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \quad (19)$$

as the partial effect on the probability of a positive outcome and

$$\begin{aligned} APE_k^+(y_{j,t-1}) &= E(y_{k,t}|y_{k,t} > 0, y_{j,t-1}^{(1)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \\ &\quad - E(y_{k,t}|y_{k,t} > 0, y_{j,t-1}^{(0)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \end{aligned} \quad (20)$$

as the partial effect conditional on the outcome being positive. Note that $\Pr(y_{k,t} = 1|y_{j,t-1}^{(m)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*)$ and $E(y_{k,t}|y_{k,t} > 0, y_{j,t-1}^{(m)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*)$ for $m = 0, 1$ are obtained directly from the marginal probit and gamma distributions respectively. The cal-

culation of $E(y_{k,t}|y_{j,t-1}^{(m)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*)$, $m = 0, 1$ involves terms from both hurdle and conditional parts of the model. Specifically

$$\begin{aligned} E(y_{k,t}|y_{j,t-1}^{(m)}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) &= \Pr(y_j = 0, y_k > 0) \times E_2(y_k|y_j = 0, y_k > 0) \\ &+ \Pr(y_j > 0, y_k > 0) \times E_k^+(y_k|y_j > 0, y_k > 0) \end{aligned}$$

In this paper, we calculate APE 's 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 calculate APE 's over the empirically observed values of $y_{j,t-1}^{(1)}$ (positive values of drug expenditure). These are reported in the 6 panels of Figure 3; APE^0 , APE^+ and APE reading from left to right. Although we display the APE 's over the entire range of $y_{j,t-1}^{(1)}$ we believe they are most reliable in the interior of the range of observations, e.g., between the 25th and 75th percentiles of observed values.

The vertical lines in the graphs mark the 25th and 75th percentiles of positive drug expenditures in the data. Within this range, estimated APE , shown in the rightmost panels, are negative for five of the six subsamples. For all six samples, the magnitude of APE decreases from the 25th to the 75th percentile. Taking the 65 and older sample as an example, previous quarter drug spending at the 25th percentile (approximately \$40) is associated with a current quarter reduction in nondrug spending of approximately \$50. Similarly previous quarter drug spending at the 75th percentile (approximately \$200) translates to a current quarter reduction in nondrug spending of approximately \$10. Estimates of APE within the 25th to 75th percentile range for the other samples are as follows: continuously insured: $-\$25$ to $-\$5$, diabetes: $-\$90$ to $-\$30$, arthritis: $-\$60$ to $-\$10$, heart condition: $-\$70$ to $-\$10$. For the mental illness sample, the change in nondrug spending switches from approximately $-\$20$ at the 25th percentile to $+\$10$ at the 75th percentile.

Estimates of APE become smaller as previous quarter drug spending increases, primarily because estimates of APE^0 , shown in the leftmost panels, are positive and increasing between the 25th and 75th percentiles of drug spending. For all six samples, previous quarter drug spending at the 25th percentile is associated with an approximate 1 percentage point increase in the probability of positive current period nondrug spending. On the other hand, previous quarter drug spending at the 75th percentile translates to an approximate

4 or 5 percentage point increase in the probability of positive current quarter nondrug spending.

For all six samples, estimates of APE^+ , which appear in the middle panels, are negative between the 25th and 75th percentiles of previous quarter drug spending. For the diabetes and heart condition samples, estimated APE^+ become larger in magnitude (more negative) as previous quarter drug spending increases. For the other four samples, the estimated APE^+ become smaller in magnitude as previous quarter drug spending increases.

The overall APE estimates suggest modest cost-offsets in nondrug spending in the quarter following an increase in drug expenditures. The only instance in which there is no cost-offset is among those with mental illnesses who experience relatively large increases in previous quarter drug spending. 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.

6.3. An alternative measure of cost-offset

The APE 's defined above estimate partial effects that are “marginal” over the distribution of current drug expenditures. 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

$$\begin{aligned} CAPE_k(y_{j,t-1}) &= E(y_{k,t}|y_{j,t-1}^{(1)}, y_{j,t}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \\ &\quad - E(y_{k,t}|y_{j,t-1}^{(0)}, y_{j,t}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*); \end{aligned} \quad (21)$$

where $j, k = 1, 2$ and $j \neq k$ and $y_{j,t-1}^{(1)}$ and $y_{j,t-1}^{(0)}$ denote values of $y_{j,t-1}$ over which the partial effect is desired. The key difference between APE described by equation (18) and $CAPE$ described by equation (21) 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 $y_{j,t}$. Analogous to APE^0 and APE^+ , we also define $CAPE^0$ and $CAPE^+$, each of which conditions additionally on

$y_{j,t}$, as

$$CAPE_k^0(y_{j,t-1}) = \Pr(y_{k,t} = 1|y_{j,t-1}^{(1)}, y_{j,t}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) - \Pr(y_{k,t} = 1|y_{j,t-1}^{(0)}, y_{j,t}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \quad (22)$$

and

$$\begin{aligned} APE_k^+(y_{j,t-1}) &= E(y_{k,t}|y_{k,t} > 0, y_{j,t-1}^{(1)}, y_{j,t}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*) \\ &\quad - E(y_{k,t}|y_{k,t} > 0, y_{j,t-1}^{(0)}, y_{j,t}, y_{-k,t-1}^*, \mathbf{y}_{t-2}^*, \mathbf{x}^*). \end{aligned} \quad (23)$$

Calculation of the conditional (on $y_{j,t}$) expectations is considerably more complicated than the unconditional expectations needed for the calculation of the APE 's. For the hurdle probabilities,

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

and

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

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,

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

which is computed using numerical integration.

$CAPE^0$ estimates, presented in the leftmost panels, suggest that conditional on positive current quarter drug spending, previous quarter drug spending between the 25th and 75th percentiles is associated with an approximate 1 percentage point reduction in the probability of current period nondrug spending. In contrast, when conditioned on zero current quarter drug spending, $CAPE^0$ estimates are positive.

$CAPE^+$ estimates, shown in the middle panels, are negative regardless of the conditioning value of current period drug spending. However, estimates are larger in magnitude (more negative) when conditioning on the 75th percentile of current quarter drug spending, compared to the 25th percentile.

Finally, overall *CAPE* estimates, shown in the rightmost panels, suggest that while the *existence* of dynamic cost-offsets is robust to conditioning on present quarter drug spending, the *magnitudes* of cost-offsets depend on the amount of present period drug spending. When conditioning on positive current period drug spending (whether at the 25th or 75th percentiles), cost-offsets are larger than dollar-for-dollar at the median of the distribution of drug spending, with the exception of the mental illness sample. Cost offsets become smaller than dollar-for-dollar as previous quarter drug spending becomes larger. Furthermore, conditioning on larger current quarter drug spending produces stronger evidence of larger-than-dollar-for-dollar cost offsets.

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 of drug and non-drug spending. 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 for median values of previous quarter drug spending, cost-offsets are larger than dollar-for-dollar for reasonably large current period drug spending (i.e., above the 25th percentile of

drug spending). However, cost-offsets are smaller than dollar-for-dollar as previous quarter drug spending become larger.

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: (1) current quarter drug spending is reasonably large and (2) previous quarter drug spending is not too large. It seems difficult to implement policies based on these conditions, as spending for certain drugs might be highly unexpected, and because new drug development and changing demographics will probably alter the distribution of drug spending in the future.

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 Bulletin of Economics and Statistics*, forthcoming.

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., E. Luciano, 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.

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

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.; Philipson, T. 2007. Integrated insurance design in the presence of multiple medical technologies. *American Economic Review*, 97(2):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(14): 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 Mullaly, 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.; Murray, L.A. 2001. Growing differences between Medicare beneficiaries with andwithout drug coverage. *Health Affairs*, 20(2):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.

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, non-linear 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

Statistic	Drug expenditure			Nondrug expenditure		
	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
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	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	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: Maximized log likelihoods for models with alternative copulas

Copula	Subsample	Hurdle lnL	Conditional lnL	Overall lnL
Models with Clayton copula				
Clayton	65 and older	-81341*	-594359	-675700*
	Continuously insured	-280333*	-1425303*	-1705636*
	Diabetes	-45777*	-347241	-393018*
	Mental illness	-81647*	-539333*	-620981*
	Arthritis	-96854*	-689050	-785904*
	Heart condition	-131270*	-916224*	-1047494*
Models with Survival Clayton copula				
Survival Clayton	65 and older	-81934	-594359	-676293
	Continuously insured	-283909	-1425319	-1709228
	Diabetes	-46009	-347249	-393258
	Mental illness	-82470	-539352	-621822
	Arthritis	-97865	-689041*	-786907
	Heart condition	-131991	-916232	-1048223
Models with Frank copula				
Frank	65 and older	-81531	-594357*	-675888
	Continuously insured	-281689	-1425343	-1707032
	Diabetes	-45822	-347237*	-393059
	Mental illness	-81887	-539344	-621231
	Arthritis	-97167	-689057	-786224
	Heart condition	-131411	-916235	-1047646
Models with Gaussian copula				
Gaussian	65 and older	-81527	-594369	-675896
	Continuously insured	-281123	-1425348	-1706471
	Diabetes	-45837	-347246	-393084
	Mental illness	-81864	-539357	-621221
	Arthritis	-97176	-689061	-786237
	Heart condition	-131448	-916243	-1047691

* denotes model with best fit for given subsample

Table 4: Bivariate two-part model: coefficients of lagged variables
 Sample: ages 65 and older

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

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table 5: Bivariate two-part model: coefficients of lagged variables
 Sample: continuously insured - medical and prescription drug

	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 (0.041)	-0.161** (0.047)	-0.215** (0.062)	-0.084 (0.093)
$1(\text{nondrug}_{t-1} > 0)$	-0.052 (0.033)	0.030 (0.038)	-0.057 (0.055)	-1.023** (0.086)
$\ln(\text{drug}_{t-1})$	0.007 (0.008)	0.047** (0.010)	0.049** (0.012)	0.004 (0.019)
$\ln(\text{nondrug}_{t-1})$	0.021** (0.005)	0.074** (0.006)	0.015 (0.008)	0.173** (0.011)
$1(\text{drug}_{t-2} > 0)$	0.006 (0.041)	-0.052 (0.047)	-0.169** (0.063)	-0.005 (0.098)
$1(\text{nondrug}_{t-2} > 0)$	0.031 (0.034)	0.169** (0.038)	-0.091 (0.055)	-0.707** (0.084)
$\ln(\text{drug}_{t-2})$	0.020* (0.008)	0.026** (0.009)	0.042** (0.012)	-0.005 (0.019)
$\ln(\text{nondrug}_{t-2})$	0.010 (0.005)	0.032** (0.006)	0.016* (0.008)	0.100** (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 $\alpha_{0j} = \alpha_j$	—	—	1.78	931.3**
$\theta_0; \theta_+$		1.062 (0.028)		0.195 (0.031)
Kendall's Tau		0.347 (0.006)		0.0890 (0.013)
lnL		-45777		-347241
N		42702		31680

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table 6: 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
$1(\text{drug}_{t-1} > 0)$	0.089** (0.032)	-0.117** (0.036)	-0.149** (0.050)	-0.046 (0.072)
$1(\text{nondrug}_{t-1} > 0)$	0.002 (0.025)	0.105** (0.028)	-0.096* (0.043)	-0.910** (0.062)
$\ln(\text{drug}_{t-1})$	0.003 (0.007)	0.038** (0.008)	0.032** (0.010)	-0.001 (0.015)
$\ln(\text{nondrug}_{t-1})$	0.022** (0.004)	0.066** (0.004)	0.022** (0.006)	0.166** (0.008)
$1(\text{drug}_{t-2} > 0)$	0.009 (0.032)	0.040 (0.036)	-0.180** (0.051)	0.023 (0.074)
$1(\text{nondrug}_{t-2} > 0)$	0.091** (0.026)	0.274** (0.028)	0.022 (0.044)	-0.458** (0.064)
$\ln(\text{drug}_{t-2})$	0.023** (0.007)	0.004 (0.008)	0.057** (0.010)	-0.004 (0.015)
$\ln(\text{nondrug}_{t-2})$	0.004 (0.004)	0.020** (0.004)	-0.009 (0.006)	0.065** (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 $\alpha_{0j} = \alpha_j$	—	—	10.7**	1811**
$\theta_0; \theta_+$	0.997 (0.020)		0.224 (0.022)	
Kendall's Tau	0.333 (0.004)		0.101 (0.009)	
lnL	-81341		-594359	
N	78162		55848	

Robust standard errors in parentheses

** p<0.01, * p<0.05

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

	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 $\alpha_{0j} = \alpha_j$	—	—	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

Table 8: 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
$1(\text{drug}_{t-1} > 0)$	0.001 (0.031)	-0.229** (0.034)	-0.145** (0.050)	0.051 (0.073)
$1(\text{nondrug}_{t-1} > 0)$	0.004 (0.026)	0.020 (0.029)	-0.174** (0.044)	-1.106** (0.067)
$\ln(\text{drug}_{t-1})$	0.029** (0.006)	0.062** (0.007)	0.032** (0.010)	-0.021 (0.015)
$\ln(\text{nondrug}_{t-1})$	0.018** (0.004)	0.085** (0.005)	0.037** (0.007)	0.206** (0.009)
$1(\text{drug}_{t-2} > 0)$	-0.012 (0.030)	-0.100** (0.033)	-0.218** (0.050)	0.074 (0.071)
$1(\text{nondrug}_{t-2} > 0)$	0.033 (0.027)	0.176** (0.029)	0.003 (0.047)	-0.532** (0.067)
$\ln(\text{drug}_{t-2})$	0.025** (0.006)	0.029** (0.007)	0.061** (0.010)	-0.026 (0.014)
$\ln(\text{nondrug}_{t-2})$	0.010* (0.004)	0.026** (0.005)	0.004 (0.007)	0.086** (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 $\alpha_{0j} = \alpha_j$	—	—	9.13**	1293**
$\theta_0; \theta_+$		1.267 (0.021)		0.246 (0.023)
Kendall's Tau		0.388 (0.004)		0.110 (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: heart condition

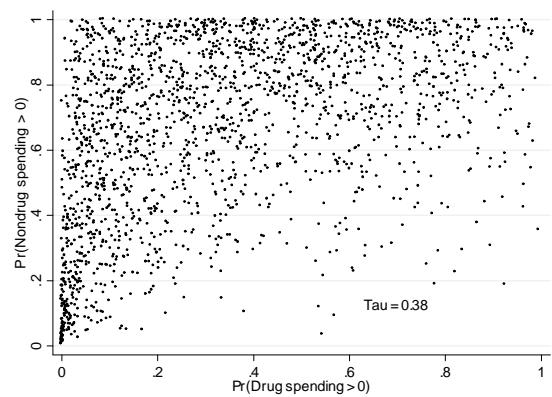
	Hurdle part		Positive spending part	
	$1(\text{drug}_t > 0)$	$1(\text{nondrug}_t > 0)$	drug_t	nondrug_t
$1(\text{drug}_{t-1} > 0)$	-0.007 (0.020)	-0.171** (0.020)	-0.228** (0.034)	-0.020 (0.053)
$1(\text{nondrug}_{t-1} > 0)$	0.077** (0.016)	0.015 (0.016)	-0.003 (0.029)	-0.995** (0.045)
$\ln(\text{drug}_{t-1})$	0.027** (0.004)	0.053** (0.004)	0.059** (0.007)	-0.015 (0.011)
$\ln(\text{nondrug}_{t-1})$	0.013** (0.003)	0.078** (0.003)	0.013** (0.005)	0.213** (0.006)
$1(\text{drug}_{t-2} > 0)$	-0.008 (0.019)	-0.082** (0.019)	-0.293** (0.035)	0.016 (0.052)
$1(\text{nondrug}_{t-2} > 0)$	0.122** (0.016)	0.249** (0.016)	0.030 (0.031)	-0.395** (0.046)
$\ln(\text{drug}_{t-2})$	0.026** (0.004)	0.031** (0.004)	0.079** (0.007)	-0.007 (0.011)
$\ln(\text{nondrug}_{t-2})$	-0.002 (0.003)	0.005 (0.003)	-0.000 (0.005)	0.075** (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 $\alpha_{0j} = \alpha_j$	—	—	11.5**	3431**
$\theta_0; \theta_+$		1.155 (0.010)		0.204 (0.017)
Kendall's Tau		0.366 (0.002)		0.093 (0.007)
lnL		-280333		-1425303
N		289374		139969

Robust standard errors in parentheses

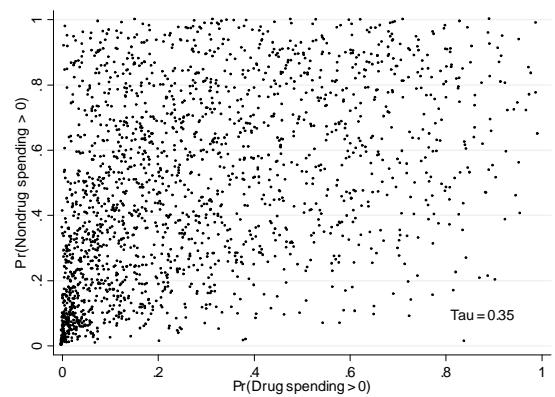
** p<0.01, * p<0.05

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

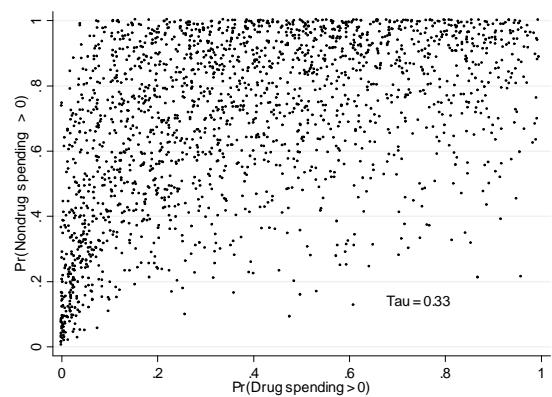
Elderly



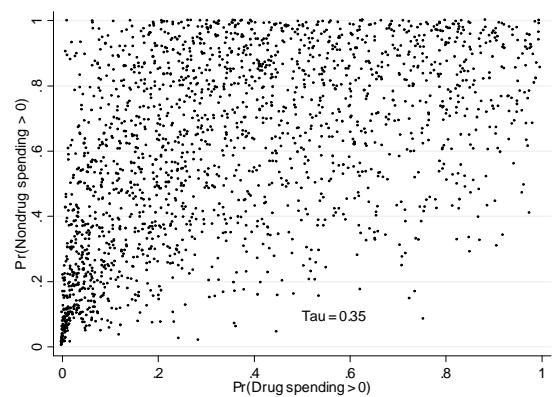
Continuously insured



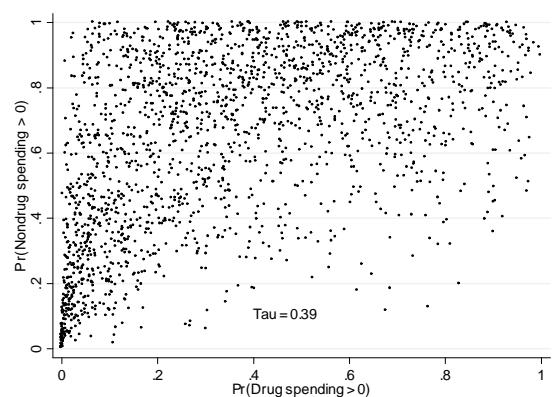
Diabetes



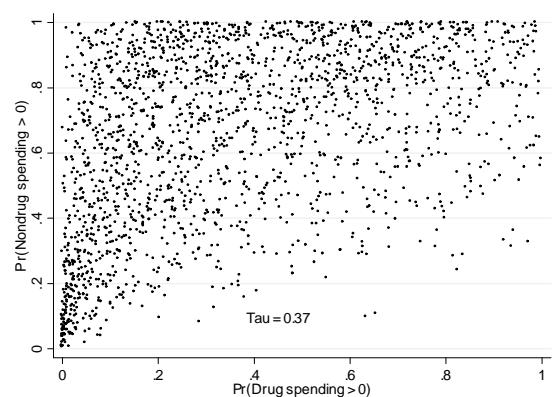
Mental Illness



Arthritis



Heart Condition



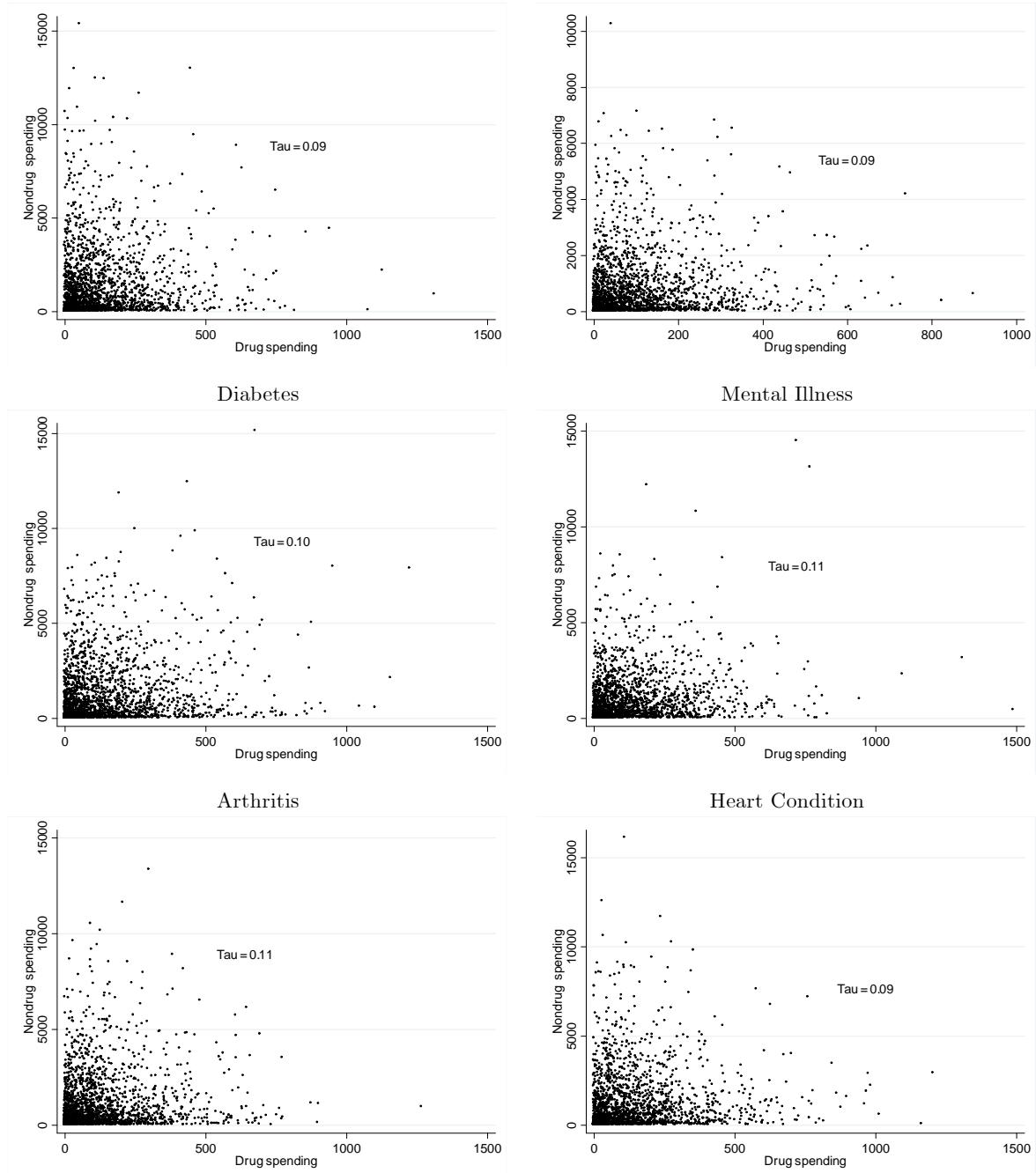
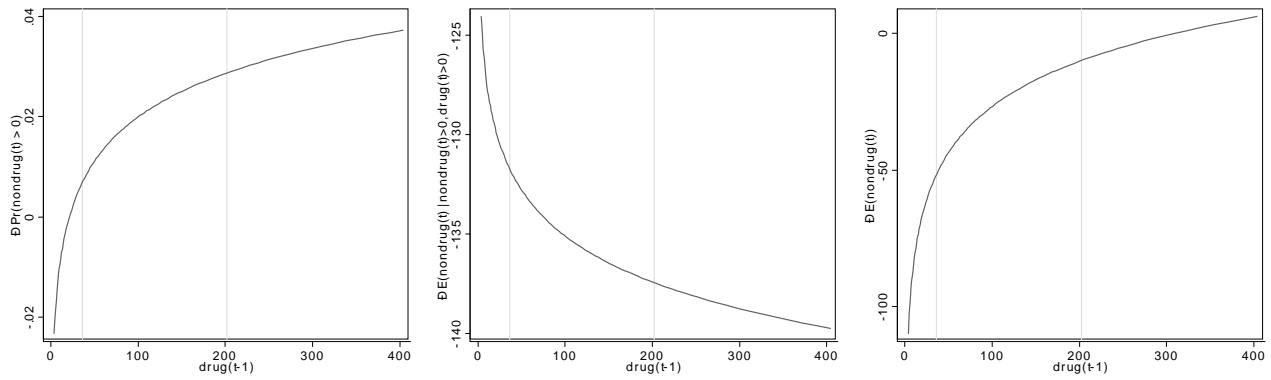


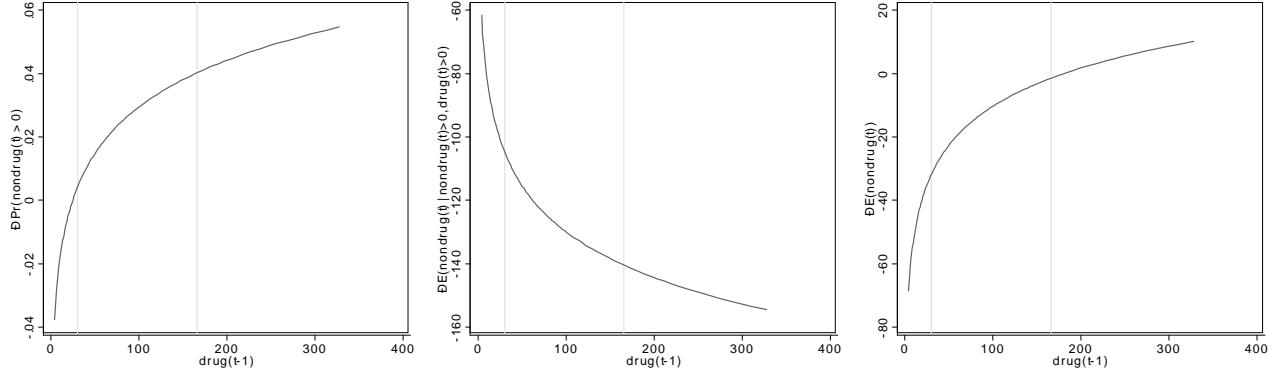
Figure 3: Average Partial Effects

Average Partial Effects: Sample of persons 65 years and older



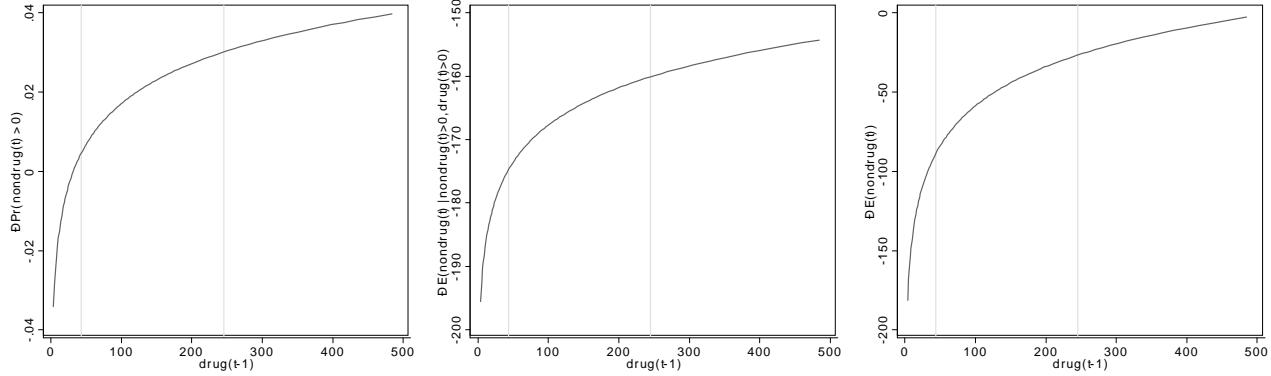
Δ denotes difference from the probability or expectation evaluated at $\text{drug}(t-1) = 0$
 vertical lines represent 25th and 75th percentiles of positive drug expenditure

Average Partial Effects: Sample of persons continuously covered by medical and Rx insurance



Δ denotes difference from the probability or expectation evaluated at $\text{drug}(t-1) = 0$
 vertical lines represent 25th and 75th percentiles of positive drug expenditure

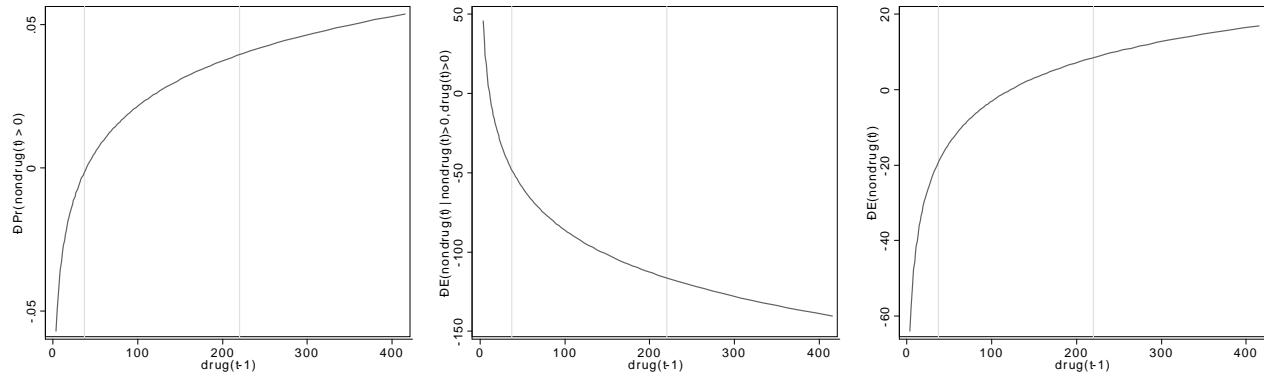
Average Partial Effects: Sample of persons with diabetes



Δ denotes difference from the probability or expectation evaluated at $\text{drug}(t-1) = 0$
 vertical lines represent 25th and 75th percentiles of positive drug expenditure

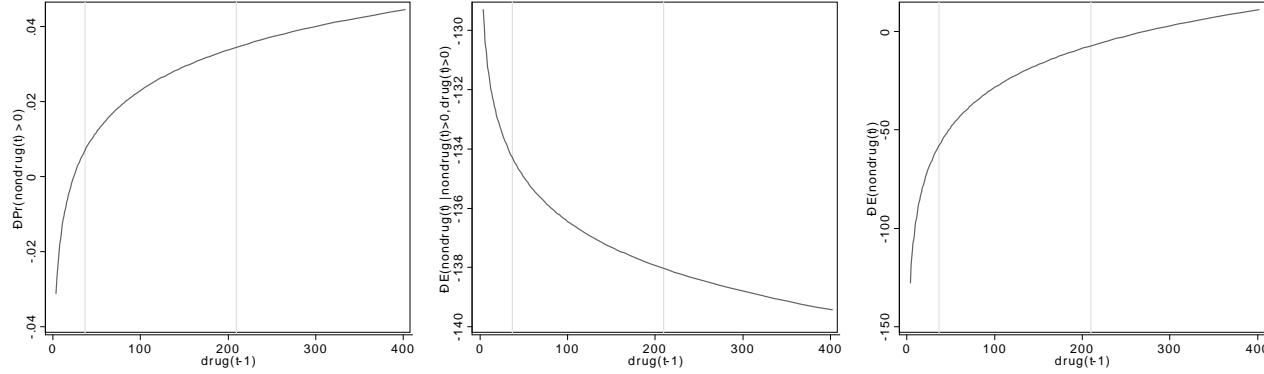
Figure 3: Average Partial Effects (cont.)

Average Partial Effects: Sample of persons with a mental illness



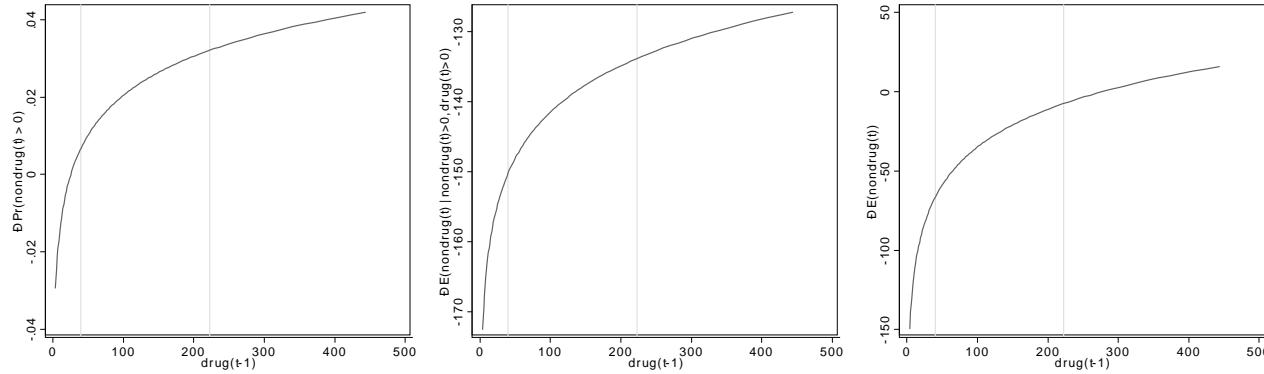
Δ denotes difference from the probability or expectation evaluated at drug(t-1)=0
vertical lines represent 25th and 75th percentiles of positive drug expenditure

Average Partial Effects: Sample of persons with arthritis



Δ denotes difference from the probability or expectation evaluated at drug(t-1)=0
vertical lines represent 25th and 75th percentiles of positive drug expenditure

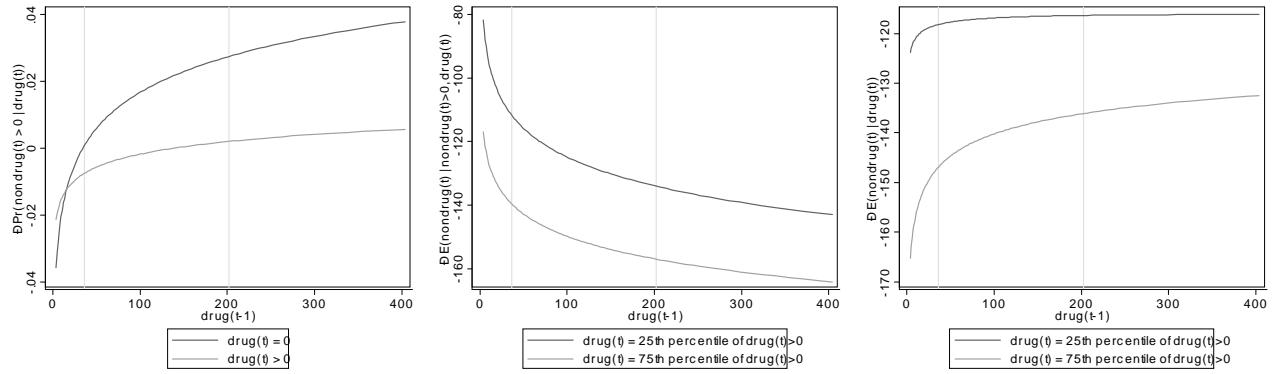
Average Partial Effects: Sample of persons with a heart condition



Δ denotes difference from the probability or expectation evaluated at drug(t-1)=0
vertical lines represent 25th and 75th percentiles of positive drug expenditure

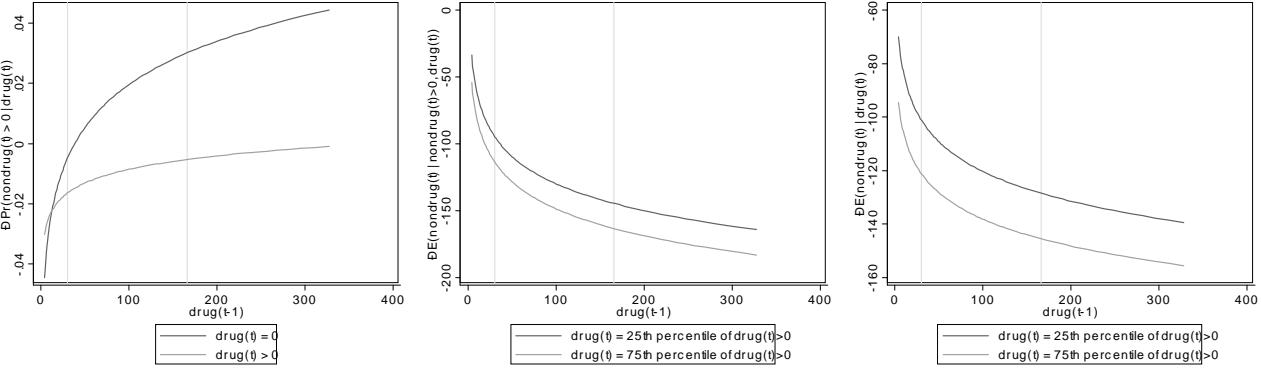
Figure 4: Conditional Average Partial Effects

Conditional Average Partial Effects: Sample of persons 65 years and older



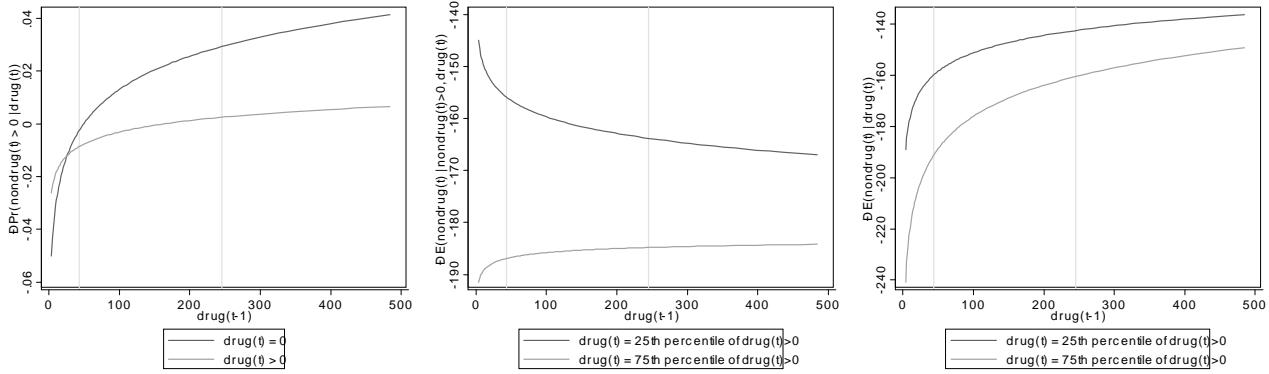
Δ denotes difference from the probability or expectation evaluated at drug(t-1)=0
vertical lines represent 25th and 75th percentiles of positive drug expenditure

Conditional Average Partial Effects: Sample of persons continuously covered by medical and Rx insurance



Δ denotes difference from the probability or expectation evaluated at drug(t-1)=0
vertical lines represent 25th and 75th percentiles of positive drug expenditure

Conditional Average Partial Effects: Sample of persons with diabetes



Δ denotes difference from the probability or expectation evaluated at drug(t-1)=0
vertical lines represent 25th and 75th percentiles of positive drug expenditure

Figure 4: Conditional Average Partial Effects (cont.)

