Did the ACA's "Guaranteed Issue" Provision Cause Adverse Selection into Nongroup Insurance? Analysis using a Copula-Based Hurdle Model

Giampiero Marra
Department of Statistical Science
University College London

Rosalba Radice
Cass Business School
City, University of London

David Zimmer
Department of Economics
Western Kentucky University

January 25, 2019

Abstract

Prior to the ACA, insurance companies could charge higher premiums – or outright denying coverage – to people with preexisting health problems. But the ACA's "guaranteed issue" provision forbids such price discrimination and denials of coverage. This paper seeks to determine whether, after implementation of the ACA, nongroup private insurance plans, have experienced adverse selection. Our empirical approach employs a copula-based hurdle regression model, with dependence modeled as a function of dimensions along which adverse selection might occur. Our main finding is that, after implementation of the ACA, nongroup insurance enrollees with preexisting health problems do not appear to exhibit adverse selection. This finding suggests that the ACA's mandate that everyone acquire coverage might have attracted enough healthy enrollees to offset any adverse selection.

JEL Classification: I13; C31

Keywords: community rating; copula, regression spline; partial effects; mixed data; hurdle model

1 Introduction

The Affordable Care Act (ACA), passed by the U.S. Congress and signed by President Obama in 2010, placed strict limitations on the types of information that private insurance companies may use to set insurance premiums. In effect, those so-called “community rating” restrictions permit insurance companies to adjust premiums along only the following dimensions: age, location, smoking status, and single/family. Insurance companies may not adjust premiums along other dimensions that might correlate with medical risk, such as gender, race, and preexisting health conditions.

The most noteworthy aspect of community rating, from an actuarial perspective, concerns preexisting health conditions. Prior to the ACA, insurance companies could charge higher premiums – including prices near infinity, effectively denying coverage – to people with preexisting health problems. But the ACA’s “guaranteed issue” provision, as it is known, forbids such price discrimination with respect to preexisting health problems. This detail of the ACA has raised concerns about adverse selection, in which people with preexisting health problems might be most likely to obtain coverage. Such adverse selection would result in a sicker pool of enrollees, eventually requiring high premiums to coverage medical risks.

To reduce the possibility of adverse selection, the ACA includes what has become its most famous, and hotly debated, provision: A mandate that everyone must acquire health insurance. For people without access to employer-based group coverage, the ACA established state- and federally-operated marketplaces, often called “exchanges,” from which people can shop for ACA-compliant plans offered by private insurance companies that participate in the exchanges. Yet despite the mandate, concerns about adverse selection remain, in part because, for many people, penalties for noncompliance fall below prices of insurance through the exchanges.

This paper seeks to determine whether, after implementation of the ACA, nongroup private insurance plans, whether obtained through the exchanges or elsewhere, have experienced adverse selection along dimensions that insurance companies are prohibited from using for premium adjustments, particularly preexisting conditions. It is not clear, *a priori*, whether such adverse selection exists. On one hand, the incentives for sick people who previously might have been denied coverage to now seek it are obvious. But on the other hand, the mandate, if effective, should draw previously uninsured *healthy* people into nongroup coverage, therefore mitigating, or even reversing, adverse selection. While plans offered through the exchanges have experienced highly-publicized year-to-year hikes in premiums, those price increases, by themselves, do not offer proof of adverse selection. To the contrary, those premium adjustments could be responses to incorrect calculations of price elasticities of medical care for those obtaining coverage through the exchanges.

Because the ACA still is a relatively new law, not many studies have investigated this topic. Among the few, Sacks (2018), in a mostly descriptive piece, argues that "it does not appear that large numbers of healthy people are exiting the Marketplaces." By contrast, Panhans (2017) reaches the opposite conclusion using exogenous premium variations in Colorado. A related strand of literature investigates the Massachusetts health insurance reform, which was enacted in 2006 and served as the template for the ACA. That literature appears to point toward *favorable* selection following the Massachusetts reform, which speaks to the effectiveness of the insurance mandate included in the Massachusetts law (Hackmann, Kolstad, and Kowalski, 2015).

In seeking to detect the presence of adverse selection, one approach is to estimate a regression of medical spending on insurance, and then sequentially add categories of control variables to determine how those controls alter the link between insurance and medical spending (Fang, Keane, and Silverman, 2008). While simple, such an exercise depends crucially on the sequence in which those controls are added, with different sequences potentially yielding conflicting findings (Gelbach, 2016).

Instead, we adopt an alternative method. Our approach jointly models nongroup insurance enrollment and medical spending using a bivariate copula. The control variables in the two marginals include only those dimensions along which the ACA permits premium adjustments. The dependence parameter from that copula blends the (causal **can we drop this word?, actually can we safely drop it from everywhere in the text?**) effect of insurance on spending and any possible selection effect. We then give the dependence parameter a regression structure based on characteristics that insurance companies are *not* permitted to use in setting premiums. The coefficients attached to those characteristics in the dependence parameter inform upon selection effects along those dimensions. Not only does this approach avoid the sequentially-added covariate problem highlighted by Gelbach (2016), but it also explicitly handles the commonly-noted, but seldomly-addressed, concern that selection effects, if present, likely occur along multiple dimensions, with potentially different magnitudes and directions (de Meza and Webb, 2001).

Our approach allows for several other flexibilities. First, because individual-level medical spending tends to show (i) large proportions of zeros and (ii) highly skewed distributional shapes among positive spenders, we adopt a semi-continuous hurdle setup in which one part models whether a person has any spending, and another part handles spending patterns among positive spenders (Shi and Zhang, 2015). That decoupling allows each part of the hurdle setup to display a separate pattern of selection into insurance. Second, to limit the possibility that parametric assumptions contaminate our main findings, we permit a wide range of flexibility regarding functional forms, while adopting a largely data-driven process for model construction. We also allow for potentially highly nonlinear covariate/response relationships by using spline-based regression functions. Those flexibilities allow us to capture the spirit of nonparametric flexibility, while still harnessing the practical benefits of parametric setups.

In summary, beyond the health policy focus of this paper, we extend somewhat the procedures discussed in Deb, Trivedi, and Zimmer (2014) and Shi and Zhang (2015) to make it possible to simultaneously model binary and semi-continuous data, capture heterogeneous dependencies between the two outcomes, allow higher moments of the distributions to depend on covariate information, and account for non-linearities of covariate effects. Computationally, we implement our approach by building upon the already-available `R` (R Core Team, 2019) package `GJRM` (Marra and Radice, 2019) which has been created to enhance reproducible research as well as with transparent and straightforward dissemination of results in mind.

Our main finding is that, after implementation of the ACA, nongroup insurance enrollees with preexisting health problems do *not* appear to exhibit adverse selection. In fact, we find some evidence, albeit only marginally significant, of *favorable* selection with respect to preexisting conditions. Therefore, despite that the law is still relatively young, and the political climate regarding its support remains somewhat in flux, we conclude that fears of adverse selection seem to have been misplaced.

2 Data

Data used in this study come from the Medical Expenditure Panel Survey (MEPS), collected and published by the Agency for Healthcare Research and Quality, a unit of the U.S. Department of Health and Human Services. The MEPS provides nationally-representative, micro-level information on medical spending, insurance status, and health conditions. We focus on the 2015 and 2016 waves of the survey, because those are the most recent waves publicly available at the time of this writing, and also because most of the ACA statutes relevant to this study were fully active by 2015.

To isolate the nongroup private insurance market of interest, the estimation sample focuses on non-elderly adults (18-64) who were never enrolled in Medicaid or Medicare, the two main publicly-operated insurance options in the U.S. The sample also eliminates subjects enrolled in group-based insurance, either through an employer or some other organization, because selection effects for group-based coverage tend to be dominated by selection into employment. Furthermore, even before the ACA, group-based plans rarely engaged in the sorts of denials of coverage central to this study. The final estimation sample includes $n = 6,014$ unique persons.

The main measure of medical spending is the sum of expenses – both out-of-pocket and insurer-reimbursed – for all office-based services during the calendar year. (Almost all non-emergency care in the U.S. requires some form of office-based contact with the medical system, so office-based services provide a broad gauge of medical usage.) The main measure of private insurance coverage is a simple binary indicator for whether the person ever had private nongroup insurance during the calendar year. Based on that measure, approximately 26 percent of the estimation sample had nongroup insurance, while 74 percent lacked any form of coverage, which speaks to the challenges the ACA has confronted in extending coverage to this historically hard-to-insure population.

Table 1 attempts to shed light on the awkward distributional shape of annual office-based spending (hereafter referred to as "spending"). First, a relatively large proportion of subjects report zero spending, especially among those lacking insurance. Second, among subjects with positive spending, the distribution shows high skewness, as evidenced by the difference between means and medians. It is those distributional quirks that our copula-based hurdle model, discussed below, seeks to address. Another noteworthy feature in Table 1 is the extent to which subjects with insurance appear to out-spend their uninsured counterparts.

Table 2 reports sample means for socioeconomic characteristics. To explore the extent to which those characteristics correlate with medical spending, the table partitions the means according to whether subjects have any spending. The top panel shows dimensions along which insurance companies are to permitted to (at least partially) adjust premiums. (Note that, for confidentiality concerns, the MEPS does not release publicly-available measures of location finer than the four broad census regions.)

More interesting to this study, the bottom panel of Table 2 shows dimensions along which insurance companies may *not* adjust premiums, even though those dimensions are widely recognized to correlate with spending. For example, females spend more than males, while blacks and Hispanics spend less than their nonblack/nonHispanic counterparts. But insurance companies must pretend to be blind to gender and race/ethnicity. Similarly positive spenders appear to have higher body mass indices (BMI) than their zero-spending counterparts.

Most striking, and perhaps least surprising, is that positive spenders report more chronic health conditions, which are typically long-lasting ailments that require ongoing medical treatment. (To calculate that measure, we sum binary measures for whether the subject has ever received a diagnosis of: a physical limitation, high blood pressure, heart disease, heart attack, stroke, high cholesterol, cancer, diabetes, arthritis, and asthma.)

3 Methodology

This section describes the adopted model (specifically, the main building blocks that make it up), parameter estimation and inference.

3.1 The model

Consider two random variables (Y_{1i}, Y_{2i}), for $i = 1, \ldots, n$, where $Y_{1i} \in \{0, 1\}$, $Y_{2i} \in \{0, \infty\}$, and n represents the sample size. Variable Y_{1i} indicates whether the person has nongroup insurance, whereas Y_{2i} denotes medical spending. Using a parametric copula $C : (0,1)^2 \to (0,1)$ the joint cumulative distribution function (cdf) of the two variables could be expressed as (e.g., Sklar, 1973; Marra and Radice, 2017; Radice, Marra and Wojtys, 2016)

$$F(y_{1i}, y_{2i}) = C(F_1(y_{1i}), F_2(y_{2i}); \theta_i), \tag{1}$$

where $F_1(y_{1i})$ and $F_2(y_{2i})$ are cdfs of the marginals of Y_{1i} and Y_{2i}, taking values in $(0, 1)$, and the association parameter θ_i describes the dependence between Y_{i1} and Y_{i2} after covariate effects at the marginal level are taken into account. Note that, as it will be made clear in the next sections, the marginal cdfs depend on distributional parameters which are in turn linked to covariates and coefficients; however, to avoid cluttering the notation we have suppressed this dependence in the notation.

Because we limit the covariates in the marginals to characteristics insurance companies may use to adjust premiums – age, location, smoking status, and single/family – remaining traits, such as preexisting conditions, become absorbed into the dependence term. Consequently, that dependence term, which appears to be positive in our data, blends together two economic phenomena: (1) the direct causal effect of insurance on spending, sometimes called "moral hazard," and (2) the indirect effect of unobserved (to the insurance company) attributes that simultaneously correlate with both insurance enrollment and medical spending. Health economists refer to that indirect effect as "adverse selection" if it makes the dependence term more positive. Being a scalar-valued parameter, the dependence parameter does not allow us to untangle moral hazard and adverse selection, nor is that the explicit goal of this paper.

However, as described in greater detail below, we can specify the dependence term as a function of unobserved (to the insurance company) attributes, in order to determine how those attributes contribute to adverse selection. To see that logic, note that the moral hazard effect is purely a consequence of insurance lowering the point-of-service price of medical services; it has nothing to do with person-specific attributes. Therefore, if, say, preexisting conditions cause the dependence term to increase, implying simultaneous increases in the probability of insurance enrollment *and* medical spending, then such a finding would offer evidence of adverse selection with respect to preexisting conditions.

Four configurations of outcomes are possible: $y_{1i} = 0, y_{2i} = 0$ (denoted by $(y_{1i}^0; y_{2i}^0)$ in what follows), $y_{1i} = 0, y_{2i} > 0$ $(y_{1i}^0; y_{2i}^+)$, $y_{1i} = 1, y_{2i} = 0$ $(y_{1i}^1; y_{2i}^0)$ and $y_{1i} = 1, y_{2i} > 0$ $(y_{1i}^1; y_{2i}^+)$, and each maps to a data distribution given by a product of a bivariate hurdle probability and a density for the positive outcomes. The joint probability mass function for the hurdle part can be described as

$$\begin{aligned} F^h(y_{1i}^1, y_{2i}^+) &= C^h(F_1^h(y_{1i}^1), F_2^h(y_{2i}^+); \theta_i^h) \\ F^h(y_{1i}^0, y_{2i}^0) &= 1 - F_1^h(y_{1i}^1) - F_2^h(y_{2i}^+) + C^h(F^h(y_{1i}^1), F^h(y_{2i}^+); \theta_i^h) \\ F^h(y_{1i}^0, y_{2i}^+) &= F_2^h(y_{2i}^+) - C^h(F_1^h(y_{1i}^1), F_2^h(y_{2i}^+); \theta_i^h) \\ F^h(y_{1i}^1, y_{2i}^0) &= F_1^h(y_{1i}^1) - C^h(F_1^h(y_{1i}^1), F_2^h(y_{2i}^+); \theta_i^h) \end{aligned} . \quad (2)$$

The term F^h is the joint probability mass function defined for the pair of binary random variables in the hurdle part, and F_1^h and F_2^h are the cdfs for the two binary outcomes, nongroup insurance and positive medical spending. The function C^h, joining the marginals, is a parametric copula with dependence parameter θ_i^h.

For subjects with positive spending, the second part of the hurdle model examines the relationship between nongroup insurance and the amount of medical expenses given positive spending. Thus, the second part of the hurdle setup involves mixed data, with both binary and continuous responses. The joint probability density function (pdf) has the following copula representation

$$\begin{aligned} f^c(y_{1i}, y_{2i}|y_{2i} > 0) = {} & h^c\left(F_1^c(y_{1i}|y_{2i} > 0), F_2^c(y_{2i}|y_{2i} > 0); \theta_i^c\right)^{y_{1i}} \\ & \times \left(1 - h^c\left(F_1^c(y_{1i}|y_{2i} > 0), F_2^c(y_{2i}|y_{2i} > 0); \theta_i^c\right)\right)^{1-y_{1i}} f_2^c(y_{2i}|y_{2i} > 0), \end{aligned} \quad (3)$$

where f^c is the joint pdf defined for the pair of mixed outcomes given positive spending; f_2^c and F_2^c denote the density and distribution functions of the positive health care expenditure respectively, and F_1^c represents the cdf of $y_{i1}|y_{i2} > 0$. Assuming that $C^c(\cdot, \cdot; \theta_i^c)$ is a copula that joins the marginals, $h^c(\cdot, \cdot; \theta_i^c)$ is then defined as

$$h^c\left(F_1^c(y_{1i}|y_{2i} > 0), F_2^c(y_{2i}|y_{2i} > 0); \theta_i^c\right) = \left(\frac{\partial C^c\left(F_1^c(y_{1i}|y_{2i} > 0), F_2^c(y_{2i}|y_{2i} > 0); \theta_i^c\right)}{\partial F_2^c(y_{2i}|y_{2i} > 0)}\right)$$

where θ_i^c is the dependence parameter that is associated with copula C^c.

Two-part hurdle models require that the mechanism that governs whether Y_{2i} is positive must remain separate from the process that determines the magnitude of Y_{2i} when it is positive. The appropriateness of that separation is especially important in the present context, where we attempt to link both parts to insurance status. To be sure, such a decoupling is not valid in all settings, especially in classic Heckman-style selection problems where the selection process correlates with magnitude, even after accounting for covariates. But two-part hurdle specifications have become a methodological cornerstone for medical spending, due to the "principal-agent" setup of the U.S. health care system, where patients (principals) typically initiate contact with physicians (agents), but then physicians determine appropriate levels of care (Zweifel, 1981; Deb and Trivedi, 2002). Manning et al. (1987, p. 109), in their seminal study of the RAND Experiment, crystallize this view by observing that "...the decision to receive some care is largely the consumer's, while the physician influences the decision about level of care." Thus, we follow conventions established in the health economics literature and assume the validity of such a decoupling.

3.1.1 Specification of marginal distributions and copula

We use probit formulations for the marginal distributions for the bivariate hurdle part of the model, as alternative setups (logit and cloglog links) appeared to offer little improvement to fit. That is, we specify $F_1^h(y_{1i}^1) = \Phi(\eta_{1i})$ and $F_2^h(y_{2i}^+) = \Phi(\eta_{2i})$, where Φ is the cdf of a standard normal distribution. Similarly, the marginal distribution for y_{1i} in the bivariate model with binary and continuous responses is modeled with a probit, $F_1^c(y_{1i}^1) = \Phi(\eta_{3i})$. The predictors $\eta_{vi} \in \mathbb{R}$, for $v = 1, 2, 3, ...$, contain covariate and coefficients and are defined in generic terms in the next section.

To accommodate the highly-skewed shape of positive expenditures, we explored several distributions, including log-normal, gamma, Dagum, Weibull, inverse Gaussian, with the Dagum appearing to offer the best fit, according to Akaike information criterion (AIC) values and residual diagnostics. The Dagum pdf is

$$f_2^c(y_{2i}|y_{2i} > 0) = \frac{a_i p_i}{y_{2i}} \left[\frac{\left(\frac{y_{2i}}{b_i}\right)^{a_i p_i}}{\left\{\left(\frac{y_{2i}}{b_i}\right)^{a_i} + 1\right\}^{p_i+1}} \right],$$

where $y_{2i} > 0$ and $b_i > 0, a_i > 0, p_i > 0$ are the related distributional parameters. The corresponding cdf is

$$F_2^c(y_{2i}|y_{2i} > 0) = \left\{1 + \left(\frac{y_{2i}}{b_i}\right)^{-a_i}\right\}^{-p_i}.$$

Note that, for the Dagum distribution, the expectation and variance of Y_{2i} are given by non-linear combinations of b_i, a_i, p_i (see, e.g., Table 2 in Marra and Radice, 2017). Also, these parameters are specified as $b_i = \exp(\eta_{4i})$, $a_i = \exp(\eta_{5i})$ and $p_i = \exp(\eta_{6i})$ which allow us to link these coefficients to regression effects. (The use of the inverse link function $\exp(\cdot)$ ensures that the parameters are always estimated as positive values.)

Our main focus is in the dependence parameters, θ_i^h and θ_i^c, which, as noted, combine the causal effect of insurance on spending and selection effects. But then, by specifying those dependence terms as regression functions of traits that insurance companies may *not* use to adjust premiums, coefficients attached to those traits inform upon whether selection effects exist with respect to those traits. Thus, we specify the dependence terms as functions of predictors: $\theta_i^h = m^h(\eta_{7i})$ and $\theta_i^c = m^c(\eta_{8i})$, where m^h and m^c are one-to-one transformations which ensure that the dependence parameters lie in their ranges (see Table 1 in Marra and Radice (2017) for the list of transformations; the table also shows the relation between θ and the Kendall's τ coefficient, which is a measure of association that lies in the customary range $[-1, 1]$). The copulae considered here include the Clayton, Frank, Gaussian, Gumbel and Joe, as well as 180 degree rotations of the Clayton, Gumbel, and Joe copulas. Note that, as pointed out for instance by Genest and Neslehova (2007), the result of Sklar (1973) for C^h and C^c can only guarantee that the copula is unique over the range of the outcomes. In a regression context, however, this potential issue is less likely to be a concern mainly because regression structures in the marginals generate additional variation in the outcomes and thus more completely cover the outcome domains (e.g., Joe, 2014; Nikoloulopoulos and Karlis, 2010; Trivedi and Zimmer, 2017).

The reader is referred to the help file of `GJRM` (Marra and Radice, 2019) for the full list of implemented marginal distributions and copulae.

3.1.2 Predictor specification

This section provides some details on the construction of the model's additive predictors. For the sake of simplicity a generic η_i is considered. Recall that the main advantages of using additive predictors are that various types of covariate effects can be dealt with, and that such effects can be flexibly determined without making strong parametric a priori assumptions regarding their forms (Wood, 2017).

We proceed by defining η_i as a function of an intercept and smooth functions of sub-vectors of a generic covariate vector called $\mathbf{z}_i$. That is,

$$\eta_i = \beta_0 + \sum_{k=1}^{K} s_k(\mathbf{z}_{ki}), \ i = 1, \dots, n, \tag{4}$$

where $\beta_0 \in \mathbb{R}$ is an overall intercept, $\mathbf{z}_{ki}$ denotes the k^{th} sub-vector of the complete covariate vector $\mathbf{z}_i$ (containing, e.g., binary, categorical, continuous, and spatial variables) and the K functions $s_k(\mathbf{z}_{ki})$ represent generic effects which are chosen according to the type of covariate(s) considered. Each $s_k(\mathbf{z}_{ki})$ can be approximated as a linear combination of J_k basis functions $b_{kj_k}(\mathbf{z}_{ki})$ and regression coefficients $\beta_{kj_k} \in \mathbb{R}$, i.e. (Wood, 2017)

$$\sum_{j_k=1}^{J_k} \beta_{kj_k} b_{kj_k}(\mathbf{z}_{ki}). \tag{5}$$

This formulation implies that the vector of evaluations $\{s_k(\mathbf{z}_{k1}), \dots, s_k(\mathbf{z}_{kn})\}^\mathsf{T}$ can be written as $\mathbf{Z}_k\boldsymbol{\beta}_k$ with $\boldsymbol{\beta}_k = (\beta_{k1}, \dots, \beta_{kJ_k})^\mathsf{T}$ and design matrix $Z_k[i, j_k] = b_{kj_k}(\mathbf{z}_{ki})$. This allows the predictor in equation (4) to be written as

$$\boldsymbol{\eta} = \beta_0 \mathbf{1}_n + \mathbf{Z}_1\boldsymbol{\beta}_1 + \ldots + \mathbf{Z}_K\boldsymbol{\beta}_K, \tag{6}$$

where $\mathbf{1}_n$ is an n-dimensional vector made up of ones. Equation (6) can also be written in a more compact way as $\boldsymbol{\eta} = \mathbf{Z}\boldsymbol{\beta}$, where $\mathbf{Z} = (\mathbf{1}_n, \mathbf{Z}_1, \dots, \mathbf{Z}_K)$ and $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_1^\mathsf{T}, \dots, \boldsymbol{\beta}_K^\mathsf{T})^\mathsf{T}$.

Each $\boldsymbol{\beta}_k$ has an associated quadratic penalty $\lambda_k \boldsymbol{\beta}_k^\mathsf{T} \mathbf{D}_k \boldsymbol{\beta}_k$ whose role is to enforce specific properties on the k^{th} function, such as smoothness. Note that $\mathbf{D}_k$ only depends on the choice of basis functions, but not on $\boldsymbol{\beta}_k$. Smoothing parameter $\lambda_k \in [0, \infty)$ controls the trade-off between fit and smoothness, and plays a crucial role in determining the shape of $\hat{s}_k(\mathbf{z}_{ki})$. The overall penalty can be defined as $\boldsymbol{\beta}^\mathsf{T}\mathbf{D}_{\boldsymbol{\lambda}}\boldsymbol{\beta}$, where $\mathbf{D}_{\boldsymbol{\lambda}} = \text{diag}(0, \lambda_1\mathbf{D}_1, \ldots, \lambda_K\mathbf{D}_K)$. Finally, the smooth functions are subject to centering (identifiability) constraints.

For parametric, linear effects, equation (5) becomes $\mathbf{z}_{ki}^\mathsf{T}\boldsymbol{\beta}_k$, and the design matrix is obtained by stacking all covariate vectors $\mathbf{z}_{ki}$ into $\mathbf{Z}_k$. No penalty is typically assigned to linear effects ($\mathbf{D}_k = \mathbf{0}$). This would be the case for binary and categorical variables.

For continuous variables the smooth functions are represented using the regression spline approach. Specifically, for each continuous variable z_{ki}, $s_k(z_{ki})$ is approximated by $\sum_{j_k=1}^{J_k} \beta_{kj_k} b_{kj_k}(z_{ki})$, where the $b_{kj_k}(z_{ki})$ are known spline basis functions. The design matrix $\mathbf{Z}_k$ comprises the basis function evaluations for each i, and hence describe J_k curves which have potentially varying degrees of complexity. We employ low rank thin plate regression splines which are numerically stable and have convenient mathematical properties, although other spline definitions and corresponding penalties are supported in our implementation. To enforce smoothness, a conventional integrated square second derivative spline penalty is typically employed (this is also the default option in the software). That is, $\mathbf{D}_k = \int \mathbf{d}_k(z_k)\mathbf{d}_k(z_k)^\mathsf{T} dz_k$, where the j_k^{th} element of $\mathbf{d}_k(z_k)$ is given by $\partial^2 b_{kj_k}(z_k)/\partial z_k^2$ and integration is over the range of z_k. The formulae used to compute the basis functions and penalties for many spline definitions are provided by Wood (2017) who also discusses their theoretical properties. This specification allows us to avoid arbitrary modeling decisions, such as choosing the appropriate degree of a polynomial or specifying cut-points, which could induce misspecification bias.

Other specifications can be employed. These include varying coefficient smooths obtained by multiplying one or more smooth components by some covariate(s), smooth functions of two or more continuous covariates, random and Markov random field smoothers.

3.2 Some estimation and inferential details

Let us define the overall quantities $\boldsymbol{\delta}^{\mathsf{T}} = (\boldsymbol{\beta}_1^{\mathsf{T}}, \ldots, \boldsymbol{\beta}_8^{\mathsf{T}})$ and $\mathbf{S}_{\boldsymbol{\lambda}} = \text{diag}(\boldsymbol{\lambda}_1\mathbf{S}_1, \ldots, \boldsymbol{\lambda}_8\mathbf{S}_8)$, where $\boldsymbol{\lambda}_v^{\mathsf{T}} = (\lambda_{vk_v}, \ldots, \lambda_{vK_v})$ for $v = 1, \ldots, 8$. Parameter vectors $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_8$ and their corresponding penalty matrices and smoothing parameter vectors are associated with $\eta_{1i} \ldots, \eta_{8i}$, respectively. Using equations (2) and (3), assuming that a random sample $(y_{1i}; y_{2i}; \mathbf{z}_i)$, $i = 1, \ldots, n$ is available, the log-likelihood function can be written as

$$\begin{aligned}\ell(\boldsymbol{\delta}) = &\sum_{y_{1i}^0 \ \& \ y_{2i}^0} \log\left(F^h(y_{1i}^0, y_{2i}^0)\right) + \sum_{y_{1i}^0 \ \& \ y_{2i}^+} \log\left(F^h(y_{1i}^0, y_{2i}^+)\right) \\ &+ \sum_{y_{1i}^1 \ \& \ y_{2i}^0} \log\left(F^h(y_{1i}^1, y_{2i}^0)\right) + \sum_{y_{1i}^1 \ \& \ y_{2i}^+} \log\left(F^h(y_{1i}^1, y_{2i}^+)\right) \\ &+ \sum_{y_{2i}^+} \log\left(f^c(y_{1i}, y_{2i} | y_{2i} > 0)\right).\end{aligned}$$

Because of the flexible predictors employed here, the use of a classic (unpenalized) optimization algorithm is likely to result in component estimates that are too rough to produce practically useful results (e.g., Wood, 2017). Therefore, we maximize $\ell_p(\boldsymbol{\delta}) = \ell(\boldsymbol{\delta}) - \frac{1}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathbf{S}_{\boldsymbol{\lambda}}\boldsymbol{\delta}$. The above log-likelihood structure suggests that parameter estimation can be carried using two separate bivariate copula models, one for the hurdle part and the other for the continuous part. Facilities to achieve this using `GJRM` are already available and are based on the works by Radice, Marra and Wojtys (2016) and Klein et al. (2019) to which we refer the reader for further details.

'Confidence' intervals for any linear and nonlinear function of $\boldsymbol{\delta}$ are obtained from a Bayesian point of view, by recalling that the penalty term associated with the smooth functions of covariates represents the prior belief that these functions are likely to be smoother rather than wiggly. This implies setting an improper multivariate Normal prior on $\boldsymbol{\delta}$, which then leads to the posterior distribution $\boldsymbol{\delta} \;\dot{\sim}\; \mathcal{N}(\hat{\boldsymbol{\delta}}, -\hat{\boldsymbol{\mathcal{H}}}_p^{-1})$, where $\boldsymbol{\mathcal{H}}_p$ is the model's penalized Hessian. The rationale for using this result post-estimation is provided, for instance, in Marra and Radice (2017). They also show that using the above posterior distribution yields confidence intervals with better frequentist properties than those obtained using a frequentist approach itself. Other advantages of using the

Bayesian result are that the distribution of nonlinear functions of $\boldsymbol{\delta}$ can easily be obtained by posterior simulation and that the resulting distribution need not be symmetric.

4 Empirical results

This section breaks the results down separately for each part of the hurdle specification. We label the part that models whether spending is positive as "Part 1" and the part that models the magnitude of positive spending as "Part 2".

4.1 Part 1

The first part of the hurdle specification jointly models the probability of nongroup insurance and the probability of positive spending. Each of those marginal probabilities relies on a probit setup, as alternative link function choices offered little improvement to fit.

For the copula linking those two probabilities, Table 3 shows AIC values for several choices of copulas, each with different dependence patterns. The Rotated Gumbel copula appears to offer the best fit, although other copulas with similar shapes, such as the Clayton and Rotated Joe, perform similarly. Those copulas all show asymmetric dependence, with dependence strongest in the lower tail, suggesting that subjects with small probabilities of having insurance also tend to have small probabilities of positive spending, but that that correlation becomes weaker for larger probabilities. (Note that, because evidence overwhelmingly suggested positive dependence, we did not consider rotations of the Clayton, Joe, and Gumbel copulas designed to accommodate negative associations.)

Having settled upon probit marginals glued together via a Rotated Gumbel copula, Table 4 presents parameter estimates, while limiting the covariates to traits that insurance companies may use to adjust premiums. Age and being married appear to positively correlate with insurance enrollment and positive medical spending. Likewise, subjects residing in the Midwest or West appear more likely to have insurance and positive expenses, relative to their counterparts residing in the (relatively poorer) South. Meanwhile, smokers are less likely to have insurance and medical expenses, suggesting that smoking status picks up some unobserved traits that tend to correlate with smoking.

The main focus of this study, appearing near the bottom of the Table 4, is the dependence parameter, which is positive and precisely estimated. As noted above, that positive dependence term combines the direct causal effect of insurance on spending and the indirect effect of unobserved (to the insurance company) attributes that simultaneously correlate with both insurance enrollment and medical spending.

Table 5 attempts to determine whether adverse selection actually exists with respect to those unobserved (to the insurance company) attributes. Our approach involves specifying the dependence term as a function of those attributes. The main results of interest, shown in the right-hand panel of Table 5, fail to find evidence of selection with respect to gender, race (black), and BMI. However, the coefficient attached to ethnicity (Hispanic) is negative and precisely estimated, suggesting favorable selection along that dimension.

Focusing on the all-important number of chronic conditions, the statistically insignificant coefficient suggests that subjects with preexisting health problems *do not* appear to adversely select into nongroup insurance. In fact, lack of statistical significance notwithstanding, the negative coefficient suggests that subjects with preexisting health problems might positively select *out of* insurance. Although somewhat counterintuitive, many widely-cited studies have reported similar findings in other types of insurance markets (Finkelstein & McGarry, 2005; Pauly, 2005; Cameron & Trivedi, 2013).

One explanation for lack of adverse selection is that risk aversion, as opposed to unobserved health problems, might represent the primary driver of insurance demand. To explore that possibility we added to the dependence parameter a binary measure of whether the person disagrees "strongly" or "somewhat" that he or she likes to take risks. We omitted this measure from our baseline models, due to the variable's highly subjective nature. Nonetheless, the coefficient of that variable failed to achieve statistical significance in the dependence parameter, casting doubt on whether risk aversion explains the lack of adverse selection.

Another explanation, and the one that seems to corroborate evidence from the Massachusetts health insurance reform (Hackmann, Kolstad, and Kowalski, 2015), is that the ACA's mandate, despite its relatively toothless penalties, might have drawn enough healthy enrollees into nongroup risk pools to offset adverse selection. This conjecture is impossible to verify without detailed panel data on pre- and post-reform enrollees, but it seems the most likely explanation for the evident lack of adverse selection.

Of course, using linear combinations of control variables might hide nonlinear relationships, especially for nonbinary variables. To explore that possibility, for "Age" in the two marginals and "Number of chronic conditions" in the dependence term, we replaced the covariate/slope terms with smooth spline functions, as described in Section 3.1.2. Figure 1 shows graphs of those splines, with 95% percent confidence bands. (All other coefficients were very similar to those reported in Table 5.) The left-hand panel of Figure 1 shows that subjects between about 25 and 45 years of age have lower probabilities of nongroup insurance enrollment, while subjects above 45 have higher probabilities. The second panel shows that the probability of positive medical spending increases in an (approximately) linear fashion with age. The right-hand panel shows the negative, but insignificant, link between chronic conditions and the dependence term. The confidence band appears to fan out as chronic conditions increase, largely because only 2 percent of subjects in the estimation sample report more than 4 chronic conditions. Nevertheless, the figure indicates that chronic conditions never appear to contribute *positively* to the dependence term, offering no evidence of adverse selection along that dimension.

4.2 Part 2

The second part focuses on subjects with positive spending. Mirroring the model selection process for the first part, we first choose appropriate marginals based on separate estimation of each marginal, before turning to the copula. We again opt for a probit specification for the probability of insurance enrollment. For positive spending, we attempted several distributions including the lognormal, Weibull, gamma and Dagum, all of which allow for the highly skewed shape of positive spending. The Dagum appeared to offer the best fit, according to AIC calculations and the residual plots reported in Figure 2. Note that only parameter b (see Section 3.1.1) was specified as function of η since including predictors in the other parameters did not lead to improvements to fit. As for the copula, Table 6 shows that AHM offers the best fit.

Having settled upon probit/Dagum marginals and the AHM copula, Table 7 presents estimates. Coefficients of control variables produce similar signs to those reported in Part 1. Shown near the bottom of the table, the dependence term, though smaller in magnitude than in Part 1, is positive and precisely estimated. Again, that positive number combines the causal effect of insurance on spending with possible selection effects.

Table 8 specifies the dependence term as a function of controls to determine the extent, if any, to which those attributes contribute to adverse selection among positive spenders. In contrast to Part 1, the coefficient of female is positive, indicating that positive-spending females might adversely select into nongroup insurance, relative to their male counterparts. Meanwhile, the coefficient of BMI suggests (marginally significant) favorable selection along that dimension. Most importantly, the coefficient attached to chronic conditions is indistinguishable from zero, which, similar to Part 1, indicates a lack of adverse selection with respect to that all-important dimension.

Figure 3 shows spline estimates for age and chronic conditions. The patterns are somewhat similar to those observed in Part 1, although less precisely estimated, perhaps due to the smaller sample size compared to Part 1. Again, most importantly, the link between chronic conditions and dependence never appears to be significantly positive.

5 Partial Effects of Insurance on Spending

Our copula-based setup allows us to recover a partial effect of insurance on spending, with particular focus on how chronic conditions alters that relationship. We focus on Part 1, because the coefficient of chronic conditions in Part 2 is nearly zero.

Therefore, using the converged parameter estimates from Part 1, we calculate

$$\Pr(y_{i2} > 0 \mid y_{i1} = 1) - \Pr(y_{i2} > 0 \mid y_{i1} = 0),$$

which would be equivalent to $\Pr(y_{2i} = 1|y_{1i} = 1) - \Pr(y_{2i} = 1|y_{1i} = 0)$, where

$$\Pr(y_{2i} = 1|y_{1i} = 1) = \frac{C\left(\Phi(\eta_{1i}), \Phi(\eta_{2i}); \theta_i\right)}{\Phi(\eta_{1i})},$$

and

$$\Pr(y_{2i} = 1|y_{1i} = 0) = \frac{\Phi(\eta_{2i}^{(y_{1i}=0)}) - C\left(\Phi(\eta_{1i}), \Phi(\eta_{2i}); \theta_i\right)}{1 - \Phi(\eta_{1i})}.$$

In other words, this provides the difference in the probability of positive spending between a person with and without nongroup insurance.

Figure 4 shows partial effects for one observation in our data; we choose a person with attributes relatively close to sample medians. That person has zero chronic conditions, but we recoded that variable several times to see how chronic conditions alter the partial effect. For zero chronic conditions, nongroup insurance correlates with an increase of 0.296 in the probability of positive spending, and that estimate differs statistically from zero. We stress that that number does *not* provide the *causal* effect of insurance on positive spending, because residual selection effects might remain. But it *does* remove selection effects stemming from variables that appear in our model: age, married, smokes, region, gender, race, ethnicity, BMI, and chronic conditions.

As the number of chronic conditions increases, the partial effect hardly budges from 0.296, indicating a lack of selection with respect to preexisting health problems. The main message from

Figure 4, and indeed the main punchline of this paper, is that we fail to uncover evidence of adverse selection with respect to preexisting health problems.

6 Conclusion

The Affordable Care Act (ACA) forbids insurance companies from adjusting premiums with respect to certain attributes that are widely believed to correlate with medical risk, a restriction known as "community rating." The most important of those attributes is preexisting health problems. This feature of the ACA, although seemingly popular with American voters, has raised concerns about adverse selection, in which people with health problems might be disproportionately likely to comply with the ACA's mandate that everyone have insurance coverage, resulting in a relatively sicker risk pool.

Focusing on the market for nongroup insurance, this paper explores whether, after implementation of the ACA, enrollees really do exhibit adverse selection with respect to attributes that insurance companies must ignore. We adopt a copula-based hurdle model with the dependence parameter specified as a function of those attributes, which allows us to determine the existence (and direction) of selection patterns stemming from those attributes. Our main finding is that nongroup insurance enrollees do not appear to exhibit adverse selection, particularly with respect to preexisting health problems. In fact, the first part of our hurdle specification finds some evidence of favorable selection with respect to preexisting conditions, although that result fails to achieve statistical significance at conventional levels.

Overall, we conclude that, at least so far, fears of community rating/guaranteed issue causing adverse selection seem to have been misplaced. The most likely explanation is that the ACA's mandate, despite its relatively small penalties for noncompliance, might have attracted enough healthy enrollees to offset any adverse selection. We stress, however, that the ACA is still relatively young, and political support for the law seems to whipsaw with respect to whichever party has political power. Furthermore, our findings apply just to nongroup private insurance markets. The ACA also introduced sweeping changes to public insurance arrangements, but investigating those likely requires a separate study.

Statistically, the copula-based hurdle model that we employ should prove useful for any outcome variable that has high probability mass at zero and a long upper tail, and where the underlying mechanism that determines whether the variable is positive can be decoupled from the process that determines its magnitude if positive. Medical spending is one such example, but other variables with similar distributional shapes, such as household income or charitable contributions, might also apply.

Brechmann, E. C. & Schepsmeier, U. (2013). Modeling dependence with c- and d-vine copulas: The R package CDVine. Journal of Statistical Software, 52(3), 1-27.

Cameron, A. C. & Trivedi, P. (2013). Regression Analysis of Count Data: Second Edition. Cambridge University Press: New York.

de Meza, D. & D. Webb (2001). “Advantageous Selection in Insurance Markets.” Rand Journal of Economics, 32, 249-262.

Deb, P. & Trivedi, P. (2002). The structure of demand for health care: latent class versus two-part models. Journal of Health Economics, 21, 601-625.

Deb, P., Trivedi, P. K., & Zimmer, D. (2014). Cost-offsets of prescription drug expenditures: Data analysis via a copula-based bivariate dynamic hurdle model. Health Economics, 23(10), 1242-1259.

Dunn, P. K. & Smyth, G. K. (1996). Randomized quantile residuals. Journal of Computational and Graphical Statistics, 5, 236-245.

Fang, H., Keane, M., & D. Silverman. 2008. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. *Journal of Political Economy*, 116: 303-350.

Finkelstein, A. & McGarry, K. (2005). Private information and its effect on market equilibrium: evidence from the long-term care insurance market. American Economic Review, 96, 938-958.

Gelbach, J. (2016). When do covariates matter? And which ones, and how much? Journal of Labor Economics, 34, 509-543.

Genest, C. & Neslehova, J. (2007). The astin bulletin. A primer on Copulas for Count Data, 37, 475-515.

Hackmann, M., Kolstad, J., & Kowalski, A. (2015). Adverse selection and an individual mandate: When theory meets practice. American Economic Review, 105, 1030-66.

Joe, H. (2014). Dependence Modelling with Copulas. CRC Press: Boca Raton, FL, USA.

Klein, N., Kneib T., Marra, G., Radice, R., Rokicki, S., McGovern, M.E. (2019). Mixed binary-continuous copula regression models with application to adverse birth outcomes. Statistics in Medicine, 38(3), 413-436.

Manning, W., Newhouse, J., Duan, N., Keeler, E., Leibowitz, A., Marquis, S. (1987). Health Insurance and the Demand for Medical Care: Evidence from a Ranomized Experiment. American Economic Review, 77, 251-277.

Marra, G. & Radice, R. (2017). Bivariate Copula Additive Models for Location, Scale and Shape. Computational Statistics and Data Analysis, 112, 99113.

Marra, G. & Radice, R. (2019). GJRM: Generalized Joint Regression Modeling. R package version 0.1-5, URL https://cran.r-project.org/package=GJRM.

Nikoloulopoulos, A. K. & Karlis, D. (2010). Regression in a copula model for bivariate count data. Journal of Applied Statistics, 37, 1555-1568.

Panhans, M. (2019). “Adverse Selection in ACA Exchange Markets: Evidence from Colorado.” American Economic Journal: Applied Economics, forthcoming.

Pauly, M. (2005). Effects of insurance coverage on use of care and health outcomes for nonpoor young women. American Economic Review: Papers and Proceedings, 95, 219–233.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/.

Radice, R., Marra, G., & Wojtys, M. (2016). Copula Regression Spline Models for Binary Outcomes Statistics and Computing, 26(5), 981-995.

Sacks, D. (2018). The health insurance marketplaces. Journal of the American Medical Association, 320, 549-550.

Shi, P. & W. Zhang (2016). Private Information in Healthcare Utilization: Specification of a Copula-Based Hurdle Model, Journal of the Royal Statistical Society: Series A, 178, 337-361.

Sklar, A. (1973). Random variables, joint distributions, and copulas. Kybernetica, 9, 449- 460.

Trivedi, P. C. & Zimmer, D. M. (2017). A Note on Identification of Bivariate Copulas for Discrete Count Data. Econometrics, 5, 10.

Wood, S. N. (2017). Generalized Additive Models: An Introduction With R: Second Edition. Chapman & Hall/CRC, London.

Zweifel, P. (1981). Supplier-Induced Demand in a Model of Physician Behavior. In Van Der Gaag, J., Perlman, M. (eds.), Health Economics, and Health Economics. North-Holland, Amsterdam, 245-267.

Table 1: Summary statistics for office-based spending (n = 6,014)

	Nongroup insurance n = 1,553	No insurance n = 4,461
Any spending?	0.64	0.33
Spending among positive spenders		
mean	$1,087	$343
median	$179	$0

Table 2: Sample means partitioned positive spending status (n = 6,014)

	Zero spending n = 3,538	Positive spending n = 2,476
Premium adjustments permitted by...		
Age	36.9	43.4
Married	0.35	0.48
Smokes	0.16	0.12
Northeast	0.11	0.11
Midwest	0.14	0.17
West	0.22	0.24
South	0.53	0.48
Premium adjustments not permitted by...		
Female	0.41	0.58
Black	0.18	0.13
Hispanic	0.55	0.41
BMI	27.9	28.6
Number of chronic conditions	0.42	1.27

Table 3: AIC values for PART 1 copula model (margins are Bernoulli with probit links)

Copula	AIC
Gaussian	13,974
Clayton	13,971
Rotated Clayton (180 degrees)	14,001
Joe	14,004
Rotated Joe (180 degrees)	13,974
Gumbel	13,982
Rotated Gumbel (180 degrees)	**13,971**
Frank	13,982
AHM	13,972
FGM	13,990
Student-t (with df = 3)	13,972
Plackett	13,979

Table 4: Rotated Gumbel copula model estimates of Part 1 (standard errors in parentheses)

	Nongroup insurance	Any spending
Age	0.016^{**}	0.024^{**}
	(0.001)	(0.001)
Married	0.246^{**}	0.160^{**}
	(0.038)	(0.036)
Smokes	-0.110^{**}	-0.197^{**}
	(0.052)	(0.048)
Northeast	-0.060	0.010
	(0.061)	(0.056)
Midwest	0.247^{**}	0.202^{**}
	(0.051)	(0.049)
West	0.202^{**}	0.100^{**}
	(0.044)	(0.041)
South (omitted)	–	–
	–	–
Constant	-1.484^{**}	-1.270^{**}
	(0.062)	(0.058)
Dependence τ (95% interval)	0.23 (0.21, 0.26)	

* p < .10; ** p < .05

Table 5: Rotated Gumbel copula model estimates of Part 1 (standard errors in parentheses). Here the dependence parameter is expressed as a function of covariates

	Nongroup insurance	Any spending		Dependence
Age	0.016^{**}	0.024^{**}	Female	-0.028
	(0.001)	(0.001)		(0.155)
Married	0.245^{**}	0.155^{**}	Black	0.036
	(0.038)	(0.036)		(0.204)
Smokes	-0.100^{*}	-0.188^{**}	Hispanic	-0.502^{**}
	(0.052)	(0.048)		(0.199)
Northeast	-0.058	0.008	BMI	0.003
	(0.061)	(0.056)		(0.015)
Midwest	0.248^{**}	0.198^{**}	Number of chronic conditions	-0.096
	(0.051)	(0.049)		(0.076)
West	0.204^{**}	0.086^{**}		
	(0.044)	(0.042)		
South (omitted)	–	–		
	–	–		
Constant	-1.489^{**}	-1.286^{**}	Constant	1.056^{**}
	(0.062)	(0.059)		(0.391)

* p < .10; ** p < .05

Figure 1. Part 1: Estimated smooth effects of age on nongroup insurance and positive spending, and of number of chronic conditions on the dependence parameter on the scale of the predictor, and associated 95% point-wise intervals. The jittered rug plot, at the bottom of the graph, shows the covariate values.

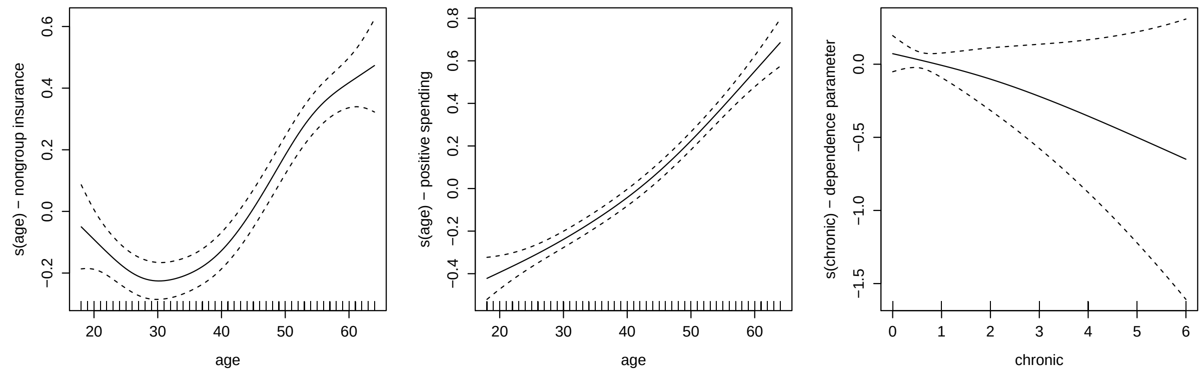

Table 6: AIC values for Part 2 copula model (margins are Bernoulli with probit link for insurance, and Dagum for spending)

Copula	AIC
Gaussian	41331.9
Clayton	41338.1
Rotated Clayton (180 degrees)	41350.7
Joe	41357.6
Rotated Joe (180 degrees)	41340.4
Gumbel	41341.8
Rotated Gumbel (180 degrees)	41330.9
Frank	41329.2
AHM	**41328.7**
FGM	41328.9
Student-t (with df = 3)	41336.6
Plackett	41329.5

Figure 2. Histogram and normal Q-Q plot of randomised normalized quantile residuals (Dunn & Smyth, 1996) for the Dagum marginal modeling positive spending.

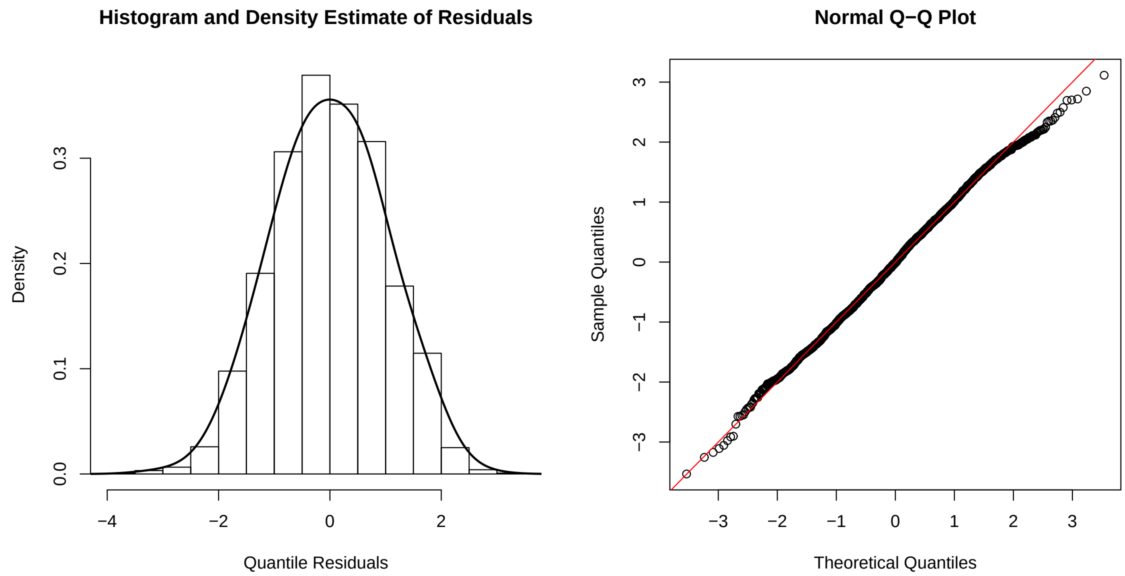

Table 7: AMH copula estimates of Part 2 (standard errors in parentheses)

	Nongroup insurance	Positive spending
Age	0.013**	0.018**
	(0.002)	(0.002)
Married	0.023	0.018
	(0.053)	(0.057)
Smokes	−0.068	−0.172**
	(0.079)	(0.084)
Northeast	−0.042	0.071
	(0.088)	(0.094)
Midwest	0.299**	−0.025
	(0.072)	(0.079)
West	0.195**	0.106
	(0.063)	(0.068)
South (omitted)	–	–
	–	–
Constant	−0.918**	4.326**
	(0.097)	(0.169)
a^2 (95% interval)	–	1.15 (1.04, 1.28)
p (95% interval)		1.81 (1.52, 2.20)
Dependence τ (95% interval)	0.15 (0.12, 0.18)	

* p < .10; ** p < .05

Figure 3. Part 2: Estimated smooth effects of age on nongroup insurance and positive values of spending, and of number of chronic conditions on the dependence parameter on the scale of the predictor, and associated 95% point-wise intervals. The jittered rug plot, at the bottom of the graph, shows the covariate values.

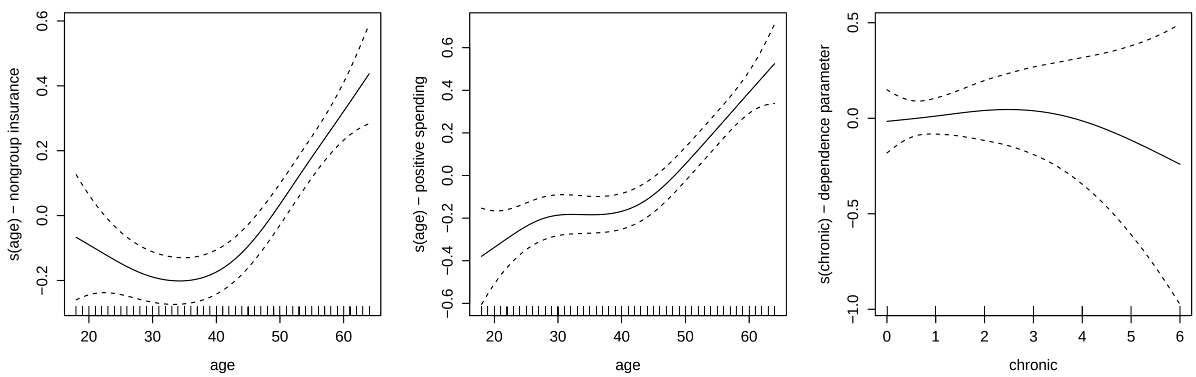

Table 8: AMH copula estimates of Part 2 (standard errors in parentheses). Here the dependence parameter is expressed as a function of covariates

	Nongroup insurance	Positive spending		Dependence
Age	0.013**	0.018**	Female	0.344**
	(0.002)	(0.002)		(0.165)
Married	0.020	0.020	Black	−0.017
	(0.053)	(0.057)		(0.250)
Smokes	−0.064	−0.174**	Hispanic	0.152
	(0.079)	(0.084)		(0.203)
Northeast	−0.060	0.066	BMI	−0.027*
	(0.088)	(0.094)		(0.014)
Midwest	0.296**	−0.025	Number of chronic conditions	0.004
	(0.072)	(0.076)		(0.063)
West	0.183**	0.105		
	(0.064)	(0.068)		
South (omitted)	–	–		
	–	–		
Constant	−0.925**	4.331**	Constant	1.200**
	(0.097)	(0.170)		(0.395)
a^2 (95% interval)	–	1.15 (1.05, 1.27)		
p (95% interval)		1.81 (1.50, 2.20)		

* p < .10; ** p < .05

Figure 4. Partial effect of insurance on the probability of positive spending (with 95% confidence band)