Uncontrolled diabetes and health care utilisation: a bivariate Latent Markov model approach

Joan Gil, University of Barcelona, Spain

Paolo Li Donni, University of Palermo, Italy

Eugenio Zucchelli[1], Lancaster University, UK; Madrid Institute for Advanced Study (MIAS), Universidad Autonoma de Madrid (UAM), Spain

30/04/2019

Abstract

While uncontrolled diabetes (UD) or poor glycaemic control is a widespread condition with potentially life-threatening consequences, there is sparse evidence of its effects on health care utilisation. We jointly model the propensities to consume health care and UD by employing an innovative bivariate Latent Markov model which allows for dynamic unobserved heterogeneity, movements between latent states and the endogeneity of UD. We estimate the effects of UD on primary and secondary health care consumption using a panel dataset of rich administrative records from Spain and measure UD using a biomarker. We find that, conditional on time-varying unobservables, UD does not have a statistically significant direct effect on health care use. Furthermore, individuals appear to move across latent classes and increase their propensities to poor glycaemic control and health care use over time. Our results suggest that by ignoring time-varying unobserved heterogeneity and the endogeneity of UD, the effects of UD on health care utilisation might be overestimated and this could lead to biased findings. Our approach reveals heterogeneity in behaviour beyond standard groupings of frequent versus infrequent users of health care services. We argue that this dynamic latent Markov approach could be used more widely to model the determinants of health care use.

Key words: diabetes; health care utilisation; unobserved heterogeneity; Latent Markov model

JEL: C35; I10; I12

---

[1] Corresponding author. Division of Health Research (DHR), Lancaster University, Furness building, LA1 4YX. Email: e.zucchelli@lancaster.ac.uk.

1. Introduction

Diabetes mellitus (DM) is a major public health concern and one of the most widespread chronic conditions worldwide with a rapidly increasing prevalence, especially among low- and middle-income countries (IDF, 2015). According to the WHO (2016), we are currently experiencing a diabetes "epidemic" where the number of adults with DM has grown from 108 million in 1980 to an estimated 422 million in 2014 and with DM projected to be the 7th leading cause of death by 2030. Furthermore, DM imposes a substantial economic burden to any health care system. Among individuals with DM, uncontrolled diabetes (UD) or poor glycaemic control is a prevalent condition, concerning between around 30 to 50 percent of patients. Glycaemic control is driven by a variety of factors such as genetic endowment; lifestyle; resistance to intensified medication (therapeutic resistance); and low treatment adherence. UD leads to sustained higher levels of blood sugar, which in turn increase the risk of life-threatening comorbidities such as kidney failure and stroke.

A related problem is the extent to which UD may cause additional, yet avoidable, health care utilisation. Indeed, the medical literature often finds that individuals with poor glycaemic control tend to use more health care resources resulting in extra utilisation and added costs for the health system (e.g. Wagner et al., 2001; Gilmer et al., 2005; Mata-Cases et al., 2016). However, the majority of these studies employ empirical approaches often based on standard (linear) regression models applied to selected samples of US health insurance enrolees. Such samples tend to limit the generalisability of the results while previous empirical approaches do not account for potentially important issues which may affect the identification of the effect of DM on health care use such as individual-level unobserved heterogeneity and endogeneity concerns.

The main objective of this paper is to jointly model the propensities towards poor glycaemic control and the use health care services via a bivariate Latent Markov (LM) model. This novel approach builds on some recent developments within the literature concerning latent class models (see Bartolucci et al., 2013 and 2014, for a comprehensive overview) and allows accounting for the endogeneity of UD and time-varying unobserved heterogeneity using dynamic latent classes. We apply this model to the relationship between UD and utilisation of primary and secondary health care services among individuals with type-2 diabetes. We employ six waves (2005-2010) of a longitudinal administrative dataset from Spain including detailed medical records of adult individuals with diabetes. We measure UD using glycated

haemoglobin (HbA1c) levels, a biomarker providing an accurate average measurement of glucose (sugar) concentration which is commonly used by physicians to diagnose and monitor diabetes and its severity (e.g. International Expert Committee, 2009; Lyons and Basu, 2012).

Differently from previous studies, we find that conditional on the dynamic latent process poor glycaemic control does not appear to increase health care use among individuals with diabetes. We also find that the association between UD and health care utilisation tend to disappear when controlling for unobserved heterogeneity. This may suggest that UD and health care consumption might be both driven by a series of unobserved factors. Furthermore, our bivariate LM model reveals the existence of a series of (latent) groups of individuals, which differ by their propensities towards poor glycaemic control and health care use. Interestingly, patients who present a high propensity to consume health care do not necessarily show a similar tendency to poor glycaemic control. Our estimates also indicate that individuals may switch groups over time often by increasing their propensities to poor glycaemic control and utilisation of primary and secondary care. This also suggests that unobserved heterogeneity is, in this case, time-varying and that ignoring this would lead to biased estimates and misleading policy implications.

This paper offers several contributions to the literature. First, to the best of our knowledge, this is one of the first studies to employ a bivariate LM model with dynamic unobserved heterogeneity to explore the determinants of health care utilisation and the first to model the propensities to consume health care and poor glycaemic control simultaneously. Secondly, our latent class model allows switching latent class over time, thus enabling changes in behaviour among individuals with diabetes while also accounting for important underlying causes of utilisation such as the ageing process. Finally, this is also the only paper employing a latent class model to investigate the relationship between UD and health care utilisation. Overall, this paper provides new evidence to the literature on (uncontrolled) diabetes and health care use while also proposing a novel dynamic latent class approach.

2. Background

Previous analyses within the medical literature consistently find that poor glycaemic control is associated with higher health care utilisation and increased medical costs (e.g. Wagner et al., 2001; Shetty et al., 2005; Oglesby et al., 2006). For instance, Wagner et al. (2001) employ linear and log-linear models on US health insurance claims data of adult patients with diabetes collected between 1992-1997 to analyse the relationship between uncontrolled diabetes, health

care use and related costs. They find that sustained reductions of blood sugar levels are correlated with lower utilisation, including hospitals' admission rates, primary and specialist visits, and significant costs savings. Similarly, Shetty et al. (2005) and Oglesby et al. (2006) use Generalised Linear Models (GLM) on data concerning patients with type-2 DM drawn from the US Health Core Managed Care Database (1998-2003). Overall, they find that, depending on their specific treatment between 41 to 66 percent of patients present a "suboptimal" control of their conditions and that the direct medical costs of treating patients with diabetes are higher among those with suboptimal glycaemic control.

However, most of these studies are based on standard regression models such as linear, log-linear and Generalised Linear Models (GLM; McCullagh and Nelder, 1989) applied to specific samples of US health insurance enrolees. These approaches tend to ignore potentially important modelling issues such as individual-level unobserved heterogeneity and the endogeneity of UD. Furthermore, with the exception of Gil et al. (2017), the economic literature has traditionally focused on the effects of DM on the labour market, including earnings, hours of work and absenteeism (Rizzo et al., 1996; Kahn, 1998; Latif, 2009; Zhang et al., 2009; Minor, 2010) while overlooking the impact of DM or UD on health care use.

The analysis of the determinants of health care utilisation has received considerable attention within the empirical economics literature. Early studies employ hurdle (e.g. Pohlmeier and Ulrich, 1995; Santos-Silva and Windmeijer, 2001; Schellhorn et al., 2000; Jimenez-Martin et al., 2002; Van Ourti, 2004) and finite mixture models (e.g. Deb and Trivedi, 1997; Deb and Holmes, 2000; Deb and Trivedi, 2002; Jimenez-Martin et al., 2002; Atella et al., 2004) to explore the factors influencing health care use mainly on cross-sectional data. More recent papers exploit panel data and propose latent class models which account for individual-level unobserved heterogeneity (e.g. Bago d'Uva, 2005) as well combinations of hurdle and latent class models (e.g. Bago d'Uva, 2006; Bago d'Uva and Jones, 2009).

Latent class models present a series of attractive modelling features. First, the use of latent classes allows modelling unobserved heterogeneity without imposing restrictive assumptions on the distribution of the individual effects. Secondly, latent classes provide information on the structure of the underlying unobserved heterogeneity by identifying two (or more) groups of individuals, which typically differ in their propensities to use health care (e.g. 'infrequent users' versus 'frequent users' as in Bago d'Uva, 2006). However, most previous latent class models used to investigate the determinants of health care use, often assume that individual-

level unobserved heterogeneity is time-invariant (with the exception of Hyppolite and Trivedi, 2012). Importantly, this implies that individuals are not allowed to move between classes over time. In this paper, we relax this assumption and employ a LM model that accounts for *time-varying* unobserved heterogeneity, movements between classes and the endogeneity of UD. More specifically, LM models can be seen an extension of latent class models for longitudinal data and their fundamental difference lies in the way the latent process is defined. Typical latent class models assume that the observed sample is drawn from a population partitioned into k latent classes identified by a discrete latent variable U with k support points (Lazarsfeld and Henry 1968; Goodman 1974). LM models additionally assume that the latent process U follows a first-order Markov chain with finite state space. We exploit this feature in our application.

Despite been used in a variety of behavioural and social sciences, models including a dynamic latent process to account for time-varying unobserved heterogeneity are not normally employed to study health care use, especially in the economics literature. Yet, previous studies have estimated such models to explore other health-related issues. For instance, Bartolucci et al. (2014) propose a flexible approach which accounts for time-varying unobservables using a mixture of latent auto-regressive AR(1) processes. They apply this specification to study self-reported health on data from the US Health and Retirement Study (HRS) and capture potential shocks changing an individual's health or its subjective perception over time. While their model is highly relevant, our empirical approach differs in that is a bivariate extension of the LM models with a discrete latent process presented in Bartolucci and Farcomeni (2009) and it explores the effects of UD on health care consumption using panel administrative data. Furthermore, Bartolucci and Farcomeni (2015) estimate an event-history variant of the LM approach on data from a medical study about primary biliary cirrhosis. Interestingly, this model allows for both time-fixed and time-varying unobserved heterogeneity while accounting for informative dropout, that is when individuals are lost to follow-up due to events (e.g. death) linked to the main outcome of interest. Finally, and similarly to Bartolucci et al. (2009), Montanari and Pandolfi (2018) analyse data drawn from a longitudinal survey on the health status of elderly patients in nursing homes within the region of Umbria (Italy). They use an extended LM model which deals with different intermitting patterns of missing data and dropout due to death and discharge of patients.

3. Empirical approach

Compared to previous studies on health care use, our LM approach presents two main innovations. First, we employ a bivariate specification to account for the endogeneity of UD via the inclusion of unobserved heterogeneity modelled as discrete latent classes. Traditionally, the majority of latent class models used to analyse health care consumption in the economics literature relied on univariate (single equation) specifications with time-invariant unobservables (e.g. Bago d'Uva, 2005, 2006; Deb and Trivedi, 1997). Secondly, we allow unobserved heterogeneity to be *time-varying* and thus enable individuals to move among latent classes over time. This accounts for changes in behaviour concerning both the propensities to use health care and poor glycaemic control between time periods. In doing so, we rely and build on the well-established LM model proposed by Bartolucci and Farcomeni (2009).

In the presence of longitudinal data, the relationship between uncontrolled diabetes, UD_{it}, and health care utilisation HC_{it}, is often modelled by a standard panel data specification such that:

$$HC_{it}^* = \alpha_i + \boldsymbol{\beta x}_{it} + \gamma UD_{it} + \eta_{it} \tag{1}$$

where HC_{it}^* is the underlying latent counterpart of a discrete observed categorical indicator of health care utilisation, HC_{it}, with $j = 0, \dots, J$ categories where $HC_{it} = j$ is the number of GP or specialist visits attended by an individual i at time t with $t = 1, \dots, T_i$; UD_{it} is a binary variable taking value 1 in the presence of uncontrolled diabetes; α_i is the time-invariant individual level unobserved heterogeneity; $\boldsymbol{x}_{it}$ is a vector of observed individual characteristics; and η_{it} captures the idiosyncratic error/shock, and could also include residual time-varying individual unobserved heterogeneity (H_{it}). If we assume, for instance, that η_{it} follows a logistic distribution, equation (1) could be estimated using either standard random or fixed effects ordered logit models.

In this case, the effect of UD on health care use, γ, could be identified by such as model if UD_{it} is uncorrelated with the time-varying individual level unobservable component, H_{it}, or simply by assuming that H_{it} is negligible. While the latter assumption is often conveniently invoked in order to employ standard panel data techniques, in most cases this may be unrealistic/untenable. Hence, it might be reasonable to assume that the unobserved time-varying factors affecting health care utilisation may also simultaneously affect UD. For

instance, the propensity to consume health care and the one leading to poor glycaemic control might be driven by a third set of common (unobserved) factors, including underlying (unobservable) health risks such as changes in the probability of developing specific health conditions as an individual ages. Hence, it might be reasonable to assume that the unobserved time-varying factors affecting health care utilisation may also simultaneously affect UD.

Importantly, if *time-varying* unobserved heterogeneity is not properly accounted for, the estimation of γ might be biased. In order to disentangle the genuine effect of UD on health care consumption and the role of time-varying unobserved heterogeneity, we jointly model the propensities to poor glycaemic control and health care use by employing a bivariate LM model (e.g. Bartolucci and Farcomeni, 2009). This allows treating time-varying unobserved heterogeneity, H_{it}, non-parametrically by including a vector of individual-specific parameters which follows a first-order Markov process. The potential endogeneity of UD is accounted for by jointly modelling UD_{it} and HC_{it}.

Bivariate Latent Markov Model
We propose a bivariate LM specification to jointly model the propensities to poor glycaemic control and health care use. Let $\boldsymbol{Y}_{it}$ denote a vector collecting values of uncontrolled diabetes and health care utilisation, UD_{it} and HC_{it}, and $\overline{\boldsymbol{Y}}_i$ the corresponding overall vector for the entire period T_i. To model the distribution of $\overline{\boldsymbol{Y}}_i$ given observable characteristics $\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{iT}$, the LM model assumes that: i) the individual-specific random parameter $\boldsymbol{\alpha}_i$ defining unobserved heterogeneity follows a first-order Markov process with latent states $\boldsymbol{U}(u)$ with $u = 1, \dots, k$; and ii) such latent states make UD_{it} and HC_{it} conditionally independent given the set of observable characteristics, $\boldsymbol{x}_{it}$. The latter is a form of local independence assumption implying that, given the set of observable covariates in equation (1), any source of residual association between UD_{it} and HC_{it} is accounted for by the latent states parameters. In this setting, the unobserved states should capture the overall propensity to poor glycaemic control and health care use. Different propensities reflect unobserved time-varying factors, such as unmeasured health risks or attitudes towards risk, which are not captured by $\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{iT}$.

Our LM model is built around three sets of parameters. The first set of parameters describes how the joint distribution of UD_{it} and HC_{it} depends on the observable characteristics $\boldsymbol{x}$ and the

latent process $\boldsymbol{\alpha}_{i1}, \dots, \boldsymbol{\alpha}_{iT}$.[2] From equation (1) the conditional distribution of UD_{it} and HC_{it} can be directly parameterised by marginal logits and log-odds ratios. In particular the first set of parameters can be described as follows:

$$\log\frac{p(UD_{it} = 1|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it,})}{p(UD_{it} = 0|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it})} = \alpha_{UD,it} + \boldsymbol{\beta}_{UD}\boldsymbol{x}_{it}$$

$$\log\frac{p(HC_{it} \geq j|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it}, UD_{it})}{p(HC_{it} < j|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it,}UD_{it})} = \alpha_{HC,j+1,it} + \boldsymbol{\beta}_{HC}\boldsymbol{x}_{it} + \gamma UD_{it} \qquad (2)$$

where $HC_{it} = j \leftrightarrow \delta_{j-1} < HC_{it}^* \leq \delta_j$, and $\alpha_{HC,j+1,it} = \alpha_{HC,1,it} + \delta_{j+1}$. Therefore if HC has for example three levels ($J = 2$), there are k random intercepts $\boldsymbol{\alpha}_{HC}$ and $J - 1 = 1$ cut-off point δ. That is, we are assuming that random intercepts and cut-off points are linearly additive. Notice that $\boldsymbol{\alpha}_{UD}$ and $\boldsymbol{\alpha}_{HC}$ capture the relative contributions of time-varying unobserved heterogeneity in determining UD_{it} and HC_{it}, respectively and γ is the remaining direct effect of uncontrolled diabetes on health care consumption (Di Mari and Bakk, 2018). Finally, to capture potential time-specific association between UD_{it} and HC_{it} (e.g. time-specific shocks that can influence both probabilities) we include a set of log-odds ratios:

$$\log\left[\frac{p(UD_{it}=1,HC_{it}\geq j|\alpha_{it},x_{it},UD_{it})p(UD_{it}=0,HC_{it}<j|\alpha_{it},x_{it},UD_{it})}{p(UD_{it}=0,HC_{it}\geq j|\alpha_{it},x_{it},UD_{it})p(UD_{it}=1,HC_{it}<j|\alpha_{it},x_{it},UD_{it})}\right] = \rho, \qquad \forall j = 1, \dots, J \qquad (3)$$

where $\boldsymbol{\alpha}_{it}$ collects the vector of intercepts in the system of equations (2). Notice that all log-odds ratios are set to be equal across different level j of health care utilisation. This implies that in our specification, the association between UD and HC is fixed across different levels of health care utilisation. While this may appear to impose a restriction, it allows us to account for any residual time-specific association after having conditioned on the latent process and substantially ease the computational burden of the model.

The two remaining sets of parameters model the underlying distribution of the random parameter vector $\boldsymbol{\alpha}_{it}$. More specifically, the second set of parameters includes the initial probabilities $\Pr(\boldsymbol{\alpha}_{i1}) = \Pr(\boldsymbol{\alpha}_{i1} = \boldsymbol{U}(u)) = \pi_{u1}$ with $u = 1, \dots, k$, which are produced by a multinomial logit specification. These can be interpreted as the standard class membership probabilities in the first time period. The third set of parameters comprises the transition

---

[2] To ease the notation, we have suppressed the suffix i in T_i. However, we should keep in mind that in this case time is individual-specific as we are employing an unbalanced panel.

probabilities between different latent states. Since this is a first-order Markov process, these depend on the previous time period such as $\pi_{v_t|u_{t-1}} = \Pr(\boldsymbol{\alpha}_{it} = \boldsymbol{U}(v)|\boldsymbol{\alpha}_{it-1} = \boldsymbol{U}(u))$, with $t = 2, \ldots, T$, and $u,v = 1, \ldots, k$. That is, differently from most previous latent class models, unobserved heterogeneity is modelled as a Markov process and this allows individuals switching between latent classes over time. The joint distribution of the latent process $\boldsymbol{U}_i$ may be written as:

$$\Pr(\boldsymbol{\alpha}_i) = \Pr(\boldsymbol{\alpha}_{i1} = \boldsymbol{U}(u)) \prod_{t>1}^{T_i} \Pr(\boldsymbol{\alpha}_{it} = \boldsymbol{U}(v)|\boldsymbol{\alpha}_{it-1} = \boldsymbol{U}(u)) = \pi_{u1} \prod_{t>1}^{T_i} \pi_{v_t|u_{t-1}} \qquad (4)$$

Equation (4) shows that individuals are allowed switching between classes. These probabilities may be computed using a forward recursion employed in the literature concerned with a specific type of Markov models termed hidden Markov models (see e.g., MacDonald and Zucchini, 1997; Bartolucci, 2006; Farcomeni, 2012). We estimate the system (2)-(4) by employing a logit model for UD_{it} and an ordered logit model for HC_{it}. More specifically, the dependent variable for UD_{it} is a binary variable defining the presence of UD among individuals with diabetes while the dependent variables for HC_{it} include the following ordered categories: 0,1,2,..>55 for GP visits and 0,1,2…> 33 for specialist visits, respectively. Here, we follow Cameron and Trivedi (1998) who suggest that ordered discrete models could be employed as an alternative to count data models. In addition, this specification can be also seen as a nonparametric alternative to count data models (e.g. Di Mari and Bakk, 2018; Vermunt and Magidson, 2005). In any case, the inclusion of a count data model such as Poisson or Negative Binomial would require extending this bivariate Latent Markov approach well beyond the scope of this paper.

The estimation is implemented by an EM algorithm (Dempster et al., 1977) and the corresponding log-likelihood of our model can be written as:

$$\ell(\boldsymbol{\theta}) = \sum_i \sum_u^k \left\{ \sum_t^{T_i} w_{itu} \log[p(\boldsymbol{y}_{it}|\boldsymbol{\alpha}_{it}, \boldsymbol{x}_{it})] + w_{i1u} \log[\pi_{u1}] + \sum_v z_{iuv} \log(\pi_{uv}) \right\} \qquad (5)$$

where $\boldsymbol{\theta}$ is the is short-hand notation for all the nonredundant model parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma$ and ρ, w_{itu} is a dummy variable which equals 1 if individual i is in latent state u at time t, while z_{iuv} equals the number of times individual i moves from state u to state v (from t >1). To estimate $\boldsymbol{\theta}$, we maximize $\ell(\boldsymbol{\theta})$ by using the EM algorithm proposed by Bartolucci and Farcomeni

9

(2009), that is a Newton–Raphson type iterative algorithm.[3] In particular, the EM algorithm alternates Expectation and Maximization steps until convergence. The E-step computes expected values of variables w_{itu} and z_{iuv} using posterior probabilities. The subsequent M-step maximises the expected values with respect to $\boldsymbol{\theta}$.[4]

4. Data

We employ individual-level longitudinal data drawn from administrative records of patients followed over six consecutive years (2005-2010) in six primary care centres and two hospitals in the municipality of Badalona (north-east of Barcelona), Spain, serving a population of around 104,000 individuals. As in Spain health care is primarily delivered by a national health care service with free universal access, data from these health care centres should be representative of the way the health system operates in the whole country.

Our data includes a rich set of information about patients' use of health care resources, including number of GP visits and specialist care. We focus on utilisation as it represents the largest component of diabetes-related medical costs. For example, Mata-Cases et al. (2016) find that in Catalonia hospital care, medications and primary care are the main drivers of costs for both patients with type-2 diabetes and non-diabetic patients. The data also encompasses information on a battery of diagnostic tests, measurements of height and weight (used to build an individual's body mass index, BMI); patient's chronic and diagnosed health conditions (classified according to the International Classification of Primary Care codes, second edition, ICPC-2); type of healthcare professional(s) contacted; and the main reason for their visit. Moreover, the dataset includes standard individual level socio-demographic characteristics such as age, gender, marital status, immigration (Spanish or EU national versus non-EU national) and employment status (active vs retired), place of birth and residence and health-behaviours (alcohol and tobacco use).

Given the main objective of our study, we focus on a sub-sample of individuals diagnosed with type-2 diabetes mellitus (DM), aged 16 years or over, who had at least one contact with the aforementioned health care centres during the study period. Diabetic patients were identified via the corresponding ICPC-2 code or alternatively through a mean value of the glycated

---

[3] We thank Bartolucci and Farcomeni for kindly providing the initial MATLAB codes, which we have amended to estimate our bivariate model. All adapted MATLAB functions are available upon request.

[4] We have also estimated an alternative specification where UD is used as a predictor of (initial) class membership probabilities. Results are very similar to the ones presented here and available upon request.

haemoglobin (HbA1c) test $\geq$ 6.5% ($\geq$ 48 mmol/mol). This test is routinely used by physicians to diagnose and monitor diabetes and provides an accurate measure of glucose concentration up until the previous 8 weeks (Goldstein et al. 2004; IEC, 2009; Lyons and Basu, 2012).

We define uncontrolled type-2 diabetes (UD) using a binary variable which equals 1 if the within year mean HbA1c level is equal or above 7.5%. This is the threshold actually employed by physicians to identify poor glycaemic control in our dataset. However, since there is no universal consensus within the medical literature about the HbA1c thresholds which identify UD, we also estimate our models using a slightly lower value of blood sugar concentration (7%).[5] Given the need to employ annual measurements of UD in the LM model, we restricted our sample to individuals with at least one within-year mean HbA1c value. Following this criteria, we obtained a sample of 2,455 patients with type-2 DM.[6]

In our models, dependent variables for health care utilisation, GP and specialist visits, are ordered variables with categories 0,1,2,..>55 and 0,1,2...> 33, respectively. Covariates include age; gender; immigration, employment and marital status; BMI; alcohol and tobacco use; and a series of diagnosed conditions: asthma, chronic obstructive pulmonary disease (COPD), dementia, psychosis, depression and cancer. In addition, all specifications include local (health) areas fixed effects and wave dummies.

Descriptive statistics

Table 1 reports variable definitions and some basic descriptive statistics for both the sample of patients with type-2 DM and the one including individuals with uncontrolled diabetes (HbA1c>=7.5%). Our pooled dataset (years 2005-2010) reveals that around 27 percent of patients are not controlling adequately their blood sugar levels, despite the well-known potential detrimental effects. In the sample of patients with diagnosed diabetes, the mean number of GP visits per year is around 17.2 while this falls to 4.9 per year for specialist visits.

---

[5] Results obtained using a lower threshold appeared to be very similar and are available upon request.

[6] Note that, given the differences between type-1 and type-2 diabetes and related treatments, we dropped all individuals with type-1 diabetes. Individuals transferred or moved to other health centres and patients from other areas were also excluded from our analysis. Since movements across centres are rare events, this exclusion is unlikely to have a major impact on our main results. We are aware that focusing only on individuals with annual values of UD may imply, at least in principle, that we might be employing those individuals with more regular or frequent measurement of blood sugar, that is individuals who could be more conscientious or risk averse (and potentially healthier). If that would be the case, then the estimates obtained by our models will likely be downwardly biased and represent the lower bound of the effects of UD on health care utilisation. However, this is unlikely to generate a systematic bias in our case as virtually all individuals with type-2 diabetes present at least an annual value of HbA1c.

As expected, these increase among individuals with UD to 18.8 and 5.3, respectively. Patients with UD appear to be more active in the labour market (23 vs 19 percent); to engage more frequently in risky-health behaviours (smoking, 21.3 vs 17.1 percent; drinking: 3.4 vs 2.6 percent); and suffer from higher levels of depression (21.5 vs 17.6) than diabetic patients with an adequate control.

(Table 1 around here)

5. Results

We estimated our Latent Markov specification for an increasing numbers of latent states for each of the two measures of health care utilisation: GP and specialist visits. We do this by first estimating a model with the no heterogeneity (i.e. k=1) followed by models estimated with an increasing number of latent states. This allows establishing the specific number of latent classes for each outcome in our final specifications. Results are reported in Table 2.

(Table 2 around here)

This table shows the log-likelihoods obtained from equation (4) together with the total number of parameters and the Bayesian Information Criterion (BIC; Schwarz, 1978) for all models estimated by imposing an increasing number of latent states. Given the nature of our response variables and the objective of our analysis, it would be more appropriate to rely on the BIC rather than the Akaike Information Criterion (AIC; Akaike, 1973), as the BIC tends to penalise model complexity more heavily.[7] According to these criteria, the specifications with the lowest BIC are the ones with four classes for GP visits and five classes for specialist. Below we report estimates for the three set of parameters produced by our LM model. These include estimates and partial effects of UD on GP and specialist visits; time-varying latent propensities to poor glycaemic control (UD) and health care use; and initial as well as transition probabilities between latent states over time.

---

[7] See, among others, McLachlan and Peel (2000) for a detailed analysis of the use of information criteria to select the number of components of a finite mixture model.

Partial effects

Tables 3 and 4 report estimated coefficients and average partial effects (APE) obtained using a bivariate Latent Markov model. The left hand-side of each table reports estimates for the logit model for UD while the right-hand side includes results for the jointly estimated ordered logit model for GP and specialist visits, respectively.

(Tables 3 and 4 around here)

Conditional on time-varying unobservables, the direct effect of UD on the number of both GP and specialist visits is positive but only weakly statistically significant (at 10% significance level, see Table 3) or not statistically significant (Table 4). This appears to suggest that low levels of glycaemic control do not increase health care use within our population of individuals with type-2 diabetes. Moreover, the absence of any residual association between UD and health care consumption (for both GP and specialist visits) implies that, after controlling for (time-varying) unobservables, the association between UD and health care use tend to disappear. This lends some further support to the idea that UD and health care utilisation might be jointly driven by a third set of common time-varying unobserved factors and that UD may not be a genuine determinant of health care use.[8]

Yet, a number of observed variables appear to be important drivers of utilisation. For example, in Table 3 (UD and GP visits), gender and age display positive and highly statistically significant coefficients with APE of around 0.056 and 0.004. This implies that women and older individuals present an increased probability of an additional GP visit of 5.6 and 0.4 percentage points (pp), respectively. As expected, a battery of health conditions including asthma, COPD, dementia, depression, and the occurrence of cancer also increase the average annual probability of an additional GP visit with the largest quantitative effects estimated for COPD and depression (14.3 and 9.3pp). Similarly, being employed decreases by 7.9pp the probability of a further annual GP visit as well as a higher BMI, although only by 0.6pp. Interestingly, only a handful of conditions such as COPD (7.7pp), depression (3.8pp) and cancer (4.8pp) seem to significantly increase the probability of a further annual visit to a specialist (Table 4). Also, being employed (-2.9pp), tobacco consumption (-2.0pp) and living within health areas 4 (-3.1pp) and 7 (-7.2pp) seem to decrease the same probability. Note that

---

[8] Following Gil et al. (2017) we also estimated our models by interacting UD with the gender variable in order to capture potential systematic gender-related differences. Results do not differ substantially and the estimated interaction terms for both models (GP and specialist visits) are not statistically significant. Estimates are available upon request.

given the complexity of the decision-making process leading to a specialist visit, which includes both a patient and the GP or the health care provider more generally, it is not surprising that the probability of an additional visit would mainly depend on the occurrence of major health conditions. Furthermore, the effects of the different administrative health areas on health care utilisation might reflect broader differences in socioeconomic status between individuals living in such areas as well as the other geographical factors, including proximity to hospitals and primary care centres.

Being a non-EU immigrant (19.2pp); living within a family or a partner (3.7pp); being a smoker (4.3pp); a higher BMI (2.9pp); and a series of diagnosed conditions such as dementia (16.8pp); depression (5.2pp); as well as living within specific health administrative areas (areas 4-7 with APE of 4.1, 6.6, 10.4 and 8.2pp, respectively) all appear to significantly increase the probability of poor glycaemic control in our LM model including UD and GP visits. Conversely, women (-2.0pp); being affected by psychosis (-8.6pp); and older age (-0.2pp) decrease the probability of poor control. As for our specification including UD and specialist visits, only variables related to specific health behaviours, such as smoking (4.3pp) and BMI (2.6pp), living within health areas 6 (6.4pp) and 7 (4.8pp), and more importantly dementia (23.2pp), increase the probability of UD. In addition, older individuals (-0.3pp) and living in health area 2 (-5.1pp) are less prone to poor glycaemic control.

Time-varying unobserved heterogeneity

Tables 5 and 6 report estimated intercepts $\boldsymbol{\alpha}_{UD}$ and $\boldsymbol{\alpha}_{HC}$ described in equation (2) for each of the estimated latent states for models with primary care (GP visits) and secondary care (specialist visits). We exploit these parameters to recover the conditional average probabilities for each model (last two columns of the tables). These could be interpreted as the propensities to poor glycaemic control (UD) and a more frequent use of health care between latent states. That is, a utilisation higher than the median quantity of health care for each latent state with the median values corresponding to 15 and 3 visits for GP and specialist visits, respectively.

(*Tables 5 and 6 around here*)

Table 5 reveals that there are four unobserved types of individuals with type-2 diabetes with varying propensities to UD and health care use (GP visits). More specifically, types 1 and 4 display similar high propensities towards health care consumption (around 57 and 59 percent, respectively). However, these two types differ markedly in their propensities to poor glycaemic control, i.e. type 4 presents a much higher propensity to UD (around 87 percent) if compared

to type 1 (about 6.5 percent) as well as types 2 and 3. Moreover, whereas types 2 and 3 share similar propensities to UD (around 32 and 23 percent), type 2 individuals appear much more likely to use health care than type 3 (95 versus 12.2 percent). This implies that not all individuals with a high propensity to GP visits present an equally high propensity to poor glycaemic control. Combinations of high propensities to health care consumption and adequate glycaemic control could point towards the presence of conscientious patients who tend to use more health care to better control their condition. Equally, types with combinations of high propensities to UD and GP visits (type 4) might include potentially less conscientious individuals who are not adequately following their prescribed health care treatments, ultimately resulting in extra utilisation.

The model estimated for secondary care presents five latent states (Table 6). Individuals in latent state 1 show relatively low propensities to both UD and specialist visits (around 11.5 and 19.2 percent) while individuals in latent states 3 and 4 present low propensities to UD (10 and 9.3 percent) but much larger ones to health care use (about 80 and 98 percent), thus also potentially revealing the presence of more conscientious individuals. Yet, type 5 includes individuals with large propensities to UD and secondary care. This might confirm the existence of patients not effectively controlling their blood sugar levels and more prone to consume secondary care. Still, it should be kept in mind that the propensities to additional GP and specialist visits should not be directly compared, as within the Spanish health care system access to secondary care would require a GP referral.

Changes between latent states over time

Our bivariate LM model enables us to investigate whether individuals move between latent states, therefore changing their propensities to UD and health care use over time. Tables 7 and 8 report estimated initial and transition probabilities between latent states across consecutive time periods. As for the initial probabilities, in both tables the largest proportion of individuals is expected to be in state 1. This is also the latent state from where patients are more likely to move out over time, as pointed out by the subsequent transition probabilities.

(Tables 7 and 8 around here)

Looking at transition probabilities (right-hand side of each table), the rows of each matrix contain previous (t -1) latent states whereas the columns represent current latent states. Both matrices show a symmetric structure, especially the one computed for the model with UD and GP visits (Table 7) i.e., individuals present high probabilities of remaining in the same latent

state in the following period with the largest "persistence" identified for individuals in state 2. However, both tables suggest the presence of transitions between latent states across time. For example, in Table 7 individuals in latent state 1 at t -1 present probabilities of moving towards latent states 2, 3 and 4 in the subsequent period of around 2, 4.3 and 5.9 percent, respectively. This implies that individuals with a high propensity to GP visits and a low propensity to poor glycaemic control (latent state 1) in one period could change their behaviour over time and present substantially higher propensities to UD and GP visits (latent states 2 and 4) as well as higher propensities to UD and a lower propensity to GP visits (latent state 3) in subsequent periods. Similarly, individuals in latent state 4 could migrate to latent states 1 (3.8), 2 (2.9) and 3 (0.7) and present a much lower propensity to poor glycaemic control while still having a high propensity to consume health care (especially in latent state 2).

Estimated transitions probabilities in Table 8 also suggest that there might be changes in the propensities to UD and specialist visits between consecutive periods. These movements appear to be concentrated among patients initially in latent state 1 (low propensities to UD and health care use) who could move to type 3 (16 percent; still low UD but very high health care use) and to a lesser extent to types 2 (7 percent; very high UD and higher health care use) and 4 (0.2 percent; low UD and also very high health care use) in subsequent periods. Type 2 individuals (very high UD and low health care use) might also fluctuate, especially towards state 5 (15.7 percent; very high UD and health care use) while type 3 (low UD and high health care use) have increased probabilities to switch to type 4 (8 percent; low UD and very high health care use). Overall, the estimated transitions appear to suggest further increases in the propensities to UD or specialist visits, mostly among patients with already higher propensities for these.

Moreover, we tested the hypothesis of diagonal transition matrices in our models, i.e. that there are no transitions between latent states (or that unobserved heterogeneity for UD and health care use is time-invariant). We followed a test specifically designed for LM models proposed by Bartolucci (2006), who showed that the asymptotic null distribution of the likelihood ratio (LR) statistic is of $\bar{\chi}^2$ type. In our case, the LR statistics for models with GP and specialist visits are equal to 141.05 and 280.03 with 12 and 20 degree of freedom, respectively. This led us to strongly reject the null hypothesis of diagonal transition matrices. This also appears to suggest that unobserved heterogeneity might not be time-invariant.

The dynamics of the time-varying unobserved heterogeneity can be explored further by computing the year-specific probabilities of each latent state (Figures 1 and 2).

(Figures 1 and 2 about here)

As for our specification including primary care (Figure 1), while probability of being in latent state 1 is the highest in all periods, it also appears to decline over time. Conversely, the probabilities of being in latent states 2, 3 and 4 appear to slightly increase over time. Overall, this appears to suggest that individuals tend to become more inclined to poor glycaemic control *and* health care use over time (apart from individuals in latent state 3 who present a higher propensity to UD, but a comparatively lower propensity to additional GP visits). The same probabilities computed for the model with UD and secondary care (Figure 2) display a substantial and constant decrease in the probability of being in latent state 1 as well as similar consistent increase in the probability of being in latent state 3 (low UD and very high health care use) over time. Furthermore, the probabilities of being in the remaining latent states also appear to increase between 2006 and 2010, especially the one for latent class 4 (very high propensities to UD and health care use).

6. Conclusions and discussion

This paper proposes a bivariate Latent Markov (LM) model to explore the relationship between uncontrolled diabetes (UD) and health care utilisation on a sample of patients with type-2 diabetes drawn from longitudinal administrative data collected in Spain. Our approach offers a series of substantial and innovative modelling features, which enable accounting simultaneously for time-varying unobserved heterogeneity; changes in individuals' behaviours over time via switches between latent classes; and the endogeneity of uncontrolled diabetes. Until now, these potentially important issues have been largely ignored, especially by the previous medical literature concerned with UD. Furthermore, whilst UD is a widespread condition with potentially severe health consequences, there is still sparse evidence on its effects on health care utilisation outside the US and within the economic literature. Finally, previous latent class models used to investigate the determinants of health care utilisation do not generally account for dynamic unobserved heterogeneity.

We estimate our LM model to quantify the effects of UD on the consumption of both primary and secondary health care. Contrary to previous evidence based on univariate models, we find that UD does not appear to have a significant direct effect on GP or specialist visits, given the dynamic latent process. Moreover, the association between UD and health care use tend to disappear when controlling for unobservable factors, suggesting that these may drive part of their otherwise perceived association. Our specifications also suggest that such unobserved

17

factors are time-varying by rejecting time-invariant unobservables and identifying the presence of individuals switching between latent classes over time. These movements between classes show a tendency towards higher propensities to poor glycaemic control and health care use, potentially reflecting changes in behaviours. Our models also reveal differences in the propensities to poor glycaemic control and health care utilisation between latent states. This suggests the presence of heterogeneity in behaviours involving UD and health care use within our sample of patients, which may go beyond the standard groupings of "*frequent*" versus "*infrequent*" users. Ultimately, by ignoring time-varying unobserved heterogeneity and the endogeneity of uncontrolled diabetes, the effects of UD on health care utilisation might be overestimated and this could lead to biased findings and misleading policy implications.

Overall, our findings appear to challenge the conventional wisdom within the current literature advocating that UD is one of main drivers of health care utilisation and related costs among patients with diabetes. This may have important implications for physicians, policy makers and more generally the development of diabetes management plans, potentially suggesting the need to reconsider the causes of excess health care utilisation.

Tables and figures

Table 1: Variables definition and descriptive statistics: pooled data 2005-2010

Variable	Description	All sample (N=2455)		Uncontrolled DM Patients (N=655)	
		Mean	SD	Mean	SD
UD	Uncontrolled Type-2 diabetes (*)	0.267	0.44	-	-
Mean HbA1c	Within year mean HbA1c value	6.983	0.95	8.47	1.00
GP visits	Number of GP visits	17.194	9.74	18.804	10.65
Spec. visits	Number of specialist visits	4.968	5.33	5.327	5.62
Gender	1= female, 0 otherwise	0.528	0.50	0.511	0.50
Age	Age in years	66.534	9.78	66.423	10.34
Immigrant	1 = Non-EU immigrant, 0 otherwise	0.007	0.08	0.013	0.11
Employment	1 = Labour market active, 0 otherwise	0.191	0.39	0.229	0.42
Marital status	1 = Not living alone, 0 otherwise	0.880	0.32	0.887	0.32
Alcohol	1 = Drink alcohol, 0 otherwise	0.026	0.16	0.034	0.18
Tobacco	1 = Smoker, 0 otherwise	0.171	0.38	0.213	0.41
BMI	Body Mass Index (BMI)	30.169	4.82	30.52	5.11
Asthma	1 = Ashma, = otherwise	0.047	0.21	0.046	0.21
COPD	1 = COPD, 0 otherwise	0.054	0.23	0.054	0.22
Dementia	1 = Dementia, 0 otherwise	0.010	0.10	0.015	0.12
Psychosis	1 = Phycosis, 0 otherwise	0.010	0.10	0.011	0.11
Depression	1 = Depression, 0 otherwise	0.176	0.38	0.215	0.41
Cancer	1 = Malignant neoplasm, 0 otherwise	0.066	0.25	0.065	0.25
Area 1	1 = Health area 1	0.083	0.28	0.073	0.26
Area 2	1 = Health area 3	0.137	0.34	0.097	0.30
Area 3	1 = Health area 8	0.164	0.37	0.158	0.37
Area 4	1 = Health area 9	0.205	0.40	0.215	0.41
Area 5	1 = Health area 10	0.121	0.33	0.120	0.32
Area 6	1 = Health area 12	0.159	0.37	0.180	0.38
Area 7	1 = Health area 15	0.132	0.34	0.157	0.36

Note: (*) Uncontrolled type-2 diabetes mellitus is defined for HbA1c values $\geq 7.5\%$.

Table 2: Model selection criteria

k	GP			Specialist		
	LL	#par	BIC	LL	#par	BIC
1	-39110.64	108	79064.31	-31957.58	86	64586.46
2	-37885.96	113	76653.98	-30884.95	91	62480.23
3	-37427.42	120	75791.56	-30269.26	98	61303.49
4	-37015.81	129	75038.58	-29959.95	107	60755.13
5	-37639.51	140	76371.84	-29747.98	118	60417.05
6				-29718.20	131	60458.96

Notes: k= latent classes; LL = log-likelihood; #par=number of estimated parameters; BIC=Bayesian Information Criterion; GP = number of GP visits; Specialist = number of specialist visits.

Table 3: Bivariate latent Markov model - UD and GP visits

	UD				GP			
	Coeff.	SE	Z	APE	Coeff.	SE	Z	APE
UD					0.4791	0.28	1.69	0.0491
Gender	-0.2047	0.10	-2.13	-0.0207	0.5471	0.07	7.31	0.0562
Age	-0.0202	0.01	-3.45	-0.0020	0.0439	0.00	9.70	0.0045
Nationality	1.5442	0.39	3.96	0.1920	-0.2048	0.30	-0.67	-0.0211
Employment	0.1539	0.15	1.06	0.0157	-0.7657	0.11	-6.84	-0.0787
Marital status	0.3898	0.15	2.67	0.0375	0.0181	0.10	0.17	0.0019
Alcohol	0.2046	0.24	0.87	0.0213	0.1384	0.17	0.80	0.0143
Tobacco	0.4079	0.12	3.53	0.0429	-0.1100	0.10	-1.15	-0.0113
BMI	0.2838	0.04	7.53	0.0287	-0.0582	0.02	-2.36	-0.0060
Asthma	-0.0959	0.21	-0.46	-0.0095	0.6242	0.18	3.54	0.0637
COPD	-0.0248	0.19	-0.13	-0.0025	1.4532	0.13	11.55	0.1438
Dementia	1.3766	0.32	4.27	0.1678	0.6296	0.30	2.12	0.0642
Psychosis	-0.9846	0.48	-2.07	-0.0858	0.0636	0.31	0.20	0.0066
Depression	0.4875	0.11	4.51	0.0517	0.9270	0.08	11.39	0.0938
Cancer	0.0822	0.15	0.55	0.0084	0.3164	0.11	2.76	0.0325
Area 2	-0.0385	0.22	-0.17	-0.0039	-0.1442	0.15	-0.93	-0.0149
Area 3	0.1632	0.20	0.80	0.0168	-0.0659	0.14	-0.47	-0.0068
Area 4	0.3962	0.19	2.04	0.0415	0.3170	0.14	2.29	0.0326
Area 5	0.6080	0.21	2.96	0.0658	0.1193	0.15	0.80	0.0123
Area 6	0.9367	0.20	4.75	0.1041	0.1995	0.14	1.38	0.0205
Area 7	0.7505	0.21	3.56	0.0823	0.4361	0.15	2.96	0.0447
Wave dummies	yes				yes			

Residual correlation

Corr.					-0.3186	0.2842	-1.1212	

Notes: this table displays coefficients and average partial effects (APE) obtained from a bivariate Latent Markov Model. UD = uncontrolled diabetes; GP = number of GP visits; SE = standard errors; Z = z-score; APE = Average Partial Effects. This specification includes the full set of covariates, health area fixed effects and time dummies.

Table 4: Bivariate latent Markov model - UD and specialist visits

	UD				Specialist			
	Coeff.	SE	Z	APE	Coeff.	SE	Z	APE
UD					0.5651	0.35	1.60	0.0477
Gender	-0.1457	0.12	-1.20	-0.0128	0.0908	0.08	1.07	0.0076
Age	-0.0293	0.01	-3.71	-0.0026	0.0084	0.01	1.56	0.0007
Nationality	1.9173	0.40	4.85	0.2293	0.5038	0.39	1.28	0.0425
Employment	-0.0438	0.17	-0.26	-0.0038	-0.3414	0.11	-3.07	-0.0287
Marital status	0.0874	0.18	0.48	0.0076	0.1770	0.11	1.61	0.0149
Alcohol	0.3888	0.31	1.27	0.0362	-0.0994	0.19	-0.53	-0.0084
Tobacco	0.4691	0.14	3.39	0.0430	-0.2376	0.10	-2.41	-0.0200
BMI	0.2997	0.04	6.85	0.0263	-0.0253	0.03	-0.90	-0.0021
Asthma	0.1853	0.24	0.77	0.0167	0.2358	0.14	1.64	0.0199
COPD	-0.1163	0.25	-0.47	-0.0101	0.9211	0.14	6.48	0.0774
Dementia	1.9392	0.34	5.78	0.2319	0.3771	0.32	1.19	0.0318
Psychosis	-0.5936	0.54	-1.10	-0.0481	-0.8028	0.34	-2.33	-0.0666
Depression	0.2235	0.14	1.63	0.0201	0.4475	0.08	5.33	0.0377
Cancer	0.2215	0.18	1.21	0.0201	0.5663	0.13	4.41	0.0478
Area 2	-0.6245	0.29	-2.17	-0.0510	-0.1759	0.15	-1.15	-0.0148
Area 3	0.0943	0.25	0.38	0.0084	-0.1526	0.16	-0.98	-0.0128
Area 4	0.1980	0.24	0.83	0.0177	-0.3770	0.15	-2.55	-0.0316
Area 5	0.2514	0.26	0.98	0.0228	-0.0679	0.15	-0.44	-0.0057
Area 6	0.6779	0.25	2.74	0.0641	-0.2175	0.15	-1.44	-0.0183
Area 7	0.5137	0.25	2.09	0.0479	-0.8700	0.16	-5.41	-0.0724
Wave dummies	yes				yes			

Residual correlation

					Corr.			
Corr.					-0.5919	0.3567	-1.6593	

Notes: this table displays coefficients and average partial effects (APE) obtained from a bivariate latent Markov Model. UD = uncontrolled diabetes; Specialist = number of specialist visits; SE = standard errors; Z = z-score; APE = Average Partial Effects. This specification includes the full set of covariates, health area fixed effects and time dummies.

Table 5: Latent states - GP visits

Latent states	Estimated intercepts				Conditional average probabilities	
	α_{UD}	SE	α_{HC}	SE	UD	GP
1	4.6249	1.19	6.3122	0.94	0.0656	0.5683
2	6.6825	1.20	9.3181	0.98	0.3181	0.9475
3	6.1784	1.21	3.6631	0.95	0.2291	0.1220
4	9.6861	1.22	6.4020	1.03	0.8721	0.5864

Notes: this table reports estimated α_{UD} and α_{HC} as described by eq. (2), and conditional average probabilities of reporting UD=1 and reporting a number of annual GP visits higher than the median value of 15 visits.

Table 6: Latent states – specialist visits

Latent states	Estimated intercepts				Conditional average probabilities	
	α_{UD}	SE	α_{HC}	SE	UD	Specialist
1	5.7392	1.40	-0.9789	0.90	0.1152	0.1926
2	10.4198	1.43	-0.4554	1.05	0.8888	0.2806
3	5.5758	1.39	2.0110	0.90	0.1007	0.7956
4	5.4764	1.39	4.6897	0.91	0.0927	0.9815
5	9.5282	1.42	3.2509	1.05	0.7812	0.9278

Notes: this table reports estimated α_{UD} and α_{HC} as described by eq. (2), and conditional average probabilities of reporting UD=1 and reporting a number of annual specialist visits higher than the median value of 3 visits.

Table 7: Initial and transition probabilities between latent states - UD and GP visits

Latent states	Initial Prob.	Transition Prob. 1	2	3	4
1	0.6149	0.8774	0.0201	0.0435	0.0590
2	0.1134	0.0003	0.9996	0.0000	0.0001
3	0.1543	0.0067	0.0000	0.9800	0.0133
4	0.1174	0.0379	0.0291	0.0073	0.9257

Table 8: Initial and transition probabilities between latent states – UD and specialist visits

Latent states	Initial Prob.	Transition Prob. 1	2	3	4	5
1	0.5270	0.7676	0.0694	0.1611	0.0020	0.0000
2	0.0805	0.0157	0.7921	0.0350	0.0000	0.1572
3	0.2413	0.0076	0.0187	0.8666	0.0791	0.0280
4	0.0804	0.0035	0.0000	0.0283	0.9268	0.0414
5	0.0708	0.0000	0.0331	0.0002	0.0268	0.9400

Figure 1: Estimated year-specific probability of each latent state – GP visits
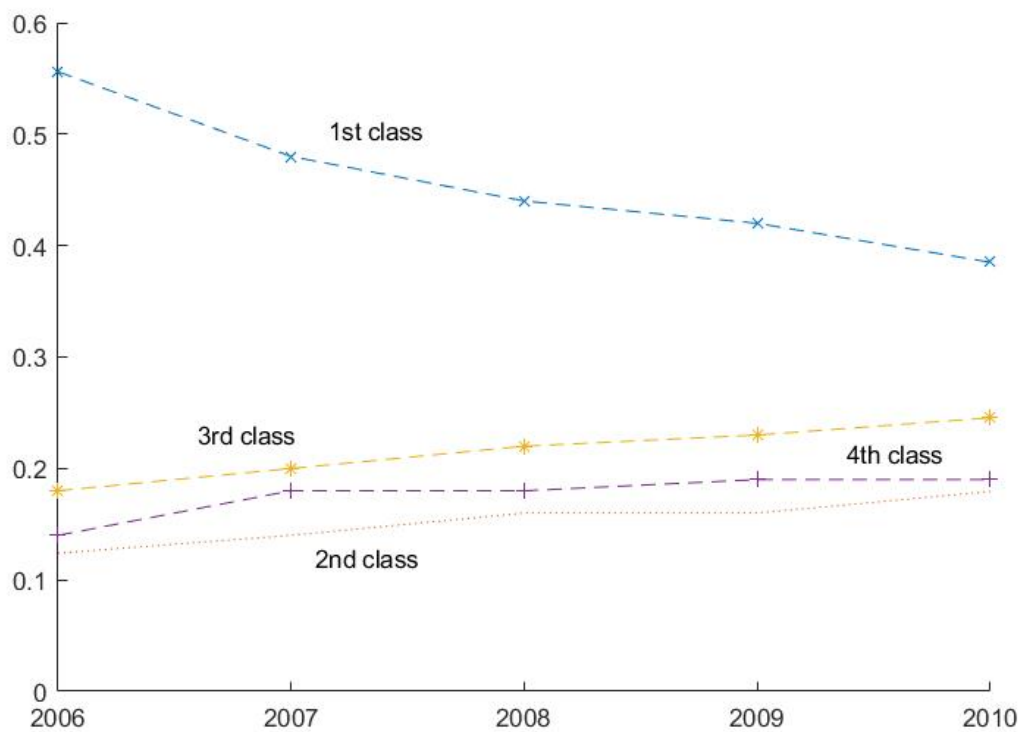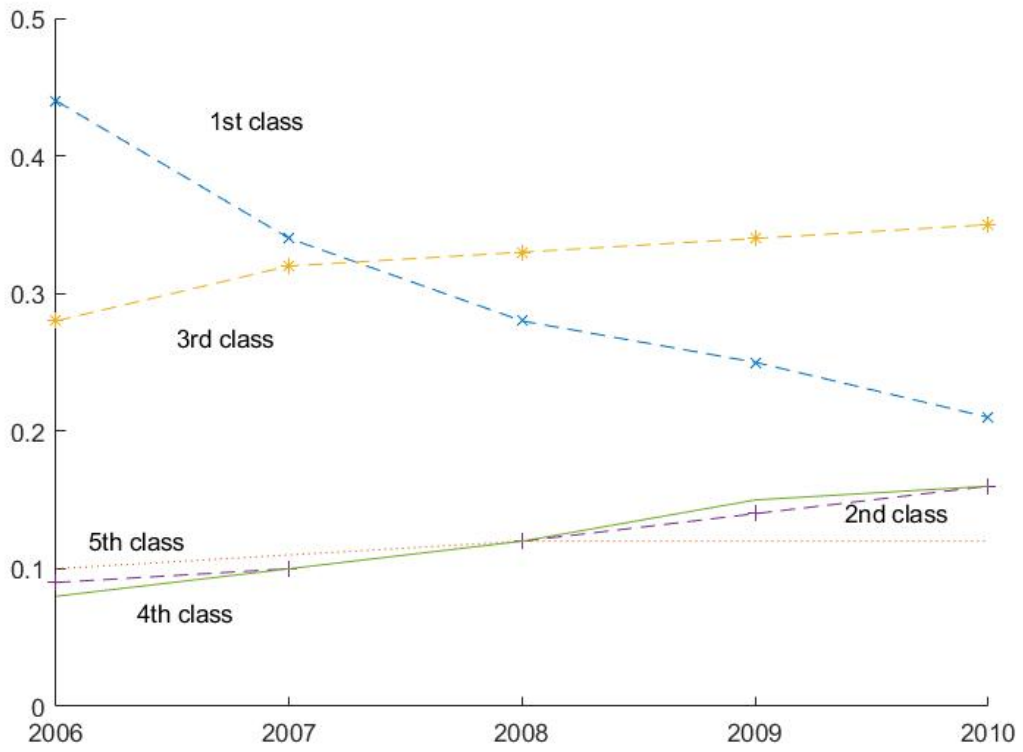
Figure 2: Estimated year-specific probability of each latent state– specialist visits

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Proc. 2nd International Symposium on Information Theory (Petrov, B.N. and Csàky, F., eds.). Budapest: Akademiai Kiàdo.

Atella, V., Brindisi, F., Deb, P., & Rosati, F.C. (2004). Determinants of access to physician services in Italy: a latent class seemingly unrelated probit approach. *Health Economics*, 13(7), 657-668.

American Diabetes Association (2013). Economic costs of diabetes in the US in 2012. *Diabetes Care*, 36(4), 1033-1046.

Bago d'Uva, T. (2005). Latent class models for use of primary care; evidence from a British panel. *Health Economics*, 14(9), 873-892.

Bago d'Uva, T. (2006). Latent class models for utilisation of health care. *Health Economics*, 15(4), 329-343.

Bago d'Uva, T., & Jones, A. (2009). Health care utilisation in Europe: new evidence from the ECHP. *Journal of Health Economics*, 28(2), 265-279.

Bartolucci, F. (2006). Likelihood inference for a class of latent Markov models under linear hypotheses on the transition probabilities. *Journal of the Royal Statistical Society:* Series B (Statistical Methodology), 68(2), 155-178.

Bartolucci, F., Lupparelli, M., & Montanari, G.E. (2009). Latent Markov model for binary longitudinal data: an application to the performance evaluation of nursing homes. *Annals of Applied Statistics*, 3(2), 611-636.

Bartolucci, F., & Farcomeni, A. (2009). A multivariate extension of the dynamic logit model for longitudinal data based on a latent Markov heterogeneity structure. *Journal of the American Statistical Association*, 104(486), 816-831.

Bartolucci, F., Farcomeni, A., & Pennoni, F. (2013). Latent Markov Models for Longitudinal Data. Boca Raton, FL: Chapman and Hall-CRC Press.

Bartolucci, F., Bacci, S., & Pennoni, F. (2014). Longitudinal analysis of self-reported health status by mixture latent auto-regressive models. *Journal of the Royal Statistical Society. Series C*, 63 (Part 2), 267-288.

Bartolucci, F., & Farcomeni, A. (2015). A discrete time-event history approach to informative drop-out in mixed latent Markov models with covariates. *Biometrics*, 71, 80-89.

Bastida, E., & Pagán, J.A. (2002). The impact of diabetes on adult employment and earnings of Mexican Americans: findings from a community based study. *Health Economics*, 11(5), 403-413.

Cameron, A.C., & Trivedi, P.K. (1998). Regression Analysis of Count Data. Econometric Society Monograph 30, Cambridge University Press, New York.

Colombi, R., & Forcina, A. (2001). Marginal regression models for the analysis of positive association of ordinal response variables. *Biometrika*, 88(4), 1007-1019.

Deb, P., & Holmes, A.M. (2000). Estimates of use and costs of behavioural health care: a comparison of standard and finite mixture models. *Health Economics*, 9(6), 475-489.

Deb, P., & Trivedi, P.K. (1997). Demand for medical care by the elderly: a finite mixture approach. *Journal of Applied Econometrics*, 12(3), 313-336.

Deb, P., & Trivedi, P.K. (2002). The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics*, 21(4), 601-625.

Dempster, A. P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*. Series B (Methodological) 39 (1), 1-38.

Di Mari, R., & Bakk, Z. (2018). Mostly harmless direct effects: a comparison of different latent Markov modelling approaches. *Structural Equation Modelling: A Multidisciplinary Journal*, 25(3), 467-483.

Farcomeni, A. (2012). Quantile regression for longitudinal data based on latent Markov subject-specific parameters. *Statistics and Computing*, 22(1), 141-152.

Gil, J., Sicras-Mainar, A., & Zucchelli, E. (2017). Uncontrolled diabetes and health care utilisation: panel data evidence from Spain. *European Journal of Health Economics*, 19(6), 785-795.

Goldstein, D. E., Little, R. R., Lorenz, R. A., Malone, J. I., Nathan, D., Peterson, C. M., & Sacks, D. B. (2004). Tests of glycaemia in diabetes. *Diabetes Care*, 27(7), 1761-1773.

Goodman, L.A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2): 215-231.

Hyppolite, J., & Trivedi, P. (2012), Alternative approaches for econometric analysis of panel count data using dynamic latent class models (with application to doctor visits data). *Health Economics*, 21(Suppl. 1), 101-128.

International Expert Committee (2009). International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care*, 32(7), 1327-1334.

International Diabetes Federation (IDF), Diabetes Atlas, 6th Edition, 2015. http://www.idf.org/diabetesatlas.

Jimenez-Martin, S., Labeaga, J.M., & Martinez-Granado, M. (2002). Latent class versus two-part models in the demand for physician services across the European Union. *Health Economics*, 11(4), 301-321.

Kahn, M. E. (1998). Health and labor market performance: the case of diabetes. *Journal of Labor Economics*, 16(4), 878-899.

Latif, E. (2009). The effect of diabetes on employment in Canada. *Health Economics*, 18(5), 577-589.

Lazarsfeld, P.F. & Henry, N.W. (1968). Latent Structure Analysis. Boston: Houghton Mifflin.

Li Donni, P. (2018). The unobserved pattern of material hardship and health among older Americans, *mimeo*.

Lyons, T.J., & Basu, A. (2012). Biomarkers in diabetes: hemoglobin A1c, vascular and tissue markers. *Translational Research*, 159(4), 303-312.

Mata-Cases, M., Casajuana, M., Franch-Nadal, J., Casellas, A., Castell, C., Vinagre, I., Mauricio, D. & Bolibar, B. (2016). Direct medical costs attributable to type 2 diabetes mellitus: a population-based study in Catalonia, Spain. *European Journal of Health Economics*, 17(8): 1001-1010.

MacDocald, I.L., & Zucchini, W. (1997). Hidden Markov and Other Models for Discrete-Valued Time Series. London: Chapman and Hall.

McCullagh, P., & Nelder, J.A. (1989). Generalised Linear Models. 2nd Edition. London: Chapman and Hall.

McLachlan, G., & Peel, D. (2000). Finite Mixture Models. New York, NY: John Wiley & Sons.

Minor, T. (2010). The effects of diabetes on female labor force decisions: new evidence from the National Health Interview Survey. *Health Economics*, 20(12), 1468-1486.

Montanari, G.E., & Pandolfi, S. (2018). Evaluation of long-term health care services through latent a Markov model with covariates. *Statistical Methods & Applications*, 27(1), 151-173.

Ng, Y.C., Jacobs, P., & Johnson, J.A. (2001). Productivity losses associated with diabetes in the US. *Diabetes Care*, 24(2), 257-261.

Oglesby A.K., Secnik, K., Barron, J., Al-Zakwani, I., & Lage, M.J. (2006). The association between diabetes related medical costs and glycemic control: a retrospective analysis. *Cost Effectiveness and Resource Allocation*, 4:1.

Pohlmeier, W. & Ulrich, V. (1995). An econometric model of the two-part decision making process in the demand for health care. *Journal of Human Resources*, 30(2), 339-361.

Rizzo, J. A., Abbott, T. A., & Pashko, S. (1996). Labour productivity effects of prescribed medicines for chronically ill workers. *Health Economics*, 5(3), 249-265.

Santos Silva, J.M.C., & Windmeijer, F. (2001). Two-part multiple spell models for health care demand. *Journal of Econometrics*, 104(1), 67-89.

Schellhorn, M., Stuck, A.E., Minder, C.E., & Beck. J.C. (2000). Health services utilisation of elderly Swiss: evidence from panel data. *Health Economics*, 9(6), 533-545.

Shetty, S., Secnik, K., & Oglesby, A.K. (2005). Relationship of glycemic control to total diabetes-related costs for managed care health plan members with type 2 diabetes. *Journal of Management Care Pharmacy*, 11(7), 559-564.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464.

Tunceli, K., Bradley C.J., Nerenz, D., Williams, L.K., Pladevall, M., & Lafata, J.E. (2005). The impact of diabetes on employment and work productivity. *Diabetes Care*, 28(11), 2662-2667.

UK Diabetes Global Health Community (2014). *The Cost of Diabetes in the UK*. (Information available at http://www.diabetes.co.uk).

Van Ourti, T. (2004). Measuring horizontal inequity in Belgian health care using a Gaussian random effects two part count data model. *Health Economics*, 13(7), 705-724.

Vermunt, J.K., & Magidson, J. (2005). Factor analysis with categorical indicators: a comparison between traditional and latent class approaches. In: Van der Ark, L., Croon, M., & Sijtsma, K. (eds.) New Developments in Categorical Data Analysis for the Social and Behavioural Sciences. Psychology Press, London.

Wagner E.H., Sandhu, N., Newton, K.M., McCulloch, D.K., Ramsey, S.D., & Grothaus, L.C. (2001). Effect of improved glycemic control on health care costs and utilization. *Journal of the American Medical Association*, 285(2), 182-189.

World Health Organisation (WHO) (2016). *Global Report on Diabetes*.

Zhang, X., Zhao, X., & Harris, A. (2009). Chronic diseases and labour force participation in Australia. *Journal of Health Economics*, 28(1), 91-108.