HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 16/07

The effects of non-adherence on health care utilisation: panel data evidence on uncontrolled diabetes

Joan Gil, Antonio Sicras-Mainar & Eugenio Zucchelli

June 2016

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

The effects of non-adherence on health care utilisation:

panel data evidence on uncontrolled diabetes

03/06/2016

Joan Gil[1]*

Antonio Sicras-Mainar[2]

Eugenio Zucchelli[3]

Abstract

Despite size and relevance of non-adherence to health treatments, robust evidence on its effects on health care utilisation is very limited. We focus on non-adherence to diabetes treatments, a widespread problem, and employ longitudinal administrative data from Spain (2004-2010) to identify and quantify the effects of uncontrolled type 2 diabetes on health care utilisation. We use a biomarker (glycated haemoglobin, HbA1c) to detect the presence of uncontrolled diabetes and explore its effects on both primary and secondary health care. We estimate a range of panel count data models, including negative binomials with random effects, dynamic and hurdle specifications to account for unobserved heterogeneity, previous utilisation and selection. We find uncontrolled diabetes in around 30% of patients of both genders. Although women appear to systematically consume more health care compared to men, their consumption levels do not appear to be influenced by uncontrolled diabetes. Conversely, among men uncontrolled diabetes increases the average number of GP visits per year by around 4%, specialist visits by 4.4% and greatly extends hospital length of stay.

Key words: non-adherence; diabetes; biomarkers; health care utilisation; panel count data

JEL codes: C23; I1; I12

Acknowledgments: We are grateful to Bruce Hollingsworth, Paolo Li Donni, Ian Walker, Will Whittaker and participants to the 2015 UK Health Economists Study Group (HESG) Conference, XXXV Spanish Health Economics Association Conference and the 11th World Congress of the International Health Economics Association (iHEA) for helpful comments. We gratefully acknowledge funds from the Catalan Government through grant (2014-SGR-1257). The usual disclaimer applies.

[1] * Corresponding author: Department of Economics and CAEPS, University of Barcelona (UB); Address: 696 Diagonal Ave., 08034 Barcelona (Spain). Email: joangil@ub.edu

[2] Badalona Serveis Assistencials (BSA), Badalona (Barcelona), Spain. Email: asicras@bsa.cat

[3] Health Economics at Lancaster (HEAL), Division of Health Research (DHR), Lancaster University. Email: e.zucchelli@lancaster.ac.uk.

1. Introduction

The medical literature consistently finds that around 20 to 50% of patients worldwide do not adhere to medical treatments. This has potential serious consequences for their health and in terms of increasing costs for the health care system (e.g. Haynes *et al.*, 2002; DiMatteo, 2004; Sokol *et al.*, 2005; Kripalani *et al.*, 2007). Non-adherence concerns about 50% of individuals suffering from chronic health conditions (see e.g. Jackevicius *et al.*, 2002, and Osterberg and Blaschke, 2005) and has been defined by the WHO (2003) as a worldwide problem of substantial importance. Furthermore, low or non-adherence may reduce benefits of health treatments and distort the assessment of their effectiveness (e.g. Vander Stichele, 1991). Among individuals with diabetes mellitus (DM), the most prevalent chronic disease in nearly all countries (International Diabetes Federation (IDF) Diabetes Altas, 2014), adherence to treatment is defined as the extent to which patients comply with the agreed recommendations on lifestyles and medications from the health care provider (WHO, 2003; Garcia-Perez *et al.*, 2013).[4] Among patients with DM non-adherence is especially common, leading to uncontrolled diabetes (i.e. poor glycaemic control resulting in higher levels of blood sugar or hyperglycemia) and potentially exposing individuals to higher risks of life-threatening comorbidities such as heart disease and stroke as well as vision problems and blindness (Ho *et al.*, 2006; Mayo Clinic, 2014).

The detrimental effects of non-adherence to diabetes treatments and the resulting condition of uncontrolled diabetes are likely to be exacerbated by the increasing prevalence of DM. According to the WHO (2016), we are currently experiencing a global diabetes epidemic with DM affecting 422 million adult individuals worldwide in 2014, compared to 108 million in 1980 and projected to be 7th leading cause of death by 2030. In the U.S. alone, the estimated total costs of diagnosed diabetes increased by 41% in a five year period, from $174 billion in 2007 to an estimated $245 billion in 2012 (American Diabetes Association, 2013). In the UK the cost of diabetes to the NHS is over £1.5m an hour or 10% of the NHS budget for England and Wales (UK Diabetes Global Health Community, 2014).

A related problem may be caused by the additional health care utilisation and in turn the extra costs caused by uncontrolled diabetes. Individuals with uncontrolled diabetes may potentially

[4] Recommendations on lifestyles often include an exercise regimen as well as a low-carbohydrate and low-fat diet. In early stages of DM, these might be followed by medications such as oral hypoglycaemic agents and later on by injectable treatments. For further information, see recommendations from the International Diabetes Federation (IDF), the umbrella organisation of 230 national diabetes associations from 170 countries, http://www.idf.org/.

use more health care resources resulting in *extra utilisation* and extra costs for the health system as well as further loss in productivity due to work absence, work limitations, lower earnings and early mortality (e.g. Ng *et al*., 2001; Bastida and Pagan, 2002; Tunceli *et al.*, 2005).

Whereas non-adherence and uncontrolled diabetes are widespread phenomena, robust evidence on their effects on health care utilisation is very limited. Furthermore, the previous medical evidence is often based on selected samples, mostly from U.S. health insurance claims, and standard methods such as linear regression models. Selected samples would limit the generalisability of the results while standard linear models would not account for potentially relevant issues such as unobserved heterogeneity and selection which may affect the estimates of the effects of non-adherence. A more accurate assessment of the health care consequences of uncontrolled diabetes would help enhancing the evidence based for physicians to better plan treatments and for policy-makers to develop cost-effective interventions to increase adherence rates and reduce excess utilisation leading to resource inefficiencies.

The main objective of this paper is to identify and quantify changes in healthcare utilisation driven by uncontrolled diabetes. We employ seven waves (2004-2010) of a longitudinal and large administrative dataset of detailed medical records of adult diabetic patients from the province of Barcelona, Spain. Our data covers a wide range of individual-level clinically assessed biomarkers and health variables. We measure uncontrolled diabetes using glycated haemoglobin (HbA1c) levels. HbA1c is a biomarker providing an accurate average measurement of glucose (sugar) concentration and is commonly used by physicians to diagnose and monitor diabetes as well as its level of severity (e.g. International Expert Committee, 2009; Lyons and Basu, 2012). We focus on the effects of uncontrolled diabetes on both primary care and secondary care (number of visits to general practitioners, GP, and specialists) as well as hospital length of stay (i.e. number of days in hospital).

We exploit the longitudinal nature of the data and estimate a succession of *panel* count data models, including random effects negative binomials as well as dynamic and hurdle specifications to account for unobserved heterogeneity, previous utilisation and potential selection issues. We find that patients with uncontrolled diabetes appear to increase their annual use of primary health care by around 4% and the one of specialist visits by around 4.4%. Uncontrolled diabetes appears to greatly extend length of stay (17.4%), conditional on

positive stays. However, these effects appear to be mainly concentrated among men and differ according to the level of uncontrolled diabetes. A series of robustness checks validate our main findings.

This paper offer several contributions to the literature. First, to the best of our knowledge, this is one of the first papers focusing on the effects of non-adherence on health care utilisation within the health economics literature. Secondly, we combine the use of longitudinal data with biomarkers from a large set of adult diabetic patients to quantify the effects of uncontrolled diabetes, a prevalent condition, on health care utilisation by employing robust panel data econometric methods. Finally, we go beyond studies based on US health insurance enrolees and present evidence based on rich administrative data from a European country. More broadly, we advance the literature by merging the medical literature on non-adherence and diabetes care with the economics literature on health care utilisation to produce a new evidence base for better informed interventions.

The paper proceeds as follows. Section 2 provides an overview of the previous literature while sections 3 and 4 describe the data and outlines our empirical approach, respectively. Section 5 presents our main results and section 6 concludes and discusses our findings.

2. Background

Wagner et al. (2001) employ US insurance claims data of adult diabetics from a large Health Maintenance Organization (HMO) in Seattle (state of Washington) between 1992-1997 to analyse the association between improved glycaemic control and health care utilisation and costs. They employ standard linear and log-linear regression models and find that sustained reduction of HbA1c levels is correlated with lower utilisation (measured by hospitalisation rates as well as primary and specialty care visits) and cost savings, although cost reductions are significant only among patients with the highest baseline HbA1c levels (over 10%) and do not appear to be affected by further baseline health care conditions. Gilmer *et al.* (2005) combine both claims and survey data of diabetic adults in Minnesota with generalised linear models (GLM) and find that coronary heart disease (CHD), hypertension and depression are stronger predictors of health care costs than high baseline HbA1c levels. Interestingly, HbA1c levels below 7.5% were not associated with increased costs. Oglesby *et al.* (2006) use similar methods and data from the U.S. Health Core Managed Care Database between 1998-2003 and observe that direct medical costs driven by type 2 diabetes were between 16% to 20% lower for individuals with good glycaemic control (HbA1c $\leq$ 7%) Similarly, Menzin *et*

al. (2010) focus on a subsample of managed-care diabetic patients from Massachusetts covering a 5-year period (2002-2006) and employ logit and GLM models. They notice that diabetes-related hospitalisations were significantly higher among patients with highly uncontrolled diabetes. Based on two-part models, hospital costs per patient were also higher with increasing uncontrolled diabetes.

Other studies within the economic literature have examined the economic consequences of diabetes by focusing mainly on its effects on the labour market. Rizzo *et al.* (1996) investigate the labour productivity of a series of chronic conditions as well as their prescribed medications. They employ U.S. data from the National Medical Expenditure Survey (NMES) and find that untreated diabetes is associated with an average of 25 days lost but that diabetes medications could save an average 16 days of work. Kahn (1998) investigates trends in diabetic employment in the U.S. by employing three different sources of data and shows that diabetic men appear to have decreased (although slightly) their participation between 1976-1992. These results are extended by Latif (2009) using data from the Canadian National Population Health Survey and Zhang *et al.* (2009) on data drawn from the Australian National Health Survey. They find that diabetes appears to have a significant negative impact on female employment and that diabetes reduces labour force participation especially among older male individuals, respectively. Findings from Minor (2010) also indicate that diabetes appears to be detrimental to a number of labour market outcomes (e.g. participation, hours of work, out-of-work-days and earnings). More recently, Alva *et al*. (2014) estimate the effects of diabetes related complications on quality of life using UK longitudinal data from the Prospective Diabetes Study. Their results highlight the importance of studying changes in quality of life over time.

Our paper builds on the previous medical evidence and exploits rich longitudinal administrative data and biomarkers combined with panel count data models to provide new evidence on the effects of a widespread condition, uncontrolled diabetes, on health care utilisation.

3. Data

We employ individual-level longitudinal data drawn from administrative records of patients followed over seven consecutive years (2004-2010) in 6 primary care centres and 2 hospitals in the north-east of Barcelona, in Spain serving more than 104,000 inhabitants. This sample

of users is mostly urban, of lower-middle socioeconomic status from a predominantly industrial area.

This dataset includes a rich set of information about patients' use of healthcare resources, including our three main outcomes of interest i.e., number of GP visits; specialist care and hospital length of stay. Analysing utilisation and its determinants is important, especially when hospitalisation costs are widely reported as the largest component of diabetes medical costs and the number of hospitalizations, re-admissions and hospital length of stay tend to increase with this condition.[5] Our data also encompasses information on clinical 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 ICPC-2); dates of hospital admission and discharge; type of healthcare professional(s) contacted; and the main reason for their visit. Moreover, the dataset includes 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 (tobacco and alcohol use).[6]

Given the purpose of this study we focus on a sub-sample of individuals with diagnosed type 2 diabetes mellitus (DM), aged 16+ who had at least one contact with the aforementioned hospitals and primary health care centres between 1 January 2004 and 31 December 2010.[7] Individuals transferred or moved to other health centres and patients from other areas were excluded from our analysis. Diabetic patients were mainly identified through the International Classification of Primary Care codes (second edition, ICPC-2) reported by physicians combined with the information provided by the glycated haemoglobin test (HbA1c).[8] This test is routinely used by physicians and provides a very accurate measure of glucose concentration up until the previous 8 weeks. Information from HbA1c levels, allowed the inclusion among our population of diabetics of those patients who might not have been reported as diabetics through the ICPC-2 codes but had a mean HbA1c level $\geq 6.5\%$ (≥ 48 mmol/mol). Following this criteria, we obtained a sample of 53,963 patients with type 2 DM.

[5] Mata-Cases *et al.* (2015) using data from a population-based study in Catalonia report that hospital care, medications and primary care appear to be the main drivers of costs in both type 2 diabetics and non-diabetics.
[6] The original dataset comprises almost 830,000 observations, including the majority of the population living in the area.
[7] The sample may include diabetics with zero utilisation in some of these years. These might be patients who had some positive use only in selected years. We dropped all individuals with type 1 diabetes.
[8] Note that in our sample there might be diabetics with no HbA1c measurement as well as patients with repeated HbA1c values.

Our main variable of interest is uncontrolled type 2 DM, that we define using a dummy variable. Following the medical literature, we assume that diabetics are not adhering adequately to health treatments and hence not fully controlling their condition, when their within year mean HbA1c level is equal or above 7.5%. Since there is no universal consensus within the medical literature about the HbA1c thresholds which identify uncontrolled diabetes, we also estimate our models using a slightly lower threshold (7%).[9]

It should be borne in mind that the dataset used in this study comes from a representative set of health care centres from Spain where health care is provided through a decentralised (at regional level) national health system, and provision is free of charge at the point of delivery with the exception of pharmaceuticals entailing some co-payments.[10]

4. Empirical approach

We focus on the effects of uncontrolled diabetes on three main outcomes: the number of GP and specialist visits and hospital length of stay. Given that these are non-negative integer outcomes and to exploit the panel element of our data, we estimate *panel* count data models that account for individual-level unobserved heterogeneity. More specifically, we present estimates on the effects of uncontrolled diabetes on the number of GP visits using negative binomial (NB) models with random effects (RE). This is a flexible specification that is often used to model health care utilisation in the economics literature as it goes beyond standard Poisson models allowing for overdispersion together with unobservables (e.g. Cameron and Trivedi, 2005; Sarma and Simpson, 2006).[11]

We model the number of specialist visits and hospital length of stay using hurdle (two-part) models. Previous literature points out that the decision to contact a physician and the one concerning the amount of visits may be the result of two distinctive decision-making processes (e.g. Pohlmeier and Ulrich, 1995; Gerdtham, 1997; Deb and Trivedi, 2002). Further, these decisions might depend on both the individual and the physician, or more generally the heath care provider, and the complexity of this process may be exacerbated by repeated decisions in the presence of longitudinal data. This might also be the case in our data

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

[10] During the period considered for this analysis (2004-2010), in Spain only individuals in employment would face a 40% copayment for prescribed medicines, whereas pensioners would be exempted. Furthermore, patients with chronic conditions would have a highly reduced copayment of 10% and civil servants would only pay 30% of the market price regardless of their employment status.

[11] We have also estimated our models using Poisson specifications. Results are similar to the ones presented here and are available upon request.

when we analyse specialist visits (i.e. the decision to see a specialist vs. the number of visits) and hospital length of stay (i.e. the decision to walk into a hospital vs. length of stay) for patients with type 2 diabetes. Accordingly, we employ two-part or hurdle specifications to account for these two separate decisions. In our case, the first part of the hurdle is a probit model that distinguishes between users and non-users while the second part estimates the intensity of use (conditional on positive use) through a NB model with RE. This approach is also motivated by the relatively high frequency of zeros (non-use) for these two outcomes. All these models exploit variations in utilisation between the years 2004-2010 and account for a number of individual observable characteristics, individual-level unobserved heterogeneity, years and geographical area (local health authorities) fixed effects. Our basic specification is:

$$y_{ijt} = UD_{ijt-1}\,\tau + X_{ijt}\theta + D_t + G_j + \mu_i + \varepsilon_{ijt} \qquad (i=1,\ldots,N;t=1,\ldots,T;j=1,\ldots,J) \qquad (1)$$

where y_{ijt} is health care utilisation, that is alternatively the number of GP visits, specialist visits or hospital length of stay of individual *i*, in (health authority) area *j* at time *t.* UD_{ijt-1} is our main variable of interest as it defines uncontrolled diabetes via the biomarker HbA1c. Following the medical literature, in our basic specification we define uncontrolled diabetes in the presence of within year average values of HbA1c ≥7.5%. This variable is lagged one period to ease concerns around endogeneity.[12] We also estimate alternative specifications including a series of binary variables defining increasing levels of uncontrolled diabetes, i.e. 7.5%≤HbA1c<8.5%; 8.5%≤HbA1c<9.5; and HbA1c≥9.5%. We do this to explore the potential presence of a gradient in health care utilisation driven by the severity of an individual's condition. X_{ijt} is a vector that includes socio-demographic individual-level observable variables as well as clinically assessed health conditions. These include age (age, age squared and cubic age to capture non-linear age effects); labour market status (being active in the labour market contrasted against being inactive); immigration status (being a non-EU immigrant versus a Spanish or EU citizen as a baseline); marital status (living alone versus married/with a partner); health-behaviours (alcohol consumption, smoking status and objectively measured BMI) and a series of diabetes-unrelated chronic health conditions (e.g. asthma, chronic obstructive pulmonary disease (COPD), dementia, psychosis, clinical

[12] We have also estimated the full set of our models using contemporaneous values of HbA1c. Main results appear to be similar and estimates are available upon request.

depression and cancer).[13] The time dummies D_t account for time trends while G_j identifies primary health authority areas (defined at geographical level) fixed effects. μ_i represents individual-level time-invariant unobserved heterogeneity and ε_{it} is the idiosyncratic error term.

Conditionally correlated random effects

To allow for correlation between observables and individual unobserved heterogeneity, we parameterise the individual effect μ_i as a function of the within individual means of the exogenous regressors (see Mundlak, 1978; Chamberlain, 1984; Cameron and Trivedi, 2013).[14] This simply translates into including among our regressors the time-average of the time-varying exogenous (continuous) variables, i.e. $\overline{X_i}$. In our case, this includes the average over the sample period of the variables defining age, BMI and the annual average value of the biomarker.

Robustness checks

We also provide a series of robustness checks to assess the validity of our main results. More specifically, we examine whether and to what extent the effects of poor glycemic control on utilisation are mediated by the number of patients' health conditions. To purge our estimates from such influences, we run our models on a sample of "healthy users", i.e., diabetics without the following diagnosed conditions: cardiovascular disease, cerebrovascular disease, neuropathy and heart failure. Furthermore, in order to separately account for individual-level unobserved heterogeneity and the effects of previous period (*t*-1) health care utilisation on current consumption, we also estimate dynamic NB models with RE. This approach extends the previous conditionally correlated random effects model already augmented by a Mundlak correction by including among our regressors values of the dependent variables lagged one period, Y_{ijt-1}, as well as initial conditions in the parameterisation of the individual effect (Wooldrige, 2005).[15] Note that these estimates are performed on a balanced panel sample

[13] Note that we exclude diabetes related health conditions (i.e. cardiovascular disease, cerebrovascular disease, neuropathy and heart failure) from our basic specifications. This is because we aim at identifying the effect of uncontrolled diabetes, together with other unrelated conditions, on utilisation. Further models reported among the robustness checks, include the full battery of health conditions.

[14] For a more recent discussion on the use of the conditionally correlated (CCR) random effects models applied to count data models, including negative binomial models, see Cameron and Trivedi (2013). Note that since the random effect in negative binomial models applies to the distribution of the dispersion parameter, in this case, this Mundlak-type correction would only concern the variance of the model.

[15] That is, we also include among our regressors initial (wave 1) values of our dependent variable, i.e. $Y_{i,j,0}$.

where patients are observed during seven consecutive years (2004-2010) to allow accounting for initial conditions and values of the lagged dependent variables.

5. Results

Descriptive statistics

The prevalence of DM in the full dataset of 104,000 patients (including diabetic and non-diabetic patients) increased from 6.4% in 2004 to 9.3% in 2010. As expected, the prevalence is substantially higher for those aged 65+ (21.1% in 2004 versus 25.3% in 2010). These figures appear similar to those at national level in Spain for diagnosed diabetes during the same years (Soriguer *et al*., 2012; Vinagre and Conget, 2013). Our sub-sample of diagnosed type 2 diabetic patients defined using both physicians' disease classification codes and blood tests reduces to 53,936 observations over the period 2004-2010.

Table 1 reports descriptive statistics for both the sub-sample of all diabetics and the one that only includes patients with uncontrolled diabetes (HbA1c>=7.5%). Both samples of diabetic patients are broken down by gender. Within the full sample of diabetics, the mean numbers of GP visits per year are 11.27 (men) and 13.68 (women), while these reduce to 3.43 (men) and 3.66 (women) in case of specialist visits and 0.56 (men) and 0.65 (women) for hospital length of stay. We also note that this sub-sample includes mainly older individuals with a mean age of 65.4 years. The mean HbA1c value in the sample is 7%, which is very close to the estimate mean figure for the whole region, Catalonia (7.15%). Importantly, and consistently with evidence reported in other developed countries, our dataset shows that a significant number of diabetic patients are not controlling adequately their condition despite the well-known potentially life-threatening health consequences caused by the diabetes-related complications: uncontrolled diabetes (HbA1c levels $\geq$ 7.5%) is as high as 30.53% (men) and 27.59% (women). For these patients health care consumption appears to be substantially higher, especially their primary care use: GP visits increase to 16.32 (men) and 20.18 (women), specialist visits to 4.68 (men) and 5.15 (women), while hospital length of stay grows only marginally to 0.63 (men) and 0.77 (women). Overall, diabetic women with and without uncontrolled diabetes appear to systematically consume more health care if compared to men and patients of both genders with uncontrolled diabetes present average HbA1c levels of around 8.7%. Men with poor glycemic control are slightly younger and more likely to be immigrants; living alone; consume alcohol and tobacco smoke; and have a higher BMI than diabetic men who better control their condition. Also, men with uncontrolled diabetes present

higher percentages of depression, heart diseases and dyslipidaemia (high values of lipids in blood such as cholesterol). Women who do not control adequately their blood sugar levels appear to be somewhat older; more likely to be outside the labour market; and also show increased percentages of alcohol and tobacco consumption and a higher BMI. This category of women also presents a higher incidence of depression, hypertension, dyslipidaemia, heart diseases and cardiovascular diseases.

Our three main outcome variables appear to be highly right skewed distributed with a long tail indicating a very large consumption of resources by a small fraction of diabetic patients.[16] Another relevant feature of the dataset is the existence of 18% of zero GP visits during the years 2004-2010, and this increases to 38% (specialist visits) and 90% (hospital length of stay). As previously mentioned, the presence of overdispersion in the data concerning utilisation justifies the use of NB models. Yet, a zero mass problem in the data is judged to be of concern only in case of specialist visits and hospital length of stay, which explains our decision to examine such outcomes through a hurdle approach.[17]

Main results

In order to quantify the effects of uncontrolled diabetes (UD) on health care utilisation, we estimate a series of (static) NB2 models with random effects (RE) that exploit the longitudinal nature of the data and account for unobserved heterogeneity.[18] All models employ conditionally correlated RE and are augmented by a Mundlak specification which includes the within mean of the continuous regressors. All models also include lagged values of the biomarker (HbA1c) defining UD to ease concerns around endogeneity and control for the full set of covariates. Since the medical literature appears to find higher percentages of poor control and non-adherence among males and this may have an impact on subsequent health care utilisation, we present separate estimates by gender on an unbalanced panel. All tables report average marginal effects.

The first two columns of Table 2 show the estimates of our variable of interest, the lagged value of the biomarker detecting UD, on the number of GP visits for men and women. We

[16] For instance, the variance of GP visits is almost 13 times greater than its mean. The same applies to specialist visits (8 times) and hospital stays (21 times).

[17] Once we condition on positive counts, the average number of specialist visits is 5.76 per year and the average length of stay is 5.9 days per year.

[18] Specifically, we estimate the variant of the NB model in which the conditional variance function is quadratic in the mean, called negative binomial 2 (NB2). See Cameron and Trivedi (1998).

find that for men the effect of UD on the use of primary care is positive and highly statistically significant with a quantitative effect of around 0.48 GP visits. This represents an increase of almost 4% of the annual GP visits of men driven by poor glycemic control. Although the effect of UD on GP visits for women also appear to be positive, it is not statistically significant.

In columns (3) and (4) we report the effects of UD on specialist visits using the hurdle RE NB2 model. Again, estimates appear to differ by gender as the quantitative effects are positive and statistically significant, pointing towards an increase in the number of specialist visits, only for diabetic men. We find that the effect of having HbA1c≥7.5% in the previous period leads an increase of specialist visits of 0.25 (at 5% significance level) conditional to positive visits. This corresponds to a percentage increase around 4.4 in the number of annual specialist visits. Finally, the last two columns of Table 2 present the effects of UD on hospital length of stay using the same hurdle approach. These models present a reduced number of observations as a consequence of a larger number of zeros in the data. Interestingly, we find again a gender pattern since UD is solely associated with a raise in the number of days in hospital among male patients. Specifically, the results indicate that one-year lagged uncontrolled DM leads to a statistically significant increase of 0.78 additional days in hospital, conditional to positive stays. This corresponds to an increase of around 17.4% in the annual length of stay.

Since reduced glycemic control increases the risk of diabetes complications and this may in turn influence health care use, Table 3 examines potential non-linearities in the effects of UD by defining increasing levels of this condition: low (7.5%≤HbA1c<8.5%); moderate (8.5%≤HbA1c<9.5) and high (HbA1c>=9.5%) UD. Our estimates show that the raise in GP visits among diabetic men appear to be mainly concentrated among those with low and moderate levels of UD with increases of around 0.4 and 0.7 GP visits respectively. Among men, we also find statistically significant effects of low and high UD on specialist visits with the largest impact observed for diabetic with the poorest control (high UD). Our estimates show that having low and high UD in the previous year lead to annual increases of 0.26 and 0.46 specialist visits respectively, conditional to positive visits. These correspond to average increases of roughly 4.5% and 8%. Men also present a positive, large and highly statistically significant effect of low UD and only a weekly significant effect (although also large) of high UD on hospital length of stay. The effect of low UD on utilisation implies an additional 1.1 day in hospital, representing a mean annual increase in the length of stay, conditional on

hospital admission, of around 25%. It should be noted that for women, none of the effects related to UD appear to be statistically significant.

Table 4 presents estimates from RE NB2 as well as hurdle RE NB2 for a sub-sample of healthy-users. In this case, we have excluded from our analysis individuals with diabetes complications as these may increase the level of health care consumption and potentially confound our effects of interest. These estimates should help us identifying the effects of UD on health care consumption purged from related health care complications. Estimates appear to substantially confirm the ones presented in Table 2. Among men, we find statistically significant effects of UD on both GP and specialist visits. These effects appear to be quantitatively larger if compared to the results obtained from the more general sample of diabetics in Table 2. This may imply that for men, increases in health care consumption might be mainly driven by UD and not necessarily its related complications. We also find a weakly significant effect of UD on hospital length of stay, conditional on admission. As for the previous models in Tables 2 and 3, we do not detect any effects for women that are statistically different from zero.

Table 5 includes results from dynamic RE NB2 and hurdle models on a balanced sample. These models were implemented with the specific purpose of separately accounting for both individual unobserved heterogeneity and the dynamics of health care consumption together with the influence of UD. Statistically significant persistence of health care consumption appears to be present for both genders and for both GP and specialist visits. However, this does not appear to play a major role among the determinants of hospital length of stay for men and also present a negative effect on consumption for women. This is expected and might be related to the different nature and frequency of health consumption identified by the three outcomes considered: hospital admission is a much rarer event if compared to GP and specialist visits. Interestingly, in this case we find statistically significant effects of UD on GP visits among women and only on specialist visits amongst men, although only significant at 10%.[19] In any case, we should be cautious in the interpretation of the effects of past heath care consumption on current use because these might be also partly driven by previous (further back in time) levels of UD. In any case, given the substantial drop in the number of observations when using a balanced sample, we should be cautious in comparing these estimates with our main results.

[19] We also run the same dynamic specifications on an unbalanced sample and we found similar effects to those reported in Table 2. These findings are also available upon request.

6. Conclusions and discussion

Very little is known on the effects of non-adherence to health treatments on health care utilisation, especially among individuals with chronic conditions and diabetes. Previous evidence is mixed, often based on limited data and standard linear models as well as self-reported information that may be plagued by reporting bias. Other literature focused on the effects on labour supply and aspects of quality of life while the consequences on health care utilisation of non-adherence and uncontrolled diabetes have not been examined. In this paper, we employ detailed longitudinal data on a large population of adult diabetic patients and use a clinically assessed biomarker to detect the presence of uncontrolled diabetes and in turn its influence on GP visits, specialist visits and hospital length of stay. We focus on these outcomes as they appear to account for the largest components of medical costs associated to diabetes complications.

We estimate a range of panel count data models, including negative binomials with random effects, dynamic and hurdle specifications to account for unobserved heterogeneity, previous utilisation and selection. The inadequate control of diabetes, the most widespread chronic condition in all developed countries, and the analysis of its multiple determinants is a key priority for health care professionals and policy makers.

Our analysis confirms that a large fraction of diabetic patients (30%) appears to have a poor control of their condition and this leads to an excess healthcare utilisation, especially among men. We find that among men, uncontrolled diabetes appears to increase the number of GP visits by around 4% and the volume of specialist visits, also by about 4.4%, after accounting for a wide set of controls and other clinically assessed conditions. Our findings also indicate that uncontrolled diabetes leads to an increase of around 17.4% in the annual hospital length of stay for men, conditional on positive stays. This may imply that, although hospitalisation is a relatively rare event for diabetic patients, uncontrolled diabetes greatly increases its length, therefore imposing additional costs to the health system. Although women with diagnosed diabetes appear to consistently consume more health care if compared to men, their consumption do not appear to be driven by uncontrolled diabetes.

Interestingly, our analysis appears to show heterogeneous effects among men. For instance, we find that the increase in GP visits among diabetic men is concentrated among those with low ($7.5\% \leq HbA1c < 8.5\%$) and moderate ($8.5\% \leq HbA1c < 9.5$) levels of uncontrolled diabetes. Yet, the annual increase in specialist visits, conditional to a positive number of visits, concern

patients having low and high (HbA1c>=9.5%) levels of inadequate control. Finally, we find that increase in the length of stay (around 25%) appears to be concentrated among patients with low levels of uncontrolled diabetes. We believe these latter findings might be particularly informative as they identify specific groups of diabetics that should be targeted and prioritised in order to use health resources more efficiently and provide better treatments. Our robustness checks appear to validate our main findings.

Overall, this study combines the use of robust panel econometric methods with rich administrative data to estimate for the first time the effects of non-adherence on health care utilisation by focusing on diabetes, one of the most widespread chronic conditions worldwide. Our results suggest that improving glycemic control would not just be beneficial to patients' wellbeing but would also substantially reduce extra health care utilisation.

References

Alva, M., Gray, A., Mihaylova, B., & Clarke, P. (2014). The Effect Of Diabetes Complications On Health-Related Quality Of Life: The Importance Of Longitudinal Data To Address Patient Heterogeneity. *Health Economics, 23*(4), 487-500.

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

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 for Count Data*. Cambridge: Cambridge University Press.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cameron, A. C., & Trivedi, P. K. (2013). Count panel data. *Oxford Handbook of Panel Data Econometrics*.

Chamberlain, G. (1984). Panel data. In Grilisches, Z. and Intriligator, M.D. (eds.) *Handbook of Econometrics*, Volume *1*, 1247-1318, Amsterdam: Elsevier.

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.

DiMatteo, M. R. (2004). Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. *Medical Care*, *42*(3), 200-209.

Garcia-Perez, L.E., Alvarez, M, Dilla, T., Gil-Guillén, V. & Orozco-Beltrán, D. (2013). Adherence to therapies in patients with type 2 diabetes, *Diabetes Therapy*, 4(2), 175-194.

Gerdtham, U-G. (1997). Equity in health care utilization: futher tests based on hurdle models and Swedish micro data. *Health Economics*, 6(3), 303-319.

Gilmer, T.P., O'Connor, P.J., Rush, W.A., Crain, A.L., Whitebird, R.R., Hanson, A.M. & Solberg, L.I. (2005). Predictors of healthcare costs in adults with diabetes. *Diabetes Care*, 28(1), 59-64.

Haynes, R. B., McDonald, H., Garg, A. X., & Montague, P. (2002). Interventions for helping patients to follow prescriptions for medications. *The Cochrane Library*.

Ho, P. M., Rumsfeld, J. S., Masoudi, F. A., McClure, D. L., Plomondon, M. E., Steiner, J. F., & Magid, D. J. (2006). Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus. *Archives of Internal Medicine*, *166*(17), 1836-1841.

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 Federation of Diabetes (IDF), Diabetes Atlas, 6th Edition, 2014. http://www.idf.org/diabetesatlas

Jackevicius, C. A., Mamdani, M., & Tu, J. V. (2002). Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *Journal of the American Medical Association*, *288*(4), 462-467.

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

Kripalani, S., Yao, X., & Haynes, R. B. (2007). Interventions to enhance medication adherence in chronic medical conditions: a systematic review. *Archives of Internal Medicine*, *167*(6), 540-549.

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

Lesniowska, J., Schubert, A., Wojna, M., Skrzekowska-Baran, I., & Fedyna, M. (2014). Costs of Diabetes and Its Complications in Poland. *European Journal of Health Economics, 15*(6), 653-660. doi:http://dx.doi.org/10.1007/s10198-013-0513-0.

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

Mata-Cases, M., Csajuana, M., Franch-Nadal, J., Casellas, A., Castell, C., Vinagre, I. & Mauricio, D. (2015). Direct medical costs attributable to type 2 diabetes mellitus: a population-based study in Catalonia, Spain. *European Journal of Health Economics* (doi 10.1007/210198-015-0742-5).

Menzin, J., Korn, J.R., Cohen, J., Lobo, F., Zhang, B., Friedman, M. & Neumann, J. (2010). Relationship between glycemic control and diabetes-related hospital costs in patients with type 1 or type 2 diabetes mellitus. *Journal of Management Care Pharmacy*, 16(4), 264-275.

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.

Minor, T. (2013). An Investigation into the Effect of Type I and Type II Diabetes Duration on Employment and Wages. *Economics & Human Biology*, 11(4), 534-544. doi:http://dx.doi.org/10.1016/j.ehb.2013.04.004

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: Journal of the Econometric Society*, 69-85.

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.

Osterberg, L., & Blaschke, T. (2005). Adherence to medication. *New England Journal of Medicine*, 353(5), 487-497.

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

Sarma S. & Simpson, W. (2006). A microeconometric analysis of Canadian health care utilization. *Health Economics* 15: 219-239.

Sokol, M. C., McGuigan, K. A., Verbrugge, R. R., & Epstein, R. S. (2005). Impact of medication adherence on hospitalization risk and healthcare cost. *Medical Care*, *43*(6), 521-530.

Soriguer, F., Goday, A., Bosch-Comas, A., Bordiú, E., Calle-Pascual, A., Carmena, R. et al. (2012). Prevalence of diabetes mellitus and impaired glucose regulation in Spain: the Di@bet.es Study. *Diabetologia* 55: 88-93.

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).

Vander Stichele, R. (1991). Measurement of patient compliance and the interpretation of randomized clinical trials. *European Journal of Clinical Pharmacology*, 41(1), 27-35.

Vinagre, I. & Conget A. (2013). Situación actual del control de la diabetes mellitus tipo 2 en España. Identificación de las principales barreras en la práctica clínica diaria. *Medicina Clínica* 141(Suppl.2): 3-6.

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.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, *20*(1), 39-54.

World Health Organisation - WHO (2003). *Adherence to Long-Term Therapies: Evidence for Action*. Geneva, Switzerland: World Health Organization; 2003.

World Health Organisation - WHO (2016). *Global Report on Diabetes*. Geneva, Switzerland: World Health Organization; 2016.

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

Table 1: Descriptive statistics: Years 2004-2010

	Entire sample of Diabetics				Diabetics with HbA1c ≥ 7.5%			
	Men	N	Women	N	Men	N	Women	N
GP Visits	11.27	26964	13.68	26972	16.32	4432	20.18	4096
Specialist visits	3.43	26964	3.66	26972	4.68	4432	5.15	4096
Hospital stays	0.56	26964	0.65	26972	0.63	4432	0.77	4096
HbA1c level	7.08	14515	7.00	14848	8.69	4432	8.68	4096
Uncontrolled DM *	30.53	14515	27.59	14848	-	4432	-	4096
Age	64.47	26964	66.42	26972	63.54	4432	67.66	4096
Immigrant	2.31	26964	3.12	26972	2.73	4432	2.37	4096
Active	32.87	26860	25.39	26789	34.46	4422	18.45	4075
Living alone	88.57	26964	83.15	26972	90.55	4432	83.76	4096
Alcohol	5.11	26964	0.56	26972	5.80	4432	0.71	4096
Tobacco	28.39	26964	8.14	26972	33.73	4432	8.64	4096
BMI	29.13	25165	30.81	24576	29.31	4352	31.73	4017
Asthma	1.86	26964	6.05	26972	1.65	4432	6.81	4096
CPOD	9.81	26964	2.68	26972	9.05	4432	2.54	4096
Dementia	2.00	26964	3.56	26972	1.65	4432	2.91	4096
Psychosis	1.11	26964	1.10	26972	1.08	4432	0.66	4096
Depression	9.55	26964	23.12	26972	10.24	4432	28.32	4096
Neoplasia	8.52	26964	7.00	26972	5.75	4432	6.91	4096
Hypertension	50.72	26964	60.91	26972	49.99	4432	68.04	4096
Dyslipidemia	47.47	26964	47.45	26972	50.10	4432	57.84	4096
Heart Diseases	16.27	26964	8.57	26972	17.89	4432	9.38	4096
Cerebrovasc D.	5.76	26964	1.81	26972	6.34	4432	3.03	4096
Heart failure	12.91	26964	12.54	26972	11.17	4432	12.06	4096
Neuro	0.92	26964	1.27	26972	0.81	4432	1.20	4096

* Uncontrolled DM is defined for HbA1c values ≥ 7.5%.

Table 2: The effects of uncontrolled diabetes on health care utilisation

			Hurdle Models			
	(1) Men GP visits	(2) Women GP visits	(3) Men Specialist	(4) Women Specialist	(5) Men Hosp. Stays	(6) Women Hosp. Stays
UD (t-1)	0.483**	0.241	0.249**	0.155	0.777**	0.111
	(0.192)	(0.230)	(0.119)	(0.127)	(0.316)	(0.369)
N	11763	12147	9235	9924	1394	1574

UD (t-1) stands for uncontrolled type 2 diabetes as measured by HbA1c>=7.5% lagged one period. Table displays average marginal effects (AME) for static RE NB2 model augmented by a Mundlak specification (unbalanced sample). All models control for the full set of covariates. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: The effects of uncontrolled diabetes on health care utilisation - HbA1c levels

			Hurdle Models			
	(1) Men GP visits	(2) Women GP visits	(3) Men Specialist	(4) Women Specialist	(5) Men Hosp. Stays	(6) Women Hosp. Stays
UD1 [7.5, <8.5%] (t-1)	0.412**	0.310	0.261**	0.122	1.099***	0.058
	(0.210)	(0.248)	(0.130)	(0.137)	(0.353)	(0.401)
UD2 [8.5, <9.5%] (t-1)	0.669**	0.343	0.125	0.245	-0.014	0.393
	(0.303)	(0.360)	(0.187)	(0.195)	(0.468)	(0.524)
UD3 [>=9.5%] (t-1)	0.582	-0.504	0.459**	0.206	1.027*	-0.335
	(0.384)	(0.449)	(0.232)	(0.244)	(0.553)	(0.680)
N	11763	12147	9235	9924	1394	1574

UD1,2,3 (t-1) stands for the uncontrolled type 2 diabetes defined within the corresponding HbA1c interval lagged one period. All models present average marginal effects (AME) for static RE NB2 models augmented by a Mundlak specification (unbalanced panel). All models control for the full set of covariates. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The effects of uncontrolled diabetes on health care utilisation – healthy users

			Hurdle Models			
	(1) Men GP visits	(2) Women GP visits	(3) Men Specialist	(4) Women Specialist	(5) Men Hosp. Stays	(6) Women Hosp. Stays
UD (t-1)	0.560**	0.097	0.286**	0.216	0.718*	0.066
	(0.219)	(0.248)	(0.139)	(0.141)	(0.368)	(0.423)
N	8179	9466	6180	7633	832	1125

UD (t-1) stands for uncontrolled type 2 diabetes as measured by HbA1c>=7.5% lagged one period. Table displays average marginal effects (AME) for static RE NB2 model augmented by a Mundlak specification (unbalanced sample). All models control for the full set of covariates. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: The effects of uncontrolled diabetes on health care utilisation – dynamic models

			Hurdle Models			
	(1) Men GP visits	(2) Women GP visits	(3) Men Specialist	(4) Women Specialist	(5) Men Hosp. Stays	(6) Women Hosp. Stays
UD (t-1)	0.431	0.871***	0.380*	0.001	-0.131	-0.523
	(0.292)	(0.310)	(0.198)	(0.183)	(0.478)	(0.495)
GP visits (t-1)	0.186***	0.106***	-	-	-	-
	(0.0137)	(0.0106)				
Spec. visits (t-1)	-	-	0.118***	0.131***	-	-
			(0.013)	(0.012)		
Hosp. stays (t-1)	-	-	-	-	0.056	-0.069**
N	3738	4704	3122	4094	468	602

UD (t-1) stands for uncontrolled type 2 diabetes as measured by HbA1c>=7.5% lagged one period; GP visits (t-1) stands for the number of GP visits lagged one period; Spec. visits (t-1) for number of specialist visits lagged one period; Hosp. Stays (t-1) for the length of hospital stays lagged one period. Table displays average marginal effects (AME) for dynamic RE NB2 model augmented by initial conditions and a Mundlak specification (balanced sample). All models control for the full set of covariates. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.