**Abstract**

A diabetes diagnosis entails important consequences for its recipients. Diagnosed patients obtain health information but also face the challenge of having to manage the condition via lifestyle adjustments, with potential consequences for—among other things—their economic activity. We investigate the causal eﬀect of a diabetes diagnosis on employment status and behavioural risk-factors, two potentially intertwined factors, using longitudinal data from the China Health and Nutrition Survey (CHNS) that cover the years 1997 to 2011. Two complementary statistical techniques—marginal structural models and fixed eﬀects panel estimation—are used for the statis-tical analysis, and generate very similar results, despite their different underlying assumptions. Both strategies find distinct patterns for males and females. They suggest a decrease in female employment chances after a diagnosis (over 11 percent-age points) and further show that women are mostly unable to positively change their behavioural risk factors by loosing weight and reducing energy intake. Men, however, do not see their employment probabilities aﬀected by diabetes and also respond to a diagnosis by losing weight and reducing energy intake as well as their intake of alcohol in ways that are sustained over time. These results suggest important inequities in the impact of diabetes between sexes in China and point to the potential of reducing behavioural risk factors for women to narrow these inequities.

**0.1 Introduction**

The effect of diabetes on employment status has received relatively little attention in middle-income countries (MICs), including China. The scarce existing evidence indicates that diabetes can aﬀect labour market outcomes in high-income countries (HICs), but also in MICs (Seuring, et al.,  [2016](#page45)). This isof growing relevance especially with diabetes appearing increasingly earlier in a person’s productive lifespan, among others due to increasing obesity at ear-lier stages of life. Importantly, once diagnosed, the onset of diabetes, and diabetes complications, strongly depend on the patient’s behaviour. Behavioural risk factors like alcohol consumption, smoking, caloric consumption and weight gain are all related to the onset of diabetes as well as ensuing diabetes complications. Research shows for instance that behaviour changes after a diabetes di-agnosis can have positive health eﬀects and reduce the risk of subsequent cardiovascular events (Long et al.,  [2014)](#page44) and may help in eﬀectively managing blood glucose levels and achieving further treatment goals (Zhou et al.,  [2016](#page46)). Consequently, if these risk factors can be reduced it may be possible to prevent some of the health and economic burden of diabetes.

Thus, it seems that a diabetes diagnosis may present an important opportunity to reduce risk factors for diabetes complications (De Fine Olivarius et al.,  [2015)](#page43) and hence also reduce the economic burden of diabetes to the individual. This raises the question how diabetes diagnosis affects both labour outcomes and health behaviour over time.

However, one of the challenges of determining a causal relationship between diabetes, employment status and changes in behavioural risk factors is their potential interrelat-edness. For example, employment status might by aﬀecting weight status by reducing the time spend on physical activity due to reductions in available leisure time, or it may promote risk factors such as smoking behaviour or energy intake that can both aﬀect the probability of developing diabetes as well as encouraging related complications, for instance by increasing stress levels. In an eﬀort to investigate the dynamic impact of unem-ployment on health behaviours, Colman and Dave  [(2014)](#page43) found heterogeneous eﬀects of unemployment, which lead to slight weight gain, a decrease in smoking and decreases in fast-food consumption. Macroeconomic evidence also indicates that job loss can lead to changes in health, especially in mental health (Charles and Decicca,  [2008),](#page43) which may have further downstream eﬀects on health behaviours.

Research on the impact of diabetes on labour market outcomes has so far ignored the potentially simultaneous relationship of diabetes with employment and behavioural diabetes risk factors. Using regression techniques such as ordinary least squares (OLS) or fixed eﬀects (FE) it is assumed that the investigated independent variables are unaﬀected by prior values of the dependent variable. However, if prior changes in employment status are causally related to a diabetes diagnosis or aﬀect the risk factors for diabetes compli-cations, not accounting for this can lead to biased estimates. Similarly, studies investigating the impact of a diabetes diagnosis on behavioural risk factors while not taking into account the eﬀect of employment status on both diabetes and these risk factors, may produce potentially biased estimates. Moreover, apart from time-varying con-founding due to observed covariates, unobserved variables present a further challenge. In particular time-invariant confounders, such as poor early life conditions or personal traits, may simulta-neously increase the probabilities to develop diabetes, be unemployed and engage in unhealthy behaviour.

The goal of this study is therefore to assess the impact of a diabetes diagnosis on both employment probabilities and behavioural risk factors while accounting for the potentially intertwined relationships between diabetes, employment and health behaviours. This is done via the use of marginal structural models (MSMs), an estimation strategy that is increasingly common in epidemiology and is able to account for time-dependent confounding across time (Robins et al.,  [2000)](#page45) when es-timating the impact of a treatment, here a diabetes diagnosis, on the outcome of interest.

2

This is by our knowledge the first time this estimation strategy is used to estimate the impact of diabetes on an individ-ual’s employment status or behavioural risk factors. We complement this strategy and test the robustness of the MSM estimates to the potential violation of one of its crucial assumptions, namely that unmeasured confounding factors are not important. To do this, we compare with FE models which—although unable to account for the potentially simultaneous relationships—account for unobserved time-invariant confounding factors in addition to confounding due to observed variables. Very diﬀerent results to the MSM would suggest a viola-tion of the assumption of no unobserved confounding. To further investigate and understand the role of confounding factors, we also estimate random eﬀects (RE) models and compare the results. Apart from these methodological contributions, the study also provides unique analytical insights. It further extends the evidence base for the impact of diabetes on labour outcomes in MICs, where currently empirical information is only available for Mexico (Seuring, et al.,  [2016](#page45)). At the same time the study provides, as far as we are aware, the first longitudinal evidence for the eﬀect of a diabetes diagnosis on behavioural risk fac-tors in any developing country .

More information about the eﬀects of a diabetes diagnosis may be particularly important for LMICs such as China, where diabetes prevalence has surged from 1% in the early 1980s to about 10% in recent years (Hu,  [2011;](#page44) NCD Risk Factor Collaboration,  [2016](#page44)). Con-fronting this diabetes epidemic puts a strain on healthcare systems (Seuring, et al.,  [2015),](#page45) increasing the need to find highly cost-eﬀective prevention and treatment op-tions in very resource constraint settings (Silink et al.,  [2010](#page45)). However, to do this it is important to assess how successful people with diabetes currently are in preventing adverse economic eﬀects and reducing their risk factors for diabetes complications.

The literature trying to identify a causal relationship between diabetes and employment has relied on instrumental variable (IV) strategies (Brown et al.,  [2005;](#page43) Latif,  [2009;](#page44) Seur-ing, et al.,  [2015)](#page45) and individual FE models (Seuring, et al.,  [2016](#page45)). However, while an IV approach could potentially account for all forms of confounding, the validity of the used instruments is at least questionable (see discussion in ChapterX). The FE model, as discussed above, also relies on important assumptions that may be violated.

Turning to the relationship between a diabetes diagnosis and behavioural risk factors, only one study has intended to causally relate a recent diabetes diagnosis with changes in health behaviours in the USA, finding positive behaviour changes shortly af-ter diagnosis. The study finds that the eﬀects were mostly short lived and tended to dissipate over time, particularly considering weight loss (Slade,  [2012](#page45)). To isolate the causal eﬀect Slade  [(2012)](#page45) created an "at risk" control group without diabetes that intended to be similar to

3

the treatment group with diabetes, apart from not having received a diagnosis. He used information on diabetes biomarkers to estimate the propensity score of those without a diabetes diagnosis to be above a specific at risk threshold, so that everybody above a certain propensity score was used to form the control group. He then estimated dynamic population averaged as well as FE models to identify a causal relationship. While this approach likely improves the control group by increasing its similarity in the diabetes risk profile to the diagnosed population, it may not have been able to suﬃciently account for the potential predetermination of the diabetes diagnosis by earlier values of the dependent variable. Further, the study did not account for employment status as one of the control variables.

A diﬀerent identification approach was used by Zhao et al.  [(2013)](#page46) when investigating the eﬀects of a hypertension diagnosis on nutritional outcomes in China. They used a regression-discontinuity design and biomarker information on blood pressure. A crucial assumption in the study was that people above the hypertension threshold were indeed informed about their hypertension while those just below the threshold were not. These two groups were then compared to isolate the particular eﬀect of the additional health information on food consumption in the following wave. The results indicated that a di-agnosis leads to reductions in fat consumption, but no other nutritional outcomes, and only for those economically better oﬀ. Several caveats exist for this study and the used approach. According to Zhao et al.  [(2013)](#page46) it was not always clear to what extend par-ticipants where informed about their hypertension status and whether they had received just the actual blood pressure measurement information, leaving the interpretation to the participants, or whether they were made explicitly aware of their hypertension (or also pre-hypertension) status. Further, the results may have limited generalisability, since the measured treatment eﬀect was a very local one, applying only to the population around the hypertension threshold. Finally, the study only provides information for a relatively short period until the first wave after diagnosis, unable to capture any changes further away from the point of diagnosis.

Accordingly, there is a need to provide new evidence on the eﬀects of a diabetes diag-nosis on employment status as well as behavioural risk behaviours that could aﬀect the development of diabetes complications, using longitudinal data and alternative estimation strategies. Thereby this study adds in several ways to the existing literature. First, it shows the impact of diabetes diagnosis on labour outcomes in China, not only over the short term, but for a period covering the entire decade of the 2000s, allowing for a more long term investigation of the eﬀects. This both confirms and extends earlier evidence for other settings and using diﬀerent methods. Second, it provides information on the eﬀect

of a diabetes diagnosis on health behaviours. Third, by considering the eﬀects over time on both employment and health behaviour, the results shed light on po-tential pathways through which the impact on employment may work. Fourth, the study provides a methodological innovation by using both MSM and FE estimation methods, oﬀering insights not only on the robustness of the MSM results, but also on the validity of some of its assumptions.

**0.2 Methods**

**0.2.1 Study sample**

The CHNS is an international collaborative project led by the Carolina Population Center at the University of North Carolina at Chapel Hill investigating nutrition and health behaviours in nine provinces of China (Zhang et al.,  [2014](#page46)). We use data from 1997 onwards, which was the first time survey participants provided diabetes information. In total we use six waves (1997, 2000, 2004, 2006, 2009 and 2011) obtained from the longitudinal dataset released in 2015. The data provide extensive information on nutrition and health, including anthropometric measures of weight and height, reducing potential measurement issues. It further provides socioeconomic information, most importantly for this study about employment. The sample is limited to the adult population from age 18–64. The sample is not nationally representative and as such does not provide sampling weights (Popkin et al.,  [2010](#page45)).

Overall, between 84% to 90% of the survey participants are followed up in the consecu-tive wave, with attrition being highest after 2006. Attrition in the CHNS due to mortality is around 1% (Popkin et al.,  [2010](#page45)). Other reasons mentioned by Popkin et al.  [(2010)](#page45) are loss in follow up due to migration, natural disasters and redevelopment of housing in the urban centres leading to relocations. We analysed if any of our variables of interest was sig-nificantly related to attrition at any wave and did only find lower calorie consumption and being unemployed to exhibit an association. Having diabetes was not related to attrition. Further, attrition was strongly related to urbanization, a higher level of education, being of younger age and having lower family income, suggesting that mostly participants of younger age, more urbanized but from less well-oﬀ households tended to leave the survey. Attrition rates between the waves are shown in Table  [0.5](#page27).

5

**0.2.2 Assessment of diabetes**

We used self-reported information on diabetes diagnosis to construct our diabetes in-dicator. We only relied on incident cases of self-reported diabetes, excluding individuals with self-reported diabetes at baseline. Given the chronic nature of diabetes, we assumed that after the initial diagnosis diabetes persists for the rest of one’s life. This is a reasonable assumption given the medical evidence. (Reference?) To construct a measure of diabetes duration for incidence cases we used self-reported information on the year of diagnosis. If we found that the year of diagnosis was reported to be before the last wave without a reported diagnosis, we used the midpoint between the last wave without diagnosis and the first wave with a diagnosis as the year of diagnosis.1

**0.2.3 Assessment of outcomes**

The economic outcome of interest is employment status, and is measured through self-reported reponse stating that the person is currently working. People who reported not to be working because they were student are excluded, while those who are not working for any other reason, such as doing housework, being disabled or being retired, were included.

The behavioural outcomes we estimate are current smoking status, if alcohol was con-sumed equal to or more than three times per week, body mass index (BMI), waist cir-cumference in centimetres and daily calorie consumption. Smoking status and alcohol consumption are self-reported, while BMI and waist circumference are based on anthro-pometric measurements, minimizing potential reporting errors. Waist circumference is reported in centimetres. Finally, daily calorie consumption is a constructed variable avail-able in the CHNS, based on the average daily consumption of carbohydrates, protein and fat of every individual in the survey, measured on three consecutive days. As robustness tests, we also considered overweight and obesity indicators instead of a continuous weight variable. These results suggest similar patterns. Since there is considerable discussion about the correct thresholds to use for Asian populations to define overweight and obesity (He et al.,  [2015;](#page43) WHO,  [2004;](#page45) Zeng et al.,  [2014](#page46)), we do not include these results in our main analysis but report them in appendix. We applied thresholds sug-gested by the China Obesity Task Force of a BMI ≥ 24 to define overweight and a BMI ≥ 28 to define obesity (China Obesity Task Force,  [2004](#page43)).

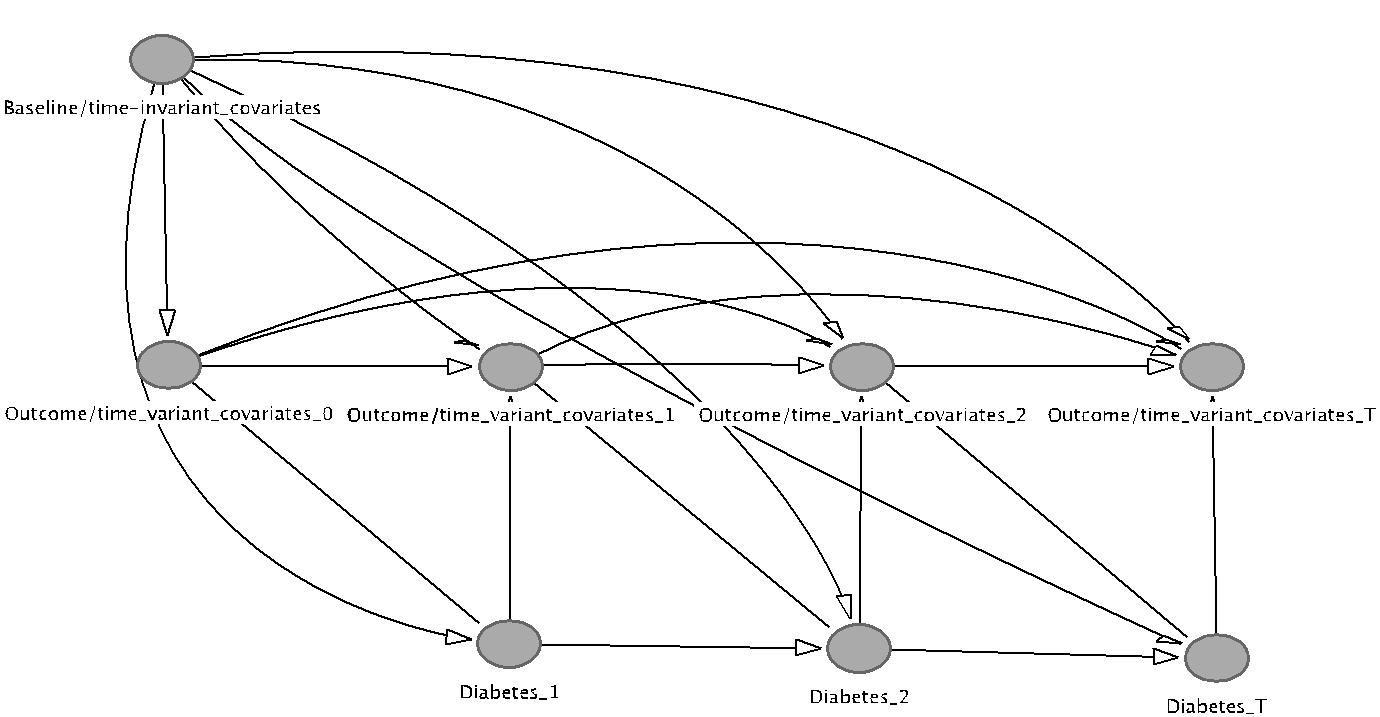
**0.2.4 Statistical analysis**

Our analysis focuses on two statistical approaches to account for potential confounding: marginal struc-tural models (MSMs) and fixed eﬀects (FE).

1The number of observations replaced at each wave was: 21 (2000), 44 (2004), 51 (2006), 78 (2009), 59 (2011). Overall it aﬀected 43% of the self-reports of the year of diagnosis.

6

Figure 0.1: DAG for marginal structural model



*Notes* MSMs assume the absence of unobserved time-invariant and unobserved time-variant confoundersbut allow the past treatments to aﬀect the current outcomes (arrows going from Diabetes to time-variant covariates) and the past outcomes to aﬀect the current treatment (arrows going from time-variant covari-ates to Diabetes). Lagged time-variant covariates, baseline and time-invariant covariates predict current diabetes status.

**Marginal structural models**

MSMs apply inverse probability weights to adjust for confounding and selection bias as a result of time-varying confounders being aﬀected by prior exposure to the treatment (Robins et al.,  [2000](#page45)). Under the assumption of the MSM(Robins et al.,  [2000](#page45))—the re-ported treatment is the treatment that has actually been received (consistency), there are no unmeasured confounders (exchangeability) and every person in the sample has a non-zero chance of receiving the treatment (positivity) (see Section  [0.4](#page22) for a discussion of the validity of these assumptions in our case)—the causal direct acyclic graph (DAG) shown in Figure  [0.1](#page7) displays the association between confounders and outcomes and a diabetes diagnosis.

In our context it seems possible that, for example, BMI could aﬀect the probability of being diagnosed with diabetes which then itself may aﬀect subsequent BMI levels, confounding the relationship between a diabetes diagnosis and BMI due to non-random selection. Similarly, employment history and current employment could aﬀect the prob-ability of a diabetes diagnosis through their impact on lifestyle and hence diabetes risk factors such as increases in weight or smoking. For example, an increase in disposable income or a reduction in leisure time as a result of a new job and the subsequent eﬀect on risk behaviours could confound the relationship between a diabetes diagnosis and employment status. MSM accounts for this by calculating weights based on the potential risk of a person being diagnosed at each point in time.

To calculate these weights we first construct unstabilized weights using baseline values of time-variant confounders, time-invariant confounders as well as time-variant confounders lagged by one period to predict the probability of developing diabetes at each wave. We use lagged time-variant confounders because current diabetes status as reported in the survey was determined at some point within the current and the previous wave that were determined before the current diabetes status, to prevent reverse causality. The used predictors are age and age squared to account for changes in risk with increasing age, an index of urbanization pre-constructed within the CHNS data, ranging from 1 to 120 as the level of urbanization increases (Zhang et al.,  [2014),](#page46) to account for the impact of urbanization on diabetes risk (Attard et al.,  [2012](#page43)). We also use secondary and university education, being married, having any medical insurance, being of Han ethnicity, living in a rural area, dummies for the diﬀerent Chinese regions and the respective survey waves as predictors. Further we use inflation adjusted per-capita household income to adjust for eﬀects of household wealth on diabetes. Finally, all outcome variables (employment status, alcohol consumption, smoking status, BMI, waist circumference and average daily calorie consumption) are used as predictors.

Because unstabilized weights can be highly variable it is recommended to stabilize the weights (Cole and Hernan,  [2008](#page43)). Using the unstabilized weights as the denominator, stabilized weights are calculated by dividing the denominator by the predicted treatment propensity from a model using only time-invariant confounders and baseline information of the time-variant confounders as predictors. Because our analysis is stratified by males and females, we create weights separately for both groups.

The MSMs are estimated using OLS for the continuous and a logistic model for the binary outcomes. For the logistic model we calculate average marginal eﬀects for greater comparability with the results of the FE models. All models are weighted by the stabilized weights constructed beforehand while adjusting for all baseline and time-invariant covari-

8

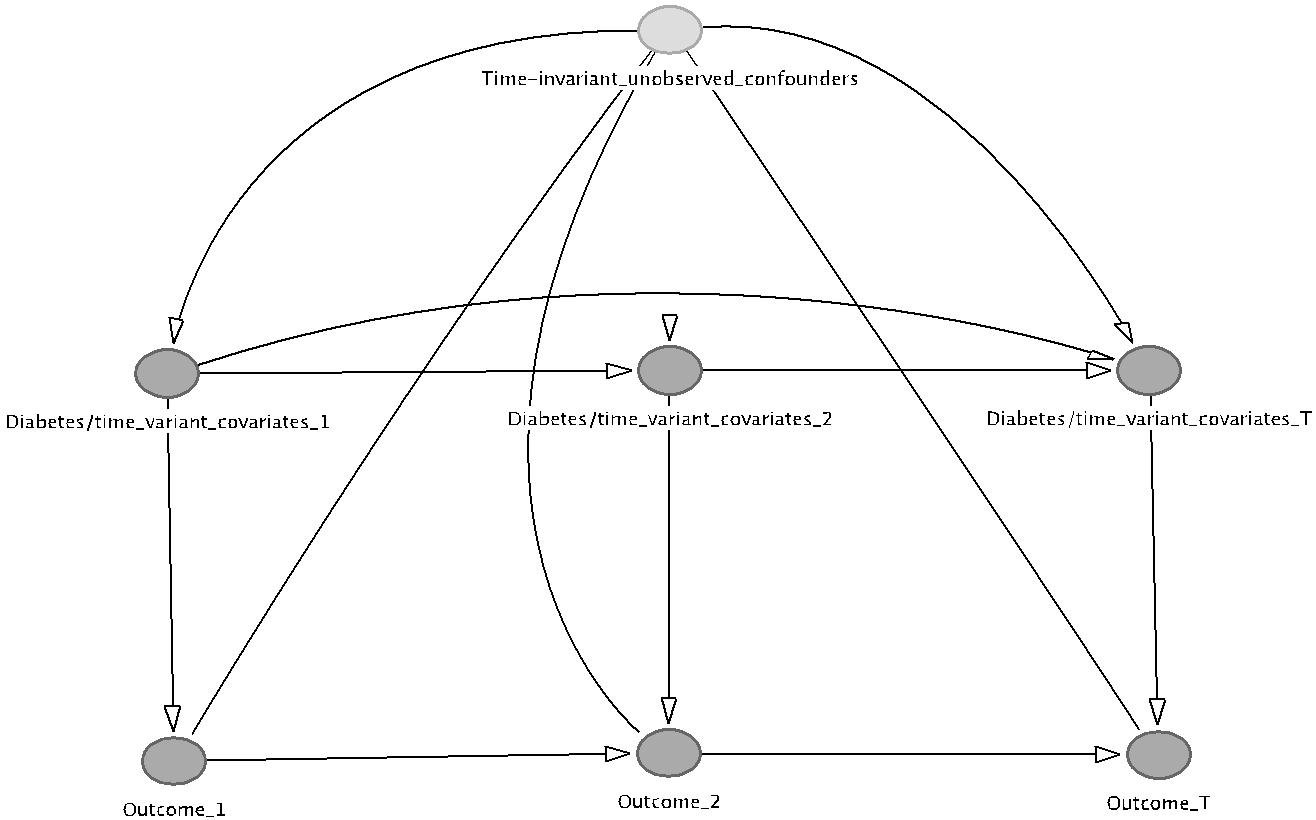
ates used in the calculation of the stabilized weights, except for the respective outcome of interest. Robust standard errors to account for intra-class correlation of repeated outcome measurements in individuals are used throughout. In our primary analysis, we present the results of the MSM with untruncated stabilized weights, as these provide theoretically unbiased estimates, albeit they likely are less eﬃcient than truncated weights (Cole and Hernan,  [2008](#page43)). The distribution of the inverse probability weights supports this decision as there are no extreme values and the mean weight is 1 (see Table  [0.6](#page28)).

**Fixed eﬀects**

While the MSM can account for pre-treatment selection on observable and time-variant confounders, it assumes that there are no unobserved time-invariant confounders such as family background, cognitive abilities, and other personal characteristics. This is a strong assumption that might be violated in practice. The individual level FE model can help remedy this problem as it is able to account for both observed time-variant and invariant variables as well as time-invariant unobserved variables as shown in the DAG in Figure  [0.2.](#page10) It does so by demeaning all covariates at each time point with the overall individual mean across all observed time points. It then uses solely the within-person variation for identification, thereby accounting for any time-invariant observed or unobserved as well as observed time-variant eﬀects.

This comes at a price: due to the demeaning, time-invariant variables such as Han ethnicity, are dropped from the model and cannot be estimated. Further, because the FE model is not able to account for any eﬀects of a diabetes diagnosis on other time-variant confounders, only a more limited set of confounders can be included compared to the MSM. Otherwise the estimates of the eﬀect of a diabetes diagnosis would likely be biased due to the inclusion of ’bad controls’ . Bad controls are control variables that have been aﬀected by the treatment itself—such as BMI or smoking status after a diabetes diagnosis—and therefore likely capture part of the causal eﬀect of diabetes on the outcome of interest, biasing the diabetes coeﬃcient (Angrist and Pischke,  [2008](#page43)). Our FE specifications thus only include controls for age, age squared, the level of urbanization, education, being married, having any medical insurance, living in a rural area, region and time dummies as well as per capita household income. For the estimation of the eﬀect of time since diagnosis, the linear age variable is dropped. In FE models, two or more variables that change at the same rate between waves cannot be separately identified. In our case this applies for age and time-dummies, as both variables increase by one each additional year (Wooldridge,  [2012](#page45)). To identify the eﬀect of diabetes duration we have to rely on the presence of people without diabetes in the sample, for which diabetes duration does not. FE models also make another assumption, which have received much less attention, namely that there is no dynamic causal relationship between treatment and outcomes, i.e. that past treatments have no direct effect on current outcomes, and that past outcomes no direct effect on current treatment (Imai and Kim 2016).

Figure 0.2: DAG for fixed eﬀects model



*Notes* FE models account for time-invariant unobserved confounding (light grey circle), but still assumethe absence of unobserved time-variant confounding. They further do not allow for past outcomes to aﬀect the current treatment, i.e. diabetes status.

increase at the same rate as time.

Because it is not possible to retrieve average marginal eﬀects from a logistic FE model, we prefer to use a linear FE model instead. It generally produces very similar estimates compared to non-linear models (Angrist and Pischke,  [2008](#page43)).

**Random eﬀects**

**Multiple imputation**

To deal with missing data, we used chained multiple imputation to impute the missing values in Stata 13 using the user written ICE command (Royston and White,  [2009](#page45)). Overall, thirty imputed datasets were created. Imputation models included all variables used in the MSMs. We imputed missing data in the same wave for which some data were recorded; we did not impute completely missing waves. Further, we did not impute missing diabetes information and instead assumed that once a diabetes diagnosis was

10

reported, the individual had diabetes in every ensuing wave, even when the observation was missing. If diabetes was never reported in any wave, we assumed that the individual never had diabetes. We then only imputed missing values for those observations that had a non-missing diabetes status. For the calculation of the marginal eﬀects in the MSM logit models, Rubin’s rules were applied using the user written Stata command mimrgns (Klein,  [2014](#page44)).

**Numbers of observations**

Because we used lagged variables to construct the stabilized weights for the MSMs, the number of observations used in the MSMs is lower than those used in the FE models, where we do not use lagged variables. The summary statistics shown in Table  [0.1](#page13) are based on the observations used in the FE models. The number of observations is stated below each table.

**Sensitivity analyses**

We conduct four additional sensitivity analyses in order to test the robustness of our results to diﬀerent assumptions and estimation strategies. First, we estimate all models using only covariate adjustment in a RE model, to investigate in how far this alternative approach diverts from the estimates generated of the FE and MSMs. These results are presented and discussed together with those of the MSM and the FE model. Second, we truncate weights at the 1st and 99th percentile to investigate the sensitivity of the MSMs to the most extreme weights. While untruncated weights provide unbiased estimates under the assumptions of the MSM, they may not be the most eﬃcient and tend to have larger standard errors (Cole and Hernan,  [2008](#page43)). Third, we estimate the FE and MSMs using the original non-imputed data to ascertain the extent to which multiple imputation aﬀected the results. Fourth, we report in apednix the estimates of models using overweight and obesity instead of BMI and waist circumference as the outcomes of interest, to investigate the eﬀect of a diabetes diagnosis on changes in the probabilities to be overweight or obese.

**0.3 Results**

From the descriptive statistics, we can observe that people with diabetes in any wave are less likely to be employed. Looking at health behaviours, the prevalence of smoking and drinking is lower for men with diabetes; they also consume fewer calories compared

11

to men without diabetes. Note that it is mainly men who smoke and report alcohol consumption while very few women do so.

Further, the diabetes group has both higher BMI and waist circumference levels. They are also older, live in more urbanized areas, are more likely to have insurance and men are somewhat better educated while women are less educated compared to their counterparts without diabetes. Both men and women report an average time since diagnosis of around 4.5 years. Looking at per capita household income, men and women with diabetes come from household with higher income levels than those without a diabetes diagnosis. Further it appears that in China it is less educated women that report a diagnosis, while men with diabetes are better educated compared to those without diabetes.

Predicting the denominator for the stabilized weights we find that for men a higher baseline BMI increases the risk of a diabetes diagnosis. Further, increases in age, waist circumference as well as urbanization levels are associated with higher chances for men to be diagnosed with diabetes throughout the survey. Interestingly becoming employed decreases the chances of being diagnosed with diabetes slightly, justifying the use of the MSM in our employment models as well (Table  [0.2](#page15)). Because these are not causal esti-mates, it may be that it is more likely for men with a lower risk of diabetes to select into employment. Interestingly, we do not find that higher household income levels are pre-dictive of a diagnosis for men or women, despite what the descriptive statistics indicated. For women, higher age and waist circumference at baseline, increases in BMI as well as living in a non-rural environment predict a diabetes diagnosis.

12

|  |
| --- |
| 13 |

Table 0.1: Sample means for males and females, by diabetes status

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Males |  |  |  | Females |  |  |
|  |  |  |  |  |  |  |  |  |
|  | No diabetes | Diabetes | p-value (t-test) |  | No diabetes | Diabetes | p-value (t-test) |  |
|  |  |  |  |  | |  |  |  |
| Employed | 82% | 68% | *<*0.001 | 67% | | 29% | *<*0.001 |  |
| Smokes | 58% | 47% | *<*0.001 | 3% | | 4% | 0.409 |  |
| Any alcohol consumption | 63% | 53% | *<*0.001 | 9% | | 4% | *<*0.001 |  |
| Daily Kcal eaten (3-day average) | 2422 | 2166 | *<*0.001 | 2068 | | 1931 | 0.001 |  |
| BMI | 22.99 | 24.90 | *<*0.001 | 23.10 | | 25.80 | *<*0.001 |  |
| Waist circ. (cm) | 82.02 | 88.81 | *<*0.001 | 78.80 | | 87.55 | *<*0.001 |  |
| Age | 42.27 | 52.76 | *<*0.001 | 43.24 | | 55.32 | *<*0.001 |  |
| Han ethnicity | 87% | 89% | 0.292 | 87% | | 93% | 0.002 |  |
| Rural area | 69% | 52% | *<*0.001 | 68% | | 51% | *<*0.001 |  |
| Married | 83% | 93% | *<*0.001 | 88% | | 87% | 0.392 |  |
| Secondary education | 65% | 68% | 0.439 | 50% | | 43% | 0.007 |  |
| University education | 5% | 11% | *<*0.001 | 4% | | 1% | 0.017 |  |
| Any health insurance | 51% | 82% | *<*0.001 | 50% | | 71% | *<*0.001 |  |
| Urbanization Index | 60.87 | 74.48 | *<*0.001 | 61.77 | | 68.68 | *<*0.001 |  |
| Per capita household income (Yuan (2011)) | 8617 | 16328 | *<*0.001 | 8581 | | 11101 | *<*0.001 |  |
| Years since diabetes diagnosis | − | 4.5 | − |  | − | 4.65 | − | |
| Observations | 23159 | 284 |  | 23369 | | 333 |  |  |
|  |  |  |  |  |  |  |  |  |

The results of our regression analysis are presented in Table  [0.3.](#page16) Both theMSM and FE model indicate that women with a diabetes diagnosis have lower probabilities of being employed than their counterparts without diabetes, with a reduction of 12 percentage points in the MSM and 11 percentage points in the FE model. This translates into a relative reduction in employment probabilities between 16–17%. For men no such eﬀect is observed.

A more ambiguous picture is painted for the eﬀect of a diabetes diagnosis on behavioural risk factor outcomes. According to the MSM, for males a diabetes diagnosis leads to smoking cessation, reductions in alcohol consumption as well as BMI, waist circumference and calorie consumption. Results for women look diﬀerent. While the point estimates indicate a reduction in all outcomes, these tend to be smaller than those for men and only exhibit strong statistical significance smoking cessation and alcohol consumptions, who already have a very low prevalence. Compared to the MSM, the FE model finds similar eﬀects for men, apart from a less important eﬀect on smoking cessation. For women, however, it finds much larger, and statistically significant, reductions in BMI and waist circumference.

The results of the RE models suggest an even stronger eﬀect of diabetes on female employment probabilities and smaller reductions in male and female BMI and waist cir-cumference, even suggesting a positive association between a diabetes diagnosis and female waist circumference. For the other outcomes, results are very similar to those from the MSMs and FE models.

Exploring the eﬀect of a diabetes diagnosis over time, we first estimate a specification using time since diagnosis as a continuous variable. The results of the MSMs (Table  [0.4](#page17)) indicate a steady reduction of female employment probabilities of close to two percentage points per year and of male alcohol consumption, BMI, waist circumference and calorie consumption. The FE model again supports the finding of the MSM, showing very similar, though somewhat larger, eﬀects in terms of size and statistical significance. The evidence for changes in risk factors for females is less consistent across models and outcomes, with the MSM suggesting almost no eﬀects while the FE model indicates a reduction in BMI. The eﬀect sizes for changes in health behaviours in women are consistently lower than those found for men.

The RE models again find larger eﬀects on female employment probabilities and a smaller impact of a diabetes diagnosis on reductions in BMI and waist circumference for both sexes.

In a second step we estimate a specification using year dummies to capture the potential non-linearity in the relationship between time since diagnosis and our outcomes. The

14

Table 0.2: Time variant and invariant predictors of a diabetes diagnosis (denominator of stabilized weights)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Males |  |  |  | Females |  |  |  |
|  |  | |  |  |  | |  |  |  |
|  | (1) | | (2) |  | (3) | | (4) |  |  |
|  | *β* | | SE |  | *β* | | SE | |  |
|  |  | |  |  |  | |  |  |  |
| Age (bl) | −.000 | | 0.001 |  | 0.004∗∗ | | 0.002 |  |  |
| Age squared (bl) | 0.000 | | 0.000 |  | −.000∗∗ | | 0.000 |  |  |
| BMI (bl) | 0.001∗∗∗ | | 0.000 | 0.001 | | | 0.000 |  |  |
| Waist circumference (cm) (bl) | 0.000 | | 0.000 |  | 0.000∗ | | 0.000 |  |  |
| 3-Day Ave: Energy (kcal) (bl) | −.000 | | 0.000 | 0.000 | | | 0.000 |  |  |
| Smoking (bl) | 0.001 | | 0.002 | 0.003 | | | 0.006 |  |  |
| Alcohol consumption (bl) | 0.003∗ | | 0.002 | 0.000 | | | 0.005 |  |  |
| Urbanization index (bl) | −.000 | | 0.000 |  | −.000 | | 0.000 |  |  |
| Secondary educ. (bl) | −.001 | | 0.003 | 0.003 | | | 0.003 |  |  |
| University educ. (bl) | −.000 | | 0.006 |  | − | |  |  |  |
| Married (bl) | −.002 | | 0.004 |  | −.000 | | 0.004 |  |  |
| Any medical insurance (bl) | 0.002 | | 0.002 |  | −.000 | | 0.002 |  |  |
| Employed (bl) | 0.002 | | 0.003 | 0.001 | | | 0.002 |  |  |
| Han ethnicity | 0.001 | | 0.003 |  | −.002 | | 0.003 |  |  |
| Rural | − | .001 | 0.002 |  | − | .005∗∗∗ | 0.002 |  |  |
| Per capita household income (2011 Yuan) (bl) | −.000 | | 0.000 |  | −.000 | | 0.000 |  |  |
| Survey year |  |  |  |  | −.001 | |  |  |  |
| 2004 | 0.002 | | 0.002 |  | 0.002 |  |  |
| 2006 | 0.003 | | 0.002 |  | −.003 | | 0.003 |  |  |
| 2009 | 0.009∗∗∗ | | 0.003 |  | −.001 | | 0.004 |  |  |
| 2011 | 0.001 | | 0.003 | 0.001 | | | 0.004 |  |  |
| Age | 0.003∗∗ | | 0.001 |  | −.002 | | 0.002 |  |  |
| Age squared | − | .000∗∗ | 0.001 | 0.000 | | | 0.000 |  |  |
| BMI | −.001 | | 0.000 |  | 0.001∗∗ | | 0.000 |  |  |
| Waist circumference (cm) | 0.000 | | 0.000 |  | −.000 | | 0.000 |  |  |
| 3-Day Ave: Energy (kcal) | −.000 | | 0.000 |  | −.000 | | 0.000 |  |  |
| Smoking | −.003 | | 0.002 | 0.000 | | | 0.006 |  |  |
| Alcohol consumption | −.004∗∗ | | 0.002 |  | −.003 | | 0.006 |  |  |
| Urbanization index | 0.000 | | 0.000 | 0.000 | | | 0.000 |  |  |
| Secondary education | 0.001 | | 0.003 | 0.000 | | | 0.003 |  |  |
| University education | 0.001 | | 0.006 |  | − | |  |  |  |
| Married | −.000 | | 0.004 |  | −.003 | | 0.004 |  |  |
| Any medical insurance | 0.001 | | 0.002 |  | −.001 | | 0.002 |  |  |
| Employed | −.004∗∗ | | 0.002 |  | −.003 | | 0.002 |  |  |
| Per capita household income (2011 Yuan) (2011 Yuan) | 0.000 | | 0.000 |  | −.000 | | 0.000 |  |  |

∗ *p <* 0*.*10,∗∗ *p <* 0*.*05,∗∗∗ *p <* 0*.*01

Results for province dummies omitted to preserve space. No observations for women with university education and diabetes.

15

Table 0.3: Analysis of the eﬀect of a diabetes diagnosis on employment status and be-havioural outcomes using MSM, FE and RE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Marginal structural model* | |  |  |  |
| Male sample | −.009 | −.070∗∗ | −.094∗∗∗ | −.735∗∗∗ | −1.887∗∗∗ | −135.061∗∗ |  |
| Diabetes |  |
| Female sample | (.026) | (.032) | (.036) | (.180) | (.574) | (58.593) |  |
| −.117∗∗∗ | −.015∗ | −.029∗∗ | −.388 | −.335 | −45.630 |  |
| Diabetes |  |
|  | (.029) | (.008) | (.012) | (.240) | (.631) | (33.530) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Fixed eﬀects* | |  |  |  |
| Male sample |  | −.023 | −.104∗∗∗ | −.715∗∗∗ | −2.217∗∗∗ | −168.297∗∗∗ |  |
| Diabetes | 0.022 |  |
| Female sample | (.030) | (.032) | (.036) | (.183) | (.610) | (62.115) |  |
| −.112∗∗∗ | −.027∗∗ | −.012 | −.644∗∗ | −1.251∗∗ | −61.175 |  |
| Diabetes |  |
|  | (.035) | (.013) | (.010) | (.263) | (.616) | (47.420) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Random eﬀects* | |  |  |  |
| Male sample | −.022 | −.064∗∗ | −.104∗∗∗ | −.379∗∗ | −.756 | −172.467∗∗∗ |  |
| Diabetes |  |
| Female sample | (.028) | (.029) | (.029) | (.177) | (.542) | (48.768) |  |
| −.152∗∗∗ | −.021∗∗ | −.019∗∗∗ | −.263 |  | −39.267 |  |
| Diabetes | 0.459 |  |
|  | (.027) | (.011) | (.006) | (.247) | (.570) | (34.256) |  |

*Notes* Standard errors in parentheses. Other control variables: age (only MSM), age squared, region, urban,education, han, marital status, urbanization index, time dummies, health insurance status, per capite house-hold income. Fixed/random eﬀects: N=23443 (male sample), N=23702 (female sample); MSM: N=16047 (male sample), N=16658 (female sample).

∗ *p <* 0*.*10,∗∗ *p <* 0*.*05,∗∗∗ *p <* 0*.*01)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 0.4: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using MSM, FE and RE |  |  |  |  |  |  |  |  |  |
|  | (1) | (2) | (3) |  | (4) | (5) | (6) |  |  |
|  | Employment | Smoking | Any alcohol |  | BMI | Waist (cm) | Calories (kcal) | |  |
|  |  |  |  | | | |  |  |  |
|  |  |  | *Marginal structural model* | | | |  |  |  |
| Male sample | −.003 | −.010∗ | −.014∗∗ |  | −.127∗∗∗ | −.340∗∗∗ | −21.770∗∗ | |  |
| Time since diagnosis |  |  |
| Female sample | (.004) | (.005) | (.007) | (.031) | | (.099) | (9.842) |  |  |
| −.017∗∗∗ | −.002 | −.004 |  | −.066∗ | −.072 | −8.735 | |  |
| Time since diagnosis |  |  |
|  | (.005) | (.001) | (.003) | (.040) | | (.109) | (5.589) |  |  |
|  |  |  |  | | |  |  |  |  |
|  |  |  | *Fixed eﬀects* | | |  |  |  |  |
| Male sample | −.001 | −.003 | −.017∗∗ |  | −.150∗∗∗ | −.520∗∗∗ | −22.286∗∗ | |  |
| Time since diagnosis |  |  |
| Female sample | (.007) | (.006) | (.007) | (.037) | | (.121) | (11.083) |  |  |
| −.019∗∗∗ | −.003 | −.000 |  | −.102∗∗∗ | −.215∗ | −6.747 | |  |
| Time since diagnosis |  |  |
|  | (.007) | (.002) | (.001) | (.039) | | (.117) | (7.028) |  |  |
|  |  |  |  | | |  |  |  |  |
|  |  |  | *Random eﬀects* | | |  |  |  |  |
| Male sample | −.006 | −.009∗ | −.015∗∗∗ |  | −.099∗∗∗ | −.269∗∗∗ | −24.703∗∗∗ | |  |
| Diabetes |  |  |
| Female sample | (.006) | (.006) | (.005) | (.035) | | (.096) | (8.655) |  |  |
| −.023∗∗∗ | −.002 | −.002∗∗ |  | −.056 |  | −6.444 | |  |
| Diabetes |  | 0.013 |  |
|  | (.006) | (.002) | (.001) | (.039) | | (.114) | (5.670) |  |  |

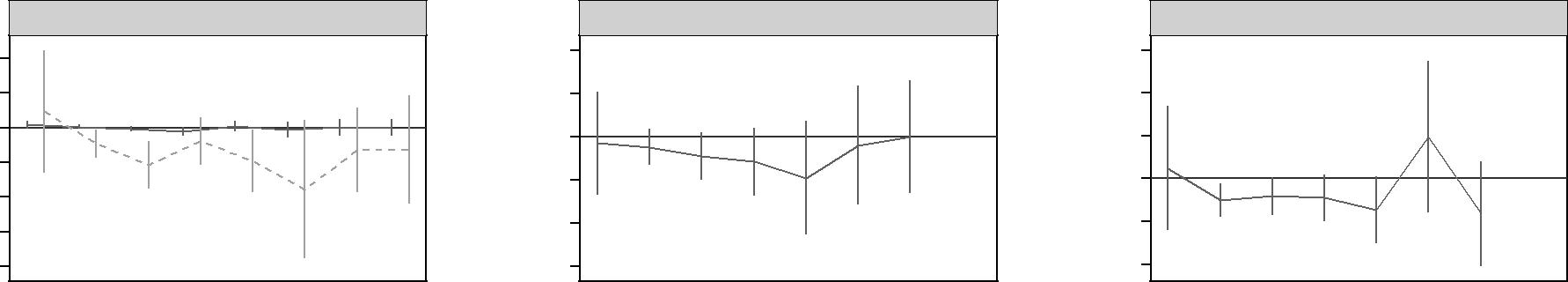
*Notes* Standard errors in parentheses. Other control variables: age (only MSM), age squared, region, urban,education, han, marital status, urbanization index, time dummies, health insurance status, household expen-ditures. Fixed/random eﬀects: N=23443 (male sample), N=23702 (female sample); MSM: N=16047 (male sample), N=16658 (female sample)

∗ *p <* 0*.*10,∗∗ *p <* 0*.*05,∗∗∗ *p <* 0*.*01)

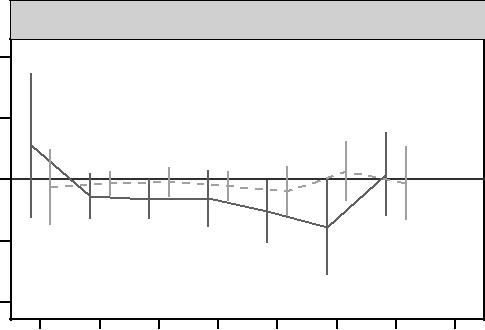
results for the diﬀerent estimation methods are visualized in Figures  [0.3,](#page19)  [0.4](#page20) and  [0.5](#page21) and presented in Tables  [0.7,](#page30)  [0.8](#page31) and  [0.9](#page32) for the MSM, FE and RE model, respectively. Despite the reduced sample size in each group and hence lower precision, the MSM model still indicates a reduction in female employment chances and male BMI, waist circumference and calorie consumption as well as smoking and alcohol consumption, especially in the first 8 to 10 years after diagnosis. Behavioural risk factors for women are again not found to be reduced consistently, apart from BMI where some trend towards a reduction over time is visible. Interestingly, female employment already decreases rapidly in the first to second year after diagnosis and it does not appear that females are able to increase their employment probabilities later on. Unfortunately it was not possible to estimate the eﬀects on female smoking and alcohol consumption due to the low prevalence of these risk factors in females, the lower sample size in the MSM and the reductions in sample size in each duration group. Using the FE model, all point estimates suggest similar eﬀects. The RE model, again suggests larger eﬀects on female employment and lower eﬀects on BMI and waist circumference than both other estimation methods.

The sensitivity analyses using truncated weights shows very similar eﬀects to those using the untruncated weights (Table  [0.10](#page33) and  [0.11),](#page34) suggesting no important loss in eﬃciency and supporting the decision to use untruncated weights. The results using non-imputed data are broadly similar (Tables  [0.12,](#page35)  [0.13,](#page36)  [0.14,](#page37)  [0.15](#page38) and  [0.16](#page39) ), in particular for the FE model, still indicating a reduction in female employment chances and male alcohol consumption, BMI and waist circumference. The coeﬃcients of the MSM still point into the same direction as those using the imputed data, but the estimated eﬀects are smaller in size and confidence intervals are relatively large. The RE model still shows a stronger eﬀect on female employment probabilities and smaller reductions in especially the weight measures BMI and waist circumference. Using overweight and obesity instead of BMI and waist circumference as indicators for weight changes, we do not find as consistent reductions in weight status for men as we did using the continuous estimates (Tables  [0.17](#page40) and  [0.18](#page41) and Figure  [0.6.](#page42) Nonetheless, the point estimates still show a reduction in obesity, in particular over time and for men, supporting the reductions found using continuous measurements. The coeﬃcients for overweight are diﬃcult to interpret as it is unclear if the negative coeﬃcient is caused by people transferring into the obesity or into normal weight.

Figure 0.3: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes (duration groups, marginal structural model)

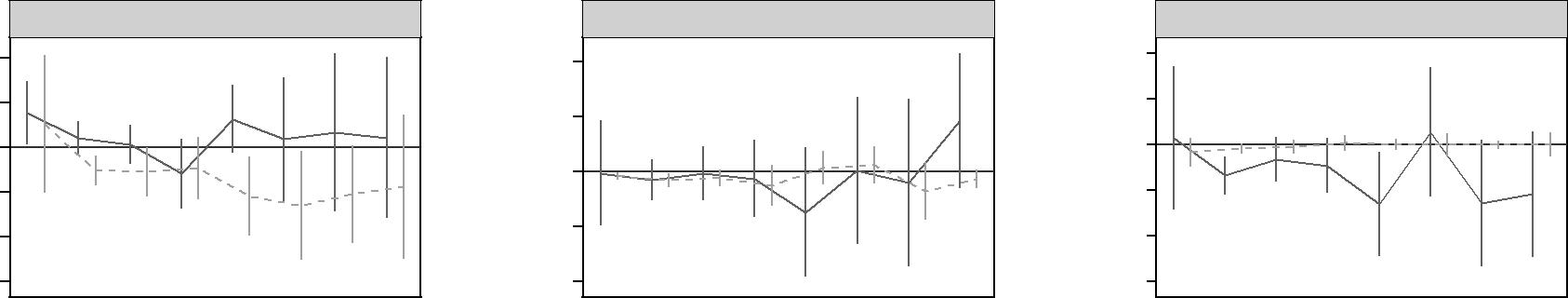


|  |
| --- |
| 19 |



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Kcal | |  |  |  |
| 1000 |  |  |  |  |  |  |  |
| 500 |  |  |  |  |  |  |  |
| 0 |  |  |  |  |  |  |  |
| −500 |  |  |  |  |  |  |  |
| −1000 |  |  |  |  |  |  |  |
| 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 |

Figure 0.4: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes (duration groups, fixed eﬀects)



|  |
| --- |
| 20 |

|  |
| --- |
| Marginal effect |

Employed

.4

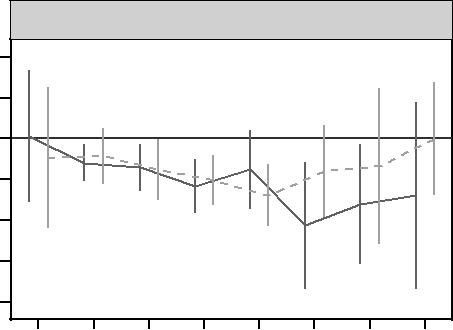
.2

0

−.2

−.4

−.6



BMI

2

1

0

−1

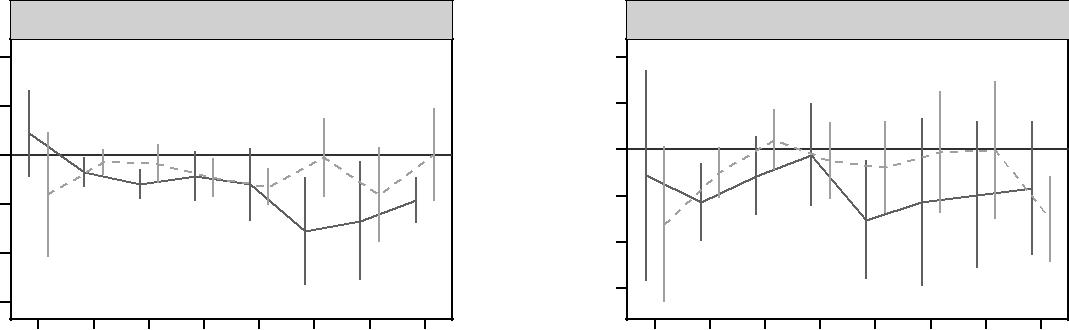
−2

−3

−4

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 |

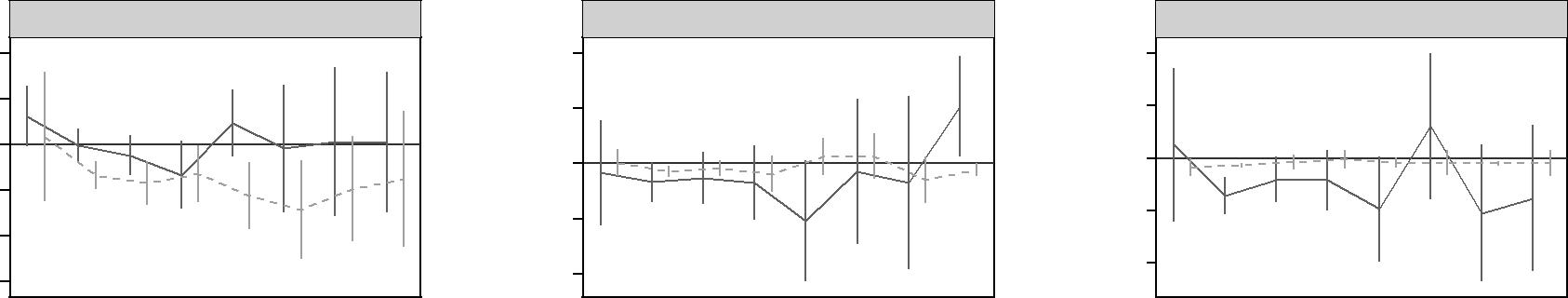
|  |  |  |
| --- | --- | --- |
| Smoking | Alcohol |  |
| .4 | .4 |  |
|  |  |
| .2 | .2 |  |
|  |  |
|  | 0 |  |
| 0 | −.2 |  |
|  |  |
| −.2 | −.4 |  |
|  |  |
| −.4 | −.6 |  |



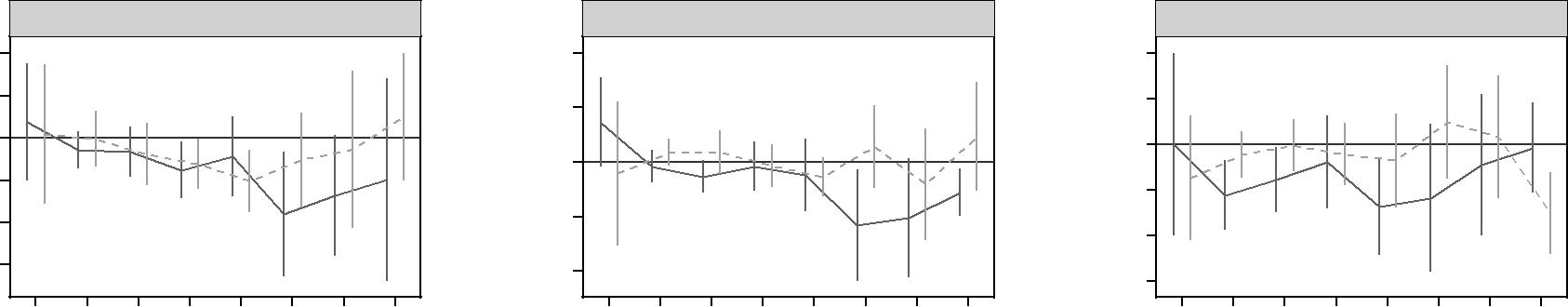
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Waist (cm) | |  |  |  |  |  |  | Kcal | |  |  |  |  |
| 10 |  |  |  |  |  |  |  | 400 |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  | 200 |  |  |  |  |  |  |  |  |
| 0 |  |  |  |  |  |  |  | 0 |  |  |  |  |  |  |  |  |
| −5 |  |  |  |  |  |  |  | −200 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| −10 |  |  |  |  |  |  |  | −400 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| −15 |  |  |  |  |  |  |  | −600 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 | 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 |  |
| Years after diagnosis | | | | | |  |  |  |  |  |  |  |  |  |  |  |
| men | |  |  |  | women | |  |  |  |  |  |  |  |  |  |  |

|  |
| --- |
| 21 |

Figure 0.5: The eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes (duration groups, random eﬀects)



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | Employed | |  |  |  |  |  |  | Smoking | |  |  |  |  |  |  | Alcohol | |  |  |  |  |
|  | .4 |  |  |  |  |  |  |  | .4 |  |  |  |  |  |  |  | .4 |  |  |  |  |  |  |  |  |
|  | .2 |  |  |  |  |  |  |  | .2 |  |  |  |  |  |  |  | .2 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0 |  |  |  |  |  |  |  | 0 |  |  |  |  |  |  |  | 0 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | −.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | −.2 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | −.2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 | 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 | 0 | 1−2 | 3−4 | 5−6 | 7−8 | 9−10 | 11−12 | 13−14 |  |
|  |  |  |  |  |  |  |  |  | Years after diagnosis | | | | | |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | men | |  |  |  | women | |  |  |  |  |  |  |  |  |  |  |



**0.4 Disussion**

The evidence for the impact of a diabetes diagnosis on employment chances and be-havioural risk factors remains scarce, in particular in MICs, where diabetes has become a mayor contributor to the burden of disease. We added to this evidence by exploring these relationships using longitudinal data from China, also improving upon previous method-ology by taking into account the potential relationship over time between these outcomes.

Our results suggest that receiving a diabetes diagnosis in China leads to a strong and lasting reduction in female, but not male employment probabilities. We also found reduc-tions in male BMI and waist circumference, alcohol and calorie consumption and poten-tially smoking. We did, however, not find similar changes in behavioural risk factors for women. Accordingly, it appears that women in China have to endure stronger adverse labour market eﬀects and at the same time are less successful then men at making risk behaviour changes to reduce their risk of diabetes complications.

The MSM models and FE models indicated very similar results suggesting that they are robust and that time-invariant confounding factors may play a limited role over and above baseline and time varying confounding factors. The MSM results suggest that in particular BMI and waist circumference levels as well as employment status can cause selection into a diabetes diagnosis and are then later themselves aﬀected by the diagnosis, justifying the use of a MSM. The robustness checks using ’naive’ regression in the form of RE models further indicated that insuﬃciently accounting for confounding can—at least in this setting—lead to an overestimation of the impact of diabetes on employment status and an underestimation of the eﬀects of a diagnosis on weight measures (BMI and waist circumference). However, confounding may only be of limited relevance for risk behaviours (smoking and alcohol consumption) and caloric intake.

**0.4.1 Limitations**

The study has several limitations. While we used two estimation methods to reduce the influence of selection bias due to unobserved confounding, they are still unable to account for all forms of selection simultaneously. Therefore a causal interpretation is only possible under restrictive assumptions, namely no unobserved time-variant confounding for the FE model and positivity, exchangeability and consistency for the MSM. The assumption of positivity is likely to hold, given that every person should have at least a small chance of receiving a diabetes diagnosis. This is also supported by the relatively small range of stabilized weights and absence of zero-weights. Exchangeability, or no unmeasured confounding could potentially be violated if not all time-invariant or

time-variant confounders were accounted for, but there is no comprehensive test. We tested for part of this assumption by estimating a FE model and given that the results remain very similar, this suggests that unobserved time-invariant confounding may be of limited relevance in this case. Consistency would have been violated if a diabetes diagnosis had been reported but the person had actually not been diagnosed with diabetes. This was likely only violated in very rare cases of misreporting, given that specificity of diabetes self-report is very high in China (Yuan et al.,  [2015](#page46)). Because we were interested in the eﬀect of a diabetes diagnosis, unobserved diabetes did not violate the consistency assumption.

A limitation of the FE model is the possibility of time-variant confounding due to prior outcomes (for example employment status) aﬀecting the current treatment (a diabetes diagnosis). Given that the FE estimates were close to those of the MSMs, it is likely that there was no strong confounding due to pre-treatment changes. Rather, the similarity of results suggests that it is important to account for the selection into diabetes due to some form of baseline values, be it via demeaning as in the FE model—and thereby accounting for all time-invariant confounding—or by using baseline values as in the MSM.

Finally, an important limitation is the that a diabetes diagnosis entails a variety of ’treatments’ that are diﬃcult to disentangle and may each have a distinct eﬀect on the explored outcomes. Currently, we are only able to observe the combined eﬀect of these treatments. Firstly, there is the provision of information at diagnosis, potentially causing increases in stress and anxiety, but may also providing an explanation for the experi-enced symptoms, both potentially aﬀecting productivity. Secondly, a diagnosis also is the starting point for medical treatment, potentially alleviating symptoms and helping with weight loss, but also posing new challenges, in particular if treatment entails the exogenous provision of insulin or strict meal plans, potentially adding to the burden of diabetes in daily life. Thirdly, adherence to medical treatment may be heterogeneous across people with diabetes, with non-adherence likely leading to a further worsening of risk factors for complications, while good adherence may prevent or delay debilitating complications. Fourthly, a diagnosis may also introduce lifestyle changes such as increasing exercise levels, eating healthier and reducing smoking or alcohol consumption, all potentially aﬀecting the risk to develop further complications and to experience changes in productivity. In the current study, it is not possible to ascertain the role of each of these factors in aﬀecting employment chances and behavioural risk factors. Only for the reductions in smoking and alcohol consumption, it seems reasonable to attribute them to diagnosis induced awareness to reduce these risk factors.

23

**0.4.2 Potential mechanisms**

There are various mechanisms that may explain the observed patterns in the effects of a diabetes diagnosis on employment and behavioural risk factors for males and females.

The found adverse eﬀect of diabetes on employment is in line with other studies on the labour market impact of diabetes that have found diabetes to reduce female employment probabilities (Harris,  [2009;](#page43) Latif,  [2009;](#page44) Minor,  [2011;](#page44) Seuring, Serneels, et al.,  [2016](#page45))—often more than for men. Most comparable to our results is likely a study from Mexico which also used FE and data for a similar time period and for a MIC that also experienced a rapid epidemiological transformation towards a very high diabetes burden (Seuring, Serneels, et al.,  [2016](#page45)). The study found significant reductions for both males and females of about 5 percentage points. Taking into account the lower overall employment rate of Mexican women compared to men, this translated into a 16% reduction in female employ-ment probabilities, a figure comparable to what Chinese women experienced. However, in Mexico also men experienced adverse eﬀects, unlike to what we found for China.

The found eﬀects on changes in behavioural risk factors can be compared to the study by (Slade,  [2012](#page45)). Slade finds reductions in alcohol consumption and smoking, though it appears that these reductions were not maintained over a longer time period. Unfortu-nately, Slade only provided information for the entire sample and the male sample, so that we cannot compare them directly with our results for women. In terms of the eﬀect on weight, again both studies cannot be directly compared because Slade investigated the ef-fect of a diagnosis on being overweight or obese, while we used continuous weight measures in our primary analysis due to the discussed diﬃculties of defining cut-oﬀ values for Asian populations. Slade found an initial reduction in weight status, but also that people with diabetes tended to become more likely to be overweight or obese after some time. Our results using overweight and obesity could tentatively be interpreted to indicate a more constant reduction in obesity over time, suggesting that reductions in weight in Chinese men may be longer lived than in the USA. Importantly—and in concordance with our

findings—he found that simple covariate adjustment led to biased estimates of the impact on weight status, indicating a positive relationship. This underlines the importance of accounting for unobserved heterogeneity.

The permanent reduction in male BMI and waist circumference we have found has also been observed in a cohort of Danish patients (De Fine Olivarius et al.,  [2015),](#page43) where weight increased the years preceding diagnosis, while after diagnosis weight decreased. The exact reasons for this decrease were unknown but attributed to motivation changes as a result of the diagnosis, concluding that time around the diagnosis may represent a window of opportunity to obtain long lasting weight change. Nonetheless, reductions in weight, as already eluded to in the limitations, may also be the result of treatment initiation with metformin or other diabetes drugs that have been shown to lead to weight reductions (Yang and Weng,  [2014](#page45)). Importantly, the reduction in male BMI levels and waist circumference were accompanied by reduced energy intake, suggesting that the changes in weight were at least partly the result of lower energy intake. Further, given that in China diabetes incidence has been especially attributed to a high accumulation of visceral fat and central obesity (Ma et al.,  [2014),](#page44) the reductions in waist circumference may have had a particular positive eﬀect on diabetes control and the prevention of comorbidities. Together, the lower levels of energy intake and waist circumference after the diagnosis allow for the interpretation that the reductions in BMI were due to fat loss and not less lean body mass (Klein et al.,  [2007](#page44)).

For women, however, we did not find similar strong evidence for reductions in BMI, waist circumference or energy intake. The relatively smaller eﬀects for women could indicate a lower ability to change behaviours supportive of weight loss. This appears to be supported by the smaller reductions in energy intake. This could have—at least partly—contributed to a higher risk for diabetes complications further down the line, also adversely aﬀecting employment probabilities. Apart from this, other explanations for the lower weight loss and larger employment penalty for women compared to men include their lower educational attainment, which has been indicated as a factor in preventing better glucose control (Luo et al.,  [2015)](#page44) and may also aﬀected the ability to successfully change behaviours. Lower income levels for females compared to men may also have negatively aﬀected the ability to receive adequate treatment following a diagnosis, limiting their ability to change health behaviours (Luo et al.,  [2015),](#page44) increasing the risk of complications. We found that women with diabetes lived in households with lower income levels compared to men with diabetes, however, these income levels were still higher then for those without diabetes. Nonetheless, it may still be the case that women were more likely to not access care due to lower income levels than men. Further, there are likely biological factors that lead to worse health outcomes for women compared to men. There is some evidence that, due to diﬀerent ways of fat storage between men and women, men tend to cross the diabetes threshold

at an earlier point in time and at a comparatively healthier metabolic state then women (Peters, Huxley, Sattar, et al.,  [2015;](#page44) Peters, Huxley, and Woodward,  [2014a,b](#page45)). Women are more likely to have spend more time in a pre-diabetes state (Bertram and Vos,  [2010](#page43)) and to cross the threshold only once the metabolic has significantly deteriorated, leading to a greater risk of cardiovascular disease and stroke (Peters, Huxley, Sattar, et al.,  [2015](#page44)). Supporting this, a study for China found a greater prevalence of diabetes comorbidities in Chinese women than men (Liu et al.,  [2010](#page44)). In this light it may not be surprising that we find more conclusive evidence of worsening employment probabilities for women than for men. If women are less likely to receive proper treatment and to change their health behaviours and at the same time have a greater risk for complications then men, the long term eﬀects of diabetes on their health are likely more severe than for men and consequently aﬀect their employment status to a greater extent.

Taken together these estimation results suggest that the effect on the probability of employment is reduced over time due to adaptions in health behaviour, while the effect for women is substantial because no such changes in behaviour take place. Further analysis is needed to test this formally, and is beyond the scope of this paper.

**0.5 Conclusion**

Our results indicate worse outcomes for women then men after a diabetes diagnosis, with women experiencing a reduction in employment probabilities accompanied by and poten-tially partly due to an inability to reduce important risk factors for diabetes complications. For males, the opposite pattern is found, as they do not experience adverse employment eﬀects and are able to achieve reductions in the investigated risk factors. These findings are robust to the application of two distinct, but complementary econometric techniques. Further research should try to unravel the mechanisms behind these diﬀerential outcomes for men and women. Overall, given the large prevalence of undiagnosed diabetes, our re-sults indicate that an early diagnosis may be a good way to foster early behaviour change that could lead to more positive health and economic outcomes for people with diabetes over time. It appears, however, that greater emphasis needs to be put on reducing the burden of diabetes for women if the observed inequities in the diabetes impact shall be reduced.

**Attrition**

Table 0.5: Attrition between waves

1997–2000 11.9% 2000–2004 13.0% 2004–2006 8.3% 2006–2009 16.2% 2009–2011 16.7% Total 10.6%

**Stabilized weights**

Table 0.6: Summary of stabilized weights

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mean | Min | Max |
|  |  |  |  |
| Untruncated (men) | 1.000515 | 0.281853 | 2.642838 |
| Untruncated (women) | 0.999907 | 0.451526 | 2.053581 |
| Truncated 1 and 99 percentile (men) | 0.999756 | 0.945491 | 1.057514 |
| Truncated 1 and 99 percentile (women) | 1.000001 | 0.960039 | 1.049472 |
|  | | |  |
| Using overweight and obesity instead of BMI and waist circumference | | |  |
| Untruncated (men) | 1.000516 | 0.232143 | 2.592925 |
| Untruncated (women) | 0.999857 | 0.251297 | 2.491703 |
| Truncated 1 and 99 percentile (men) | 0.999794 | 0.944632 | 1.058910 |
| Truncated 1 and 99 percentile (women) | 0.999782 | 0.932321 | 1.077095 |
|  |  |  |  |

**Duration groups results**

29

Table 0.7: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using marginal structural models (duration groups)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |
| Male sample |  | −.031 |  | −1.138∗∗ | −.728 |  |  |
| 0 | 0.088 | 0.049 | 278.504 |  |
|  | (.059) | (.122) | (.147) | (.530) | (1.927) | (301.190) |  |
| 1-2 | 0.024 | −.049 | −.102∗∗ | −.485∗ | −1.261 | −133.527 |  |
|  | (.034) | (.042) | (.040) | (.260) | (.876) | (96.402) |  |
| 3-4 | −.033 | −.091 | −.082∗ | −.665∗∗ | −2.505∗∗∗ | −160.612∗ |  |
|  | (.042) | (.056) | (.045) | (.309) | (.814) | (84.241) |  |
| 5-6 | −.110 | −.116 | −.090 | −.917∗∗ | −1.009 | −156.064 |  |
|  | (.068) | (.080) | (.056) | (.384) | (.980) | (117.322) |  |
| 7-8 | 0.044 | −.191 | −.146∗ | −.833∗ | −1.590 | −260.923∗∗ |  |
|  | (.076) | (.134) | (.079) | (.467) | (2.276) | (130.336) |  |
| 9-10 | −.052 | −.040 | 0.197 | −2.198∗∗∗ | −6.075∗∗ | −386.292∗ |  |
|  | (.117) | (.140) | (.181) | (.765) | (2.591) | (199.311) |  |
| 11-12 | 0.013 | −.001 | −.165 | −.881 | −3.505 | 40.936 |  |
|  | (.120) | (.132) | (.125) | (.708) | (2.522) | (174.858) |  |
| 13-14 | 0.004 |  |  |  |  |  |  |
|  | (.124) |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Female sample |  |  |  |  | −1.210 | −59.570 |  |
| 0 | 0.078 |  |  | 0.099 |  |
|  | (.139) |  |  | (1.021) | (3.866) | (157.723) |  |
| 1-2 | −.085∗∗ |  |  | −.191 | −.303 | −32.947 |  |
|  | (.040) |  |  | (.352) | (.724) | (50.797) |  |
| 3-4 | −.202∗∗∗ |  |  | −.411 | 0.591 | −21.502 |  |
|  | (.067) |  |  | (.461) | (1.232) | (62.460) |  |
| 5-6 | −.070 |  |  | −.475 | −.187 | −53.234 |  |
|  | (.066) |  |  | (.337) | (1.055) | (61.737) |  |
| 7-8 | −.180∗∗ |  |  | −1.049∗∗ | −1.787∗ | −94.532 |  |
|  | (.088) |  |  | (.426) | (1.057) | (105.698) |  |
| 9-10 | −.329∗ |  |  | −1.054 | 0.324 | 66.951 |  |
|  | (.168) |  |  | (.822) | (2.538) | (125.902) |  |
| 11-12 | −.119 |  |  | −.554 | −3.906 | −29.022 |  |
|  | (.120) |  |  | (1.089) | (2.464) | (152.223) |  |
| 13-14 | −.117 |  |  |  |  |  |  |
|  | (.154) |  |  |  |  |  |  |

*Notes* Other control variables: age, age squared, region, urban, education, han, marital status, urban-ization index, time dummies, health insurance status, household expenditures. N=16047 (male sample), N=16658 (female sample).

∗ *p <* 0*.*10,∗∗ *p <* 0*.*05,∗∗∗ *p <* 0*.*01)

Table 0.8: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using fixed eﬀects (duration groups)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |
| Male sample |  | −.005 |  |  |  | −112.476 |  |
| 0 | 0.151∗∗ | 0.027 | 0.064 | 2.200 |  |
|  | (.072) | (.097) | (.161) | (.822) | (2.257) | (232.264) |  |
| 1-2 | 0.040 | −.029 | −.137∗∗∗ | −.598∗∗∗ | −1.714∗∗ | −228.738∗∗∗ |  |
|  | (.038) | (.038) | (.042) | (.230) | (.784) | (85.913) |  |
| 3-4 | 0.010 | −.007 | −.066 | −.706∗∗ | −2.992∗∗∗ | −113.409 |  |
|  | (.044) | (.051) | (.050) | (.296) | (.797) | (86.909) |  |
| 5-6 | −.118 | −.026 | −.093 | −1.164∗∗∗ | −2.191∗ | −22.369 |  |
|  | (.079) | (.072) | (.062) | (.341) | (1.309) | (112.692) |  |
| 7-8 | 0.126 | −.147 | −.262∗∗ | −.750 | −3.009 | −302.744∗∗ |  |
|  | (.078) | (.120) | (.116) | (.493) | (1.886) | (131.910) |  |
| 9-10 | 0.036 | 0.004 | 0.054 | −2.123∗∗∗ | −7.756∗∗∗ | −228.356 |  |
|  | (.141) | (.138) | (.145) | (.788) | (2.799) | (184.833) |  |
| 11-12 | 0.066 | −.042 | −.256∗ | −1.604∗∗ | −6.693∗∗ | −195.061 |  |
|  | (.180) | (.156) | (.141) | (.742) | (3.094) | (160.761) |  |
| 13-14 | 0.042 | 0.186 | −.218 | −1.389 | −4.626∗∗∗ | −167.675 |  |
|  | (.183) | (.126) | (.140) | (1.168) | (1.190) | (147.716) |  |
|  |  |  |  |  |  |  |  |
| Female sample |  | −.015∗∗ | −.035 | −.468 | −4.036 | −322.767∗ |  |
| 0 | 0.102 |  |
|  | (.157) | (.007) | (.032) | (.884) | (3.229) | (171.460) |  |
| 1-2 | −.104∗∗∗ | −.031∗∗ | −.019∗ | −.419 | −.727 | −98.608∗ |  |
|  | (.034) | (.013) | (.011) | (.349) | (.683) | (56.443) |  |
| 3-4 | −.110∗∗ | −.022 | −.012 | −.756∗∗ | −.896 | 42.743 |  |
|  | (.056) | (.015) | (.016) | (.378) | (1.000) | (67.154) |  |
| 5-6 | −.095 | −.049 | 0.007 | −1.012∗∗∗ | −2.293∗∗ | −49.270 |  |
|  | (.072) | (.038) | (.018) | (.309) | (1.021) | (84.604) |  |
| 7-8 | −.219∗∗ | 0.014 | −.000 | −1.385∗∗∗ | −3.238∗∗∗ | −76.316 |  |
|  | (.090) | (.032) | (.013) | (.391) | (.962) | (102.021) |  |
| 9-10 | −.261∗∗ | 0.024 | −.001 | −.794 | −.240 | −12.562 |  |
|  | (.124) | (.035) | (.025) | (.572) | (2.056) | (134.903) |  |
| 11-12 | −.209∗ | −.070 | −.002 | −.676 | −4.068∗ | −2.327 |  |
|  | (.111) | (.053) | (.009) | (.973) | (2.462) | (152.643) |  |
| 13-14 | −.178 | −.026 | −.001 | −.001 | 0.056 | −301.362∗∗∗ |  |
|  | (.164) | (.018) | (.027) | (.708) | (2.411) | (94.674) |  |

*Notes* Other control variables: age squared, region, urban, education, han, marital status, urbanization index,time dummies, health insurance status, household expenditures. N=23443 (male sample), N=23702 (female sample).

∗ *p <* 0*.*10,∗∗ *p <* 0*.*05,∗∗∗ *p <* 0*.*01)

Table 0.9: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using random eﬀects (duration groups)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |
| Male sample |  | −.034 |  |  |  |  |  |
| 0 | 0.123∗ | 0.051 | 0.381 | 3.652∗ | 2.069 |  |
|  | (.068) | (.097) | (.150) | (.707) | (2.075) | (203.971) |  |
| 1-2 | −.005 | −.067∗ | −.142∗∗∗ | −.276 | −.392 | −223.036∗∗∗ |  |
|  | (.038) | (.037) | (.036) | (.224) | (.766) | (78.475) |  |
| 3-4 | −.048 | −.052 | −.081∗ | −.316 | −1.318∗ | −155.191∗∗ |  |
|  | (.044) | (.048) | (.045) | (.304) | (.769) | (72.913) |  |
| 5-6 | −.133∗ | −.071 | −.084 | −.759∗∗ | −.403 | −75.706 |  |
|  | (.076) | (.069) | (.058) | (.344) | (1.148) | (104.001) |  |
| 7-8 | 0.093 | −.208∗ | −.194∗ | −.434 | −1.172 | −272.523∗∗ |  |
|  | (.075) | (.112) | (.102) | (.485) | (1.703) | (109.241) |  |
| 9-10 | −.018 | −.028 | 0.122 | −1.804∗∗ | −5.786∗∗ | −234.745 |  |
|  | (.142) | (.134) | (.142) | (.749) | (2.609) | (166.358) |  |
| 11-12 | 0.012 | −.071 | −.209 | −1.360∗ | −5.108∗ | −90.369 |  |
|  | (.166) | (.160) | (.132) | (.726) | (2.790) | (158.103) |  |
| 13-14 | 0.008 | 0.206∗∗ | −.152 | −.985 | −2.776∗∗ | −14.049 |  |
|  | (.157) | (.093) | (.142) | (1.225) | (1.122) | (101.033) |  |
|  |  |  |  |  |  |  |  |
| Female sample |  |  | −.035∗∗ |  | −1.037 | −145.397 |  |
| 0 | 0.034 | 0.003 | 0.097 |  |
|  | (.145) | (.025) | (.017) | (.842) | (3.375) | (139.781) |  |
| 1-2 | −.135∗∗∗ | −.028∗∗∗ | −.026∗∗∗ | −.025 | 0.857 | −44.182 |  |
|  | (.031) | (.011) | (.004) | (.337) | (.631) | (52.022) |  |
| 3-4 | −.169∗∗∗ | −.018 | −.015 | −.379 | 0.901 | −3.834 |  |
|  | (.049) | (.014) | (.014) | (.372) | (1.005) | (57.700) |  |
| 5-6 | −.129∗∗ | −.038 | −.005 | −.612∗∗ | −.317 | −43.769 |  |
|  | (.063) | (.033) | (.018) | (.305) | (.992) | (69.632) |  |
| 7-8 | −.225∗∗∗ | 0.024 | −.018∗ | −1.015∗∗∗ | −1.357 | −69.287 |  |
|  | (.075) | (.034) | (.010) | (.377) | (.908) | (105.179) |  |
| 9-10 | −.286∗∗ | 0.026 | −.018 | −.515 | 1.421 | 98.605 |  |
|  | (.111) | (.042) | (.024) | (.572) | (1.937) | (127.672) |  |
| 11-12 | −.195∗ | −.060 | −.020∗∗∗ | −.265 | −2.043 | 31.945 |  |
|  | (.117) | (.043) | (.005) | (.948) | (2.622) | (137.113) |  |
| 13-14 | −.152 | −.022∗ | −.018 | 0.503 | 2.325 | −301.291∗∗∗ |  |
|  | (.152) | (.013) | (.026) | (.773) | (2.541) | (91.369) |  |

*Notes* Other control variables: age squared, region, urban, education, han, marital status, urbanization index,time dummies, health insurance status, household expenditures. N=23443 (male sample), N=23702 (female sample).

∗ *p <* 0*.*10,∗∗ *p <* 0*.*05,∗∗∗ *p <* 0*.*01)

**Robustness checks**

**MSMs using truncated weights**

Table 0.10: Analysis of the eﬀect of a diabetes diagnosis on employment status and be-havioural outcomes using marginal structural models with truncated stabilized weights at 1st and 99th percentile

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | | (2) | | (3) | | (4) | | (5) |  |  | (6) |  |
|  | Employment | | Smoking | | Any alcohol | | BMI | | Waist (cm) | Calories (kcal) | | |  |
|  |  |  |  |  |  |  | |  |  |  |  |  |  |
|  |  |  |  |  |  | *Diabetes* | |  |  |  |  |  |  |
| Male sample |  | .022 |  | .070∗∗ |  | .094∗∗∗ |  | .732∗∗∗ | 1.637∗∗∗ | 175.662∗∗∗ | | |  |
| Diabetes | − | − | − | − |  |
|  |  |  |  |  | − | − |  |  |  |
| Female sample | (.023) | | (.032) | | (.036) | | (.179) | | (.532) | (51.574) | | |  |
| −.132∗∗∗ | | −.015∗ | | −.029∗∗ | | −.178 | |  | −47.980 | | |  |
| Diabetes | 0.186 |  |
|  | (.029) | | (.008) | | (.012) | | (.248) | | (.638) | (34.319) | | |  |
|  |  |  |  |  |  | | | |  |  |  |  |  |
|  |  |  |  |  | *Years since diagnosis* | | | |  |  |  |  |  |
| Male sample |  | .006 |  | .010∗∗ |  | .016∗∗ |  | .133∗∗∗ | .326∗∗∗ |  | 26.261∗∗∗ | |  |
| Time since diagnosis | − | − | − | − | − |  |
|  |  |  |  |  | − |  |  |  |
| Female sample | (.004) | | (.005) | | (.006) | | (.033) | | (.095) |  | (9.160) | |  |
|  | .019∗∗∗ |  | .002 |  | .004 |  | .044 | .016 |  |  | 9.096 |  |
| Time since diagnosis | − | − | − | − | − | |  |
|  |  |  |  |  | − |  |  |
|  | (.006) | | (.001) | | (.003) | | (.042) | | (.112) |  | (5.681) | |  |

*Notes* Standard errors in parentheses. Other control variables: age squared, region, urban, education, han, maritalstatus, urbanization index, time dummies, health insurance status, household expenditures. N=16047 (male sample), N=16658 (female sample).

Table 0.11: Eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using MSM with truncated stabilized weights (1st and 99th per-centile; imputed)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |
| Male sample |  | −.047 |  | −1.107∗∗ | −.326 |  |  |
| 0 | 0.089 | 0.031 | 83.518 |  |
|  | (.061) | (.135) | (.143) | (.522) | (1.909) | (236.282) |  |
| 1-2 | −.002 | −.072∗ | −.121∗∗∗ | −.472∗ | −.962 | −197.071∗∗ |  |
|  | (.034) | (.041) | (.033) | (.254) | (.843) | (82.739) |  |
| 3-4 | −.042 | −.073 | −.088∗∗ | −.654∗∗ | −2.113∗∗∗ | −189.546∗∗ |  |
|  | (.038) | (.050) | (.040) | (.299) | (.693) | (77.787) |  |
| 5-6 | −.107∗ | −.091 | −.094∗ | −1.022∗∗∗ | −.954 | −151.346 |  |
|  | (.063) | (.074) | (.053) | (.360) | (1.013) | (107.678) |  |
| 7-8 | 0.054 | −.222∗ | −.127 | −.863∗ | −2.157 | −264.374∗∗ |  |
|  | (.063) | (.118) | (.078) | (.462) | (2.034) | (115.620) |  |
| 9-10 | −.075 | −.024 | 0.122 | −2.270∗∗∗ | −5.774∗∗ | −289.988∗ |  |
|  | (.117) | (.136) | (.148) | (.700) | (2.424) | (174.301) |  |
| 11-12 | −.024 | −.028 | −.167 | −.888 | −3.275 | −8.651 |  |
|  | (.126) | (.127) | (.112) | (.713) | (2.467) | (163.025) |  |
| 13-14 | −.053 |  |  |  |  |  |  |
|  | (.142) |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Female sample |  |  |  |  |  | −102.210 |  |
| 0 | 0.068 |  |  | 0.541 | 0.219 |  |
|  | (.134) |  |  | (1.136) | (4.359) | (139.467) |  |
| 1-2 | −.114∗∗∗ |  |  | 0.130 | 0.472 | −28.298 |  |
|  | (.040) |  |  | (.359) | (.723) | (53.113) |  |
| 3-4 | −.208∗∗∗ |  |  | −.298 | 0.866 | −31.300 |  |
|  | (.064) |  |  | (.457) | (1.193) | (61.496) |  |
| 5-6 | −.097 |  |  | −.319 | 0.103 | −60.088 |  |
|  | (.063) |  |  | (.347) | (1.084) | (66.056) |  |
| 7-8 | −.184∗∗ |  |  | −.979∗∗ | −1.522 | −94.059 |  |
|  | (.089) |  |  | (.449) | (1.074) | (107.062) |  |
| 9-10 | −.344∗∗ |  |  | −.975 | 0.637 | 71.060 |  |
|  | (.168) |  |  | (.827) | (2.541) | (133.178) |  |
| 11-12 | −.119 |  |  | −.432 | −3.355 | −12.232 |  |
|  | (.113) |  |  | (1.070) | (2.603) | (141.560) |  |
| 13-14 | −.106 |  |  |  |  |  |  |
|  | (.152) |  |  |  |  |  |  |

*Notes* Standard errors in parentheses. Other control variables: age squared, region, urban, education,Han, marital status, urbanization index, time dummies, health insurance status, household expenditures. N=16047 (male sample), N=16658 (female sample).

**Results using non-imputed data**

Table 0.12: Analysis of the eﬀect of a diabetes diagnosis on employment status and be-havioural outcomes using MSM, FE and RE (no imputation)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Marginal structural model* | |  |  |  |
| Male sample |  | −.054 | −.118∗∗ | −.601∗∗∗ | −1.290 | −205.746∗ |  |
| Diabetes | 0.049 |  |
| Female sample | (.043) | (.040) | (.053) | (.229) | (.859) | (109.375) |  |
| −.087∗ | −.026∗ |  | −.637 | −1.043 | −45.166 |  |
| Diabetes | 0.000 |  |
|  | (.047) | (.016) | (.) | (.402) | (.865) | (56.543) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Fixed eﬀects* | |  |  |  |
| Male sample |  | −.004 | −.103∗∗∗ | −.844∗∗∗ | −2.463∗∗∗ | −152.316∗∗ |  |
| Diabetes | 0.024 |  |
| Female sample | (.030) | (.033) | (.036) | (.169) | (.508) | (67.898) |  |
| −.110∗∗∗ | −.024∗∗ | −.015 | −.634∗∗ | −1.105∗ | −81.340∗ |  |
| Diabetes |  |
|  | (.034) | (.012) | (.012) | (.288) | (.636) | (49.016) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Random eﬀects* | |  |  |  |
| Male sample | −.023 | −.045 | −.109∗∗∗ | −.569∗∗∗ | −1.163∗∗ | −143.470∗∗∗ |  |
| Diabetes |  |
| Female sample | (.027) | (.030) | (.029) | (.166) | (.482) | (51.625) |  |
| −.164∗∗∗ | −.020∗∗ | −.021∗∗∗ | −.309 |  | −59.269∗ |  |
| Diabetes | 0.494 |  |
|  | (.026) | (.009) | (.005) | (.269) | (.583) | (35.037) |  |

*Notes* Standard errors in parentheses. Other control variables: age (only MSM), age squared, region, ur-ban, education, han, marital status, urbanization index, time dummies, health insurance status, household expenditures. FE/RE: N=22135 (male sample), N=23143 (female sample), MSM: N=10006 (male sample), N=11471 (female sample).

Table 0.13: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using MSM, FE and RE (non-imputed)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Marginal structural model* | |  |  |  |
| Male sample |  | −.019 | −.036∗ | −.203∗∗ | −.550∗ | −85.203∗∗ |  |
| Time since diagnosis | 0.019 |  |
| Female sample | (.017) | (.015) | (.022) | (.081) | (.310) | (38.378) |  |
| −.028 | −.008 |  | −.338∗ | −.579∗ | −14.298 |  |
| Time since diagnosis | 0.000 |  |
|  | (.017) | (.006) | (.) | (.178) | (.333) | (21.193) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Fixed eﬀects* | |  |  |  |
| Male sample | −.001 |  | −.016∗∗ | −.158∗∗∗ | −.516∗∗∗ | −18.202 |  |
| Time since diagnosis | 0.003 |  |
| Female sample | (.007) | (.006) | (.007) | (.039) | (.118) | (12.059) |  |
| −.023∗∗∗ | −.002 | −.001 | −.103∗∗ | −.177 | −9.987 |  |
| Time since diagnosis |  |
|  | (.008) | (.002) | (.001) | (.045) | (.127) | (7.788) |  |
|  |  |  |  | |  |  |  |
|  |  |  | *Random eﬀects* | |  |  |  |
| Male sample | −.007 | −.003 | −.015∗∗∗ | −.120∗∗∗ | −.317∗∗∗ | −20.749∗∗ |  |
| Time since diagnosis |  |
| Female sample | (.006) | (.006) | (.006) | (.038) | (.101) | (9.382) |  |
| −.026∗∗∗ | −.002 | −.003∗∗∗ | −.065 |  | −7.041 |  |
| Time since diagnosis | 0.043 |  |
|  | (.006) | (.002) | (.001) | (.044) | (.124) | (6.479) |  |

*Notes* Standard errors in parentheses. Other control variables: age (only MSM) age squared, region, urban,education, han, marital status, urbanization index, time dummies, health insurance status, household expendi-tures. FE/RE: N=22117 (male sample), N=23130 (female sample), MSM: N=10028 (male sample), N=11465 (female sample).

Table 0.14: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using marginal structural models (duration groups) (non-imputed)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |
| Male sample |  |  |  | −.942 |  |  |  |
| 0 | 0.119∗ | 0.053 | 0.010 | 0.596 | 459.443 |  |
|  | (.070) | (.170) | (.156) | (.589) | (.934) | (474.665) |  |
| 1-2 | 0.026 | −.055 | −.137∗∗∗ | −.571∗∗ | −1.270 | −182.199 |  |
|  | (.044) | (.046) | (.043) | (.273) | (1.040) | (121.087) |  |
| 3-4 | 0.000 | −.043 | 0.131 | −1.013∗∗ | −3.347 | −782.090∗∗∗ |  |
|  | (.) | (.153) | (.156) | (.450) | (2.116) | (177.206) |  |
|  |  |  |  |  |  |  |  |
| Female sample |  |  |  | −.136 | −1.772 | −101.086 |  |
| 0 | 0.123 | 0.000 | 0.000 |  |
|  | (.188) | (.) | (.) | (1.488) | (5.608) | (203.293) |  |
| 1-2 | −.083 | −.018∗∗ | −.053∗ | −.613 | −.685 | −40.447 |  |
|  | (.067) | (.009) | (.028) | (.489) | (1.026) | (65.853) |  |
| 3-4 | 0.000 | 0.000 | 0.000 | −5.530∗ | −8.510∗∗∗ | 0.676 |  |
|  | (.) | (.) | (.) | (3.260) | (1.787) | (257.875) |  |

*Notes* Due to Standard errors in parentheses. Other control variables: Age, age squared, region, urban,education, han, marital status, urbanization index, time dummies, health insurance status, household ex-penditures. N=10028 (male sample), N=11465 (female sample).

Table 0.15: Analysis of the eﬀect of time since diabetes diagnosis on employment status and behavioural outcomes using fixed eﬀects (duration groups) (non-imputed)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |  |
|  | Employment | Smoking | Any alcohol | BMI | Waist (cm) | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |
| Male sample |  | −.013 |  | −.013 |  | −268.541 |  |
| 0 | 0.126∗ | 0.081 | 1.444 |  |
|  | (.073) | (.084) | (.156) | (.704) | (1.883) | (213.448) |  |
| 1-2 | 0.046 | −.019 | −.135∗∗∗ | −.817∗∗∗ | −2.298∗∗∗ | −225.905∗∗ |  |
|  | (.039) | (.039) | (.042) | (.199) | (.637) | (90.437) |  |
| 3-4 | 0.013 | 0.035 | −.052 | −.786∗∗ | −3.016∗∗∗ | −107.317 |  |
|  | (.046) | (.054) | (.055) | (.325) | (.819) | (98.624) |  |
| 5-6 | −.134∗ | 0.028 | −.134∗∗ | −1.159∗∗∗ | −1.715 | 34.167 |  |
|  | (.079) | (.077) | (.065) | (.343) | (1.178) | (117.774) |  |
| 7-8 | 0.162∗∗ | −.138 | −.270∗∗ | −.692 | −2.555 | −305.553∗∗ |  |
|  | (.078) | (.117) | (.117) | (.429) | (1.726) | (133.202) |  |
| 9-10 | −.018 | 0.044 | 0.082 | −1.938∗∗∗ | −8.278∗∗∗ | −196.802 |  |
|  | (.136) | (.123) | (.131) | (.667) | (2.262) | (201.492) |  |
| 11-12 | 0.063 | 0.089 | −.177∗∗ | −1.743∗∗ | −5.843∗∗ | −22.708 |  |
|  | (.178) | (.134) | (.082) | (.736) | (2.828) | (140.771) |  |
| 13-14 | 0.060 | 0.222∗∗ | −.164 | −1.508 | −4.207∗∗∗ | −119.852 |  |
|  | (.194) | (.113) | (.111) | (1.202) | (1.063) | (178.187) |  |
|  |  |  |  |  |  |  |  |
| Female sample |  | −.014∗∗ | −.046 | −.778 | −3.920 | −358.037∗∗ |  |
| 0 | 0.101 |  |
|  | (.154) | (.007) | (.040) | (.909) | (3.420) | (173.529) |  |
| 1-2 | −.100∗∗∗ | −.029∗∗ | −.023∗ | −.329 | −.558 | −118.162∗∗ |  |
|  | (.033) | (.012) | (.012) | (.363) | (.671) | (56.839) |  |
| 3-4 | −.148∗∗ | −.017 | −.025∗ | −.822∗ | −.824 | 49.550 |  |
|  | (.059) | (.013) | (.014) | (.442) | (1.148) | (82.984) |  |
| 5-6 | −.122∗ | −.043 | 0.002 | −1.028∗∗∗ | −1.616 | −69.012 |  |
|  | (.073) | (.041) | (.020) | (.325) | (1.016) | (96.779) |  |
| 7-8 | −.235∗∗∗ | 0.023 | −.004 | −1.327∗∗∗ | −3.174∗∗∗ | −90.185 |  |
|  | (.090) | (.027) | (.008) | (.390) | (.978) | (111.004) |  |
| 9-10 | −.247∗∗ | 0.031 | −.010 | −.981 | −.260 | −64.808 |  |
|  | (.118) | (.039) | (.009) | (.621) | (2.131) | (134.146) |  |
| 11-12 | −.239∗∗ | −.070 | −.005 | −.715 | −3.440 | −25.527 |  |
|  | (.103) | (.056) | (.009) | (1.021) | (2.512) | (173.367) |  |
| 13-14 | −.199 | −.023 | −.008 | −.111 | 0.693 | −366.259∗∗∗ |  |
|  | (.166) | (.018) | (.009) | (.665) | (2.153) | (87.213) |  |

*Notes* Standard errors in parentheses. Other control variables: age squared, region, urban, education, han,marital status, urbanization index, time dummies, health insurance status, household expenditures. N=22117 (male sample), N=23130 (female sample).

Table 0.16: Analysis of the eﬀect of time since diabetes diagnosis on employment sta-tus and behavioural outcomes using random eﬀects (duration groups) (non-imputed)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | | (2) | | (3) | |  | (4) | |  | (5) | (6) |  |
|  | Employment | | Smoking | | Any alcohol | | BMI | | | Waist (cm) | | Calories (kcal) |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Male sample |  |  | −.043 | |  |  |  |  |  |  |  | −28.615 |  |
| 0 | 0.094 | | 0.065 | |  | 0.148 | |  | 2.276 |  |
| 1-2 | (.069) | | (.087) | | (.144) | |  | (.610) | | (1.683) | | (188.201) |  |
| − | .008 | − | .053 | − | .144∗∗∗ | − | | .533∗∗∗ | − | 1.045 | 203.986∗∗ |  |
|  |  |  |  |  |  | − |  |
| 3-4 | (.038) | | (.038) | | (.036) | |  | (.195) | |  | (.658) | (80.054) |  |
| − | .041 | − | .007 | − | .070 | − | | .493 | − | 1.730∗∗ | 140.623 |  |
|  |  |  |  |  |  | − |  |
|  | (.045) | | (.051) | | (.051) | |  | (.336) | |  | (.809) | (87.834) |  |
| 5-6 | −.159∗∗ | | −.012 | | −.120∗∗ | | −.866∗∗∗ | | | −.330 | | −69.752 |  |
| 7-8 | (.077) | | (.073) | | (.060) | |  | (.333) | | (1.054) | | (115.094) |  |
| 0.114 | | − | .213∗∗ | − | .215∗∗ | − | | .473 | − | 1.072 | 243.936∗∗ |  |
|  |  |  |  |  |  |  | − |  |
| 9-10 | (.074) | | (.108) | | (.097) | |  | (.431) | | (1.538) | | (105.320) |  |
| − | .070 | 0.001 | | 0.127 | | − | 1.803∗∗∗ | | − | 7.021∗∗∗ | 173.366 |  |
|  |  |  |  |  |  |  |  |  | − |  |
|  | (.134) | | (.118) | | (.132) | |  | (.620) | | (2.127) | | (167.349) |  |
| 11-12 | 0.005 | | 0.060 | | −.160 | | −1.446∗ | | | −4.339 | | 92.244 |  |
|  | (.159) | | (.144) | | (.100) | |  | (.767) | | (2.681) | | (148.282) |  |
| 13-14 | 0.029 | | 0.234∗∗∗ | | −.118 | | −1.101 | | | −2.531∗∗∗ | | 38.227 |  |
|  | (.161) | | (.083) | | (.128) | | (1.263) | | |  | (.931) | (100.439) |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Female sample | 0.025 | | 0.003 | |  | .039∗∗ |  |  | .238 |  | 1.178 | 123.300 |  |
| 0 | − | − | | − |  |
|  |  |  |  |  |  |  |  | − |  |
|  | (.145) | | (.025) | | (.016) | |  | (.874) | | (3.554) | | (139.671) |  |
| 1-2 | −.142∗∗∗ | | −.028∗∗∗ | | −.028∗∗∗ | |  | 0.001 | |  | 0.848 | −66.418 |  |
|  | (.031) | | (.010) | | (.004) | |  | (.349) | |  | (.622) | (49.483) |  |
| 3-4 | −.195∗∗∗ | | −.020∗ | | −.028∗∗∗ | | −.481 | | |  | 1.064 | 43.196 |  |
|  | (.052) | | (.012) | | (.005) | |  | (.433) | | (1.090) | | (68.580) |  |
| 5-6 | −.159∗∗ | | −.034 | | −.007 | | −.647∗∗ | | |  | 0.445 | −52.781 |  |
|  | (.063) | | (.035) | | (.021) | |  | (.315) | |  | (.981) | (77.715) |  |
| 7-8 | −.247∗∗∗ | | 0.029 | | −.022∗∗∗ | | −1.073∗∗∗ | | | −1.501∗ | | −90.408 |  |
|  | (.070) | | (.031) | | (.003) | |  | (.368) | |  | (.886) | (116.975) |  |
| 9-10 | −.286∗∗∗ | | 0.029 | | −.024∗∗∗ | | −.748 | | |  | 1.422 | 124.263 |  |
|  | (.099) | | (.046) | | (.003) | |  | (.605) | | (1.900) | | (156.687) |  |
| 11-12 | −.214∗ | | −.062 | | −.022∗∗∗ | | −.335 | | | −1.482 | | 49.789 |  |
| 13-14 | (.114) | | (.046) | | (.005) | | (1.000) | | | (2.752) | | (155.171) |  |
| − | .176 | − | .022∗ | − | .024∗∗∗ |  | 0.298 | |  | 2.665 | 332.344∗∗∗ |  |
|  |  |  |  |  |  |  |  |  | − |  |
|  | (.153) | | (.012) | | (.006) | |  | (.755) | | (2.407) | | (99.899) |  |

*Notes* Standard errors in parentheses. Other control variables: age squared, region, urban, education, han,marital status, urbanization index, time dummies, health insurance status, household expenditures. N=22117 (male sample), N=23130 (female sample).

**Overweight and obesity results**

Table 0.17: Analysis of the eﬀect of a diabetes diagnosis on overweight and obesity

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Males |  |  | Females |  |  |
|  |  |  |  |  |  |  |
|  | (1) | (2) |  | (3) | (4) |  |
|  | Overweight | Obese |  | Overweight | Obese |  |
|  |  | | | |  |  |
|  | *Marginal structural model* | | | |  |  |
| Diabetes | −.000 | −.024 |  | −.031 | −.009 |  |
|  | (.031) | (.015) | (.034) | | (.014) |  |
|  |  |  | | |  |  |
|  |  | *Fixed Eﬀects* | | |  |  |
| Diabetes | −.041 | −.035 |  | −.095∗∗∗ | −.034 |  |
|  | (.035) | (.025) | (.036) | | (.027) |  |
|  |  |  | | |  |  |
|  |  | *Random Eﬀects* | | |  |  |
| Diabetes | 0.014 | −.006 |  | −.070∗∗ | 0.028 |  |
|  | (.030) | (.023) | (.030) | | (.024) |  |

*Notes* Standard errors in parentheses. Other control variables:Age squared, region, urban, education, han, marital status, ur-banization index, time dummies, health insurance status, house-hold expenditures. FE/RE: N=23443 (male sample), N=23702 (female sample). MSM: N=16047 (male sample), N=16658 (fe-male sample).

Table 0.18: Analysis of the eﬀect of time since diagnosis on overweight and obesity

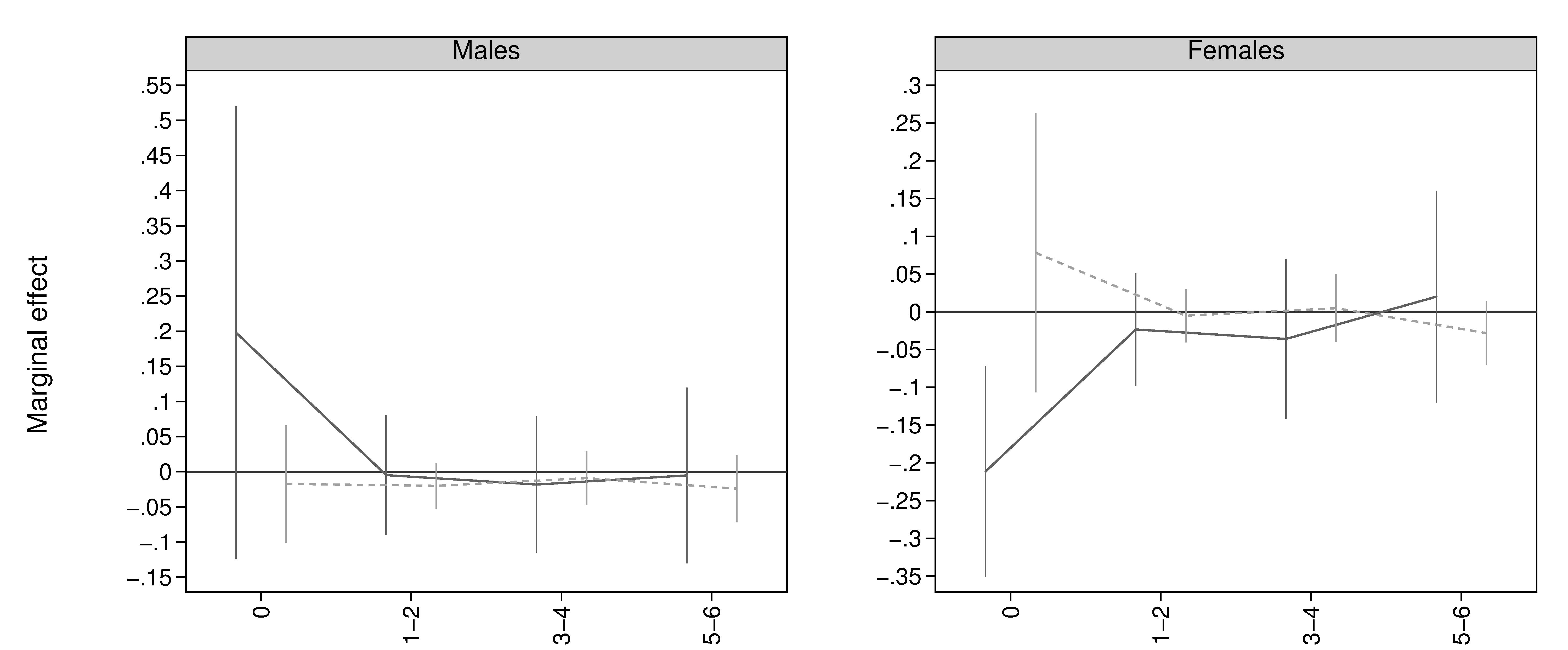
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Males |  |  | Females | |  |
|  |  |  |  |  |  |  |
|  | (1) | (2) |  | (3) | (4) |  |
|  | Overweight | Obese |  | Overweight | Obese |  |
|  | *Marginal structural model* | | | |  |  |
| Time since diagnosis | −.001 | −.005∗ |  | −.003 | −.003 |  |
|  | (.005) | (.003) | (.005) | | (.002) |  |
|  |  |  | | |  |  |
|  |  | *Fixed Eﬀects* | | |  |  |
| Time since diagnosis | −.006 | −.007∗ |  | −.006 | −.009∗ |  |
|  | (.007) | (.004) | (.006) | | (.005) |  |
|  |  |  | | |  |  |
|  |  | *Random Eﬀects* | | |  |  |
| Time since diagnosis | 0.002 | −.003 |  | −.006 | −.001 |  |
|  | (.006) | (.003) | (.005) | | (.004) |  |

*Notes* Standard errors in parentheses. Other control variables: Agesquared, region, urban, education, han, marital status, urbanization index, time dummies, health insurance status, household expenditures. FE/RE: N=23443 (male sample), N=23702 (female sample). MSM: N=16047 (male sample), N=16658 (female sample).

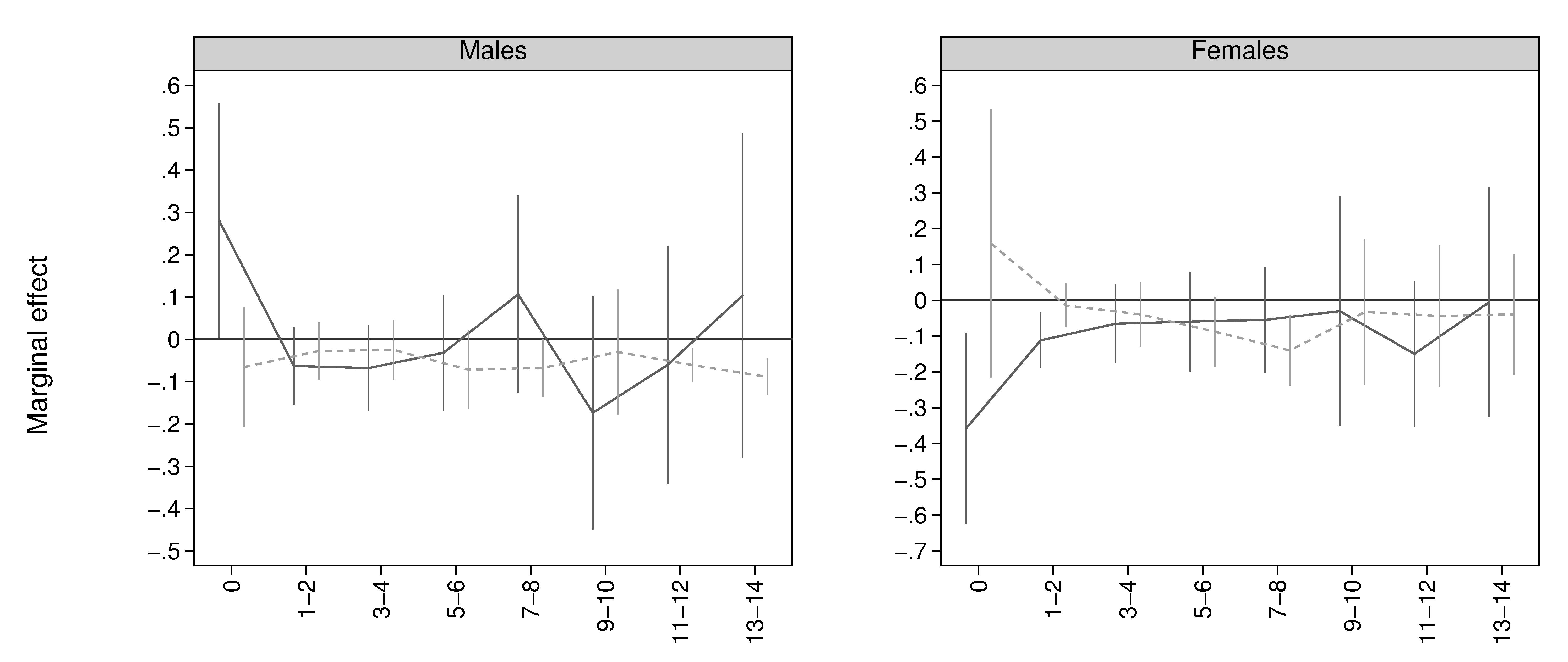
Figure 0.6: Analysis of the eﬀect of time since diabetes diagnosis on overweight and obesity

(duration groups)

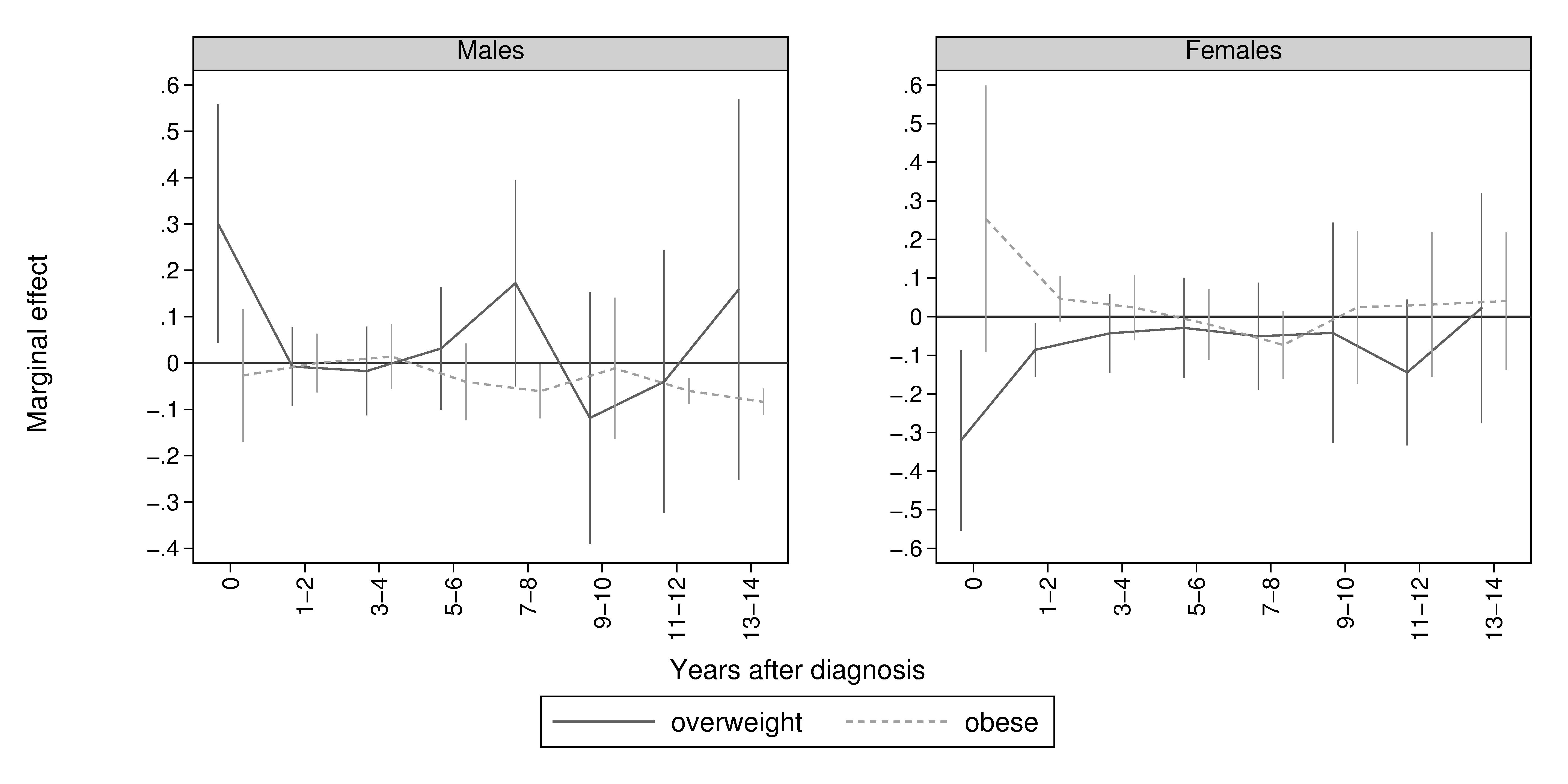
Marginal structural models



Fixed eﬀects



Random eﬀects



Notes: For MSM, eﬀects after 6 years could 42not be estimated due to too few observations.

**Bibliography**

Angrist, J. and Pischke, J. (2008). *Mostly Harmless Econometrics: An Empiricist’s Com-panion*. Princeton University Press.

Attard, S. M., Herring, a. H., Mayer-Davis, E. J., Popkin, B. M., Meigs, J. B., and Gordon-Larsen, P. (2012). “Multilevel examination of diabetes in modernising China: what ele-ments of urbanisation are most associated with diabetes?” *Diabetologia* 55 (12), 3182– 92.

Bertram, M. Y. and Vos, T. (2010). “Quantifying the duration of pre-diabetes.” *Australian* *and New Zealand journal of public health* 34 (3), 311–314.

Brown, H. S., Pagán, J. A., and Bastida, E. (2005). “The Impact of Diabetes on Employ-ment: Genetic IVs in a Bivariate Probit.” *Health Economics* 14 (5), 537–544.

Charles, K. K. and Decicca, P. (2008). “Local labor market fluctuations and health: Is there a connection and for whom?” *Journal of Health Economics* 27, 1532–1550.

China Obesity Task Force (2004). “Body mass index reference norm for screening over-weight and obesity in Chinese children and adolescents.” *Chinese Journal of Epidemi-ology* 25 (2), 97–102.

Cole, S. R. and Hernan, M. A. (2008). “Constructing Inverse Probability Weights for Marginal Structural Models.” *American Journal of Epidemiology* 168 (6), 656–664.

Colman, G. and Dave, D. (2014). “Unemployment and Health Behaviors Over the Business Cycle: a Longitudinal View.”

De Fine Olivarius, N., Siersma, V. D., Køster-Rasmussen, R., Heitmann, B. L., and Wal-dorﬀ, F. B. (2015). “Weight changes following the diagnosis of type 2 diabetes: The impact of recent and past weight history before diagnosis. Results from the Danish Diabetes Care in General Practice (DCGP) Study.” *PLoS ONE* 10 (4), 1–14.

Harris, A. (2009). “Diabetes, Cardiovascular Disease and Labour Force Participation in Australia: An Endogenous Multivariate Probit Analysis of Clinical Prevalence Data.” *Economic Record* 85 (271), 472–484.

He, W., Li, Q., Yang, M., Jiao, J., Ma, X., Zhou, Y., Song, A., Heymsfield, S. B., Zhang, S., and Zhu, S. (2015). “Lower BMI cutoﬀs to define overweight and obesity in China.” *Obesity* 23 (3), 684–691.

43

Hu, F. B. (2011). “Globalization of diabetes: the role of diet, lifestyle, and genes.” *Diabetes* *care* 34 (6), 1249–57.

Klein, D. (2014). *MIMRGNS: Stata module to run margins after mi estimate*. Statistical Software Components, Boston College Department of Economics.

Klein, S., Allison, D. B., Heymsfield, S. B., Kelley, D. E., Leibel, R. L., Nonas, C., and Kahn, R. (2007). “Waist circumference and cardiometabolic risk: A consensus state-ment from Shaping America’s Health: Association for Weight Management and Obesity Prevention; NAASO, the Obesity Society; the American Society for Nutrition; and the American Diabetes Associat.” *Diabetes Care* 30 (6), 1647–1652.

Latif, E. (2009). “The impact of diabetes on employment in Canada.” *Health Economics* 18 (5), 577–589.

Liu, Z., Fu, C., Wang, W., and Xu, B. (2010). “Prevalence of chronic complications of type 2 diabetes mellitus in outpatients - a cross-sectional hospital based survey in urban China.” *Health and quality of life outcomes* 8, 62.

Long, G. H., Cooper, A. J., Wareham, N. J., Griﬃn, S. J., and Simmons, R. K. (2014). “Healthy Behavior Change and Cardiovascular Outcomes in Newly Diagnosed Type 2 Diabetic Patients: A Cohort Analysis of the ADDITION-Cambridge Study.” *Diabetes* *Care* 37 (6), 1712–1720.

Luo, X., Liu, T., Yuan, X., Ge, S., Yang, J., Li, C., and Sun, W. (2015). “Factors Influencing Self-Management in Chinese Adults with Type 2 Diabetes: A Systematic Review and Meta-Analysis.” *International Journal of Environmental Research and Public Health*

12 (9), 11304–11327.

Ma, R. C. W., Lin, X., and Jia, W. (2014). “Causes of type 2 diabetes in China.” *The* *Lancet Diabetes & Endocrinology* 2 (12), 980–991.

Minor, T. (2011). “The eﬀect of diabetes on female labor force decisions: new evidence from the National Health Interview Survey.” *Health Economics*. 15th Annual International Meeting of the International Society for Pharmacoeconomics and Outcomes Research, ISPOR 2010 Atlanta, GA United States. Conference Start: 20100515 Conference End: 20100519 20 (12), 1468–1486.

NCD Risk Factor Collaboration (2016). “Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants.” *The Lancet* 387, 1513–1530.

Peters, S. A. E., Huxley, R. R., Sattar, N., and Woodward, M. (2015). “Sex Diﬀerences in the Excess Risk of Cardiovascular Diseases Associated with Type 2 Diabetes: Potential Explanations and Clinical Implications.” *Current Cardiovascular Risk Reports* 9 (7), 1– 7.

44

Peters, S. A. E., Huxley, R. R., and Woodward, M. (2014a). “Diabetes as a risk factor for stroke in women compared with men: A systematic review and meta-analysis of 64 cohorts, including 775 385 individuals and 12 539 strokes.” *The Lancet* 383 (9933), 1973–1980.

– (2014b). “Diabetes as risk factor for incident coronary heart disease in women com-pared with men: A systematic review and meta-analysis of 64 cohorts including 858,507 individuals and 28,203 coronary events.” *Diabetologia* 57 (8), 1542–1551.

Popkin, B. M., Du, S., Zhai, F., and Zhang, B. (2010). “Cohort profile: The China Health and Nutrition Survey-monitoring and understanding socio-economic and health change in China, 1989-2011.” *International Journal of Epidemiology* 39 (6), 1435–1440.

Robins, J. M., Hernan, M. Á., and Brumback, B. (2000). “Marginal Structural Models and Causal Inference in Epidemiology.” *Epidemiology* 11, 550–560.

Royston, P. and White, I. (2009). “Multiple imputation by chained equations (MICE): Implementation in Stata.” *Journal of Statistical Software* 45 (4), pages. arXiv:  [arXiv](http://arxiv.org/abs/arXiv:1501.0228):

[1501.022](http://arxiv.org/abs/arXiv:1501.0228)8.

Seuring, T., Archangelidi, O., and Suhrcke, M. (2015). “The Economic Costs of Type 2 Diabetes: A Global Systematic Review.” *PharmacoEconomics* 33 (8), 811–831.

Seuring, T., Goryakin, Y., and Suhrcke, M. (2015). “The impact of diabetes on employment in Mexico.” *Economics & Human Biology* 18, 85–100.

Seuring, T., Serneels, P., and Suhrcke, M. (2016). “The Impact of Diabetes on Labor Market Outcomes in Mexico: A Panel Data and Biomarker Analysis.”

Silink, M., Tuomilehto, J., Mbanya, J. C., Narayan, K. M. V., Fradkin, J., and Roglic, G. (2010). “Research priorities: Prevention and Control of Diabetes with A Focus on Low and Middle Income Countries.” *WHO Meetings on Development of A Prioritized*

*Research Agenda for Development of Prevention and Control of Noncommunicable Dis-ease* 6.

Slade, A. N. (2012). “Health Investment Decisions in Response to Diabetes Information in Older Americans.” *Journal of Health Economics* 31 (3), 502–520.

WHO (2004). “Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies.” *The Lancet* 363, 157–163.

Wooldridge, J. (2012). *Introductory Econometrics. A Modern Approach*. 5th ed. Cengage Learning.

Yang, W. and Weng, J. (2014). “Early therapy for type 2 diabetes in China.” *The Lancet* *Diabetes & Endocrinology* 2 (12), 992–1002.

45

Yuan, X., Liu, T., Wu, L., Zou, Z.-Y., and Li, C. (2015). “Validity of self-reported dia-betes among middle-aged and older Chinese adults: the China Health and Retirement Longitudinal Study.” *British Medical Journal Open* 5 (4), e006633–e006633.

Zeng, Q., He, Y., Dong, S., Zhao, X., Chen, Z., Song, Z., Chang, G., Yang, F., and Wang, Y. (2014). “Optimal cut-oﬀ values of BMI, waist circumference and waist:height ratio for defining obesity in Chinese adults.” *The British journal of nutrition* 112 (10), 1735– 44.

Zhang, B., Zhai, F. Y., Du, S. F., and Popkin, B. M. (2014). “The China Health and Nutrition Survey, 1989-2011.” *Obesity Reviews* 15 (S1), 2–7. arXiv:  [NIHMS15000](http://arxiv.org/abs/NIHMS150003)3.

Zhao, M., Konishi, Y., and Glewwe, P. (2013). “Does information on health status lead to a healthier lifestyle? Evidence from China on the eﬀect of hypertension diagnosis on food consumption.” *Journal of Health Economics* 32 (2), 367–385.

Zhou, X., Ji, L., Ran, X., Su, B., Ji, Q., Pan, C., Weng, J., Ma, C., Hao, C., Zhang, D., and Hu, D. (2016). “Prevalence of Obesity and Its Influence on Achievement of Cardiometabolic Therapeutic Goals in Chinese Type 2 Diabetes Patients: An Analysis of the Nationwide, Cross-Sectional 3B Study.” *PLOS ONE* 11 (1). Ed. by J. Devaney, e0144179.

46