University of Pretoria
Department of Economics Working Paper Series

The Impact of Diabetes on Labour Market Outcomes

Steven F. Koch
University of Pretoria
Evelyn Tshela
University of Pretoria and South Africa and Council for Medical Schemes

Working Paper: 2020-109
December 2020

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

The Impact of Diabetes on Labour Market Outcomes*

Steven F. Koch† Evelyn Tshela‡

18 December, 2020

Abstract

This study estimates the effect of diabetes on labour market outcomes (employment, unemployment and labour force paricipation) in South Africa using data from the South African General Household Survey (2018). We first examine the possibility that diabetes status is endogenous, through the application of heteroscedasticity-based instruments. Internal instruments meet the underlying diagnostic expectations, but do not consistently accept the endogeneity hypothesis. Thus, we turn to multinomial logit models, ignoring endogeneity, to estimate the effect of diabetes. Our findings indicate that diabetes has differential effects for men and women, where the magnitude of the effect tends to be larger (in absolute value) for women.

1 Introduction

South Africa is on the verge of a Type II Diabetes mellitus (diabetes) crisis. Diabetes is one of the leading causes of mortality in the country, accounting for 5.4 percent of deaths in 2015 (Statistics South Africa 2014). According to the World Health Organization, approximately 9.8 percent of the population suffers from Diabetes, and the prevalence rate is expected to rise to 30 percent by the year 2030. In recognition of the substantial individual and government costs of the disease, diabetes is a 'national health priority area' in South Africa (National Planning Commssion 2012). To curb the prevalence of diabetes, the South African government adopted a policy in April 2018 to tax sugar-sweetened beverages (SSBs), on the assumption that increased SSB prices will deter consumers from purchasing these beverages, thereby reducing sugar consumption. Since the consumption of SSBs has been linked to non-communicable diseases, such as Type II Diabetes (Basu et al. 2014; Malik et al. 2010), it is further assumed that the reduction in sugar consumption

*The authors would like to thank John Ataguba, Heinrich Bohlmann, Frikkie Booysen and Zoë McLaren for their comments on earlier drafts of this research. However, we remain responsible for any and all remaining errors.

†Corresponding author: Department of Economics, University of Pretoria, Private Bag X20, Hatfield, South Africa; +27-12-420-5285; steve.koch@up.ac.za

‡Department of Economics, University of Pretoria, Private Bag X20, Hatfield, South Africa and Council for Medical Schemes, Block A, Eco Glades 2 Office Park, 420 Witch-Hazel Avenune, Eco Park, Centurion, South Africa, e.thsehla@medicalschemes.com

will yield concomitant reductions in diabetes, or at least slow its expected rise and help reduce individual and government costs associated with the disease. Furthermore, as we describe more fully below, reductions in diabetes could also yield improvements in the labour market, which would yield additional improvements for government (increased tax revenues) and the economy.

Globally, the number of people living with diabetes and the number of deaths due to diabetes has also increased over the years. An estimated 422 million adults worldwide were living with diabetes in 2014 as compared to 108 million in 1980. In 2012, diabetes was responsible for 1.5 million deaths in the world (World Health Organization 2016). Most of the deaths attributable to high blood glucose or diabetes occur prior to age 70 and the mortality rate is higher in low- and middle-income countries than in high-income countries (World Health Organization 2016).

In addition to imposing large morbidity and mortality costs, diabetes and its complications impose substantial economic costs on people with diabetes and their families, as well as to health systems and national economies (Kirigia et al. 2009). Diabetes exerts three broad categories of economic cost; direct healthcare costs, indirect healthcare costs and intangible costs (Jonsson 1998). The direct healthcare costs include medications and consultations and hospitalisation for the condition, itself, and for its complications. The indirect healthcare costs include other expenses, such as travelling to health care facilities, productivity loss associated with morbidity, as well as the loss of income due to mortality, morbidity and disability associated with diabetes and its complications. The intangible costs include losses due to physical and psychological pain (Kirigia et al. 2009), which is difficult to quantify. According to the International Diabetes Federation (IDF), estimated global diabetes health care costs were at least 376 billion USD in 2010. In South Africa, the national economic burden was forecast to be 865 thousand USD in 2011. The IDF estimates that approximately seven percent of total health expenditure went to diabetes care amongst adults aged 20 – 79 in 2010. It is further estimated that expenditure will rise by 30 – 34 percent by 2030 in developing countries. In South Africa, these costs will rise to between 1.1 – 2 billion USD by 2030 (P. Zhang et al. 2010).

Over the years, numerous studies have investigated the direct and indirect cost of diabetes. Most of the studies use the human capital approach to determine direct costs of illness, lost income and lost hours, due to morbidity and mortality. A review of the global evidence argues that the studies in low and middle-income countries mostly focus on the direct costs of the disease, whereas evidence on labour market effects remains scarce (Seuring, Goryakin, and Suhrcke 2015). Mutyambizi et al. (2018) agree; their review of diabetes costs in Africa suggests a similar focus on the direct cost of diabetes. Globally, the impact of diabetes on labour market outcomes is primarily available for upper-income countries (Brown III, Pagan, and Bastida (2005); Latif (2009); Seuring, Archangelidi, and Suhrcke (2015)). Thus, there is a need to learn more about labour market effects in less developed countries that, as in the case of South Africa, have relatively high levels of diabetes.

A number of studies have investigated the relationship between poor health outcomes and labour force outcomes in several countries (X. Zhang, Zhao, and Harris 2009; J. Brown et al. 2012; Pharr, Moonie, and Bungum 2012; Norström et al. 2019). Current evidence suggests that poor health is associated with unemployment whilst healthier people are more likely to gain and retain employment. Whilst the correlation between health and employment has been well established, the causal relationship seems to work in both directions. Evidence from research shows that being employed can lead to better health in two ways. Firstly, individuals who are employed have income which provides financial means to access better healthcare, nutritious food and better sanitation which are all determinants of health (Schaller and HuffStevens 2019; Zavras et al. 2019). On the other side individuals who are unemployed have increased risk for mortality, physical illness, psychological stress and family breakdown (Wilson and Walker 1993). Secondly being employed has been associated with psychological benefits such as better self-esteem and a sense of purpose and identity (Noordt et al. 2014; Cottini and Ghinetti 2018). Furthermore, having work has been shown to facilitate recovery from illness and enhance mental health (Modini et al. 2016).

Focusing more closely on diabetes, most of the studies show that diabetes has a negative impact on labour outcomes, and some evidence, but not consistent evidence of endogeneity. Pedron et al. (2019) argues that few studies include co-morbidities, control for endogeneity or differentiate between type I and type II diabetes, which may lead to over-estimated impacts. For those that do consider endogeneity, the methods rely on genetic instrumentation; however, the results from that literature is also not consistent. Brown III, Pagan, and Bastida (2005) use family diabetes history as an instrumental variable, finding a substantial negative effect on employment for men but not for women. Furthermore, diabetes is endogenous for women, but not men. Latif (2009) follows a similar approach, but is able to more finely distinguish between the possible genetic pathways.[1] Those results point to a significant negative impact on female employment, but no significant impact on the employment of non-white Canadians. Furthermore, there are endogeneity differences by gender, such that assuming exogeneity yields over-estimates of the diabetes impact on male employment. In Taiwan, Lin (2011) finds that diabetes has a negative and significant effect on employment, the magnitude is larger for men than for women and the null hypothesis of exogeneity is rejected.

The preceding research has focused exclusively on developed economies. We were only able to find studies for one other developing country. Seuring, Goryakin, and Suhrcke (2015) analyses data from the Mexican Family

[1] The five instrumental variables incorporate paternal Diabetes (as well as having died from the disease), maternal diabetes (also, having died from it) and sibling diabetes.

Life Survey with mother and father diabetes status as instruments. Their analysis finds that Diabetes reduces employment for men (about 10 percentage points) and women (about 4.5 percentage points), although they find no indication that diabetes is endogenous. Seuring, Serneels, and Suhrcke (2019) apply fixed effects methods, finding employment reductions around 5.5 percentage points for both men and women.

Although there are a number of studies in the literature, they rarely consider developing countries. As noted previously, diabetes affects about 10% of the South African population, currently, and is expected to affect close to 30% of the population within the next 15 years. Moreover, the majority of individuals with diabetes are less than 64 years of age, with the highest incidences (54%) reported between the ages of 35 and 54 years (Bertram et al. 2013), the "most economically active" population in most countries, including South Africa. Furthermore, in Africa, approximately 76% of diabetes deaths occur in people younger than 60 years of age, compared to the global proportion of 49% (International Diabetes Federation, 2013). The situation is further exacerbated by the prevalence of undiagnosed diabetes, which could be as high as 46% in upper-middle income countries (Peer et al. 2014).

The high prevalence of diabetes is concerning, given that healthy workers are an important economic asset for a nation. When people cannot work, due to serious health problems, they cannot fully support their families, let alone a nation, because they do not generate economic output or pay taxes on earnings (Davis et al. 2005). Not only that, when people are healthy, the government incurs smaller health-related expenditure, and, therefore it is able to invest in education or to upgrade infrastructure, which will enhance overall economic productivity.

Given the prevalence of diabetes in the working age group (Bertram et al. 2013), and its expected future prevalence, as well as other features of the South African economy, it is possible that the effects of diabetes are more nefarious in South Africa than in other countries. It is not known, however, whether the effects of diabetes are worse in South Africa than in other countries previously investigated, or even whether diabetes can be tied causally to labour market outcomes in South Africa. Therefore, we add to the literature by investigating the effect of diabetes on the labour market in a developing country that has both high unemployment and high diabetes prevalence rates. We do so using very recent data, which comes from the 2018 General Household Survey (@ Statistics South Africa 2019a). As with previous literature, we consider the possibility that diabetes and labour market outcomes are endogenously determined; however, because we do not have typical instruments, we, instead, consider identification through heteroscedasticity (Lewbel 2012, 2018).

We are not the first researchers to consider health and labour market issues in South Africa. Given the rate of

HIV/AIDS in the country, understandably, South African research has often focused on it (Arndt and Lewis 2000; Levinsohn et al. 2013; Young 2005), although more recent research considers obesity (Some, Rashied, and Ohonba 2016) and self-assessed poor health (Nwosu and Woolard 2017). Other research considering health and the labour market includes Lawana et al. (2020), who consider a number of non-communicable diseases, including diabetes, within a recursive system that assumes normality. Lawana et al. (2020) also focus exclusively on participation. Thus, we complement this local literature by offering a more nuanced examination of the labour market and its endogeneity.

2 Methodology

Our interest is in the effect of diabetes on the labour market, specifically whether diabetes increases or decreases the probability of labour force participation and/or employment (amongst participants). Thus, we estimate models, such as the following:

$$\text{prob}(\ell = j) = f(d\tau_j + x\beta_j + u_j), \tag{1}$$

where d is a dummy variable denoting diabetes status, x represents a series of control variables, which we discuss below, and ℓ represents the labor market outcome of interest. Labour market outcomes are multinomial: employed, unemployed or non-labour force participant, although it is possible to look at subsets of these outcomes.

$$j = \begin{cases} \text{n} & \text{if not a labour force participant} \\ \text{u} & \text{if unemployed} \\ \text{e} & \text{if employed} \end{cases} \tag{2}$$

Given the multinomial structure, we would prefer to specify the labour outcomes following a multinomial logit model, such that

$$\text{prob}(\ell = j) = \frac{d\tau_j + x\beta_j}{\sum_k \exp\left(d\tau_k + x\beta_k\right)}. \tag{3}$$

However, we since there might be endogeneity in diabetes, we consider other specifications.

2.1 Potential endogeneity

A major concern in the analysis is that diabetes status might be endogenous. It could be measured with error, as individuals may not be aware of their status. It is also possible that early life health issues, such as those related to diabetes, impacted education success, and, therefore, labour market prospects. Also, as noted by Minor (2011), diabetes is likely to affect work decisions, while work decisions are likely to influence both diet and exercise, and, therefore, the ability to either contract or contain diabetes.

Solving the endogeneity issue, whether estimation follows a multinomial structure or not, requires instrumentation. In the literature, it is common to instrument with genetic indicators of diabetes (Brown III, Pagan, and Bastida 2005; Latif 2009; Lin 2011). The South African General Household Survey, which underpins this analysis, is not designed to capture parental information, and, therefore, we are not able to use genetic instruments. Specifically, the survey only captures data as parental data for parents living with their children. Relatively few parents are living with their working age children, and, in fact, elderly adults in need of care may be more likely to be living with their children, if their children are in a position to care for them. In other words, the survey is more likely to capture children (of rather elderly adults) who are more likely to be working, and, therefore, the genetic instrument would not be randomly distributed.

Since the outcomes are multinomial, and the multinomial model is nonlinear, direct application of instrumental variable methods through, for example, two-stage least squares is not appropriate. One solution in such a setting is to follow Imbens and Newey (2009), which is a control function approach requiring the estimation of a conditional cumulative distribution function; it also requires an instrument. Koch and Tshiswaka-Kashalala (2018) apply this in estimating the demand for contraceptive efficacy. Unfortunately, as previously implied, an instrument is not readily available in our data. Instead, one could consider a method proposed by Dong (2010), which is similar in spirit to Imbens and Newey (2009), but does not require an instrument. However, Dong (2010) does not explicitly allow for binary endogenous variables; rather, it requires the underlying support of the estimated error associated with the endogenous variable to be large.[2]

Given the aforementioned issues with the availability of a suitable instrument, as well as the inappropriateness of Dong's (2010) simple estimator, we were left with the option of applying Lewbel (2012) and Lewbel (2018), identifying the endogenous effect through heteroscedasicity within a linear setting. Lewbel (2018) proves that the method can be used when the endogenous regressor is binary, although the proof does not guarantee its validity. Following this approach, we have a series of triangular linear models like the following, where k represents any binary outcome pair that can be derived from the set: employed, unemployed, non-participant.

$$\ell_k = x\beta_k + D\tau_k + v_k \tag{4}$$

$$D = x\delta + u \tag{5}$$

[2]In preliminary analysis, we followed Li and Racine (2004), which is implemented by the np package (Hayfield and Racine 2008) in R (R Core Team 2020), to estimate the residual from the first stage. We applied a Gaussian kernel for the continuous variables and the Li and Racine (2007) kernel, which is an extension of the kernel proposed by Wang and van Ryzin (1981) that works well for both ordered and unordered discrete variables. Our estimated residual was neither continuous nor offered a large support, and, therefore, we did not consider this avenue any further.

The underlying identification assumptions are (i) second and third moments of x, ℓ_k and d are identified by the data and (ii) x is independent of both u and v_k, while the $\text{cov}(Z, v_k u) = 0$, for some Z, and $\text{cov}(Z, v_k^2) = 0$. The variables Z can be a subset of variables from x and/or a more standard instrument. The identification of second and third moments (distributional features) are not limiting in the LPM analysis; rather, other concerns arise. For instance, a typical instrumental variable identification concern arises from the assumption that the remaining x variables are exogenous. We test for the appropriateness of these instruments using standard tests. Furthermore, in this setting, there is also a concern that the zero heteroscedasticity identification assumption holds. We use mean-centered values of wealth levels, marital status, ethnic group, location, level of education and the number of children (young and old, separately) in our Z. Given that the method does not require standard instruments, it is recommended when traditional instruments are not available or it is suspected that a proposed instrument is too weak for identification. The approach has been used extensively in the literature, see T. T. Brown (2014), Awaworyi Churchill, Ocloo, and Siawor-Robertson (2017), J. R. Brown, Coile, and Weisbenner (2010) and Drichoutis, Nayga, and Lazaridis (2012) for a few examples applied to health related questions.

Therefore, when considering endogeneity, rather than specifying a multinomial structure, we treat the multinomial outcomes as a set of binary outcomes, instead, such as: (i) employed vs. unemployed, employed vs. not active and unemployed vs not active, and (ii) employed vs. not employed, active vs. not active and unemployed vs. not unemployed. One advantage of the linear probability model (LPM) structure in equation (4) over the multinomial (MNL) model, other than being able to apply internal instrumental variables, is that consistent MNL estimates arise only if the model, which includes both the linear index specification and the underlying error distribution assumption, are correctly specified. Potential specification issues include missing variables and independence of irrelevant alternatives (IIA) – the idea that any binary comparison that can be derived from the multinomial outcomes would be the same regardless of which other outcomes remained in the model – as well as an incorrectly specified error distribution.

The LPM structure offers a solution to these problems, as it does not require a specific error distribution. However, it is mis-specified, since it does not account for the discrete nature of the outcome, while the error term necessarily depends on the independent variables (Lewbel, Dong, and Yang 2012). Furthermore, although it is often claimed that the LPM offers a reasonable estimate of the treatment effect at the mean (Angrist and Pischke 2009), Lewbel, Dong, and Yang (2012) offer a simple counterexample showing where the estimated sign of the treatment effect is incorrect. Given that we find only limited support for the endogeneity hypothesis, we present both LPM and MNL estimates that assume exogeneity, the similarity of the results across these models offers corroborating evidence on the effect of diabetes on the labour market.

2.2 Data source

The present study uses the 2018 General Household Survey (GHS) conducted by Statistics South Africa (2019a). The GHS is a nationally representative household survey conducted yearly. It targets households and workers' hostels, but does not attempt to capture military barracks or student hostels. A total of 20 908 households (including multiple households) were successfully interviewed face-to-face; within those households 71 137 individuals stayed together at least four nights per week, and, therefore, were included in the sample. Accounting for our age limitation and missing information for some of the covariates that we discuss, below, our analysis sample consisted of 27 843. Interviews were conducted with a knowledgeable household member, and information on a wide range of demographic, social, economic and health related topics was collected. For a summary of survey outcomes, see Statistics South Africa (2019b).

In 2018, the response rate was 88.6% and there were extensive differences by location. For example, the response rate was 98.8% in Limpopo, a relatively rural and poor province, but was only 68.0% in the City of Johannesburg and 71.9% in the City of Tshwane. These two cities are located in Gauteng, the business and government epicenter of the country, such that many households cannot be reached during the day. Survey weights are available to make the data representative and those weights account for non-response; however, they are meant to match on five-year age bands, ethnicity and gender. Given that neither labour market outcomes nor diabetes outcomes are randomly distributed across these categories, it may not be appropriate to use weights in the analysis. Thus, we do not use weights in the analysis. It should be noted, however, that all interpretations should remain limited to the sample. Given our interest, we further limit the sample to individuals aged from 25 to 60. Accounting for our age limitation and incomplete information for some of the covariates, our analysis sample consisted of 33 604 observations.

2.3 Control variables

We consider a number of covariates shown to be of importance in the literature. Diabetes risk increases with age (Fletcher, Gulanick, and Lamendola 2002; Kasiam et al. 2009). Furthermore, employment rates have been shown to differ by age group in South Africa (Statistics South Africa 2018). We further include ethnicity indicators, because diabetes prevalence often differs by ethnic group (Fletcher, Gulanick, and Lamendola 2002; Klimenditis et al. 2011). South Africa's apartheid past has led to wide differences in human capital accumulation by population group (Gamede 2017), as well as labour market access and success (Burger and Jafta 2006). Thus, ethnicity is also expected to affect employment, possibly through labour market discrimination, as well as possible differences in genetic susceptibility to diabetes.

The level of education is included in our analysis, since education is related to diabetes (Bachmann et al. 2003; Lee, Kim, and Han 2013), health seeking behaviour and labour market outcomes. In South Africa, education attainment has been shown to significantly improve employment prospects (Branson and Leibbrandt 2013). For the analysis, we use dummy variables for primary, secondary, certificate, degree or postgraduate qualification, with 'no schooling' as the base measure of education.

We include a number of additional sociodemographic characteristics to capture their relationship with employment status and lifestyle habits which could result in Diabetes. A number of these are more carefully examined by Lawana et al. (2020). One is marital status (using indicators for single, the reference category, married and divorced/widowed). A number of studies find increased labour force participation (perhaps employment), as well as higher earnings, for married males. However, there is decreased labour force participation, employment and earnings for females (Ntuli 2007). Furthermore, marital status is a risk factor for Type II Diabetes in men (Cornelis et al. 2014).

We also separately include the number of children under the age of 6, which is the age children generally begin school, as well as under the age of 18, residing in the household, because child bearing increases the risk of gestational diabetes and ultimately Type II Diabetes in women (Bellamy et al. 2009). In addition, studies find reduced employment for women (Troske and Voicu 2009) compared to men (Cools, Markussen, and Strøm 2017), when there are young children in the household.

Further variables relate to dwelling and location. South Africa is a large and diverse country with regional socioeconomic differences, and, therefore, we also account for potential provincial differences in diabetes prevalence (Maier et al. 2013) and employment (Statistics South Africa 2014) with provincial indicators; Limpopo is our reference province.

To account for socioeconomic status, we incorporate wealth in the analysis. The association between wealth and health has been studied extensively, the causal direction between wealth and health is, however, uncertain (Smith 1999). In terms of diabetes, some studies show a positive association between increased wealth and diabetes prevalence (Hosseinpoor et al. 2012), whilst others find an inverse relationship between wealth and prevalence (Tanaka, Gjonca, and Gulliford 2011). Wealth is also correlated with attachment to the labour force. Previous studies consider the association between wealth and labour force participation using, housing/non-housing (Amedah and Fougère 2017; Fu, Liao, and Zhang 2016) inheritance (Amedah and Fougère 2017; J. R. Brown, Coile, and Weisbenner 2010) lottery winnings (Imbens, Rubin, and Sacerdote 2001) and rental subsidies (Jacob and Ludwig 2012) finding a negative relationship between wealth and labour supply. Similar, and potentially exogenous, wealth data is not available for us. Instead, households are asked to supply their self-assessed wealth on a six-point scale from wealthy to very poor. Although self-assessed responses may not be objectively accurate, we believe they will reasonably reflect the relative conditions of each household.

We consider one additional control variable, which relates to a slightly different literature that examines, the impact of chronic diseases, such as diabetes, hypertension and cardiovascular disease on labour outcomes (Kouwenhoven-Pasmooij et al. 2016; Unmuessig et al. 2016; de Boer et al. 2018). From this literature, we know that individuals with chronic disease have a high probability of early retirement and unemployment (de Boer et al. 2018) and that co-morbidities might also influence employment outcomes (Pedron et al. 2019). Therefore, we include hypertension, because of its high prevalence in individuals with diabetes (Nowakowska et al. 2019; Long and Dagogo-Jack 2011). According to Long and Dagogo-Jack (2011), 75% of individuals with diabetes have hypertension. These two conditions have the same risk factors and complications, such that the existence of one is a predisposition to the other.

3 Results

We present our findings in the following subsections, beginning with a brief overview of the analysis variables, followed by our analysis of endogeneity via a series of linear probability models. These models suggest that endogeneity is not observed across all outcomes. For that reason, we return to the multinomial model to examine the relationship between diabetes, age and labour market outcomes for both men and women.

3.1 Descriptive statistics

Summary statistics are presented in Table 1, separated by diabetes status. Explicitly, that is determined by whether or not the individual in the household has been tested for diabetes, and the household survey respondent knows. Thus, although we accept that there might be measurement error in the reported value, and, therefore, we do consider that possibility in the analysis, we report the data, as it is. Less than 1 000 of the 33 604 are reported to have diabetes, and whether it is Type I or Type II is not available in the data.

Table 1: Summary statistics of analysis data: GHS 2018 Note: Numbers listed for categorical variables represent the number of observations of that category; the rounded percent of observations in that subgroup is presented in parenthesis. For age, the mean is presented with its standard deviation, separated by ±.

Variable	Men w/ Diabetes	Men w/o	Women w/ Diabetes	Women w/o

Labour Force Participation				
Employed	219 (68)	8,361 (67)	262 (45)	7,162 (50)
Unemployed	20 (6)	2,030 (16)	22 (4)	2,446 (17)
Inactive	84 (26)	2,133 (17)	294 (51)	4,810 (33)
Possible Co-Morbidities				
Hypertension	142 (44)	706 (6)	355 (61)	1,647 (11)
Ethnic Group				
African	200 (62)	10,258 (82)	421 (73)	11,856 (82)
Mixed	79 (24)	1,171 (9)	100 (17)	1,421 (10)
Asian	19 (6)	293 (2)	27 (5)	260 (2)
White	25 (8)	802 (6)	30 (5)	881 (6)
Marital Status				
Living Together	200 (62)	10,258 (82)	421 (73)	11,856 (82)
Living Apart	79 (24)	1,171 (9)	100 (17)	1,421 (10)
Never Married	19 (6)	293 (2)	27 (5)	260 (2)
Education				
No Schooling	12 (4)	357 (3)	35 (6)	526 (4)
Some Schooling	78 (24)	2,445 (20)	198 (34)	2,625 (18)
Completed Grade 9	26 (8)	957 (8)	47 (8)	1,026 (7)
Completed Grade 10	40 (12)	1,574 (13)	58 (10)	1,629 (11)
Completed Grade 11	33 (10)	1,763 (14)	52 (9)	2,150 (15)
Completed Grade 12	69 (21)	3,713 (30)	117 (20)	4,330 (30)
First Year University	9 (3)	317 (3)	6 (1)	373 (3)
Second Year University	26 (8)	467 (4)	29 (5)	659 (5)
Completed University	17 (5)	573 (5)	22 (4)	708 (5)
Completed Honours	10 (3)	210 (2)	10 (2)	225 (2)
Completed Masters	3 (1)	120 (1)	1 (0)	141 (1)
Completed Doctorate	0 (0)	28 (0)	3 (1)	26 (0)
Wealth Status				
Wealthy	2 (1)	115 (1)	2 (0)	137 (1)
Very comfortable	18 (6)	564 (5)	29 (5)	658 (5)

Reasonably comfortable	92 (28)	2,340 (19)	130 (22)	2,701 (19)
Just getting by	152 (47)	6,219 (50)	279 (48)	7,199 (50)
Poor	47 (15)	2,464 (20)	106 (18)	2,896 (20)
Very Poor	12 (4)	822 (7)	32 (6)	827 (6)
Residence				
Rural	59 (18)	3,809 (30)	175 (30)	4,998 (35)
Urban	264 (82)	8,715 (70)	403 (70)	9,420 (65)
Province				
Western Cape	68 (21)	1,403 (11)	100 (17)	1,532 (11)
Eastern Cape	30 (9)	1,455 (12)	86 (15)	1,871 (13)
Northern Cape	27 (8)	573 (5)	36 (6)	672 (5)
Free State	20 (6)	678 (5)	33 (6)	852 (6)
KwaZulu-Natal	51 (16)	2,146 (17)	139 (24)	2,534 (18)
North West	19 (6)	823 (7)	28 (5)	861 (6)
Gauteng	74 (23)	3,217 (26)	82 (14)	3,315 (23)
Mpumalanga	17 (5)	1,051 (8)	38 (7)	1,222 (8)
Limpopo	17 (5)	1,178 (9)	36 (6)	1,559 (11)
Age				
mean (sd)	49.80 (8.22)	39.12 (9.88)	51.27 (7.45)	39.91 (10.12)
Children Under 18				
mean (sd)	1.25 (1.52)	1.18 (1.53)	1.59 (1.61)	1.86 (1.76)
Children under 6				
mean (sd)	0.35 (0.67)	0.40 (0.71)	0.50 (0.80)	0.62 (0.85)

The descriptive statistics suggest a high level of hypertension co-morbidity, as expected, and that those with diabetes are relatively more likely to be older females of mixed race, who are living apart from their spouse (either due to death, separation or divorce) and generally less educated; they also tend to lie within the upper-middle portion of the wealth distribution, at least in their own view. They are also relatively more urban, although less likely to be from Gauteng, a largely urban province and the locale for both Johannesburg and Pretoria. We also see that they are relatively more likely to be non-labour force participants and less likely to be unemployed than those who do not have diabetes.

3.2 Diabetes and endogeneity

Given our goal of estimating the impact of diabetes on common labour market outcomes, as well as the possibility that diabetes status is mismeasured or endogenous for other reasons, we begin by examining the exogeneity of diabetes status within the triangular linear system outlined in (4). We undertake that analysis applying heteroscedastic instrumental variables regression, as outlined by Lewbel (2012) and Lewbel (2018). We estimate the model using the REndo package (Gui et al. 2020) in R (R Core Team 2020), which also accommodates heteroscedasticiy-consistent standard errors via the sandwich package (Zeileis 2004, 2006) and standard instrumental variable (IV) diagnostic tests, including a test of endogeneity - our main interest - a test of weak instruments and a test of overidentification, via the AER package (Kleiber and Zeileis 2008). Although our main interest is in endogeneity, we also discuss the estimated diabetes effect on labour market outcomes.

We undertook two separate analyses. In the first, we take subsamples, such that the linear probability model accounting for endogeneity compares: (i) employed and unemployed, in which case the non-participants are ignored; (ii) employed and non-participation, which ignores the unemployed; and (iii) the unemployed and non-participants ignoring the employed. Across these subsets, we estimated both heteroscedasticity IV linear probability models, as well as uncorrected linear probability models. Those results are presented in Table 2, which is limited to the estimated diabetes effect (corrected and uncorrected), along with model diagnostics. The full set of results are available in Appendix Table A.1, which assumes exogenous diabetes, and A.3, which does not. We find some evidence of endogeneity in this analysis; however, it is only for men and only for the effect of diabetes on employment relative to non-participation.[3]

Our results suggest that women are statistically significantly more likely - approximately 6% - to be employed than unemployed, if they have diabetes. Furthermore, they are statistically significantly less likely (ranging from 5–9%) to be employed or unemployed, as opposed to not participating in the labour force. As noted above, these results are based on exogenous diabetes, given that we do not find evidence that diabetes is endogenous for women. For men, there is no statistically significant effect of diabetes on employment, relative to unemployment. However, there is an estimated 13% statistically significant reduction, at the mean, in their probability of employment relative to non-participation, arising from endogenous diabetes; it is estimated to only be around 5% if diabetes is not endogenous. For unemployment, relative to non-participation, the exogenous effect (which cannot be rejected) is approximately 8%, while the endogenous effect is estimated to lie between 13% and 14%.

[3] In additional analysis not reported, we examined the sensitivity of our endogeneity conclusion. If we limited our internal instruments to only household wealth, no endogeneity was uncovered; for all other combinations of internal instruments, endogeneity was robust.

Table 2: Linear probability model of the effect of diabetes on labour market outcomes, where the endogeneity is identified from heteroscedasticity, along with instrument diagnostics

	Emp(=1) v Unemp(=0)		Emp(=1) v NLFP(=0)		Unemp(=1) v NLFP(=0)	
	Males	Females	Males	Females	Males	Females
Diabetes (Unorrected)	0.0165	0.0593^{c}	-0.0498^{c}	-0.0467^{c}	-0.0837^{d}	-0.0848^{a}
	(0.024)	(0.025)	(0.022)	(0.020)	(0.048)	(0.026)
Diabetes (Corrected)	0.0508	0.0270	-0.1294^{a}	-0.0744	-0.1355^{a}	-0.0787^{c}
	(0.031)	(0.040)	(0.041)	(0.055)	(0.046)	(0.036)
Weak Instruments	282.533^{a}	104.625^{a}	336.728^{a}	103.195^{a}	283.247^{a}	111.826^{a}
Wu-Hausman	1.100	0.376	8.558^{a}	0.326	1.665	0.017
Sargan	14.309	15.961	29.167^{d}	22.954	13.940	10.091

The dependent variable is binary, with values and lables listed in the column headers. Results include exogenous (Uncorrected) and those that allow for the endogeneity of diabetes (Corrected) using mean-centered values of wealth levels, ethnic group, marital status, location, education level and number of children (young and old) for identification. Instrumental variable diagnostic tests include: a test of weak instruments, endogeneity (Wu-Hausman) and overidentification (Sargan). In each case, the test statistic is provided along with its significance, i.e., the probability that the null hypothesis is true, where the following notation and probability levels are assumed: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1. Emp is employed, Unemp is unemployed and NLFP is non-participant in the labour force. Heteroscedasticity corrected standard errors reported for the diabetes estimate.

In the second analysis, we use all the data and focus on simple labour market indicators, such as: (i) employed and not employed, in which case the non-participants are also not employed; (ii) unemployed and not unemployed, where the employed are lumped with non-participants; and (iii) participants and non-participants. For each of these comparisons, we estimated an uncorrected linear probability model, as well as a heteroscedasticity IV linear probability model. A subset of results from these regressions are presented in Table 3. As before, it is limited to the estimated diabetes effect (both corrected and uncorrected), along with model diagnostics. The full set of results are available in Appendix Table A.2, which assumes exogeneity, and Table A.4 for the heteroscedastic IV models. We also find evidence of endogeneity in this analysis, and, again, it is only for men and only for labour force participant compared to non-participant. As we found above, the estimated corrected diabetes effect is more than double the uncorrected estimate; however, given the IV application, the underlying standard errors are also much larger.

Women are found less likely to be unemployed than not, as well as less likely to be participating in the labour force than not. The statistically significant estimates range from 4%-7%, assuming exogeneity, which cannot be rejected. For men, diabetes is only found to be endogenous and statistically significant, when comparing labour force participation and non-participation. In that regard, diabetes is found to reduce participation by approximately 14%. For both men and women, diabetes is found to be both exogenous in the employed-v-not-employed comparison and statistically insignificant.

Table 3: Linear probability model of the effect of diabetes on on employment, unemployment and non-participation, where the endogeneity is identified from heteroscedasticity, along with instrument diagnostics

	Emp(=1) v Not Emp(=0)		LFP(=1) v NLFP(=0)		Unemp(=1) v Not Unemp(=0)	
	Males	Females	Males	Females	Males	Females
Diabetes (Unorrected)	-0.0361	-0.0266	-0.0507^{c}	-0.0680^{a}	-0.0146	-0.0414^{b}
	(0.025)	(0.020)	(0.021)	(0.019)	(0.020)	(0.016)
Diabetes (Corrected)	-0.0849^{c}	-0.0468	-0.1374^{a}	-0.0853^{d}	-0.0525^{c}	-0.0384^{d}
	(0.041)	(0.052)	(0.039)	(0.051)	(0.021)	(0.022)
Weak Instruments	390.264^{a}	136.833^{a}	390.264^{a}	136.833^{a}	390.264^{a}	136.833^{a}
Wu-Hausman	2.508	0.188	11.368^{a}	0.151	2.210	0.007
Sargan	26.772	22.929	32.423^{c}	26.200	10.185	10.100

The dependent variable is binary with values and lables listed in the column headers. Results include exogenous (Uncorrected) and those that allow for the endogeneity of diabetes (Corrected) using mean-centered values of wealth levels, ethnic group, marital status, location, education level and number of children (young and old) for identification. Instrumental variable diagnostic tests include: a test of weak instruments, endogeneity (Wu-Hausman) and overidentification (Sargan). In each case, the test statistic is provided along with its significance, i.e., the probability that the null hypothesis is true, where the following notation and probability levels are assumed: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1. Emp is employed, Not Emp is not employed (i.e., unemployed or NLFP), Unemp is unemployed, Not Unemp is not unemployed, LFP is labour force participant and NLFP is non-participant in the labour force. Heteroscedasticity corrected standard errors reported for the diabetes estimate.

A relatively small proportion of the sample are reported to have diabetes, 2.5% of the sampled men and 3.9% of the sampled women. In terms of our results, for both men and women with diabetes, employment is less likely than non-employment, while unemployment is also less likely than non-unemployment. The estimates are not precise in all cases. However, for labour force participation relative to non-participation, there is strong evidence of a statistically significant diabetes effect; it ranges from 5–14% for men (exogenous v endogenous) to 7–9% for women (exogenous v endogenous). Given current labour force participation rates for those with diabetes (74.2% for men and 49.0% for women), the complete elimination of diabetes is estimated to increase labour force participation rates to 84.6% for men and 52.4% for women (for those who previously had diabetes). For comparison, the participation rates for men without diabetes is only 83%, while that for women without diabetes is 66.7%; see Table 1. In other words, the linear probability model results appear to overstate male labour force reductions arising from diabetes, while understating the female reductions.

3.3 Mutinomial results

Although there is some evidence of endogeneity, our analysis suggests it is only for men in our sample and only for results related to non-participation. However, the suggestion that diabetes might be endogenous for men, when it comes to one of three possible labour market outcomes, but not the others, is not logically consistent; all labour market outcomes arise from the same mutually exclusive set of outcomes. Furthermore, the linear probability models could be (over-) under-stating the importance of diabetes.

One criticism of the preceding analysis is that it has relied on linear probability models, which are, by assumption, mis-specified and may not even offer accurate treatment effects (Lewbel, Dong, and Yang 2012). Another criticism implied by Lewbel (2018) is that although it is possible to meet the underlying IV-heteroscedasticity identification assumptions, the criteria is not obvious or intuitive. For those reasons, we also estimate a simple multinomial logit model assuming there is no endogeneity; we estimate this model in R (R Core Team 2020) using the nnet package (Venables and Ripley 2002). We report the full results in Appendix Table A.5. We include the same set of controls as before.

However, parameter estimates from a multinomial logit do not offer much interpretive information, and, therefore, researchers often present marginal effects. We present something different, although similarly motivated. Specifically, we present three illustrations of the net conditional probability that a male/female is in a particular state of the labour market against their age. It is a net conditional probability, because we estimate the difference in labour state probabilities by diabetes status. Specifically, at all ages, we subtract the non-diabetes predicted conditional probability from the same prediction for those with diabetes. Our illustration offers more interpretational insight than a simple mean marginal effect.

In order to develop the estimates, we had to address another issue that arises in nonlinear marginal effects estimation. All of the variables in the model affect the estimate, and, therefore, a **ceteris paribus** assumption must be made, which requires a decision on the level/value of the remaining covariates. In this case, we chose the median for continuous variables (except for age) and the mode of the specific distribution, in the case of categorical variables. For this analysis, the modes imply that our comparison individuals are Africans, who are married and living with their spouse, located in an urban setting in Gauteng, have completed matric (high school) and have one child between the age of 5 and 17. Furthermore, given that hypertension is so common amongst those with diabetes, we make one further distinction between those with and without diabetes: those with are assumed to have hypertension as a co-morbidity, while those without do not.

Given the results above, which suggest that diabetes lowers labour force participation, we initially focus our attention there; see Figure 1. For both males and females and across all ages, the estimated net probability of non-participation is positive, which means that those with diabetes are less likely to participate. We also see that the estimated difference is larger for females than males at most ages, although the 95% confidence intervals first intersect around the age of 40. Finally, there is a U-shape to that difference: labour force participation peaks somewhere around the age of 40 to 45. By assumption, these marginal effects assume exogeneity of diabetes, although they do acknowledge that multinomial outcomes, which are non-linear in nature, should be estimated via non-linear models.

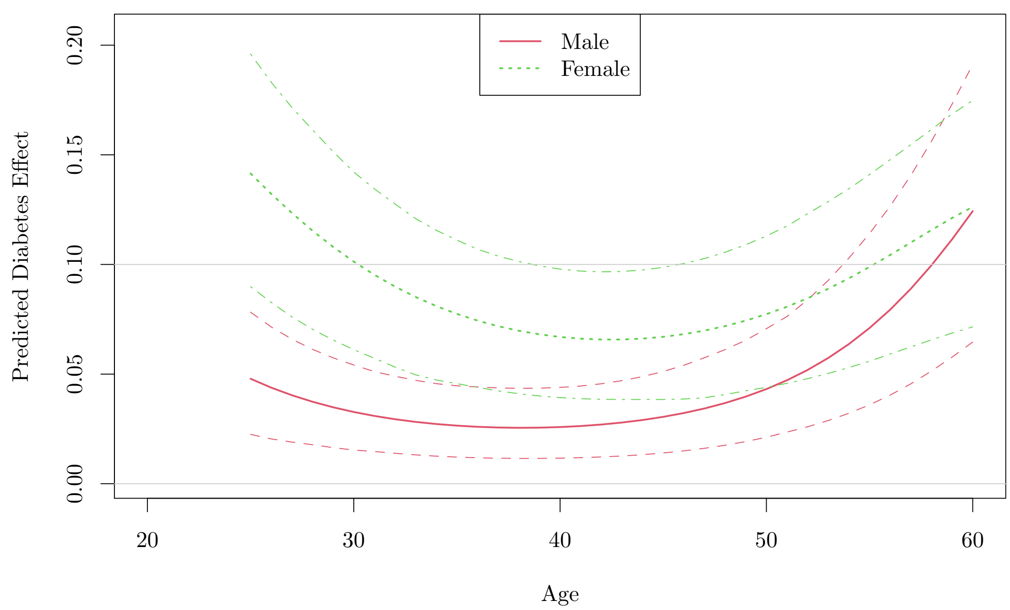

Figure 1: Marginal Effects of Diabetes on Non-participation in the Labour Force from Multinomial Logit. The illustration depicts the difference in the predicted probability of employment for those with and without diabetes. The differences are calculated across the ages 20 to 60 and for men and women. Standard errors are taken from 399 bootstrap replications of the differences.

Subject to our conditioning strategy, we find only limited statistically significant evidence that diabetes affects the probability of employment by age - see Figure 2 - and only near the age of 60. For both men and women, diabetes lowers the probability of employment at ages near 60. The larger affect appears to be on unemployment, especially for women. Women with diabetes are much less likely to be unemployed than those without, and that is across all age groups; see Figure 3. For men, on the other hand, there is no statistically significant evidence that diabetes is affecting unemployment, again, subject to our conditioning strategy

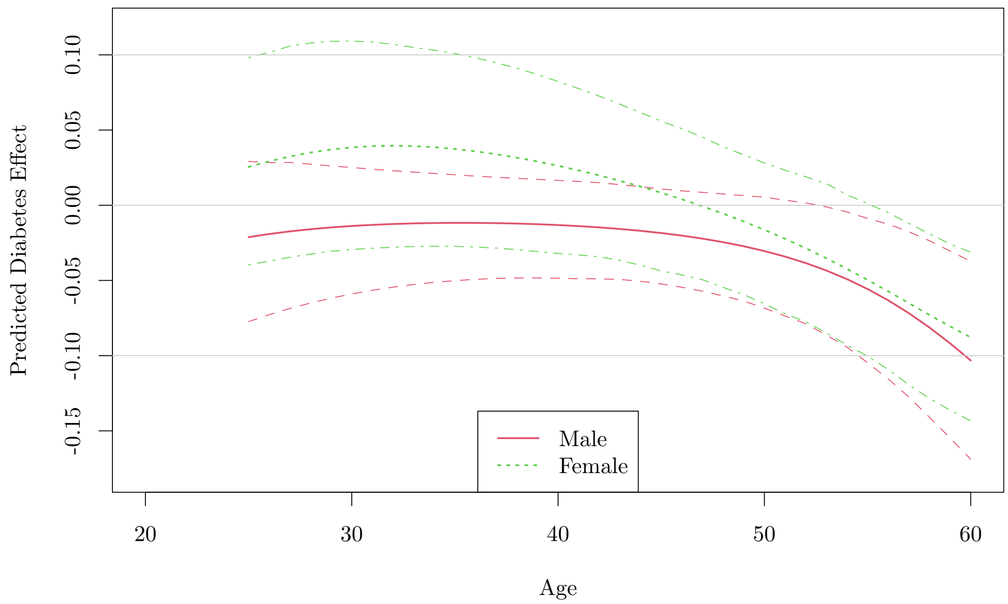

Figure 2: Marginal Effects of Diabetes on Employment from Multinomial Logit. The illustration depicts the difference in the predicted probability of employment for those with and without diabetes. The differences are calculated across the ages 20 to 60 and for men and women. Standard errors are taken from 399 bootstrap replications of the differences.

Three basic conclusions can be deduced from these figures: (1) Employment is not particularly affected by diabetes; (2) the unemployment conditional net probability for women with diabetes is lower, although that reduction is mitigated by age; (3) the non-participation conditional net probability follows a U-shape with both the youngest and oldest with diabetes more likely to be non-participants. From our illustration, conditional net non-participation amongst women is around 3 to 12 percentage points higher, depending on age, for those with diabetes than without. In our sample, non-participants are about 30% of the non-diabetic sample and 42% of the diabetic sample. Therefore, within the multinomial model, diabetes explains between 7.5% and 30% of that 12 point gap (3% and 12% of 30 are 0.9 and 3.6, respectively, while 0.9/12 and 3.6/12 yield 12.5% and 30%, respectively).

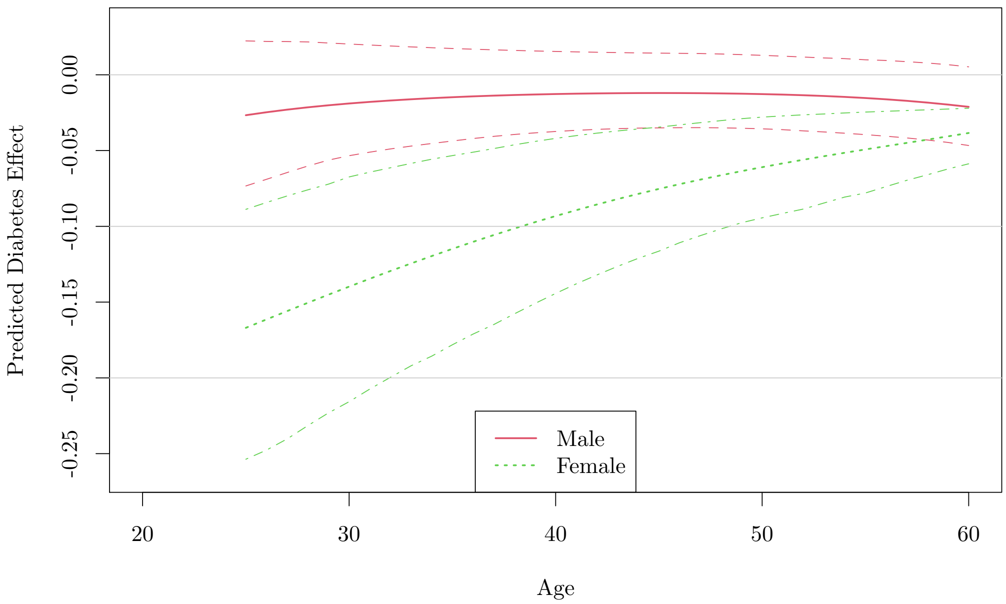

Figure 3: Marginal Effects of Diabetes on Unemployment from Multinomial Logit. The illustration depicts the difference in the predicted probability of employment for those with and without diabetes. The differences are calculated across the ages 20 to 60 and for men and women. Standard errors are taken from 399 bootstrap replications of the differences.

4 Discussion

As noted at the outset, we are not the first researchers to consider health and labour market issues in South Africa. That literature has mostly examined HIV/AIDS (Arndt and Lewis 2000; Levinsohn et al. 2013; Young 2005), although more recent research considers obesity (Some, Rashied, and Ohonba 2016), self-assessed poor health (Nwosu and Woolard 2017) and non-communicable diseases, including diabetes (Lawana et al. 2020). Lawana et al. (2020) present evidence related to behaviour and its pathways to non-communicable diseases and labour force participation. Their model relies on normality and assumes that lifestyle behaviours (such as smoking, drinking, exercise and obesity) can be excluded from the labour force participation equation. A relatively extensive literature looks at the direct relationship between such behaviours and labour market outcomes – Neumann (2013), Larose et al. (2016), and Böckerman et al. (2018) – suggesting that these exclusions may not be appropriate. However, their focus was on the estimation of total effects through a recursive structure and labour force participation. Their total results suggested relatively large negative female labour market effects, about a 14% reduction, and no effects for men. Our multinomial results that do not correct for endogeneity are qualitatively similar - which suggest that effects for females are larger (in absolute value) than for men; however our linear probability models controlling for endogeneity suggest the opposite. Thus, we complement this local literature by offering a more nuanced examination of the labour market, as well as a further look at endogeneity.

In the international literature, diabetes has a negative impact on labour outcomes, either reduced participation or employment; furthermore, the evidence on endogeneity is not consistent. Pedron et al. (2019) argues that too few studies include co-morbidities, control for endogeneity or differentiate between type I and type II diabetes, which may lead to over-estimated impacts. In this analysis, in addition to accounting for endogeneity, we were able to account for co-morbidities; however, we were not able to differentiate between Type I and II diabetes. Although endogeneity has typically been addressed through genetic instrumentation (Brown III, Pagan, and Bastida 2005; Latif 2009), which was not available to us, we did address it in the analysis. In previous literature, exogeneity is either rejected (Lin 2011) or rejected for men, but not for women (Brown III, Pagan, and Bastida 2005). However, in our analysis, exogeneity is only rejected for men, and only for labour force participation.

Our analysis points to fairly large negative effects of diabetes on labour force participation, regardless of exogeneity. However, we do not find negative effects for employment. Instead, we find negative effects on unemployment. Unemployment rates are high in our sample (16–17%), although lower than reported by the official statistics.[4] Thus, it would appear that South African diabetics are bimodal. They are either managing their condition and, therefore, able to remain employed, or they are not able to manage their condition and they are not healthy enough to remain in the labour force. Our results suggest that the former group gets smaller with age, which is expected, given that diabetes worsens with age.

Although there are a number of studies in the literature, they rarely consider developing or even middle income countries. Our results are rather different to those available in another middle income country, Mexico. Seuring, Goryakin, and Suhrcke (2015) and Seuring, Serneels, and Suhrcke (2019) find 4-6% reductions in employment, arising from diabetes for both men and women, compared to our negligible, at best, effects on employment. Our results are similar in that they did not uncover endogeneity with respect to employment; we only uncovered evidence of endogeneity for men and for labour force participation.

Given the prevalence of diabetes in the working age group (Bertram et al. 2013), and its expected future prevalence, as well as other features of the South African economy, it is possible that the effects of diabetes are more nefarious in South Africa than in other countries. Our results, however, suggest that it may not be more of a problem here than in any other country. Furthermore, we find that the effects, themselves, appear to be quite nuanced, and in need of further analysis.

[4] The official statistics point to unemployment rates closer to 27% (http://www.statssa.gov.za/?p=11897)[http://www.statssa.gov.za/?p=11897], and these figures are based on labour force surveys, rather than the general household survey used for this analysis.

5 Conclusion

In this research, we examined the effect of diabetes on labour market outcomes using recent data from South Africa. Due to the fact that endogeneity is a concern in any analysis of health effects, although it has not always been considered in the literature (Pedron et al. 2019), we applied internal instrumental variable methods, as outlined in Lewbel (2012) and Lewbel (2018). In order to do so, we were forced to restructure the multinomial labour market outcomes into a series of binary comparisons. The results from that analysis pointed to some endogeneity effects. However, it was only uncovered for men and only for labour force participation. If accounted for in the analysis, it was found that diabetes explained the entire difference in participation rates between men with and without diabetes. For women, no evidence of endogeneity was uncovered, such that the effect of diabetes explained a rather smaller amount of the difference in female participation rates across diabetes status.

The lack of support for endogeneity, as well as general issues that arise from linear probability models, led us to also apply multinomial models. We presented marginal effects from that analysis, focusing on the conditional difference in the probability of being in any particular state of the labour market for those with and without diabetes. As with much of the previous literature in this area, as well as what we have already noted, we found differences between men and women with women generally more affected (at least in the exogenous cases). However, we found little evidence to suggest that employment is reduced, because of diabetes. Instead, we find that it is unemployment that is reduced, while non-participation is higher. These results suggest that the employed are able to manage their diabetes and remain in the workplace. However, we are not able to offer any analysis on mechanisms that might drive such a result. We leave that as an open question for future research.

References

Amedah, Sid Ali, and Maxime Fougère. 2017. "Impact of Inheritance and Housing Wealth on Retirement Decision." *Open Journal of Social Sciences* 05 (07): 1–12. https://doi.org/10.4236/jss.2017.57001.

Angrist, Josh, and Stephen-Jorn Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ, US: Princeton University Press.

Arndt, Channing, and Jeffrey D Lewis. 2000. "The macro implications of HIV/AIDS in South Africa: A preliminary assessment." *South African Journal of Economics* 68 (5): 856–87. https://doi.org/10.1111/j.1813-6982.2000.tb01283.x.

Awaworyi Churchill, Sefa, Janet Exornam Ocloo, and Diana Siawor-Robertson. 2017. "Ethnic Diversity and Health Outcomes." *Social Indicators Research* 134 (3): 1077–1112. https://doi.org/10.1007/s11205-016-1454-7.

Bachmann, M O, J Eachus, C D Hopper, S G Davey, C Propper, N J Pearson, S Williams, D Tallon, and

S Frankel. 2003. “Socio-economic inequalities in diabetes complications, control, attitudes and health service use: a cross-sectional study.” *Diabetic Medicine* 20 (11): 921–29. https://doi.org/10.1046/j.1464-5491.2003.01050.x.

Basu, Sanjay, Sukumar Vellakkal, Sutapa Agrawal, David Stuckler, Barry Popkin, and Shah Ebrahim. 2014. “Averting obesity and type 2 diabetes in India through sugar-sweetened beverage taxation: an economic-epidemiologic modeling study.” *PLoS Med* 11 (1): e1001582. https://doi.org/10.1371/journal.pmed.1001582.

Bellamy, Leanne, Juan-Pablo Casas, Aroon D Hingorani, and David Williams. 2009. “Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis.” *The Lancet* 373 (9677): 1773–79. https://doi.org/10.1016/s0140-6736(09)60731-5.

Bertram, Melanie Y, Aneil VS Jaswal, Victoria Pillay Van Wyk, Naomi S Levitt, and Karen J Hofman. 2013. “The non-fatal disease burden caused by type 2 diabetes in South Africa, 2009.” *Global Health Action* 6: 19244. https://doi.org/10.3402/gha.v6i0.19244.

Böckerman, Petri, Ari Hyytinen, Jaakko Kaprio, and Terhi Maczulskij. 2018. “If You Drink, Don’t Smoke: Joint Associations Between Risky Health Behaviors and Labor Market Outcomes.” *Social Science & Medicine* 207: 55–63. https://doi.org/https://doi.org/10.1016/j.socscimed.2018.04.039.

Branson, Nicola, and Murray Leibbrandt. 2013. “Educational Attainment and Labour Market Outcomes in South Africa, 1994-2010,” no. 1022. https://doi.org/https://doi.org/https://doi.org/10.1787/5k4c0vvbvv0q-en.

Brown III, H. Shelton, José A Pagan, and Elena Bastida. 2005. “The impact of diabetes on employment: genetic IVs in a bivariate probit.” *Health Economics* 14 (5): 537–44. https://doi.org/10.1002/hec.942.

Brown, Jeffrey R, Courtney C Coile, and Scott J Weisbenner. 2010. “The Effect of Inheritance Receipt on Retirement.” *The Review of Economics and Statistics* 92 (2): 425–34. https://doi.org/10.1162/rest.2010.11182.

Brown, Judith, Evangelia Demou, Madeleine Ann Tristram, Harper Gilmour, Kaveh A Sanati, and Ewan B Macdonald. 2012. “Employment Status and Health: Understanding the Health of the Economically Inactive Population in Scotland.” *BMC Public Health* 12 (327): 481–90. https://doi.org/10.1186/1471-2458-12-327.

Brown, Timothy Tyler. 2014. “How effective are public health departments at preventing mortality?” *Economics & Human Biology* 13 (March): 34–45. https://doi.org/10.1016/j.ehb.2013.10.001.

Burger, Ronelle, and Rachel Jafta. 2006. “Returns to Race: Labour Market Discrimination in Post-Apartheid South Africa.” PhD thesis, Stellenbosch: University of Stellenbosch.

Cools, Sara, Simen Markussen, and Marte Strøm. 2017. “Children and Careers: How Family Size Affects Parents’ Labor Market Outcomes in the Long Run.” *Demography* 54 (5): 1773–93. https://doi.org/10.1007/s13524-017-0612-0.

Cornelis, M C, S E Chiuve, M M Glymour, S C Chang, E J Tchetgen Tchetgen, L Liang, K C Koenen, E B Rimm, I Kawachi, and L D Kubzansky. 2014. “Bachelors, divorcees, and widowers: does marriage protect men from type 2 diabetes?” *Plos One* 9 (9): e106720. https://doi.org/10.1371/journal.pone.0106720.

Cottini, Elena, and Paolo Ghinetti. 2018. “Employment Insecurity and Employees’ Health in Denmark.” *Health Economics* 27 (2): 426–39. https://doi.org/10.1002/hec.3580.

Davis, Karen, Sara R Collins, Michelle M Doty, Alice Ho, and Alyssa Holmgren. 2005. “Health and productivity among U.S. workers.” *Issue Brief (Commonwealth Fund)*, no. 856 (August): 1–10. http://www.ncbi.nlm.nih.gov/pubmed/16138438.

de Boer, Angela G. E. M., Goedele A. Geuskens, Ute Bültmann, Cécile R. L. Boot, Haije Wind, Lando L. J. Koppes, and Monique H. W. Frings-Dresen. 2018. “Employment Status Transitions in Employees with and Without Chronic Disease in the Netherlands." *International Journal of Public Health* 63 (6): 713–22. https://doi.org/10.1007/s00038-018-1120-8.

Dong, Yingyang. 2010. "Endogenous Regressor Binary Choice Models Without Instruments, with an Application to Migration." *Economics Letters* 107 (1): 33–35. https://doi.org/10.1016/j.econlet.2009.12.017.

Drichoutis, Andreas C., Rodolfo M. Nayga, and Panagiotis Lazaridis. 2012. "Food away from home expenditures and obesity among older Europeans: are there gender differences?" *Empirical Economics* 42 (3): 1051–78. https://doi.org/10.1007/s00181-011-0453-4.

Fletcher, Barbara, Meg Gulanick, and Cindy Lamendola. 2002. "Risk Factors for Type 2 Diabetes Mellitus." *The Journal of Cardiovascular Nursing* 16: 17–23.

Fu, Shihe, Yu Liao, and Junfu Zhang. 2016. "The Effect of Housing Wealth on Labor Force Participation: Evidence from China." *Journal of Housing Economics* 33: 59–69. https://doi.org/https://doi.org/10.1016/j.jhe.2016.04.003.

Gamede, Ntombifuthi W. 2017. "Human Capital Development in South Africa: Perspectives on Education in the Post-Apartheid Era." Master's thesis, Pretoria: University of South Africa.

Gui, Raluca, Markus Meierer, Rene Algesheimer, and Patrik Schilter. 2020. *REndo: Fitting Linear Models with Endogenous Regressors Using Latent Instrumental Variables*. https://CRAN.R-project.org/package=REndo.

Hayfield, Tristen, and Jeffrey S Racine. 2008. "Nonparametric Econometrics: The Np Package." *Journal of Statistical Software* 27 (5): 1–32.

Hosseinpoor, Ahmad R, Nicole Bergen, Shanthi Mendis, Sam Harper, Emese Verdes, Anton Kunst, and Somnath Chatterji. 2012. "Socioeconomic inequality in the prevalence of noncommunicable diseases in low- and middle-income countries: results from the World Health Survey." *BMC Public Health* 12: 474. https://doi.org/10.1186/1471-2458-12-474.

Imbens, Guido W, and Whitney K Newey. 2009. "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity." *Econometrica* 77 (5): 1481–1512. https://doi.org/10.3982/ECTA7108.

Imbens, Guido W, Donald B Rubin, and Bruce I Sacerdote. 2001. "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players." *The American Economic Review* 91 (4). https://doi.org/10.1257/aer.91.4.778.

Jacob, Brain A, and Jens Ludwig. 2012. "The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery." *The American Economic Review* 102 (1): 272–304. https://doi.org/10.1257/aer.102.1.272.

Jonsson, Bengt. 1998. "The Economic Impact of Diabetes." *Diabetes Care* 21 (Supplement 3): C7–10. https://doi.org/10.2337/diacare.21.3.C7.

Kasiam, Lasi O, B Longo-Mbenza, Okwe A Nge, Kabangu N Kangola, Fuele S Mbungu, and Dipa G Milongo. 2009. "Classification and dramatic epidemic of diabetes mellitus in Kinshasa Hinterland: the prominent role of type 2 diabetes and lifestyle changes among Africans." *Nigerian Journal of Medicine* 18: 311–20. https://doi.org/10.4314/njm.v18i3.51202.

Kirigia, Joses M, Hama B Sambo, Luis G Sambo, and Saidou P Barry. 2009. "Economic burden of diabetes mellitus in the WHO African region." *BMC International Health and Human Rights* 9 (1): 6. https://doi.org/10.1186/1472-698X-9-6.

Kleiber, Christian, and Achim Zeileis. 2008. *Applied Econometrics with R*. New York: Springer-Verlag. https://CRAN.R-project.org/package=AER.

Klimenditis, Yann C, Marshall Abrams, Jelai Wang, José R Fernandez, and David B Allison. 2011. "Natural selection at genenomic regions associated with obesity and type-2 diabetes. East Asians and sub-Saharan

Africans exhibit high levels of differentiation at type 2 diabetes region." *Human Genetics* 129: 407–18. https://doi.org/10.1007/s00439-010-0935-z.

Koch, Steven F, and Gauthier Tshiswaka-Kashalala. 2018. "The Demand for Reproductive Health Care." *Journal of African Economies* 27 (4): 405–29. https://doi.org/10.1093/jae/ejx042.

Kouwenhoven-Pasmooij, T A, A Burdorf, J W Roos-Hesselink, M G M Hunink, and S J W Robroek. 2016. "Cardiovascular Disease, Diabetes and Early Exit from Paid Employment in Europe; the Impact of Work-Related Factors." *International Journal of Cardiology* 215 (332–7): 1–17. https://doi.org/10.1016/j.ijcard.2016.04.090.

Larose, Samantha L., Koffi A. Kpelitse, M. Karen Campbell, Gregory S. Zaric, and Sisira Sarma. 2016. "Does Obesity Influence Labour Market Outcomes Among Working-Age Adults? Evidence from Canadian Longitudinal Data." *Economics & Human Biology* 20: 26–41. https://doi.org/https://doi.org/10.1016/j.ehb.2015.09.007.

Latif, Ehsan. 2009. "The impact of diabetes on employment in Canada." *Health Economics* 18 (5): 577–89. https://doi.org/10.1002/hec.1390.

Lawana, Nozuko, Frederik Booysen, Asrat Tsegaye, Forget Mingiri Kapingura, and Charles Hongoro. 2020. "Lifestyle Risk Factors, Non-Communicable Diseases and Labour Force Participation in South Africa." *Development Southern Africa* 37 (3): 446–61. https://doi.org/10.1080/0376835X.2019.1678459.

Lee, D S, Y J Kim, and H R Han. 2013. "Sex differences in the association between socio-economic status and type 2 diabetes: data from the 2005 Korean National Health and Nutritional Examination Survey (KNHANES)." *Public Health* 127 (6): 554–60. https://doi.org/10.1016/j.puhe.2013.02.001.

Levinsohn, James A, Zoë McLaren, Olive Shisana, and Khangelani Zuma. 2013. "HIV Status and Labor Market Participation in South Africa." *The Review of Economics and Statistics* 95 (1): 98–108. https://doi.org/https://doi.org/10.1162/REST_a_00237.

Lewbel, Arthur. 2012. "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models." *Journal of Business & Economic Statistics* 30 (1): 67–80. https://doi.org/10.1080/07350015.2012.643126.

———. 2018. "Identification and Estimation Using Heteroscedasticity Without Instruments: The Binary Endogenous Regressor Case." *Economics Letters* 165 (1): 10–12. https://doi.org/https://doi.org/10.1016/j.econlet.2018.01.003.

Lewbel, Arthur, Yingyang Dong, and Thomas Tao Yang. 2012. "Comparing Features of Convenient Estimators for Binary Choice Models with Endogenous Regressors." *Canadian Journal of Economics* 45 (3): 809–29. https://doi.org/10.1111/j.1540-5982.2012.01733.x.

Li, Qi, and Jeffrey S Racine. 2004. "Cross-Validated Local Linear Nonparametric Regression." *Statistica Sinica* 14 (2): 485–512.

———. 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton, NJ, US: Princeton University Press.

Lin, Shin-Jong. 2011. "Estimating the impact of diabetes on employment in Taiwan." *Economics Bulletin* 31 (4): 3089–3102.

Long, Amanda N., and Samuel Dagogo-Jack. 2011. "Comorbidities of Diabetes and Hypertension: Mechanisms and Approach to Target Organ Protection." *Journal of Clinical Hypertension* 13 (4): 244–51. https://doi.org/10.1111/j.1751-7176.2011.00434.x.

Maier, W, R Holle, M Hunger, A Peters, C Meisinger, K H Greiser, A Kluttig, et al. 2013. "The impact of regional deprivation and individual socio-economic status on the prevalence of Type 2 diabetes in Germany. A pooled analysis of five population-based studies." *Diabetic Medicine* 30 (3): e78–86. https://doi.org/10.1111/dme.12062.

Malik, Vasanti S., Barry M. Popkin, George A. Bray, Jean Pierre Després, Walter C. Willett, and Frank B. Hu. 2010. "Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: A meta-analysis." *Diabetes Care* 33 (11): 2477–83. https://doi.org/10.2337/dc10-1079.

Minor, Travis. 2011. "The Effect of Diabetes on Female Labor Force Decisions: New Evidence from the National Health Interview Survey." *Health Economics* 20 (12): 1468–86. https://doi.org/doi.org/10.1002/hec.1685.

Modini, Matthew, Sadhbh Joyce, Arnstein Mykletun, Helen Christensen, Richard A Bryant, Philip B Mitchell, and Samuel B Harvey. 2016. "The Mental Health Benefits of Employment: Results of a Systematic Meta-Review." *Australasian Psychiatry* 24 (4): 331–36. https://doi.org/10.1177/1039856215618523.

Mutyambizi, Chipo, Milena Pavlova, Lumbwe Chola, Charles Hongoro, and Wim Groot. 2018. "Cost of diabetes mellitus in Africa: a systematic review of existing literature." *Globalization and Health* 14 (1). https://doi.org/10.1186/s12992-017-0318-5.

National Planning Commssion. 2012. "National Development Plan." Pretoria: Department: The Presidency.

Neumann, Todd C. 2013. "The Effect of Drinking and Smoking on the Labour Market Outcomes of Low-Income Young Adults." *Applied Economics* 45 (5): 541–53. https://doi.org/10.1080/00036846.2011.579063.

Noordt, Maaike van der, Helma IJzelenberg, Mariël Droomers, and Karin I Proper. 2014. "Health Effects of Employment: A Systematic Review of Prospective Studies." *Occupational and Environmental Medicine* 71 (10): 730–36. https://doi.org/10.1136/oemed-2013-101891.

Norström, Fredrik, Anna-Karin Waenerlund, Lars Lindholm, Rebecka Nygren, Klas-Göran Sahlén, and Anna Brydsten. 2019. "Does Unemployment Contribute to Poorer Health-Related Quality of Life Among Swedish Adults?" *BMC Public Health* 19 (457). https://doi.org/10.1186/s12889-019-6825-y.

Nowakowska, Magdalena, Salwa S. Zghebi, Darren M. Ashcroft, Iain Buchan, Carolyn Chew-Graham, Tim Holt, Christian Mallen, et al. 2019. "The Comorbidity Burden of Type 2 Diabetes Mellitus: Patterns, Clusters and Predictions from a Large English Primary Care Cohort." *BMC Medicine* 17 (145). https://doi.org/10.1186/s12916-019-1373-y.

Ntuli, Miracle. 2007. "Determinants of South African Women'Labour Force Participation, 1995-2004." PhD thesis, University of Cape.

Nwosu, Chijioke O., and Ingrid Woolard. 2017. "The Impact of Health on Labour Force Participation in South Africa." *South African Journal of Economics* 85 (4): 481–90. https://doi.org/10.1111/saje12163.

Pedron, Sara, Karl Emmert-Fees, Michael Laxy, and Lars Schwettmann. 2019. "The impact of diabetes on labour market participation: a systematic review of results and methods." *BMC Public Health* 19 (25): 25. https://doi.org/10.1186/s12889-018-6324-6.

Peer, Nasheeta, Andre-Pascal Kengne, Ayesha A Motala, and Jean Claude Mbanya. 2014. "Diabetes in the Africa region: An update." *Diabetes Research and Clinical Practice* 103 (2): 197–205. https://doi.org/10.1016/j.diabres.2013.11.006.

Pharr, Jennifer R., Sheniz Moonie, and Timothy J. Bungum. 2012. "The Impact of Unemployment on Mental and Physical Health, Access to Health Care and Health Risk Behaviors." *International Scholarly Research Notices Public Health* 2012 (483432): 7. https://doi.org/10.5402/2012/483432.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Schaller, Jessamyn, and Ann HuffStevens. 2019. "Short-Run Effects of Job Loss on Health Conditions, Health Insurance, and Health Care Utilization." *Journal of Health Economics* 43 (1): 190–203. https://doi.org/10.1016/j.jhealeco.2015.07.003.

Seuring, Till, Olga Archangelidi, and Marc Suhrcke. 2015. "The Economic Costs of Type 2 Diabetes: A Global Systematic Review." *Pharmacoeconomics* 33 (8): 811–31. https://doi.org/10.1007/s40273-015-0268-9.

Seuring, Till, Yevgeniy Goryakin, and Marc Suhrcke. 2015. "The impact of diabetes on employment in Mexico." *Economics & Human Biology* 18: 85–100. https://doi.org/10.1016/j.ehb.2015.04.002.

Seuring, Till, Pieter Serneels, and Marc Suhrcke. 2019. "The Impact of Diabetes on Labour Market Outcomes in Mexico: A Panel Data and Biomarker Analysis." *Social Science & Medicine* 233: 252–61. https://doi.org/https://doi.org/10.1016/j.socscimed.2019.05.051.

Smith, J P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status." *Journal of Economic Perspectives* 13 (2): 144–66.

Some, M., N. Rashied, and A. Ohonba. 2016. "The Impact of Obesity on Employment in South Africa." *Studies in Economics and Econometrics* 40 (2): 87–104. https://doi.org/10.1080/10800379.2016.12097299.

Statistics South Africa. 2014. "Labour market dynamics in South Africa, 2014." 02-11-02(2014). Pretoria: Statistics South Africa.

———. 2018. "Mortality and causes of death in South Africa, 2016: Findings from death notification." Pretoria: Statistics South Africa. http://www.statssa.gov.za/publications/P03093/P030932016.pdf.

———. 2019a. "General Household Survey 2018 [Dataset]." Pretoria: Statistics South Africa [producer], 2019. Cape Town: DataFirst [distributor], 2019. doi.org/10.25828/9tmn-fz97.

———. 2019b. "General Household Survey 2018: Statistical Release P0318." Pretoria, South Africa: Statistics South Africa.

Tanaka, T, E Gjonca, and M C Gulliford. 2011. "Income, wealth and risk of diabetes among older adults: cohort study using the English longitudinal study of ageing." *The European Journal of Public Health* 22 (3): 310–17. https://doi.org/10.1093/eurpub/ckr050.

Troske, Kenneth R, and Alexandru Voicu. 2009. "The Effect of Children on the Level of Labor Market Involvement of Married Women: What Is the Role of Education?" Germany: IZA Discussion Paper No. 4074.

Unmuessig, Victoria, Paul A. Fishman, Hubertus J. M. Vrijhoef, Arianne M. J. Elissen, and David C. Grossman. 2016. "Association of Controlled and Uncontrolled Hypertension with Workplace Productivity." *The Journal of Clinical Hypertension* 18 (3): 217–22. https://doi.org/https://doi.org/10.1111/jch.12648.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. Fourth. New York: Springer. http://www.stats.ox.ac.uk/pub/MASS4.

Wang, Min-Chiang, and John van Ryzin. 1981. "A Class of Smooth Estimators for Discrete Distributions." *Biometrika* 68 (1): 301–9. https://doi.org/10.2307/2335831.

Wilson, S H, and G M Walker. 1993. "Unemployment and Health: A Review." *Public Health* 107 (3): 153–62. https://doi.org/10.1016/S0033-3506(05)80436-6.

World Health Organization. 2016. "Global report on diabetes." France: World Health Organization.

Young, Alwyn. 2005. "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generation." *Quarterly Journal of Economics* 120 (2): 423–66. https://doi.org/10.1093/qje/120.2.424.

Zavras, Dimitris, Athanasios I. Zavras, Ilias-Ioannis Kyriopoulos, and John Kyriopoulos. 2019. "Economic Crisis, Austerity and Unmet Healthcare Needs: The Case of Greece." *BMC Health Services Research* 16 (309). https://doi.org/10.1186/s12913-016-1557-5.

Zeileis, Achim. 2004. "Econometric Computing with HC and HAC Covariance Matrix Estimators." *Journal of Statistical Software* 11 (10): 1–17. https://doi.org/10.18637/jss.v011.i10.

———. 2006. "Object-Oriented Computation of Sandwich Estimators." *Journal of Statistical Software* 16 (9): 1–16. https://doi.org/10.18637/jss.v016.i09.

Zhang, Ping, Xinzhi Zhang, Jonathan Brown, Dorte Vistisen, Richard Sicree, Jonathan Shaw, and Gregory Nichols. 2010. "Global Healthcare Expenditure on Diabetes for 2010 and 2030." *Diabetes Research and Clinical Practice* 87: 293–301. https://doi.org/10.1016/j.diabres.2010.01.026.

Zhang, Xiaohui, Xueyan Zhao, and Anthony Harris. 2009. "Chronic diseases and labour force participation in Australia." *Journal of Health Economics* 28 (1): 91–108. https://doi.org/10.1016/j.jhealeco.2008.08.001.

A Full Empirical Results

In this appendix, we present the full set of empirical results, beginning with the linear probability model in Table A.1 and A.2. These linear probability models assume that diabetes is exogenous. In the first of these, we present three sets of results, comparing (i) employment to unemployment, (ii) employment to non-participation and (iii) unemployment to non-participation. The second of these considers (i) employed or not (ii) unemployed or not and (iii) active in the labour market or not.

Table A.1: Linear probability model marginal effect of diabetes on employment, unemployment and non-participation

	Emp(=1) v Unemp(=0)		Emp(=1) v NLFP(=0)		Unemp(=1) v NLFP(=0)	
	Males	Females	Males	Females	Males	Females
Intercept	0.4831^{a}	0.0537	0.1411^{d}	-0.8098^{a}	0.0758	-0.1008
	(0.081)	(0.094)	(0.078)	(0.087)	(0.152)	(0.108)
Diabetes	0.0165	0.0593^{c}	-0.0498^{c}	-0.0467^{c}	-0.0837^{d}	-0.0848^{a}
	(0.024)	(0.025)	(0.022)	(0.020)	(0.048)	(0.026)
Age	0.0271^{a}	0.0346^{a}	0.0447^{a}	0.0696^{a}	0.0291^{a}	0.0218^{a}
	(0.003)	(0.004)	(0.003)	(0.004)	(0.006)	(0.004)
Age Sq	-0.0003^{a}	-0.0003^{a}	-0.0006^{a}	-0.0008^{a}	-0.0005^{a}	-0.0004^{a}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mixed Ethnic	-0.1826^{a}	-0.0105	-0.1956^{a}	0.0827^{a}	-0.0559^{a}	0.0934^{a}
	(0.008)	(0.009)	(0.008)	(0.008)	(0.018)	(0.011)
Asian	-0.1243^{a}	0.0351	-0.0740^{b}	0.1270^{a}	0.0670	0.0915^{b}
	(0.026)	(0.027)	(0.027)	(0.027)	(0.054)	(0.035)
White	-0.0735^{a}	-0.0128	0.0552^{b}	0.0508^{c}	0.1472^{a}	0.0296
	(0.026)	(0.029)	(0.021)	(0.022)	(0.038)	(0.024)
Married Apart	-0.0812^{a}	-0.0419	0.0827^{a}	0.0953^{a}	0.1773^{a}	0.0751^{b}
	(0.028)	(0.032)	(0.024)	(0.025)	(0.043)	(0.028)
Single	-0.1128^{a}	-0.0701^{c}	0.0604^{b}	0.0905^{a}	0.1915^{a}	0.0976^{a}
	(0.027)	(0.030)	(0.023)	(0.024)	(0.040)	(0.027)
Some Schooling	-0.0900^{a}	-0.0660^{c}	0.1164^{a}	0.1287^{a}	0.2413^{a}	0.1190^{a}
	(0.026)	(0.030)	(0.023)	(0.023)	(0.041)	(0.026)
Grade 9	-0.0687^{b}	-0.0217	0.1464^{a}	0.2298^{a}	0.2742^{a}	0.1812^{a}

	(0.026)	(0.029)	(0.022)	(0.022)	(0.039)	(0.025)
Grade 10	-0.1023^{a}	0.0423	0.1381^{a}	0.2769^{a}	0.3359^{a}	0.1366^{a}
	(0.033)	(0.037)	(0.030)	(0.033)	(0.061)	(0.042)
Grade 11	-0.0510^{d}	0.0550^{d}	0.1618^{a}	0.3589^{a}	0.2681^{a}	0.2019^{a}
	(0.030)	(0.033)	(0.027)	(0.028)	(0.063)	(0.040)
Grade 12	-0.0277	0.0751^{c}	0.1694^{a}	0.4050^{a}	0.2147^{a}	0.2102^{a}
	(0.030)	(0.033)	(0.026)	(0.028)	(0.067)	(0.047)
Uni Year 1	-0.0601^{d}	0.1086^{b}	0.1705^{a}	0.4153^{a}	0.2542^{c}	-0.0112
	(0.036)	(0.040)	(0.033)	(0.036)	(0.107)	(0.089)
Uni Year 2	-0.0152	0.1080^{c}	0.2185^{a}	0.4033^{a}	0.4763^{b}	0.0970
	(0.042)	(0.045)	(0.040)	(0.043)	(0.177)	(0.106)
Bachelor Degree	-0.0968	0.0419	0.2018^{b}	0.5266^{a}	0.4096	0.7054^{d}
	(0.074)	(0.080)	(0.075)	(0.087)	(0.266)	(0.426)
Honours Deg	-0.0036	0.0412^{c}	-0.0378^{b}	-0.0721^{a}	-0.0894^{a}	-0.1131^{a}
	(0.014)	(0.016)	(0.014)	(0.016)	(0.030)	(0.021)
M Degree	0.0422^{d}	0.0326	-0.0442^{d}	-0.2525^{a}	-0.1961^{a}	-0.2433^{a}
	(0.023)	(0.034)	(0.023)	(0.029)	(0.058)	(0.037)
Doctorate	0.0461^{a}	0.0546^{a}	-0.0220	-0.0702^{a}	-0.2868^{a}	-0.2197^{a}
	(0.015)	(0.018)	(0.015)	(0.017)	(0.052)	(0.030)
Rural	-0.0075	-0.0130	-0.0816^{a}	-0.0913^{a}	-0.1078^{a}	-0.0867^{a}
	(0.009)	(0.011)	(0.009)	(0.011)	(0.018)	(0.013)
Eastern Cape	-0.0547^{a}	-0.0107	-0.1114^{a}	-0.1248^{a}	-0.0524	-0.1080^{a}
	(0.016)	(0.018)	(0.016)	(0.018)	(0.035)	(0.024)
Northern Cape	-0.0814^{a}	-0.0275	-0.0344^{d}	-0.0589^{b}	0.0790^{c}	-0.0244
	(0.019)	(0.022)	(0.019)	(0.022)	(0.040)	(0.028)
Free State	-0.0789^{a}	-0.0672^{a}	-0.0448^{c}	-0.1007^{a}	0.0443	-0.0383
	(0.019)	(0.021)	(0.019)	(0.022)	(0.042)	(0.028)
KwaZulu-Natal	-0.0493^{a}	-0.0259	-0.0496^{a}	-0.0649^{a}	0.0272	-0.0140
	(0.015)	(0.018)	(0.015)	(0.018)	(0.035)	(0.024)
North West	-0.0373^{c}	-0.0252	-0.0086	-0.1515^{a}	0.0257	-0.1177^{a}
	(0.019)	(0.023)	(0.019)	(0.022)	(0.041)	(0.028)
Gauteng	-0.0591^{a}	-0.0507^{a}	-0.0105	-0.0376^{c}	0.1007^{a}	0.0386^{d}

	(0.014)	(0.016)	(0.014)	(0.016)	(0.034)	(0.023)
Mpumalanga	-0.0829^a	-0.1016^a	0.0300^d	-0.0381^d	0.1918^a	0.0839^a
	(0.017)	(0.020)	(0.018)	(0.021)	(0.041)	(0.027)
Limpopo	0.0429^c	0.0542^b	-0.0504^a	-0.1259^a	-0.1689^a	-0.1795^a
	(0.018)	(0.021)	(0.017)	(0.020)	(0.040)	(0.027)
Kids under 5	0.0542	0.0220	-0.0078	0.0645	-0.1781^d	0.0542
	(0.039)	(0.045)	(0.039)	(0.043)	(0.098)	(0.065)
Kids under 18	0.0322	-0.0062	-0.0203	0.0385	-0.1400	0.0982^d
	(0.037)	(0.043)	(0.037)	(0.040)	(0.087)	(0.059)
Very comfortable	0.0050	-0.0582	-0.0257	0.0226	-0.0857	0.1366^c
	(0.037)	(0.042)	(0.037)	(0.039)	(0.086)	(0.057)
Reasonable comfort	-0.1212^a	-0.1894^a	-0.1058^a	-0.0596	-0.0309	0.1730^a
	(0.037)	(0.043)	(0.037)	(0.040)	(0.086)	(0.058)
Getting by	-0.2381^a	-0.2824^a	-0.1932^a	-0.1138^b	-0.0205	0.2026^a
	(0.039)	(0.046)	(0.039)	(0.043)	(0.088)	(0.060)
Poor	0.0104	-0.0009	0.0057	-0.0442^a	0.0094	-0.0389^a
	(0.007)	(0.007)	(0.007)	(0.007)	(0.013)	(0.007)
Very poor	-0.0412^a	-0.0303^a	-0.0466^a	-0.0263^a	-0.0165^b	0.0049
	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.004)
Hypertension	0.0031	-0.0035	-0.0535^a	-0.0504^a	-0.0885^a	-0.0123
	(0.016)	(0.014)	(0.015)	(0.013)	(0.031)	(0.016)

Full set of parameter estimates from uncorrected linear probabality models estimated separately for men and women. Each set of results examines subsets of the sample, where the dependent variable is binary, with values and lables listed in the column headers. Thus, positive estimates refer to a positive effect on the xxx(=1) outcome. Results assume exogeneity of diabetes: Emp is employed, Unemp is unemployed and NLFP is non-participant in the labour force. Significance levels, based on heteroscedasticity corrected standard errors: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1.

Table A.2: Linear probability model marginal effect of diabetes on employment, unemployment and non-participation

	Emp(=1) v Not Emp(=0)	LFP(=1) v NLFP(=0)	Unemp(=1) v Not Unemp(=0)

	Males	Females	Males	Females	Males	Females
Intercept	-0.0577	-0.8880^a	0.2800^a	-0.4747^a	0.3378^a	0.4132^a
	(0.083)	(0.083)	(0.069)	(0.079)	(0.068)	(0.064)
Diabetes	-0.0361	-0.0266	-0.0507^c	-0.0680^a	-0.0146	-0.0414^b
	(0.025)	(0.020)	(0.021)	(0.019)	(0.020)	(0.016)
Age	0.0529^a	0.0685^a	0.0360^a	0.0555^a	-0.0169^a	-0.0131^a
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age Sq	-0.0007^a	-0.0008^a	-0.0005^a	-0.0007^a	0.0002^a	0.0001^c
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mixed Ethnic	-0.2617^a	0.0564^a	-0.1426^a	0.0862^a	0.1191^a	0.0298^a
	(0.009)	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)
Asian	-0.1459^a	0.1090^a	-0.0524^c	0.1177^a	0.0935^a	0.0087
	(0.028)	(0.026)	(0.024)	(0.025)	(0.023)	(0.020)
White	0.0072	0.0401^d	0.0830^a	0.0652^a	0.0757^a	0.0251
	(0.024)	(0.021)	(0.020)	(0.020)	(0.019)	(0.016)
Married Apart	0.0173	0.0688^b	0.1083^a	0.1185^a	0.0910^a	0.0497^b
	(0.026)	(0.025)	(0.022)	(0.023)	(0.022)	(0.019)
Single	-0.0198	0.0603^b	0.1000^a	0.1336^a	0.1198^a	0.0733^a
	(0.025)	(0.023)	(0.021)	(0.022)	(0.020)	(0.018)
Some Schooling	0.0323	0.0851^a	0.1431^a	0.1631^a	0.1107^a	0.0780^a
	(0.025)	(0.023)	(0.021)	(0.022)	(0.020)	(0.018)
Grade 9	0.0683^a	0.1684^a	0.1679^a	0.2411^a	0.0996^a	0.0727^a
	(0.024)	(0.022)	(0.020)	(0.021)	(0.020)	(0.017)
Grade 10	0.0395	0.2385^a	0.1679^a	0.2665^a	0.1284^a	0.0280
	(0.033)	(0.032)	(0.027)	(0.030)	(0.027)	(0.024)
Grade 11	0.0999^a	0.3253^a	0.1810^a	0.3426^a	0.0812^a	0.0173
	(0.030)	(0.027)	(0.025)	(0.026)	(0.025)	(0.021)
Grade 12	0.1273^a	0.3902^a	0.1849^a	0.3891^a	0.0575^c	-0.0011
	(0.029)	(0.027)	(0.024)	(0.026)	(0.024)	(0.021)
Uni Year 1	0.1136^a	0.4378^a	0.1955^a	0.3983^a	0.0819^b	-0.0396
	(0.037)	(0.037)	(0.031)	(0.035)	(0.031)	(0.028)
Uni Year 2	0.1844^a	0.4187^a	0.2340^a	0.3837^a	0.0496	-0.0350

	(0.045)	(0.044)	(0.037)	(0.042)	(0.037)	(0.034)
Bachelor Degree	0.1131	0.4986^{a}	0.2354^{a}	0.5246^{a}	0.1223^{d}	0.0261
	(0.083)	(0.088)	(0.069)	(0.084)	(0.069)	(0.068)
Honours Deg	-0.0313^{c}	-0.0326^{c}	-0.0378^{a}	-0.0756^{a}	-0.0065	-0.0430^{a}
	(0.015)	(0.015)	(0.013)	(0.014)	(0.013)	(0.012)
M Degree	0.0019	-0.1860^{a}	-0.0547^{b}	-0.2605^{a}	-0.0567^{b}	-0.0745^{a}
	(0.025)	(0.029)	(0.021)	(0.027)	(0.021)	(0.022)
Doctorate	0.0314^{d}	-0.0081	-0.0227	-0.0735^{a}	-0.0541^{a}	-0.0655^{a}
	(0.017)	(0.017)	(0.014)	(0.017)	(0.014)	(0.013)
Rural	-0.0589^{a}	-0.0693^{a}	-0.0726^{a}	-0.0918^{a}	-0.0137^{d}	-0.0225^{a}
	(0.010)	(0.010)	(0.008)	(0.009)	(0.008)	(0.008)
Eastern Cape	-0.1153^{a}	-0.0965^{a}	-0.0938^{a}	-0.1199^{a}	0.0216	-0.0234^{d}
	(0.017)	(0.017)	(0.014)	(0.016)	(0.014)	(0.013)
Northern Cape	-0.0812^{a}	-0.0602^{a}	-0.0143	-0.0432^{c}	0.0669^{a}	0.0170
	(0.021)	(0.021)	(0.017)	(0.020)	(0.017)	(0.016)
Free State	-0.0903^{a}	-0.1033^{a}	-0.0259	-0.0760^{a}	0.0643^{a}	0.0274^{d}
	(0.021)	(0.021)	(0.017)	(0.019)	(0.017)	(0.016)
KwaZulu-Natal	-0.0678^{a}	-0.0594^{a}	-0.0298^{c}	-0.0444^{b}	0.0379^{b}	0.0150
	(0.017)	(0.017)	(0.014)	(0.016)	(0.014)	(0.013)
North West	-0.0316	-0.1165^{a}	-0.0043	-0.1361^{a}	0.0273^{d}	-0.0196
	(0.020)	(0.021)	(0.017)	(0.020)	(0.016)	(0.016)
Gauteng	-0.0482^{a}	-0.0586^{a}	0.0014	-0.0224	0.0496^{a}	0.0362^{a}
	(0.015)	(0.016)	(0.012)	(0.015)	(0.012)	(0.012)
Mpumalanga	-0.0376^{c}	-0.0750^{a}	0.0466^{a}	0.0033	0.0842^{a}	0.0783^{a}
	(0.019)	(0.019)	(0.016)	(0.019)	(0.016)	(0.015)
Limpopo	-0.0156	-0.0707^{a}	-0.0664^{a}	-0.1409^{a}	-0.0508^{a}	-0.0702^{a}
	(0.019)	(0.019)	(0.016)	(0.018)	(0.016)	(0.015)
Kids under 5	0.0333	0.0683	-0.0136	0.0655	-0.0469	-0.0028
	(0.043)	(0.043)	(0.036)	(0.041)	(0.035)	(0.033)
Kids under 18	0.0051	0.0226	-0.0247	0.0446	-0.0298	0.0220
	(0.040)	(0.040)	(0.033)	(0.038)	(0.033)	(0.031)
Very comfortable	-0.0228	-0.0147	-0.0253	0.0431	-0.0025	0.0578^{d}

	(0.040)	(0.039)	(0.033)	(0.038)	(0.033)	(0.030)
Reasonable comfort	-0.1537^a	-0.1206^a	-0.0641^d	0.0048	0.0896^b	0.1254^a
	(0.040)	(0.040)	(0.034)	(0.038)	(0.033)	(0.031)
Getting by	-0.2727^a	-0.1798^a	-0.1138^a	-0.0134	0.1588^a	0.1664^a
	(0.042)	(0.042)	(0.035)	(0.040)	(0.035)	(0.033)
Poor	0.0110	-0.0297^a	0.0058	-0.0433^a	-0.0051	-0.0135^a
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Very poor	-0.0538^a	-0.0293^a	-0.0325^a	-0.0152^a	0.0213^a	0.0140^a
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Hypertension	-0.0393^c	-0.0440^a	-0.0513^a	-0.0475^a	-0.0121	-0.0035
	(0.016)	(0.012)	(0.014)	(0.012)	(0.013)	(0.010)

Full set of parameter estimates from uncorrected linear probability models estimated separately for men and women. Each set of results examines subsets of the sample, where the dependent variable is binary, with values and lables listed in the column headers. Thus, positive estimates refer to a positive effect on the xxx(=1) outcome. Results assume exogeneity of diabetes: Emp is employed, Not Emp is not employed (i.e., unemployed or NLFP), Unemp is unemployed, Not Unemp is not unemployed, LFP is labour force participant and NLFP is non-participant in the labour force. Significance levels, based on heteroscedasticity corrected standard errors: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1.

The second set of linear probability model results relaxes the exogeneity assumption. The full set of results for the heteroscedastic-IV models (Lewbel 2012, 2018), are presented in Tables A.3 and A.4. These tables also contain the same weak instrument test results, along with Wu-Hausman and Sargan test results. As with the first two appendix tables, the first of these compares (i) employment to unemployment, (ii) employment to non-participation and (iii) unemployment to non-participation, while the second considers (i) employed or not (ii) unemployed or not and (iii) active in the labour market or not.

Table A.3: Linear probability model marginal effect of diabetes on employment, unemployment and non-participation

	Emp(=1) v Unemp(=0)		Emp(=1) v NLFP(=0)		Unemp(=1) v NLFP(=0)	
	Males	Females	Males	Females	Males	Females
Intercept	0.4745^a	0.0428	0.0638	-0.8990^a	-0.0254	-0.1881^d
	(0.083)	(0.090)	(0.083)	(0.088)	(0.154)	(0.105)

Diabetes	0.0508	0.0270	-0.1294[a]	-0.0744	-0.1355[a]	-0.0787[c]
	(0.031)	(0.040)	(0.041)	(0.055)	(0.046)	(0.036)
Age	0.0272[a]	0.0344[a]	0.0443[a]	0.0694[a]	0.0288[a]	0.0219[a]
	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)	(0.004)
Age Sq	-0.0003[a]	-0.0003[a]	-0.0006[a]	-0.0008[a]	-0.0005[a]	-0.0004[a]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mixed Ethnic	-0.0044	0.0416[b]	-0.0352[c]	-0.0719[a]	-0.0881[a]	-0.1130[a]
	(0.014)	(0.015)	(0.014)	(0.016)	(0.030)	(0.021)
Asian	0.0417[c]	0.0333	-0.0419[c]	-0.2514[a]	-0.1928[a]	-0.2435[a]
	(0.019)	(0.028)	(0.020)	(0.028)	(0.056)	(0.029)
White	0.0466[a]	0.0542[a]	-0.0233[c]	-0.0706[a]	-0.2875[a]	-0.2195[a]
	(0.009)	(0.012)	(0.012)	(0.016)	(0.043)	(0.024)
Married Apart	-0.1825[a]	-0.0108	-0.1964[a]	0.0825[a]	-0.0569[a]	0.0934[a]
	(0.008)	(0.008)	(0.008)	(0.008)	(0.018)	(0.010)
Single	-0.1237[a]	0.0346	-0.0756[a]	0.1272[a]	0.0666	0.0913[c]
	(0.029)	(0.027)	(0.025)	(0.028)	(0.053)	(0.036)
Some Schooling	-0.0736[a]	-0.0127	0.0558[c]	0.0513[c]	0.1478[a]	0.0294
	(0.024)	(0.030)	(0.026)	(0.022)	(0.032)	(0.019)
Grade 9	-0.0818[a]	-0.0412	0.0841[a]	0.0958[a]	0.1772[a]	0.0749[a]
	(0.027)	(0.033)	(0.029)	(0.026)	(0.039)	(0.025)
Grade 10	-0.1131[a]	-0.0697[c]	0.0616[c]	0.0909[a]	0.1920[a]	0.0974[a]
	(0.025)	(0.031)	(0.027)	(0.025)	(0.036)	(0.023)
Grade 11	-0.0904[a]	-0.0655[c]	0.1176[a]	0.1293[a]	0.2419[a]	0.1189[a]
	(0.025)	(0.031)	(0.027)	(0.024)	(0.036)	(0.023)
Grade 12	-0.0690[a]	-0.0210	0.1473[a]	0.2305[a]	0.2743[a]	0.1810[a]
	(0.023)	(0.030)	(0.026)	(0.023)	(0.034)	(0.022)
Uni Year 1	-0.1028[a]	0.0427	0.1400[a]	0.2774[a]	0.3366[a]	0.1364[a]
	(0.031)	(0.037)	(0.032)	(0.033)	(0.058)	(0.043)
Uni Year 2	-0.0521[c]	0.0560[d]	0.1653[a]	0.3598[a]	0.2711[a]	0.2017[a]
	(0.026)	(0.032)	(0.029)	(0.027)	(0.057)	(0.040)
Bachelor Degree	-0.0280	0.0758[c]	0.1702[a]	0.4055[a]	0.2147[a]	0.2101[a]
	(0.025)	(0.031)	(0.028)	(0.025)	(0.057)	(0.045)

Honours Deg	-0.0609^{c}	0.1089^{a}	0.1729^{a}	0.4160^{a}	0.2581^{c}	-0.0121
	(0.027)	(0.031)	(0.030)	(0.029)	(0.111)	(0.076)
M Degree	-0.0152	0.1078^{a}	0.2192^{a}	0.4030^{a}	0.4832^{a}	0.0972
	(0.029)	(0.034)	(0.028)	(0.034)	(0.156)	(0.118)
Doctorate	-0.0963^{c}	0.0446	0.2007^{a}	0.5289^{a}	0.4092	0.7052
	(0.046)	(0.045)	(0.041)	(0.037)	(0.439)	(20.667)
Rural	0.0074	0.0132	0.0821^{a}	0.0916^{a}	0.1084^{a}	0.0867^{a}
	(0.010)	(0.011)	(0.010)	(0.011)	(0.019)	(0.013)
Eastern Cape	-0.0544^{a}	-0.0110	-0.1114^{a}	-0.1250^{a}	-0.0533	-0.1078^{a}
	(0.015)	(0.017)	(0.016)	(0.017)	(0.035)	(0.024)
Northern Cape	-0.0817^{a}	-0.0281	-0.0339^{d}	-0.0591^{b}	0.0776^{d}	-0.0244
	(0.020)	(0.021)	(0.020)	(0.023)	(0.041)	(0.028)
Free State	-0.0787^{a}	-0.0679^{a}	-0.0445^{c}	-0.1014^{a}	0.0438	-0.0381
	(0.019)	(0.021)	(0.019)	(0.022)	(0.042)	(0.030)
KwaZulu-Natal	-0.0494^{a}	-0.0260	-0.0490^{a}	-0.0646^{a}	0.0266	-0.0140
	(0.014)	(0.017)	(0.015)	(0.017)	(0.036)	(0.025)
North West	-0.0369^{c}	-0.0257	-0.0087	-0.1520^{a}	0.0255	-0.1175^{a}
	(0.018)	(0.023)	(0.018)	(0.022)	(0.042)	(0.029)
Gauteng	-0.0590^{a}	-0.0513^{a}	-0.0104	-0.0383^{c}	0.0992^{a}	0.0389
	(0.012)	(0.015)	(0.012)	(0.015)	(0.034)	(0.024)
Mpumalanga	-0.0826^{a}	-0.1016^{a}	0.0295^{d}	-0.0384^{d}	0.1911^{a}	0.0841^{a}
	(0.016)	(0.020)	(0.015)	(0.020)	(0.041)	(0.028)
Limpopo	0.0430^{b}	0.0540^{a}	-0.0504^{a}	-0.1263^{a}	-0.1696^{a}	-0.1793^{a}
	(0.016)	(0.019)	(0.017)	(0.019)	(0.040)	(0.026)
Kids under 5	0.0541	0.0226	-0.0081	0.0653	-0.1783^{d}	0.0539
	(0.035)	(0.033)	(0.033)	(0.041)	(0.102)	(0.057)
Kids under 18	0.0317	-0.0053	-0.0197	0.0394	-0.1409	0.0980^{d}
	(0.034)	(0.032)	(0.031)	(0.039)	(0.094)	(0.051)
Very comfortable	0.0048	-0.0575^{d}	-0.0258	0.0232	-0.0869	0.1364^{b}
	(0.034)	(0.032)	(0.031)	(0.038)	(0.093)	(0.050)
Reasonable comfort	-0.1214^{a}	-0.1888^{a}	-0.1060^{a}	-0.0589	-0.0321	0.1729^{a}
	(0.036)	(0.034)	(0.032)	(0.039)	(0.093)	(0.051)

Getting by	-0.2380^a	-0.2817^a	-0.1938^a	-0.1130^b	-0.0221	0.2025^a
	(0.039)	(0.039)	(0.036)	(0.042)	(0.095)	(0.053)
Poor	0.0104	-0.0010	0.0055	-0.0442^a	0.0092	-0.0389^a
	(0.008)	(0.007)	(0.008)	(0.007)	(0.013)	(0.007)
Very poor	-0.0413^a	-0.0303^a	-0.0464^a	-0.0263^a	-0.0164^a	0.0049
	(0.004)	(0.004)	(0.004)	(0.003)	(0.006)	(0.004)
Hypertension	-0.0009	0.0001	-0.0431^b	-0.0465^a	-0.0818^a	-0.0131
	(0.013)	(0.014)	(0.017)	(0.015)	(0.028)	(0.014)
Weak Instruments	282.533^a	104.625^a	336.728^a	103.195^a	283.247^a	111.826^a
Wu-Hausman	1.100	0.376	8.558^a	0.326	1.665	0.017
Sargan	14.309	15.961	29.167^d	22.954	13.940	10.091

The dependent variable is binary, with values and lables listed in the column headers. Results allow for the endogeneity of diabetes using mean-centered values of wealth levels, ethnic group, marital status, location, education level, and number of children (young and old) for identification. Instrumental variable diagnostic tests include: a test of weak instruments, endogeneity (Wu-Hausman) and overidentification (Sargan). In each case, the test statistic is provided along with its significance, i.e., the probability that the null hypothesis is true, where the following notation and probability levels are assumed: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1. Emp is employed, Unemp is unemployed and NLFP is non-participant in the labour force. Heteroscedasticity corrected standard errors reported for all estimates.

Table A.4: Linear probability model marginal effect of diabetes on employment, unemployment and non-participation

	Emp(=1) v Not Emp(=0)		LFP(=1) v NLFP(=0)		Unemp(=1) v Not Unemp(=0)	
	Males	Females	Males	Females	Males	Females
Intercept	-0.1137	-0.9554^a	0.2126^a	-0.5649^a	0.3264^a	0.3905^a
	(0.086)	(0.082)	(0.074)	(0.080)	(0.069)	(0.061)
Diabetes	-0.0849^c	-0.0468	-0.1374^a	-0.0853^d	-0.0525^c	-0.0384^d
	(0.041)	(0.052)	(0.039)	(0.051)	(0.021)	(0.022)
Age	0.0527^a	0.0684^a	0.0356^a	0.0553^a	-0.0171^a	-0.0131^a
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)

Age Sq	-0.0006^{a}	-0.0008^{a}	-0.0005^{a}	-0.0007^{a}	0.0002^{a}	0.0001^{b}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mixed Ethnic	-0.0299^{d}	-0.0325^{c}	-0.0354^{b}	-0.0755^{a}	-0.0055	-0.0431^{a}
	(0.015)	(0.015)	(0.013)	(0.014)	(0.012)	(0.011)
Asian	0.0033	-0.1852^{a}	-0.0524^{b}	-0.2598^{a}	-0.0557^{a}	-0.0746^{a}
	(0.023)	(0.028)	(0.019)	(0.027)	(0.016)	(0.016)
White	0.0306^{c}	-0.0084	-0.0241^{c}	-0.0738^{a}	-0.0547^{a}	-0.0654^{a}
	(0.013)	(0.016)	(0.011)	(0.015)	(0.008)	(0.009)
Married Apart	-0.2621^{a}	0.0562^{a}	-0.1434^{a}	0.0861^{a}	0.1187^{a}	0.0298^{a}
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)
Single	-0.1467^{a}	0.1091^{a}	-0.0538^{b}	0.1178^{a}	0.0929^{a}	0.0087
	(0.029)	(0.026)	(0.021)	(0.025)	(0.026)	(0.020)
Some Schooling	0.0076	0.0404^{d}	0.0836^{a}	0.0655^{a}	0.0760^{a}	0.0250^{d}
	(0.025)	(0.021)	(0.025)	(0.022)	(0.016)	(0.013)
Grade 9	0.0180	0.0693^{b}	0.1096^{a}	0.1189^{a}	0.0915^{a}	0.0497^{a}
	(0.028)	(0.025)	(0.027)	(0.025)	(0.019)	(0.017)
Grade 10	-0.0192	0.0607^{b}	0.1011^{a}	0.1339^{a}	0.1202^{a}	0.0732^{a}
	(0.026)	(0.023)	(0.026)	(0.024)	(0.018)	(0.015)
Grade 11	0.0330	0.0855^{a}	0.1442^{a}	0.1635^{a}	0.1112^{a}	0.0780^{a}
	(0.026)	(0.023)	(0.025)	(0.023)	(0.018)	(0.015)
Grade 12	0.0688^{b}	0.1689^{a}	0.1687^{a}	0.2415^{a}	0.0999^{a}	0.0726^{a}
	(0.025)	(0.022)	(0.025)	(0.022)	(0.016)	(0.014)
Uni Year 1	0.0405	0.2389^{a}	0.1696^{a}	0.2668^{a}	0.1291^{a}	0.0279
	(0.033)	(0.032)	(0.029)	(0.030)	(0.025)	(0.023)
Uni Year 2	0.1018^{a}	0.3260^{a}	0.1845^{a}	0.3432^{a}	0.0827^{a}	0.0172
	(0.029)	(0.026)	(0.027)	(0.025)	(0.020)	(0.018)
Bachelor Degree	0.1278^{a}	0.3906^{a}	0.1857^{a}	0.3895^{a}	0.0579^{a}	-0.0011
	(0.028)	(0.025)	(0.026)	(0.024)	(0.018)	(0.016)
Honours Deg	0.1150^{a}	0.4383^{a}	0.1979^{a}	0.3987^{a}	0.0829^{a}	-0.0397^{c}
	(0.031)	(0.029)	(0.029)	(0.029)	(0.021)	(0.016)
M Degree	0.1848^{a}	0.4185^{a}	0.2347^{a}	0.3835^{a}	0.0499^{c}	-0.0350^{d}
	(0.031)	(0.034)	(0.027)	(0.033)	(0.024)	(0.021)

Doctorate	0.1125^{c}	0.5002^{a}	0.2343^{a}	0.5261^{a}	0.1218^{b}	0.0258
	(0.048)	(0.044)	(0.043)	(0.035)	(0.047)	(0.036)
Rural	0.0592^{a}	0.0695^{a}	0.0731^{a}	0.0919^{a}	0.0139^{d}	0.0224^{a}
	(0.010)	(0.010)	(0.009)	(0.010)	(0.008)	(0.008)
Eastern Cape	-0.1156^{a}	-0.0967^{a}	-0.0943^{a}	-0.1201^{a}	0.0213^{d}	-0.0234^{d}
	(0.017)	(0.017)	(0.014)	(0.016)	(0.013)	(0.012)
Northern Cape	-0.0811^{a}	-0.0604^{b}	-0.0141	-0.0434^{c}	0.0670^{a}	0.0170
	(0.021)	(0.022)	(0.017)	(0.020)	(0.017)	(0.015)
Free State	-0.0903^{a}	-0.1038^{a}	-0.0261	-0.0764^{a}	0.0643^{a}	0.0274^{d}
	(0.021)	(0.021)	(0.017)	(0.019)	(0.017)	(0.016)
KwaZulu-Natal	-0.0677^{a}	-0.0593^{a}	-0.0296^{c}	-0.0444^{a}	0.0380^{a}	0.0150
	(0.016)	(0.017)	(0.013)	(0.015)	(0.013)	(0.013)
North West	-0.0318	-0.1169^{a}	-0.0047	-0.1364^{a}	0.0271^{d}	-0.0195
	(0.019)	(0.021)	(0.016)	(0.020)	(0.016)	(0.016)
Gauteng	-0.0484^{a}	-0.0591^{a}	0.0010	-0.0228^{d}	0.0494^{a}	0.0363^{a}
	(0.014)	(0.015)	(0.011)	(0.014)	(0.011)	(0.012)
Mpumalanga	-0.0379^{c}	-0.0752^{a}	0.0460^{a}	0.0031	0.0839^{a}	0.0783^{a}
	(0.018)	(0.019)	(0.014)	(0.018)	(0.015)	(0.016)
Limpopo	-0.0158	-0.0710^{a}	-0.0667^{a}	-0.1412^{a}	-0.0509^{a}	-0.0702^{a}
	(0.018)	(0.019)	(0.016)	(0.018)	(0.013)	(0.013)
Kids under 5	0.0332	0.0688^{d}	-0.0138	0.0659^{d}	-0.0469	-0.0029
	(0.038)	(0.040)	(0.030)	(0.039)	(0.031)	(0.023)
Kids under 18	0.0055	0.0232	-0.0240	0.0451	-0.0295	0.0220
	(0.036)	(0.038)	(0.029)	(0.037)	(0.031)	(0.022)
Very comfortable	-0.0228	-0.0143	-0.0253	0.0435	-0.0025	0.0577^{b}
	(0.036)	(0.038)	(0.028)	(0.037)	(0.031)	(0.022)
Reasonable comfort	-0.1538^{a}	-0.1202^{a}	-0.0642^{c}	0.0052	0.0895^{a}	0.1253^{a}
	(0.037)	(0.038)	(0.029)	(0.038)	(0.032)	(0.023)
Getting by	-0.2730^{a}	-0.1793^{a}	-0.1145^{a}	-0.0130	0.1585^{a}	0.1663^{a}
	(0.039)	(0.040)	(0.032)	(0.040)	(0.034)	(0.026)
Poor	0.0109	-0.0297^{a}	0.0056	-0.0433^{a}	-0.0052	-0.0135^{b}
	(0.008)	(0.006)	(0.007)	(0.006)	(0.007)	(0.005)

Very poor	-0.0537^a	-0.0293^a	-0.0324^a	-0.0152^a	0.0213^a	0.0140^a
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Hypertension	-0.0331^c	-0.0414^a	-0.0404^b	-0.0453^a	-0.0073	-0.0039
	(0.017)	(0.014)	(0.016)	(0.014)	(0.011)	(0.009)
Weak Instruments	390.264^a	136.833^a	390.264^a	136.833^a	390.264^a	136.833^a
Wu-Hausman	2.508	0.188	11.368^a	0.151	2.210	0.007
Sargan	26.772	22.929	32.423^c	26.200	10.185	10.100

Full set of parameter estimates from uncorrected linear probabality models estimated separately for men and women. Each set of results examines subsets of the sample, where the dependent variable is binary, with values and lables listed in the column headers. Thus, positive estimates refer to a positive effect on the xxx(=1) outcome. Results allow for the endogeneity of diabetes using mean-centered values of wealth levels, ethnic group, marital status, location, education level and number of children (young and old) for identification. Emp is employed, Not Emp is not employed (i.e., unemployed or NLFP), Unemp is unemployed, Not Unemp is not unemployed, LFP is labour force participant and NLFP is non-participant in the labour force. Significance levels, based on heteroscedasticity corrected standard errors: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1.

The final set empirical results are those for the mulitnomial logit.

Table A.5: Multinomial probability model parameter estimates for unemployment and non-participation, relative to employment.

	Probability Unemployed		Probability Not Active	
	Males	Females	Males	Females
Intercept	-0.0014	1.2503^a	2.7419^a	6.6444^a
	(0.002)	(0.001)	(0.003)	(0.001)
Diabetes	-0.1208^a	-0.7100^a	0.3732^a	0.2539^a
	(0.004)	(0.002)	(0.019)	(0.011)
Age	-0.1733^a	-0.1268^a	-0.3250^a	-0.3520^a
	(0.005)	(0.004)	(0.005)	(0.003)
Age Sq	0.0018^a	0.0008^a	0.0042^a	0.0042^a
	(0.000)	(0.000)	(0.000)	(0.000)
Mixed Ethnic	0.0430	-0.2296^a	0.3800^a	0.3602^a

	(0.051)	(0.030)	(0.057)	(0.048)
Asian	-0.5922^a	-0.1985^a	0.3822^a	1.3140^a
	(0.004)	(0.002)	(0.008)	(0.008)
White	-1.3974^a	-1.0284^a	0.1103^a	0.3872^a
	(0.002)	(0.002)	(0.010)	(0.012)
Married Apart	1.3587^a	0.0691	1.4990^a	-0.4456^a
	(0.061)	(0.050)	(0.061)	(0.040)
Single	1.0327^a	-0.1982^a	0.6919^a	-0.6824^a
	(0.008)	(0.004)	(0.005)	(0.005)
Some Schooling	0.6044^a	0.0199	-0.2790^a	-0.2557^a
	(0.059)	(0.058)	(0.053)	(0.043)
Grade 9	0.6458^a	0.1168^a	-0.3703^a	-0.4258^a
	(0.073)	(0.024)	(0.071)	(0.043)
Grade 10	0.9032^a	0.2706^a	-0.1913^a	-0.4389^a
	(0.063)	(0.055)	(0.062)	(0.047)
Grade 11	0.7124^a	0.1818^a	-0.5990^a	-0.5899^a
	(0.061)	(0.053)	(0.064)	(0.046)
Grade 12	0.5865^a	-0.0127	-0.8733^a	-1.0732^a
	(0.050)	(0.045)	(0.054)	(0.039)
Uni Year 1	0.8462^a	-0.4235^a	-0.8751^a	-1.3172^a
	(0.007)	(0.004)	(0.004)	(0.003)
Uni Year 2	0.3361^a	-0.6490^a	-1.1678^a	-1.8776^a
	(0.008)	(0.007)	(0.005)	(0.006)
Bachelor Degree	-0.1543^a	-1.0510^a	-1.2773^a	-2.3963^a
	(0.005)	(0.004)	(0.006)	(0.005)
Honours Deg	0.0658^a	-2.9705^a	-1.5033^a	-2.4686^a
	(0.002)	(0.000)	(0.002)	(0.001)
M Degree	-0.3789^a	-1.9197^a	-2.7817^a	-2.4205^a
	(0.001)	(0.000)	(0.000)	(0.001)
Doctorate	0.6334^a	-1.1957^a	-1.9028^a	-12.6187^a
	(0.000)	(0.000)	(0.000)	(0.000)
Rural	0.0581	0.0260	0.5113^a	0.4572^a

	(0.058)	(0.050)	(0.057)	(0.046)
Eastern Cape	0.5361[a]	0.0995[a]	0.7711[a]	0.6381[a]
	(0.075)	(0.035)	(0.067)	(0.045)
Northern Cape	0.7120[a]	0.2736[a]	0.3222[a]	0.3049[a]
	(0.034)	(0.014)	(0.026)	(0.020)
Free State	0.7401[a]	0.4362[a]	0.4312[a]	0.5283[a]
	(0.062)	(0.034)	(0.055)	(0.054)
KwaZulu-Natal	0.4913[a]	0.2428[a]	0.4067[a]	0.3162[a]
	(0.061)	(0.056)	(0.059)	(0.046)
North West	0.3671[a]	0.2077[a]	0.1622[b]	0.7673[a]
	(0.057)	(0.028)	(0.058)	(0.054)
Gauteng	0.5678[a]	0.3992[a]	0.0590	0.2023[a]
	(0.056)	(0.050)	(0.063)	(0.047)
Mpumalanga	0.6942[a]	0.6401[a]	-0.2080[a]	0.1997[a]
	(0.067)	(0.036)	(0.051)	(0.045)
Limpopo	-0.3192[a]	-0.3503[a]	0.4542[a]	0.6179[a]
	(0.041)	(0.025)	(0.065)	(0.049)
Kids under 5	-0.0288	0.0310	-0.0724	0.2282[a]
	(0.053)	(0.040)	(0.050)	(0.033)
Kids under 18	0.2443[a]	0.1550[a]	0.3040[a]	0.1293[a]
	(0.024)	(0.019)	(0.023)	(0.016)
Very comfortable	-0.6981[a]	-0.1553[a]	0.0770[a]	-0.3885[a]
	(0.004)	(0.003)	(0.009)	(0.007)
Reasonable comfort	-0.3044[a]	0.3704[a]	0.2626[a]	-0.2134[a]
	(0.061)	(0.049)	(0.057)	(0.039)
Getting by	0.0080	0.7228[a]	0.3133[a]	-0.1330[a]
	(0.043)	(0.037)	(0.043)	(0.032)
Poor	0.7820[a]	1.3597[a]	0.8184[a]	0.2675[a]
	(0.050)	(0.044)	(0.052)	(0.038)
Very poor	1.3989[a]	1.7589[a]	1.3416[a]	0.5222[a]
	(0.067)	(0.030)	(0.067)	(0.039)
Hypertension	-0.0660[b]	0.1336[a]	0.3945[a]	0.2462[a]

	(0.025)	(0.017)	(0.089)	(0.053)

Full set of parameter estimates from multinomial logit probabality model, with the probability of employment as the reference group; thus parameters are relative to this reference. Thus, positive estimates refer to an increased probability of unemployment/ non-participation over employment. Results assume exogeneity and standard errors are weighted to account for sample structure. Significance levels, based on corrected standard errors: ^{a} - 0.005, ^{b} - 0.01, ^{c} - 0.05, ^{d} - 0.1.