CAUSES AND CONSEQUENCES OF LABOR MARKET DECISIONS: AN EMPIRICAL INVESTIGATION

BY

Travis Joseph Minor

A Dissertation Submitted to the Graduate School at Middle Tennessee State University in Partial Fulfillment of the Requirement for the Degree

Doctor of Philosophy/Economics

Murfreesboro, TN

August 2009

APPROVAL PAGE

CAUSES AND CONSEQUENCES OF LABOR MARKET DECISIONS: AN EMPIRICAL INVESTIGATION

BY

Travis Joseph Minor

A Dissertation Submitted to the Graduate School at Middle Tennessee State University in Partial Fulfillment of the Requirement for the Degree Doctor of Philosophy/Economics

Murfreesboro, TN

August 2009

Approved by:

Mark F. Owens

Dr. Mark F. Owens, Committee Chair

Reuben Kyle

Dr. Reuben Kyle, Committee Member

Adam Rennhoff

Dr. Adam Rennhoff, Committee Member

Gregory Givens

Dr. Gregory Givens, Graduate Director, Economics and Finance

Charles L. Baum

Dr. Charles Baum, Department Chair, Economics and Finance

Michael D. Allen

Dr. Michael D. Allen, Dean, College of Graduate Studies

To my family

I could not have done this without all of you.

ACKNOWLEDGEMENTS

I would like to thank all of my committee members: Mark F. Owens (chair), Adam Rennhoff, Charles L. Baum, and Reuben Kyle. I would also like to thank Brandeanna Allen, Gregory Givens, Adam Hogan, Pam Morris, John Nunley, Alan Seals, Rachel Wilson, and Joachim Zietz for their comments, suggestions, and editorial advice.

ABSTRACT

Collected in this dissertation are three separate works that examine several different factors in an individual's wage determination. Chapter 1 looks at the effect of diabetes on an individual's employment decision and wage rate. Estimates show the importance of a continuously-specified diabetes measure, as opposed to the static measure estimated by previous studies. Additionally, numerous sources of statistical bias are accounted for utilizing the panel data available for this study. Chapter 2 explores labor market similarities and differences of type-I and type-II diabetes. Results show that type-I diabetes is detrimental to most labor market outcomes, accounting for an average loss in earnings of about 17 %; and that the effects of type-II diabetes are similar, though not as large, with an average loss of 8 %. Chapter 3 takes a different approach by analyzing the importance of factors that influence a state's decision to adopt an above-federal minimum wage level. Results indicate that state political leanings are the primary significant factor in explaining differences in state minimum wage laws since 1991.

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

CHAPTER 1

AN INVESTIGATION INTO THE EFFECT OF DIABETES'S DURATION ON EMPLOYMENT AND WAGES: A PANEL DATA ANALYSIS

1.1 INTRODUCTION

According to the Center for Disease Control and Prevention (CDC), about 10.8 percent of Americans over the age of twenty have diabetes. This statistic has alarmed public health officials because: First, it shows a very high concentration of diabetics in the eligible labor force, which could be having a significant impact on the labor market. Second, this is the result of a five percent annual growth rate in the incidence of diabetes since 1990 (CDC, 2007), which indicates that any observed effect of diabetes could be amplified if this trend were to continue. In fact, the CDC estimates that the incidence of diabetics in America could double by 2050, and a more recent study by Wild, et al. (2004) suggests that this increase could happen as early as 2030. In either case, the growing incidence of diabetes within the eligible American labor force is of considerable concern to both potential employers and employees.

In addition to the large and growing prevalence of diabetes in the American population, diabetes and its associated problems may worsen as the patient ages. Fox et al. (2004), Nichols et al. (2001), and Ivers et al. (2001) find significant negative health effects due to diabetes duration. These range from increased risk of bone fracture to a heightened mortality risk over the course of the disease. Diabetes duration may also have a limited range of benefits. Donaghue et al. (2003) suggest that patients diagnosed young, although they have worse health outcomes than the general population, may be better able to manage their condition later in life. These studies indicate that the true effect of diabetes may be changing over the course of the disease.

While the economic literature has established an overall negative *average* effect on employment and wages, the incremental effects of an additional year of diabetes may differ drastically across the population. Estimates of diabetes in this study are not simply the average effect; instead, the effect of diabetes is allowed to change, conditional on the length of time a person has had the disease. Also, unlike previous studies that utilize a restricted sample of the population, the data used in this study comes from the National Longitudinal Survey of the Youth 1979 (NLSY79), which means the estimated effects should be current and representative. Finally, this study attempts to account for numerous sources of statistical bias not previously accounted for in the literature. Specifically, unobserved heterogeneity is taken into account in all estimations; additionally selection bias is considered specifically in the estimation of wages. Also, an attempt to control endogeneity bias is performed utilizing a respondent's sibling diabetes information to instrument their own diabetes.

The purpose of this study is to estimate the effect of diabetes, not only as a static disease, but as one which is allowed to change over time. Results suggest the significant negative effects on wages and employment estimated by previous studies are likel y derived from the most severe cases of diabetes, in which numerous other medical complications may be producing a significant effect not entirely attributable to diabetes alone. A continuous measure of diabetes duration more closely reflects the true impact of diabetes on an individual. Diabetes and diabetes duration are shown to have no significant impact on wages once the decision to work has been taken into account. This is somewhat in contrast to other studies which show a negative effect of diabetes on an individual's wage. Results also suggest that although there appears to be no significant effect on wages, diabetes duration does significantly lower an individual's probability of selecting into the labor market.

The rest of the paper is organized as follows. Section II provides a brief description of the economic literature on diabetes. Section III presents the model and estimation methodology. Section IV provides a description of the data used in analysis. Section V presents results, and Section VI concludes.

1.2 Literature Review

The recent interest in diabetes's impact on the labor market is likely due to two primary factors: the high and rising incidence of the disease in the eligible American workforce (ADA, 2008), and the amount of relatively new data available on the topic (Pango, 1999). Papers estimating the costs of diabetes to an employer have established that employers face higher medical costs (ADA, 2007) and experience diminished productivity due to diabetic workers (Lavigne, 2003).[1] These studies estimate diabetes as a singular indicator variable, where all diabetics are combined to show the marginal effect of diabetes. Ramsey, et al., (2002) estimates a per-employee cost of about $4,671 annually for employees aged 18–35 and $4,369 for those aged 56–64 years. Although they are looking at the average effect of diabetes on these different age groups, there is

[1]Due to the design of their survey, they only examine a small group of New York residents.

some indication, by the changing dollar costs, that diabetes changes over the lifetime of the patient.[2]

Studies that estimate the cost of diabetes on the individual are somewhat more varied. Kahn (1998) estimates that the negative effect of diabetes on productivity is actually decreasing over time. This is probably due to the sample period, in which diabetes was not growing at current rates.[3] Therefore, his results may not be indicative of the contemporary effect diabetes has on the labor market. Vijan, et al. (2004) and Tuncli, et al. (2005) estimate a reduction in earnings due to diabetes. However, both papers are limited by the data from the Health and Retirement Survey (HRS), which samples only people between 51 and 61 years of age, and they estimate only the singular effect of diabetes as an indicator variable. Brown, et al. (2005) suggests that diabetes may be endogenous with respect to work. Using genetic information about a respondent's parents, they find a negative effect of diabetes on employment and an indication of endogeneity bias. All of these studies estimate diabetes's average effect, combining all diabetics into one broad category, regardless of the duration of diabetes. If, however, there is a change in the effect of diabetes with duration, these studies may have misrepresented the true impact of diabetes.

This paper extends the current economic literature by estimating diabetes not only as a singular impact, but also as a continuous measure of diabetes duration. Second, numerous sources of statistical bias, such as selection and unobserved heterogeneity,

[2] From their results it is not clear if they are identifying an effect of increasing age of the patient or increasing diabetes duration, as the two will be highly correlated and are not separated in their estimates.
[3] The sample period consists of the years 1976, 1989, and 1992. His results suggest that the overwhelming increase in technology over the time period may have overshadowed the relatively small growth in the incidence of diabetes.

previously unaccounted for in the literature, are taken into consideration in a panel data framework. Third, an instrumental variable technique is implemented to account for the potential endogenous relationship diabetes has with an individual's wage rate.

1.3 Empirical Methodology

The purpose of this paper is to estimate the effect of diabetes on employment and wages. The decision to select into or out of the labor market can be represented by:

$$s_{i,t} = \alpha_0 + \alpha_1 Diabetes_{i,t} + \alpha_2' X_{i,t} + c_i + \eta_{i,t}. \qquad (1)$$

where $s_{i,t}$ is a zero/one indicator variable equal to one if person i is employed at time t; α_0 is a constant term; $Diabetes_{i,t}$ is a variable containing information on person i's diabetes status in year t; $X_{i,t}$is a is a vector of person-year specific variables that control for all other observable influences on the employment decision, e.g. age, education, family size, industry, etc.; c_i represents a time invariant unobserved characteristic; and $\eta_{i,t}$ is the error term for each individual, i, in time t.

Because diabetes, among other observable characteristics, will influence a person's decision to work, and because wages are only observed for those people who chose to work, the decision to enter the labor market should be included in the estimation of wages. This decision can be represented by:

$$s_{i,t} = \alpha_0 + \alpha_1 Diabetes_{i,t} + \alpha_2' X_{i,t} + c_i + \eta_{i,t}. \qquad (2)$$

$$w_{i,t} = \beta_0 + \beta_1 Diabetes_{i,t} + \beta_2' X_{i,t} + \beta_3' \lambda_{i,t} + c_i + \varepsilon_{i,t}. (where\ w_{i,t} > 0\ iff\ s_{i,t} = 1). \qquad (3)$$

where $w_{i,t}$ is a log of the real hourly wage for person i in time t; β_0 is the constant term; $Diabetes_{i,t}$ contains information on diabetes status; $X_{i,t}$is a is a vector of person-year specific variables; c_i represents the time invariant unobserved component; and $\varepsilon_{i,t}$ is the residual error term. Equation (3) will only be observed if a person decides to select into working; that is, a positive value for $w_{i,t}$ is only observed when $s_{i,t}$ is equal to one. From the estimation of equation (2) a probability of employment can be calculated and this probability enters equation (3) as $\lambda_{i,t}$, commonly referred to as the Inverse of the Mill's Ratio (IMR) (Heckman, 1976).

Potential problems arise when attempting to estimate equations (2) and (3) with panel data. First, the components of c_i will bias estimates if they are not properly accounted for in the estimation.[4] To control for this unobserved heterogeneity, a 'fixed-effect' term, comprised of the time invariant means of all observable characteristics, is included in all estimations. Additionally, the decision to work is not made once; rather a person continually chooses whether to remain in or out of the labor market, so the selection equation must be estimated in every time period for every person.[5] The econometric specification used in all estimations is according to Jackle & Himmler (2007) and Wooldridge (1995). After incorporating the two techniques addressed above, equation (2) and (3) can be estimated by pooled OLS, where standard errors must be clustered at the individual level and bootstrapped.[6] Equations (2) and (3) provide the first set of results presented in this paper; the coefficients of interest are α_1 and β_1.

---

[4] For example a person could simply be unproductive or place a low value on working. In either case these factors could affect both the decision to work and the wage level.

[5] For each person i there will be an Inverse of the Mill's Ratio (IMR) generated for every time period that is included in the estimation of wages. This specification assumes that the decision to work in each time period is independent of the decision in all other time periods. Joint tests of all IMR and 'fixed-effect' terms are included in the results table.

[6] Clustered standard errors allow every individual within the data to have their own error term, independent of other observations, and bootstrapping will account for the fact that all IMRs and time invariant means included in estimation are calculated variables, instead of true data points. Because the actual probability

Due to the nature of diabetes, there is a concern that it may be endogenous with respect to wages and employment. If a person's wage increases, they could buy healthier foods or join a gym because they have more discretionary income, reducing the probability of contracting diabetes. Similarly, if diabetes is endogenous with respect to work, our results will be biased. For example, it could be that a person loses their job, and because of this they no longer get as much physical exercise or eat cheaper, less nutritious food, increasing the probability of contracting diabetes. These are simple examples, and it is easy to imagine a case where the opposite is true. E ither way, employment, wages, or both may affect the probability of diabetes, thus producing biased results. Additionally, it is possible that there exist some omitted variables that influences both diabetes and wages. If this were the case diabetes would be absorbing some of that omitted variable's effect. This endogeneity bias can be overcome with instrumentation.

An individual's instrumented employment decision can be represented by:

$$s_{i,t} = \alpha_0 + \alpha_1' \widehat{Diabetes}_{i,t} + \alpha_2' Y_{i,t} + \alpha_3' \bar{Y}_{i,t} + \eta_{i,t}, \qquad (4)$$

where $Y_{i,t}$ is a vector of observable variables that have a correlation with an individual *i*'s diabetes; $\bar{Y}_{i,t}$is the time invariant mean of all observable characteristics plus the time invariant mean of all instruments; and $\widehat{Diabetes}_{i,t}$ are the instrumented, predicted values for diabetes;. Equation (4) can then be consistently estimated using a panel probit approach.

The instrumented effect of diabetes on wages takes the form:

$$s_{i,t} = \alpha_0 + \alpha_1' Z_{i,t} + \alpha_2' Y_{i,t} + \alpha_3' \bar{Y}_{i,t} + \eta_{i,t}, \qquad (5)$$

---

distributions of these created terms are unknown, the standard errors will approach their true values after this replication (Efron and Tibshirani 1986)

$$w_{i,t}= \beta_0 + \beta_1 \widehat{Diabetes}_{i,t} + \beta_2' Y_{i,t} + \beta_3' \bar{Y}_{i,t} + \beta_4' \lambda_{i,t} + \epsilon_{i,t} \; (where\ w_{i,t} > 0 \ iff \ s_{i,t} = 1). \quad (6)$$

where $Z_{i,t}$ is a vector of instruments that have a correlation with an individual *i*'s diabetes but are uncorrelated with that individual's wage level; $\widehat{Diabetes}_{i,t}$ are the predicted values for diabetes; $Y_{i,t}$ is a vector of observables for person *i* that does not include the suspected endogenous regressor; and $\bar{Y}_{i,t}$ is the time invariant mean of all observable characteristics plus the time invariant mean of all instruments. Equation (6) can then be consistently estimated using a set of IMRs, $\lambda_{i,t}$, and explanatory variables, $Y_{i,t}$ and $\bar{Y}_{i,t}$, that are uncorrelated with the error term, $\epsilon_{i,t}$.[7]

The exact causes of diabetes are still somewhat unknown, but research indicates that there is a strong genetic link in the contraction of diabetes (CDC, 2007). If this is the case, a person with a sibling who has diabetes may be more likely than a person with healthy siblings to contract the disease due to genetic predisposition.[8] Therefore, a sibling's diabetes information is used to instrument each respondent's own diabetes. This is similar in theory and implementation to the estimation process used by Cawley (2004) which estimates the effect of obesity on wages. Sibling diabetes follows the same specification as a person's own diabetes in all estimations. When own diabetes is a zero/one indicator variable, so too is sibling diabetes. Additionally, when diabetes is

[7] This specification is from Jackle & Himmler (2007) and Wooldridge (1995). Semykina and Wooldridge (2006) show that the FE-2SLS estimator is consistent, even when the instrument is correlated with selection and the unobserved effect.

[8] There is some concern that this instrument may not be picking up entirely genetic predisposition but rather behavioral or family upbringing. This should not be a problem as either of these factors will be accounted for by the fixed-effect term, leaving only genetics to identify diabetes. Additionally, the fixed-effect term could be absorbing any genetic characteristic that do not change over time, leaving sibling diabetes no predictive power in first stage estimations. We can conclude this is not the case due to first stage significance tests of the instrument, presented with all instrumented results. Also diabetes itself, even if you are predisposed to it genetically will not be constant over time in either specification. That is, even if your family is genetically more likely to contract diabetes, this may not be observed in the data until part of the way through the sample.

specified as a linear duration variable, sibling diabetes will take the same form. This allows the endogenous variable, a person's own diabetes, to be identified in all estimations, and each specification of the instrument provides a strong theoretical and statistical link to an individual's own diabetes.

Identification in each case comes from the fact that a diabetic is much more likely to have a sibling with diabetes, regardless of the variable's specification, than a non-diabetic. For the 0/1 specification of diabetes it is likely that a diabetic will also have a sibling with diabetes, and therefore receive a positive value for their sibling's diabetes measure in some time periods. When diabetes is specified as a linear measure of duration, identification works in much the same way. A person with a positive value for diabetes duration is more likely than a non-diabetic person to a have a sibling that has or will contract diabetes during the sample. Also, it is reasonable to assume that this lengths will be somewhat correlated. If a person has had diabetes for a long amount of time, it is likely that their siblings are also predisposed to contract diabetes early on. Conversely, a person who recently contracted diabetes may have siblings that also recently contacted the disease or that may be likely to contract it in the near future.

A valid instrument also should be uncorrelated with the error term in the equation of interest, and there is statistical and theoretical evidence to support exogeneity of sibling diabetes.[9] Previous economic literature suggests that diabetes does affect numerous aspects of a person's labor market decisions. However, it is unlikely that

[9] Test for instrument validity are presented along with all instrumented results.

simply having a sibling with diabetes, or a sibling that has had diabetes for some number of years, directly impacts their hourly wage or employment decision.[10]

1.4 Data

The data used in this analysis are from the NLSY79, which is collected by the Bureau of Labor Statistics (BLS). This is an ongoing survey that gathers information on the same individuals from 1979 until the most recent year of data, 2006.[11] When the survey first began it included 12,686 men and women between the ages of 14 and 22. The NLSY79 collects detailed information on employment, wages, work history, and numerous other labor market characteristics of interest, but it was not until 2006 that they began collecting specific information on health characteristics. With the addition of the supplemental 40 and over health questionnaire, the NLSY79 asks respondents if they have diabetes and in what year they were first diagnosed. From this information a zero/one indicator and a linear specification of diabetes duration can be created to examine the overall effect of diabetes and the incremental impact each additional year of diabetes has on an individual's wage. Figure 1 illustrates the number of reported cases of diabetes. Just as in the national statistics, there is a very high growth rate in the incidence of diabetes beginning in the early 1990s.

---

[10] Some might make the argument that an ill person in the household could cause the other members to work more; this still would be uncorrelated with the hourly wage. Although you might work more hours to provide supplemental income, it is probably not the case that a sick sibling results in a higher or lower wage rate. Also, the benefit of this sample is that siblings likely have not lived in the same household for quite some time, lowering the probability that their disease would influence the others employment decision.

[11] Until 1994 the survey is conducted annually. After 1994 data is collected biennially.

Table 1 presents summary statistics for the variables used in estimation. Key variables are employment, wages, and diabetes. The real average hourly wage is $11.48 for the 73 percent of the working sample, and about one percent of the entire sample has diabetes.[12] This seems low, as the national incidence of diabetes in the American labor force is about ten percent (CDC, 2007), but this average may be misleading as some diabetics in the sample do not contract the disease until later in their lifetime. In the last year of the sample, 221 people have diabetes, out of 4,079, or over five percent.[13] Other demographic controls are nationally representative.[14]

Table 2 presents a comparison of summary statistics for diabetics and non-diabetics. Diabetics report a higher average wage and a lower incidence of employment than non-diabetics. However, this higher wage may be due primarily to type II diabetes which accounts for over 90 percent of all reported diabetes cases and typically develops in older adults, who have higher wages than younger workers. The higher average wage may simply be due to 'age-effects'. Figure 2 shows the average hourly wages for diabetics and non-diabetics over the entire sample. Here we see that after 1984 non-diabetics consistently earn a higher wage and experience faster wage growth overall, than the diabetic population.[15] In fact, the most recent year of data indicates diabetics have a mean wage of 15.38 which is well below that of non-diabetics, who report an hourly wage of 19.11. The average age of each subgroup also supports the existence of 'age-effects' in the summary statistics. Diabetics on average are about 37, where non-diabetics

[12] All wages are presented in year 2000 dollars, according to a deflator estimated by the BLS
[13] This is still somewhat below the national average, but is more in line with what we might expect.
[14] Also included in estimation, but not reported in Table 1, are individual year and industry indicator variables. Table A-1 presents a breakdown of the industry variables for diabetics and non-diabetics.
[15] The likely reason wages for diabetics are higher for diabetics before 1984 is because this sample is very small. Figure 1 indicates that the large growth in the incidence of diabetes dose not begin until this time.

are on average about 29. A difference between diabetics and non-diabetics also appears in job tenure and work experience. Specifically, diabetics seem to work at a job longer on average, and, conditional on working, they tend to remain in the labor force longer. This may indicate that diabetics are sorting into particular jobs and remaining there longer than the non-diabetic population.[16] The average length that respondents have had diabetes is 8.9 years, but this ranges from one to 44 years of the disease. Lastly, the incidence of a sibling having diabetes is much higher for diabetics than non-diabetics, which supports its use in the instrumentation process.

Figure 3 and Figure 4 present local polynomial regressions of years on employment and the wage level, respectively, for diabetics and non-diabetics.[17] Figure 3 shows the dramatic difference in labor force participation between the two samples. Initially diabetics are engaged in the work force in much higher percentage, likely due to the relatively small amount of diabetics in the early sample period. Over time, the change in work force participation indicates that diabetics are leaving the work force, where non-diabetics continue to grow in numbers, suggesting a fundamental difference in employment decisions for the two samples over time. Figure 4 indicates that not only do diabetics have a lower wage level, but they also experience slower wage growth and see negative growth, relative to non-diabetics. Figure 5 and Figure 6 present local polynomial regressions of work experience on employment and the wage level of diabetics and non-

---

[16] It is also very plausible that this is another 'age-effect' being picked up in the data. That is, diabetics are simply older thus they have more experience and tenure than the rest of the population.

[17] Local polynomial regressions estimate the log of the real hourly wage conditional only on one explanatory variable, year and experience, respectively in this analysis. Local polynomial regressions perform a locally weighted regression to smooth the estimates, so that a clear linear relationship can be extracted from the data (Fox 2004). These figures are generated using the locpoly command for Stata. Results show the dependence of wages on the explanatory variables where no inference is made on the specific function that relates the two variables.

diabetics. Figure 5, again, shows a fundamental difference between diabetics and non-diabetics regarding the employment decision and years of experience. For non-diabetics the returns to additional work experience largely raise their likelihood of working, this is not the case for diabetic individuals. A negative slope in this figure shows that even as their work experience goes up diabetics are more likely to exit the labor force. Figure 6 illustrates that over their working lives diabetics and non-diabetics experience different returns to their experience. Non-diabetics have a higher wage level for all levels of experience and even have a steeper function which indicates a higher growth rate for each additional year of work. Later in their working life diabetics begin to close the wage gap, but they never achieve the same level as non-diabetics. These figures suggest that the length of time an individual has diabetes may matter as the slopes of the lines differ, indicating different rates of growth for diabetics and non-diabetics over the last 30 years.

1.5 Results

Table 3 presents the effect of diabetes on employment for the entire sample and males and females separately.[18] All models contain corrections for unobserved heterogeneity.[19] In Model 1, diabetes is defined as a zero/one indicator variable, equal to one if the person i has diabetes at time t. This specification, which has been used in previous economic studies of diabetes, estimates the average marginal effect of contracting diabetes. Model 1 shows that diabetes has a highly significant negative effect

[18] Tables 3 and 4 are panel probit estimations including a fixed-effect.

[19] Semykina and Wooldridge (2006) provide a straightforward way to test for unobserved heterogeneity. A joint test of significance the fixed-effect terms will indicate the presence of unobserved heterogeneity. The null hypothesis of this test is that fixed-effects are not necessary, and a rejection of this indicates unobserved heterogeneity in the data.

on employment, reducing the probability of working by about 16 percent for the entire sample.

Model 2 uses duration of diabetes in years as the principal explanatory variable. Results show the effect each consecutive year of diabetes has over its duration. Estimated coefficients of the entire sample show that the disease causes a reduction of about 2.2 percent annually holding all else constant. This indicates that a diabetic does not actually see a consistent reduction in their probability of working over the lifetime of the disease; rather, diabetes's effect is growing with the duration of the disease. Model 2 constrains diabetes to have the same impact each year. However, this is probably not the case; it is easy to imagine that if diabetes does have a larger impact over its lifetime that this growth rate might not be constant over the duration of diabetes.

Model 3 attempts to address this by including a quadratic specification of diabetes duration. Results suggest a changing effect over the course of diabetes. In Model 3, the negative effect of diabetes drops in size to an initial penalty of 0.45 percent for the entire sample, annually. However, a negative quadratic term indicates that as diabetes progresses the negative effect of employment is exacerbated. Although the estimated coefficient on diabetes length alone is insignificant, a joint test indicates that both terms are jointly significant beyond the one percent level and affect a person's employment decision.

Table 4 presents the effect of diabetes on employment, using an instrumental variable technique. All models are estimated as before, but now an individual's own diabetes information is instrumented with information on a sibling's diabetes. Model 1 shows that diabetes has no significant effect on employment, indicating that endogeneity bias causes an overstatement of the negative average impact of diabetes. This is in direct contrast to previous estimates, which attributed a significant and negative effect to diabetes.

Validity test of the instrument are included for all specifications. The first test statistic presented is the Durbin–Wu–Hausman Test for endogeneity. A rejection of the null hypothesis indicates that diabetes is endogenous with respect to employment and previous estimates were biased. Second, a first stage test of sibling diabetes's predictive power with respect to a person's own diabetes is presented. A rejection of the null indicates that sibling diabetes is significantly different from zero in the prediction of an individual's own diabetes. Lastly, the instrument is tested on the outcome variable, employment. The null hypothesis in this case is that sibling diabetes has no correlation with the error term in the prediction on an individual's employment decision.

For Model 1 there is an indication of endogeneity bias in the previous estimates, and sibling diabetes predicts a person's own diabetes. Instrumented results suggest that diabetes has no average effect on a person's employment decision.[20] Model 2, where diabetes is estimated as a continuous variable, shows that the disease causes no significant reduction in employment probability over time. However, there is no statistical indication of endogeneity bias in these results, so previous estimates (Table 3), where diabetes causes a 2.2 percent annual reduction in employment probability, are preferred for efficiency. A significant effect of diabetes length on employment indicates that contracting diabetes does influence an individual's decision to enter or exit the labor

[20] Although there is an indication of endogeneity, this result from Table 4 may not be preferred to Table 3. This is because the instrument utilized, sibling diabetes, has a statistically significant correlation with the error term in the equation of interest.

force. Model 3 shows no significant effect of diabetes length or diabetes length squared, and this specification is estimated to be endogenous with respect to the employment decision.

Results change when the sample is partitioned by gender.[21] Male estimates, presented in Table 3 and 4, indicate no significant effect on male wages. However, a test for endogeneity reveals the presence of endogeneity bias, so the instrumented results (Table 4) are preferred. Examining only the female sample, Table 3 suggests that all linear measures of diabetes are significant with respect to the employment decision. Because there is no indication of endogeneity bias, un-instrumented results are preferred. Findings show a significant 2.9 percent reduction in the probability of employment each year for females, and cubic estimates indicate a smaller initial reduction but one that increases exponentially with diabetes duration.

Theoretically, the negative effect of diabetes on employment could stem from any one of numerous factors: A diabetic person may not be well enough to perform any serious labor activity and therefore select out of the market entirely. Perhaps, diabetics are rejected from the labor market by prospective employers more often than healthy employees. Or, a diabetic may reasonably expect to earn a lower wage rate than their colleagues, and therefore select out of the labor market. To investigate specifically the last hypothesis, we next examine the effect of diabetes on an individual's wage rate controlling for any unobserved heterogeneity and the decision to select into work.[22]

---

[21] Appendix Tables A-4 through A-9 provide estimates also partitioned by race.

[22] Semykina and Wooldridge (2006) provide a straightforward way to test for selection bias. A joint test of significance of the IMR indicates selection bias within the data. Here the null hypothesis is no selection bias, and rejection implies that IMRs are needed for correction. Tests for selection and unobserved heterogeneity are included in Tables 5 and 6.

Table 5 presents results of diabetes's effect on wages. For the whole sample, diabetes has a significant negative impact on wages in every specification. Diabetes on average (Model 1) causes a 13 percent reduction in an individual's wages. When we examine diabetes duration (Model 2), these numbers decrease in size, indicating that each consecutive year of diabetes reduces a person's wages by about one percent. A reduction in wages of 13 percent (the amount estimated by Model 1 and the method other studies have utilized) occurs only after their thirteenth year of diabetes, and after that length of time the wage penalty becomes more pronounced. Once again, these estimates indicate that diabetes does not have a singular, static impact on wages, but its negative consequences grow with the duration of the disease. Lastly, diabetes duration estimated with a quadratic term (Model 3) indicates a smaller annual effect of about 0.76 percent, but the negative term on diabetes duration squared shows that this effect is growing more negative with duration. According to these estimates, a reduction of 13 percent would occur around the sixteenth year of the disease and continue to grow more severe thereafter.

Table 6 presents the instrumented results of diabetes effect on wages.[23] Results overwhelmingly show no significant effect of diabetes on an individual's wage, regardless of the form diabetes takes. Examining the test statistics, diabetes is endogenous with respect to wages in every specification, and the instrument, sibling diabetes, passes all validity tests. Due to the bias, instrumented results from Table 6 are preferred over our previous estimates and diabetes is estimated to have no true causal

---

[23] This estimation methodology follows Jackle & Himmler (2007) and is extension of Wooldridge (1995). Just as in previous estimations, sibling diabetes takes on the same specification as a person's own diabetes. This allows the construction of an IMR without the inclusion of the suspected endogenous variable

impact on wages. This finding is contrary to many pervious studies of diabetes's impact on wages and suggests that previous estimates may have overstated the effect of diabetes. However, a negligible effect of diabetes is not entirely unexpected, as many related medical conditions associated with diabetes may be included in the non-instrumented results, these may include kidney disease, heart failure, stroke, blindness, or high blood pressure (CDC, 2007).

Results change when the sample is partitioned by gender.[24] Male estimates, presented in Table 6, indicate no significant effect on male wages. However, a test for endogeneity reveals no statistical bias, so the un-instrumented results are preferred for efficiency. Examining only the female sample suggests that only the zero/one indicator variable measure of diabetes is significant in the determination of wages. A positive coefficient on diabetes indicates that females receive a wage increase from contracting diabetes. This change in sign is suspect, and when tests for endogeneity are examined, they reveal that un-instrumented results are preferred for efficiency. Endogeneity tests for the other two specifications reveal that previous estimates were subject to an endogeneity bias, and diabetes does not significantly affect female wages.[25]

Instrumented results overwhelmingly show that diabetes has no causal impact on an individual's wage. This is in contrast to previous findings, which attributed a significant wage penalty to diabetes. Those results which did not account for an endogeneity bias may have overstated the negative impact of diabetes, probably confounding some of the serious related medical conditions and attributing their negative

---

[24] Appendix Tables A-10 through A-15 provide me estimates also partitioned by race.

[25] Although there is an indication of endogeneity, female results from Table 6 may not be preferred to Table 5, because the instrument utilized, sibling diabetes, has a statistically significant correlation with the error term in the equation of interest for the female sample.

effects to diabetes. This finding indicates that contracting diabetes alone has no significant effect on an individual's wage level once unobserved fixed-effects and the decision to select into work have been accounted for statistically.

Taken wholly, the results from this paper suggest that diabetes has no causal impact on an individual's wages. However, there is some evidence that diabetes does impact the decision of whether or not to enter the labor market. Previous estimates of diabetes's impact on wages may have misrepresented the actual impact for two reasons: First, tests show that selection bias is present in the data, and if the decision to work is not included in the estimation of wages, results may incorrectly reflect this decision's impact. Second, tests reveal that diabetes is endogenous with respect to wages, and if this is unaccounted for estimated results will not show the true impact of diabetes on wages.

1.6 Conclusion

This paper estimates the effect of diabetes on employment and wages. Initial findings indicate that a diabetic can expect a 16 percent lower probability of working and a 13 percent lower wage on average over their lifetime. However, this study shows that these effects are probably due to the severe medical conditions associated with diabetes and there is no significant effect on wages simply from the contraction of the disease itself. With the considerable amount of time and money that both individuals and employers spend on diabetes prevention, medication, and education, these results are somewhat concerning. If the negative effects traditionally associated with diabetes are, in fact, due to some other factors, it seems that some of the resources spent on diabetes could be utilized in a more efficient manner to prevent the true cause of the loss in productivity in this market.

Additionally, estimates indicate that diabetics do not receive a uniform penalty, as others have estimated, but the impact of diabetes changes with diabetes duration. Specifically, diabetes is estimated to have a negative impact on employment during the first few years after diagnosis. This could be due to a person struggling to cope with the new symptoms and complications that diabetes presents. But, even as a person learns to control their diabetes the negative impact continues to grow, becoming more pronounced in the later stages of the disease. This is likely due to the inherent nature of the disease. Diabetes, once contracted, is not a disease that typically gets better as the person ages. Rather, its penalty is felt much more in the elderly, and even when the patient is on a steady treatment, the course of the disease could worsen over time as would the negative effect. Once, a diabetic person has made the decision to enter or exit the workforce, findings show that they earn no less statistically than their contemporaries, due only to diabetes. It is likely that negative wage results found in previous estimations are due to the severe related medical conditions associated with diabetes and not the disease itself.

Previous studies on diabetes may have misrepresented the impact of the disease by not properly accounting for the decision to work or not work. It is also likely that some amount of an unobservable 'fixed-effect', unaccounted for in other studies, influences the probability of contracting diabetes along with an individual's wage rate and their decision to enter the work force. Corrections for selection bias and unobserved heterogeneity are shown to be statically significant in the estimation of diabetes's impact on wages, and if unaccounted for, will produce biased results. Similarly, 'fixed-effects' are statistically significant in the estimation of a person's employment decision. Finally, endogeneity bias is shown to be of concern in the estimation process. Primarily in the estimation of diabetes's effect on wages, this is likely due to the simultaneous relationship between diabetes and wages or some omitted explanatory variables that are correlated with both wages and diabetes. The use of an instrumental variable technique shows the impact of diabetes on wages and employment, rather than a correlation between the two, which previous papers have identified.

As mentioned earlier, there could be an 'age-effect' of diabetes that has yet to be established. A pseudo-panel approach to estimation would show if, in fact, diabetes affects the young differently than it does the old. Also, the separation of type-one and type-two diabetes could be significant, and it would be interesting to see if the two diseases have different effects over their durations.[26] Although the effect of diabetes on a person's own wage has been examined in the literature, little attention has been paid to the effect on a spouse's wage and employment decision or benefit packages offered to employees. It is reasonable to assume that diabetes affects not only your own work decision but also that of your spouse. Also it is likely that diabetics place a greater emphasis on benefits, such as health insurance, than the general population. This could alter their work decisions; perhaps making a diabetic more likely to stay a job with good health insurance even though the pay is somewhat lower. It could also make diabetics less likely to enter into a job search, knowing they have the security of their current job.

[26] A panel data set with a larger incidence of diabetics would be necessary to perform this estimation. Currently, the NLSY79 does not contain enough diabetics to accurately separate and estimate the continuous effects of both type-one and type-two diabetes.

Although these questions are interesting, with the currently utilized data they are impossible to ascertain.

References

American Diabetes Association. (2008). Economic Costs of Diabetes in the U.S. in 2007. *Diabetes Care, 31* (3), 1-20.

Brown, H. S., Pagan, J. A., & Bastida, E. (2005). The Impact of Diabetes on Employment: Geneitic IVs in a Bivariate Probit. *Health Economics, 14* (5), 537-544.

Cawley, J. (2004). The Impact of Obesity on Wages. *Journal of Human Resources, 39* (2), 451-474.

Centers for Disease Control and Prevention, A. G. (2007). National Diabetes Fact Sheet.

Donaghue, K. C., Hing, S., Fairchild, J. M., Cutler, L. R., Craig, M. E., Howard, N. J., et al. (2003). Do All Prepubertal Years of Diabetes Duration Equally Contribute to Diabetes Complications? *Diabetes Care, 24* (4), 1224-1229.

Efron, B., & Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracty. *Statictical Science*, 1(1), 54-77.

Fox, C. S., D'Agostino, R. B., Sullivan, L., & Wilson, P. W. (2004). The Significant Effect of Diabetes Duration on Coronary Heart Disease Mortality. *Diabetes Care, 27* (3), 704-708.

Fox, J. (2004) Nonparametric Regression. *Unpublished Manuscript*, McMaster University.

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

Heckman, J. J. (1976). The Common Structure of Statitical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement, 5* (4), 475-492.

Ivers, R. Q., Mitchell, P., Cumming, R. G., & Peduto, A. J. (2001). Diabetes and Risk of Fracture. *Diabetes Care, 24* (7), 1198-1203.

Jackle, R., & Himmler, O. (2007). Health and Wages: Panel Data Estimates Considering Selection and Endogeneity. *Munich Personal RePEc Archive Working Paper No.11578* .

Kahn, M. E. (1998). Health and Labor Market Performance: The Case of Diabetes. *Journal of Labor Economics, 16* (4), 878-889.

Lavigne, J. E. (2003). Reductions in Individual Work Productivity Associated with Type 2 Daibetes Mellitus. *PharmacoEconomics, 21* (15), 1123-1134.

Nichols, G. A., Erbey, J. R., Hiller, T. A., & Brown, J. B. (2001). Congestive Heart Failure in Type 2 Diabetes: Prevalence, Incidence, and Risk Factors. *Diabetes Care, 24* (9), 1614-1619.

Olivia, J., Lobo, F., Molina, B., & Monero, S. (2004). Direct health care costs of diabetic patients in Spain. *Diabetes Care*, 27, 2616-2621.

Pango, E. (1999). Cost of Diabetes: A Methodological Analysis of the Literature. *PharmacoEconomics, 15* (6), 583-595.

Ramsey, S., Summers, K. H., Leong, S. A., Birnbaum, H. G., Kemner, J. E., & Greenburg, P. (2002). Productivity and Medical Costs of Diabetes in a Large Employer Population. *Diabetes Care, 25* (1), 23-29.

Semykina, A., & J. M. Wooldridge (2006). Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application. *Unpublished Manuscript, Michigan State Univeristy.*

Tuncli, K., Bradley, C. J., Nernez, D., Williams, L., Keoki, P., Manel, L., et al. (2005). The Impact of Diabetes on Employment and Work Productivity. *Diabetes Care, 28* (11), 2662-2667.

Vijan, S., Hayward, R. A., & Langa, K. M. (2004). The Impact of Diabetes on Workforce Participation: Results from a National Househould Sample. *Health Services Research, 39* (6), 1653-1669.

Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004). Global Prevalence of Diabetes: Estimates for the Year 2000 and Projections for 2030. *Diabetes Care, 27* (5), 1047-1053.

Wooldridge, J. M. (1995). Selection Correction for Panel Data Models Under Conditional Mean Independance Assumption. *Journal of Econometrics, 68*, 115-132.

FIGURE 1. INCIDENCE OF DIABETES FROM 1980-2006

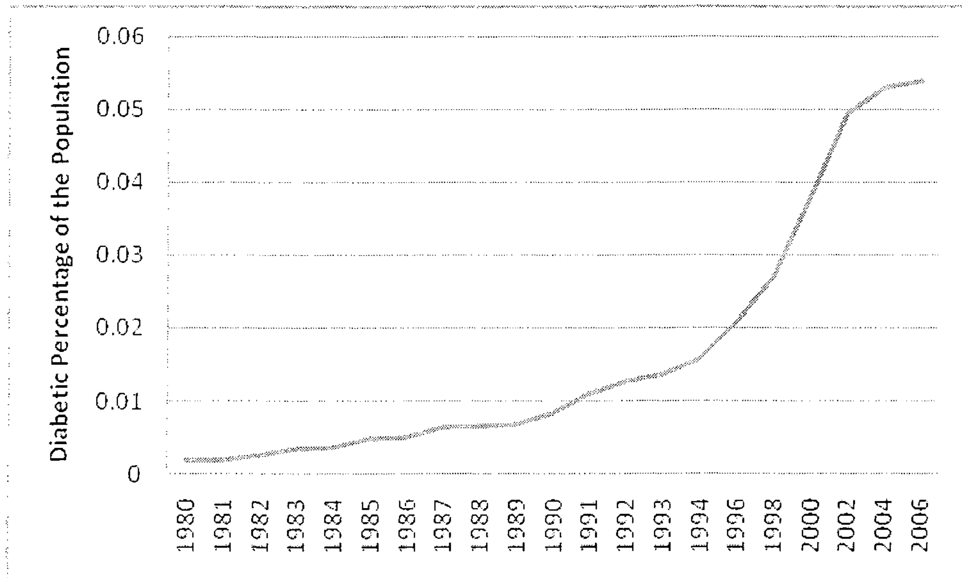


FIGURE 2. COMPARISON OF HOURLY WAGES FOR DIABETICS AND NON-DIABETICS

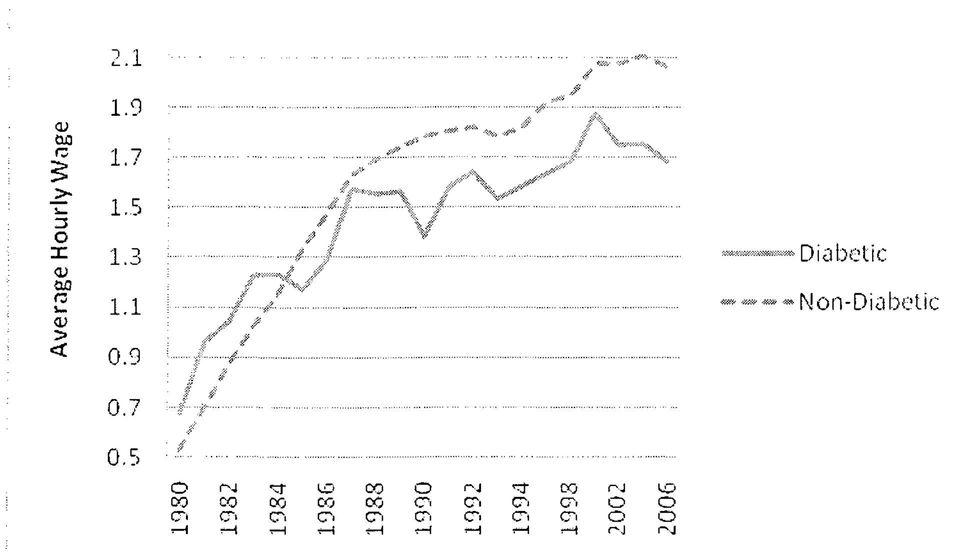
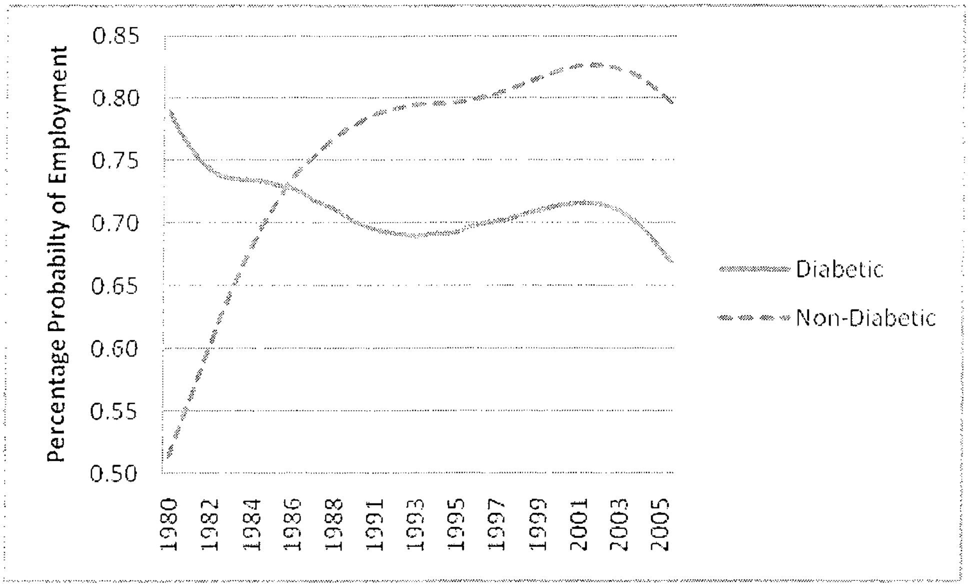
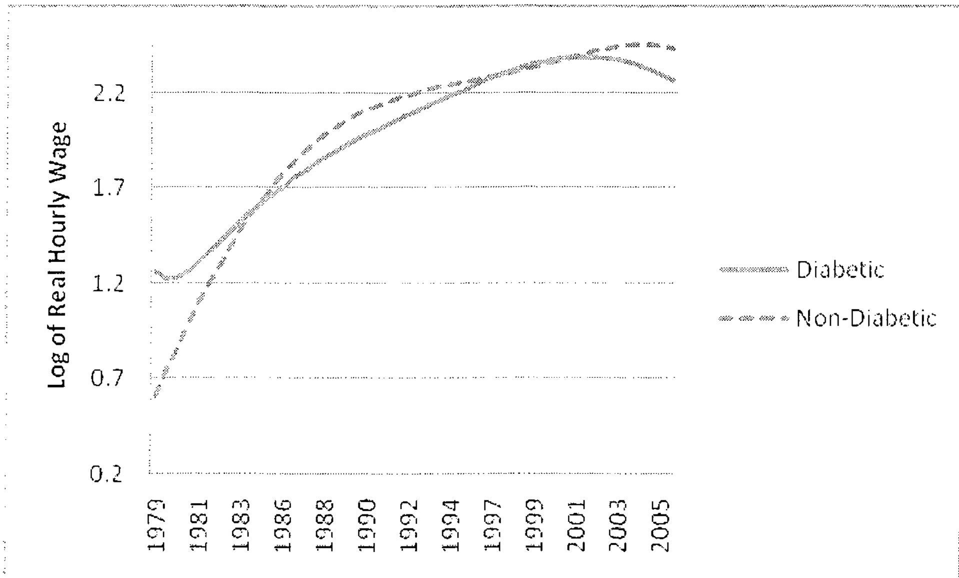


FIGURE 3. LOCAL POLYNOMIAL REGRESSION OF EMPLOYMENT OVER TIME



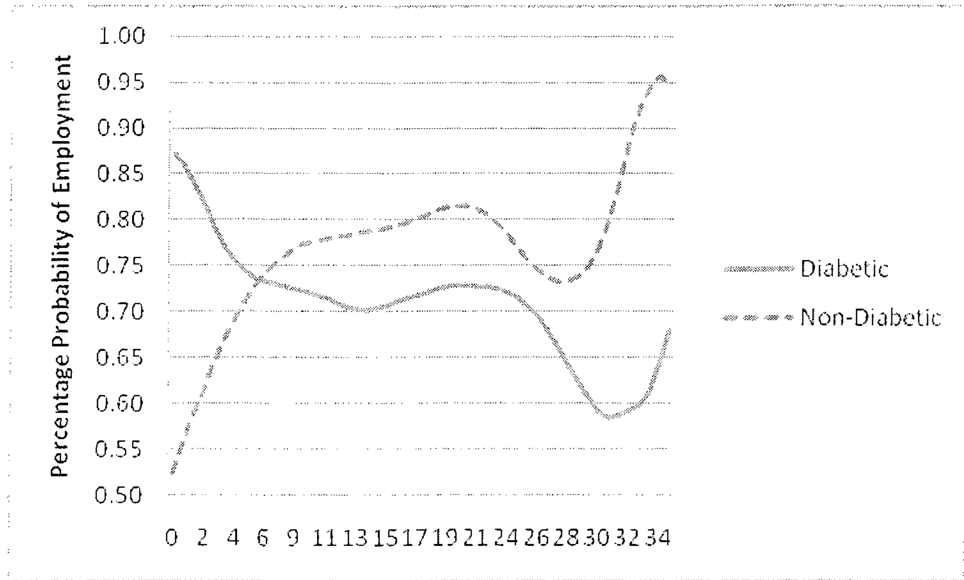
Notes: Graph is the result of a local polynomial regression of employment over time for diabetics and non-diabetics separately.

FIGURE 4. LOCAL POLYNOMIAL REGRESSION OF WAGES OVER TIME



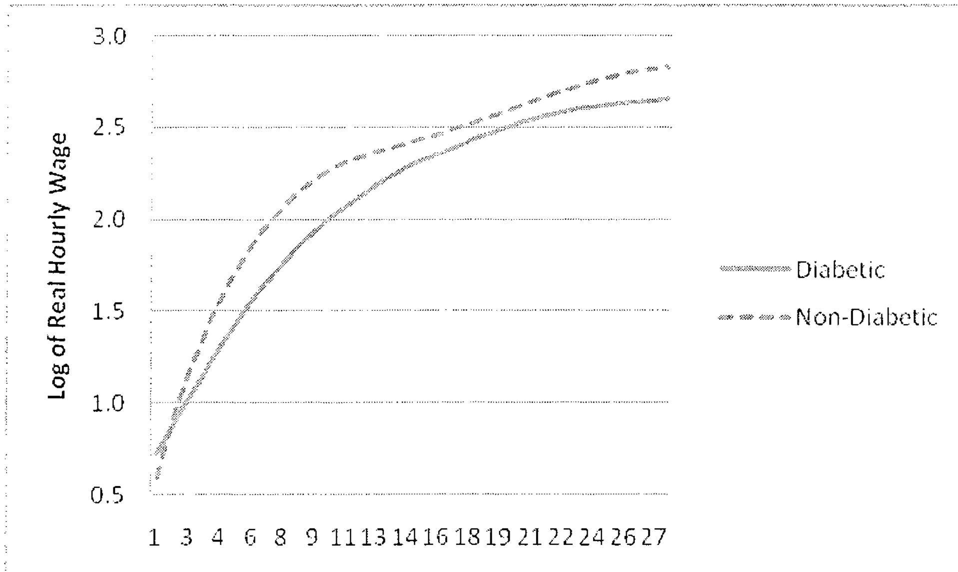
Notes: Graph represents only working individuals and is the result of a local polynomial regression of the log of real hourly wage level over time for diabetics and non-diabetics separately.

FIGURE 5. LOCAL POLYNOMIAL REGRESSION OF EMPLOYMENT BY WORK EXPERIENCE



Notes: Graph is the result of a local polynomial regression of work experience on employment for diabetics and non-diabetics separately.

FIGURE 6. LOCAL POLYNOMIAL REGRESSION OF WAGES BY WORK EXPERIENCE



Notes: Graph represents only working individuals and is the result of a local polynomial regression of work experience on the log of real hourly wage for diabetics and non-diabetics separately.

TABLE 1. SUMMARY STATISTICS

	Mean	Standard Deviation	Minimum	Maximum
Key Variables				
Hourly Wage	11.78	14.77	1	500
Employment*	0.73	0.44	0	1
Diabetes	0.01	0.12	0	1
Diabetes Length	0.13	1.40	0	44
Demographic				
Male	0.56	0.50	0	1
Age	29.76	7.55	15	49
Family Size	3.32	1.87	1	15
Number of Children	0.83	1.13	0	9
Black	0.17	0.38	0	1
Hispanic	0.27	0.45	0	1
Married	0.44	0.50	0	1
Separated	0.04	0.18	0	1
Divorced	0.08	0.27	0	1
Widowed	0.00	0.06	0	1
Regional				
Urban	0.79	0.41	0	1
Northeast	0.18	0.39	0	1
South	0.38	0.49	0	1
West	0.18	0.38	0	1
Employment				
Job Tenure	3.99	4.70	0	28
Work Experience	9.80	6.49	1	28
Part Time	0.17	0.37	0	1
School				
High School Graduate	0.21	0.41	0	1
Some College	0.15	0.36	0	1
College Graduate	0.05	0.22	0	1
Attending School	0.10	0.30	0	1
Instruments				
Sibling Diabetes	0.05	0.21	0	1
Sibling Diabetes Length	0.10	1.28	0	47

Notes: Summary statistics are for all working individuals and contain 70,810 observations. *Except employment which contains .96,401

TABLE 2. COMPARATIVE STATISTICS OF DIABETICS AND NON-DIABETICS

	Diabetic				Non-Diabetic			
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Key Variables								
Hourly Wage	13.48	10.98	1	123	11.75	14.81	1	500
Employment*	0.71	0.46	0	1	0.73	0.44	0	1
Diabetes	1.00	0.00	1	1	--	--	--	--
Diabetes Length	8.90	7.44	1	44	--	--	--	--
Demographic								
Male	0.46	0.50	0	1	0.56	0.50	0	1
Age	36.73	6.96	16	49	29.66	7.51	15	49
Family Size	3.44	1.71	1	10	3.32	1.87	1	15
Number of Children	1.28	1.23	0	5	0.83	1.13	0	9
Black	0.23	0.42	0	1	0.17	0.38	0	1
Hispanic	0.30	0.46	0	1	0.27	0.45	0	1
Married	0.58	0.49	0	1	0.44	0.50	0	1
Separated	0.03	0.18	0	1	0.04	0.18	0	1
Divorced	0.13	0.34	0	1	0.08	0.27	0	1
Widowed	0.01	0.12	0	1	0.00	0.06	0	1
Regional								
Urban	0.75	0.43	0	1	0.79	0.41	0	1
Northeast	0.19	0.40	0	1	0.18	0.39	0	1
South	0.42	0.49	0	1	0.38	0.49	0	1
West	0.17	0.38	0	1	0.18	0.38	0	1
Employment								
Job Tenure	6.68	6.41	0	28	3.95	4.66	0	28
Work Experience	15.31	6.86	1	28	9.72	6.45	1	28
Part Time	0.13	0.34	0	1	0.17	0.38	0	1
School								
High School Graduate	0.23	0.42	0	1	0.21	0.41	0	1
Some College	0.15	0.36	0	1	0.15	0.36	0	1
College Graduate	0.06	0.23	0	1	0.05	0.22	0	1
Attending School	0.05	0.22	0	1	0.10	0.30	0	1
Instruments								
Sibling Diabetes	0.13	0.34	0	1	0.05	0.21	0	1
Sibling Diabetes Length	0.81	3.02	0	23	0.09	1.24	0	47

Notes: Summary statistics are for all working individuals. Statistics for diabetics contain 1,045 observations and non-diabetics contain 69,765 observations. * Except employment which contains 1,482 diabetics and 94,919 non-diabetics.

TABLE 3. EFFECT OF DIABETES ON EMPLOYMENT

	Whole Sample			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.1597**	-0.0220***	-0.0045	-0.1732	-0.0122	-0.0142	-0.1411	-0.0290***	-0.0128
	(0.0779)	(0.0086)	(0.0142)	(0.1196)	(0.0142)	0.0274	(0.1032)	(0.0110)	(0.0178)
Diabetes^2			-0.0010*			0.0000			-0.0008
			(0.0006)			(0.0015)			(0.0007)
Log Likelihood	-26199	-26199	-26197	-12993	-12993	-12990	-12896	-12894	-12892
Wald Statistic	$X^2(121)$= 28593***	$X^2(121)$= 28591***	$X^2(123)$= 28593***	$X^2(120)$= 13375***	$X^2(120)$= 13376***	$X^2(122)$= 13373***	$X^2(120)$= 14906***	$X^2(120)$= 14901***	$X^2(122)$= 14900***
Joint Tests									
Diabetes			10.23***			1.00			9.14***
p-value			(0.006)			(0.606)			(0.010)
Fixed-effect	1431***	1433***	1437***	867***	867***	871***	656***	658***	661***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 96,332 observations; male estimations contain 51,001 observations; and female estimations contain 45,331 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity.

TABLE 4. INSTRUMENTED EFFECT OF DIABETES ON EMPLOYMENT

	Whole Sample			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	0.3419	-0.0283	0.2096	0.0595	-0.0626	-0.5009	0.5471	-0.0022	0.0420
	(1.1435)	(0.0704)	(0.6525)	(1.7503)	(0.1111)	(1.5328)	(1.5349)	(0.0934)	(0.8102)
Diabetes^2			-0.0131			0.0221			-0.0023
			(0.0353)			(0.0837)			(0.0438)
Log Likelihood	-26202	-26202	-26200	-12994	-12993	-12986	-12898	-12898	-12898
Wald Statistic	$X^2(121)=$ 28603***	$X^2(121)=$ 28602***	$X^2(123)=$ 28600***	$X^2(120)=$ 13379***	$X^2(120)=$ 13379***	$X^2(122)=$ 13368***	$X^2(120)=$ 14913***	$X^2(120)=$ 14912***	$X^2(122)=$ 14913***
Joint Tests									
Diabetes			0.35			0.95			0.00
p-value			(0.839)			(0.622)			(0.999)
Fixed-effect	1432***	1432***	1436***	866***	866***	875***	657***	657***	658***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Tests									
Endogeneity	5.04**	2.50	3.20*	5.41**	9.30***	9.23***	0.78	0.51	0.33
p-value	(0.025)	(0.114)	(0.074)	(0.020)	(0.002)	(0.002)	(0.376)	(0.474)	(0.566)
Diabetes	11.31***	10.60***	124.28***	8.17***	5.48**	130.59***	3.63*	5.07*	47.82***
p-value	(0.001)	(0.001)	(0.000)	(0.004)	(0.019)	(0.000)	(0.057)	(0.024)	(0.000)
Employment	3.43*	0.00	4.17	0.39	1.26	1.31	6.58***	6.80***	6.82**
p-value	(0.064)	(0.947)	(0.124)	(0.532)	(0.261)	(0.520)	(0.010)	(0.009)	(0.033)

Notes: Whole sample estimations contain 96,332 observations; male estimations contain 51,001 observations; and female estimations contain 45,331 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's employment or the error term from estimation of employment.

TABLE 5. EFFECT OF DIABETES ON WAGES

	Whole Sample			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.1299***	-0.0108***	-0.0076	-0.1361*	-0.0065	-0.0157	-0.0979	-0.0119**	-0.0006
	(0.0459)	(0.0039)	(0.0091)	(0.0711)	(0.0075)	(0.0130)	(0.0674)	(0.0051)	(0.0106)
Diabetes^2			-0.0001			0.0005			-0.0005
			(0.0003)			(0.0006)			(0.0005)
R^2	0.6986	0.6987	0.6987	0.6708	0.6709	0.6707	0.7227	0.7229	0.7230
Wald Statistic	$X^2(142)$= 42718***	$X^2(142)$= 42719***	$X^2(144)$= 42821***	$X^2(141)$= 23583***	$X^2(141)$= 23607***	$X^2(143)$= 23726***	$X^2(141)$= 22774***	$X^2(141)$= 22756***	$X^2(143)$= 22742***
Joint Tests									
Diabetes			5.76*			1.48			3.98
p-value			(0.056)			(0.478)			(0.137)
IMR	286***	304***	305***	202***	142***	144***	181***	232***	222***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	4845***	5727***	4696***	3870***	2589***	3643***	2905***	2322***	2040***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 70,766 observations; male estimations contain 39,655 observations; and female estimations contain 31,111 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively.

TABLE 6. INSTRUMENTED EFFECT OF DIABETES ON WAGES

	Whole Sample			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	0.1224	-0.0132	0.8833	-1.1085	-0.1338	4.7756	1.3871**	0.0954	-4.0483
	(0.5335)	(0.0589)	(4.3583)	(0.8644)	(0.0812)	(5.8908)	(0.6802)	(0.0623)	(4.4515)
Diabetes^2			-0.0586			-0.3152			0.2675
			(0.2833)			(0.3833)			(0.2909)
R^2	0.6987	0.6987	0.6988	0.6710	0.6710	0.6713	0.7230	0.7231	0.7231
Joint Tests									
Diabetes			0.04			1.62			2.16
p-value			(0.979)			(0.446)			(0.339)
IMR	291***	375***	324***	139***	144***	202***	147***	149***	168***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	5036***	8164***	5919***	3259***	2297***	2334***	2191***	2516***	3174***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Tests									
Endogeneity	3.73*	3.41*	4.00**	1.30	0.02	0.02	2.05	6.94***	7.66***
p-value	(0.054)	(0.065)	(0.045)	(0.254)	(0.888)	(0.891)	(0.152)	(0.008)	(0.006)
Diabetes	11.31***	10.60***	124.28***	8.17***	5.48**	130.59***	3.63*	5.07**	47.82***
p-value	(0.001)	(0.001)	(0.000)	(0.004)	(0.019)	(0.000)	(0.057)	(0.024)	(0.000)
Wages	0.21	0.01	0.18	0.86	1.98	1.90	4.14**	3.33*	5.08*
p-value	(0.646)	(0.939)	(0.915)	(0.355)	(0.160)	(0.388)	(0.042)	(0.068)	(0.079)

Notes: Whole sample estimations contain 70,766 observations; male estimations contain 39,655 observations; and female estimations contain 31,111 observations. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Wages) is the instruments are not correlated with a person's wages or the error term from estimation of wages.

APPENDIX

TABLE A-1. BREAKDOWN OF INDUSTRY CLASSIFICATIONS FOR DIABETICS AND NON-DIABETICS

	Diabetic		Non-Diabetic	
	Mean	S.D.	Mean	S.D.
Agriculture	0.013	0.115	0.025	0.156
Mining	0.003	0.054	0.005	0.072
Utilities	0.006	0.076	0.001	0.037
Construction	0.035	0.185	0.067	0.250
Manufacturing	0.149	0.357	0.165	0.371
Retail	0.128	0.335	0.188	0.391
Transportation	0.038	0.192	0.053	0.223
Information	0.014	0.119	0.003	0.054
Finance	0.056	0.229	0.053	0.224
Real Estate	0.007	0.082	0.002	0.045
Management	0.023	0.150	0.006	0.076
Education	0.048	0.214	0.011	0.105
Social Services	0.061	0.240	0.017	0.130
Entertainment	0.005	0.069	0.013	0.113
Food	0.015	0.123	0.005	0.070
Other	0.011	0.102	0.005	0.073
Public	0.027	0.162	0.044	0.206
Professional	0.128	0.335	0.143	0.350
Business	0.033	0.180	0.056	0.229
Personal	0.016	0.127	0.035	0.184
Total	0.817		0.897	

Notes: Summary statistics are for all working individuals. Statistics for diabetics contain 1,045 observations and non-diabetics contain 69,765 observations.

TABLE A-2. COMPARISON OF DIFFERENT ESTIMATION METHODS OF DIABETES

	OLS	FE	FE-IMR	IV
Diabetes (0/1)	-0.1153***	-0.1250**	-0.1299***	0.1224
	(0.0442)	(0.0498)	(0.0493)	(0.5335)
R^2	0.6394	0.6980	0.6986	0.6987
Wald Statistic	$X^2(62)=$ 32960***	$X^2(121)=$ 40142***	$X^2(142)=$ 42718***	$X^2(142)=$ 42614***
Joint Tests				
Fixed-effect		1751***	1383***	291***
p-value		(0.000)	(0.000)	(0.000)
IMR			288***	287.48***
p-value			(0.000)	(0.000)
Endogeneity				3.73*
p-value				(0.054)

Notes: *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. Standard errors are reported in parentheses, except where otherwise noted. The null for both the *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively. The null of the Durbin–Wu–Hausman Test for Endogeneity is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages.

TABLE A-3. 'FULL' ESTIMATION OF WAGES

	Estimate	S.D.
Diabetes	-0.130***	(0.046)
Male	0.200***	(0.010)
Age	0.008	(0.009)
Family Size	-0.019***	(0.002)
Number of Children	0.005	(0.006)
Black	-0.017	(0.012)
Hispanic	-0.127***	(0.012)
Married	0.101***	(0.014)
Separated	0.051**	(0.026)
Divorced	0.077***	(0.021)
Widowed	-0.100	(0.092)
Urban	0.017	(0.015)
Northeast	-0.037	(0.046)
South	-0.002	(0.033)
West	0.074*	(0.039)
Job Tenure	0.017***	(0.001)
Work Experience	0.048***	(0.003)
Part Time	-0.879***	(0.010)
High School Graduate	-0.083***	(0.016)
Some College	-0.060***	(0.019)
College Graduate	-0.082***	(0.020)
Attending School	-0.331***	(0.013)
R^2	0.6986	
Wald Statistic	$X^2(142)=$ 42718***	

Notes: Estimation contains 70,766 observations. Standard errors are presented in parenthesis. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. Also included in estimation, but not reported, are indicator variables for industry and year; IMRs for each year; and time invariant means of all explanatory variables.

TABLE A-4. EFFECT OF DIABETES ON EMPLOYMENT FOR BLACK SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.1515	-0.0190	-0.0257	-0.2633	0.0027	-0.0160	-0.0273	-0.0417	-0.0176
	(0.1473)	0.0174	(0.0324)	(0.2105)	(0.0239)	(0.0444)	(0.2138)	(0.0262)	(0.0525)
Diabetes^2			0.0003			0.0009			-0.0016
			(0.0017)			(0.0022)			(0.0031)
Log Likelihood	-4677	-4679	-4675	-2375	-2376	-2375	-2191	-2191	-2188
Wald Statistic	$X^2(119) = 5140$***	$X^2(119) = 5140$***	$X^2(121) = 5138$***	$X^2(118) = 2309$***	$X^2(118) = 2313$***	$X^2(120) = 2313$***	$X^2(118) = 2654$***	$X^2(118) = 2654$***	$X^2(120) = 2649$***
Joint Tests									
Diabetes			1.59			0.18			2.67
p-value			(0.451)			(0.914)			(0.263)
Fixed-effect	265***	265***	270***	169***	172***	172***	147***	145***	150***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 16,777 observations; male estimations contain 9,152 observations; and female estimations contain 7,625 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity and selection bias, respectively.

TABLE A-5. INSTRUMENTED EFFECT OF DIABETES ON EMPLOYMENT FOR BLACK SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-3.4769	-0.2539	-0.9049	-6.0884*	-0.1205	-7.5483*	-2.1196	-0.4035*	2.1202
	(2.1797)	(0.1623)	(3.0561)	(3.3608)	(0.2738)	(4.0467)	(3.0038)	(0.2142)	(3.4840)
Diabetes^2			0.0373			0.4109*			-0.1374
			(0.1703)			(0.2227)			(0.1923)
Log Likelihood	-4679	-4678	-4678	-2374	-2376	-2374	-2192	-2190	-2190
Wald Statistic	$X^2(119)$= 5144***	$X^2(119)$= 5138***	$X^2(121)$= 5139***	$X^2(118)$= 2311***	$X^2(118)$= 2311***	$X^2(120)$= 2309***	$X^2(118)$= 2656***	$X^2(118)$= 2655***	$X^2(120)$= 2654***
Joint Tests									
Diabetes			2.59			3.58			4.02
p-value			(0.274)			(0.167)			(0.134)
Fixed-effect	268***	270***	265***	177***	174***	172***	146***	147***	147***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Tests									
Endogeneity	2.00	0.00	1.89	1.92	0.59	0.00	3.10*	0.32	0.26
p-value	(0.157)	(0.991)	(0.170)	(0.166)	(0.442)	(0.951)	(0.078)	(0.573)	(0.610)
Diabetes	2.45	7.31***	952.40***	3.93**	11.33***	4636.79***	0.14	1.84	215.66***
p-value	(0.117)	(0.007)	(0.000)	(0.047)	(0.001)	(0.000)	(0.710)	(0.175)	(0.000)
Employment	3.55*	1.76	5.51*	0.68	0.17	41.89***	0.62	0.72	0.81
p-value	(0.060)	(0.184)	(0.064)	(0.408)	(0.680)	(0.000)	(0.431)	(0.397)	(0.668)

Notes: Whole sample estimations contain 12,072 observations; male estimations contain 7,102 observations; and female estimations contain 4,970 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's wages or the error term from estimation of wages. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's employment or the error term from estimation of employment.

TABLE A-6. EFFECT OF DIABETES ON EMPLOYMENT FOR HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.1062	-0.0266*	0.0241	0.1563	0.0073	0.0474	-0.3280**	-0.0507**	0.0338
	(0.1316)	(0.0157)	(0.0257)	(0.2223)	(0.0272)	(0.0463)	(0.1644)	(0.0206)	(0.0379)
Diabetes^2			-0.0028**			-0.0025			-0.0055***
			(0.0011)			(0.0023)			(0.0021)
Log Likelihood	-8404	-8402	-8399	-4523	-4523	-4522	-3791	-3788	-3784
Wald Statistic	$X^2(119)$= 9457***	$X^2(119)$= 9453***	$X^2(121)$= 9448***	$X^2(118)$= 4755***	$X^2(118)$= 4753***	$X^2(120)$= 4752***	$X^2(118)$= 4624***	$X^2(118)$= 4617***	$X^2(120)$= 4608***
Joint Tests									
Diabetes			10.14***			1.26			13.13***
p-value			(0.006)			(0.532)			(0.001)
Fixed-effect	539***	540***	541***	356***	356***	357***	239***	241***	245***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 29,228 observations; male estimations contain 15,456 observations; and female estimations contain 13,772 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity and selection bias, respectively.

TABLE A-7. INSTRUMENTED EFFECT OF DIABETES ON EMPLOYMENT FOR BLACK SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.2114	-0.0216	-0.8248	-0.7394	-0.2834*	-0.6547	0.5849	0.4020*	-10.3988
	(2.0344)	(0.1208)	(1.7399)	(2.9126)	(0.1557)	(1.8225)	(2.8958)	(0.2192)	(7.4369)
Diabetes^2			0.0440			0.0194			0.6069
			(0.0950)			(0.0983)			(0.4248)
Log Likelihood	-8405	-8404	-8400	-4523	-4521	-4518	-3796	-3794	-3791
Wald Statistic	$X^2(119)=$ 9462***	$X^2(119)=$ 9461***	$X^2(121)=$ 9453***	$X^2(118)=$ 4752***	$X^2(118)=$ 4751***	$X^2(120)=$ 4746***	$X^2(118)=$ 4631***	$X^2(118)=$ 4630***	$X^2(120)=$ 4624***
Joint Tests									
Diabetes			0.25			3.86			2.99
p-value			(0.884)			(0.145)			(0.225)
Fixed-effect	538***	539***	544***	359***	362***	363***	239***	240***	243***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Tests									
Endogeneity	0.01	1.11	0.03	0.30	0.27	2.64	0.27	0.11	0.11
p-value	(0.930)	(0.292)	(0.864)	(0.585)	(0.605)	(0.105)	(0.602)	(0.741)	(0.740)
Diabetes	3.62*	1.99	21.69***	3.73*	2.05	37.03***	0.27	0.16	0.98
p-value	(0.057)	(0.159)	(0.000)	(0.053)	(0.153)	(0.000)	(0.604)	(0.691)	(0.419)
Employment	2.07	4.43**	5.62*	4.67**	5.63**	2.08	1.70	2.45	2.98
p-value	(0.150)	(0.035)	(0.060)	(0.031)	(0.018)	(0.353)	(0.193)	(0.118)	(0.225)

Notes: Whole sample estimations contain 29,228 observations; male estimations contain 15,456 observations; and female estimations contain 13,772 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's wages or the error term from estimation of wages. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's employment or the error term from estimation of employment.

TABLE A-8. EFFECT OF DIABETES ON EMPLOYMENT FOR NON-BLACK NON-HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.2835**	-0.0247*	-0.0188	-0.4119**	-0.0403	-0.0906*	-0.1020	-0.0177	-0.0054
	(0.1311)	(0.0129)	(0.0219)	(0.2022)	(0.0245)	(0.0509)	(0.1732)	(0.0156)	(0.0263)
Diabetes^2			-0.0003			0.0028			-0.0005
			(0.0009)			(0.0031)			(0.0009)
Log Likelihood	-12959	-12960	-12960	-5947	-5948	-5945	-6755	-6755	-6754
Wald Statistic	$X^2(119)$= 13692***	$X^2(119)$= 13694***	$X^2(121)$= 13695***	$X^2(118)$= 6040***	$X^2(118)$= 6039***	$X^2(120)$= 6034***	$X^2(118)$= 7434***	$X^2(118)$= 7433***	$X^2(120)$= 7434***
Joint Tests									
Diabetes			3.76			4.80*			1.65
p-value			(0.152)			(0.091)			(0.437)
Fixed-effect	711***	712***	712***	451***	450***	453***	367***	368***	369***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 50,327 observations; male estimations contain 26,393 observations; and female estimations contain 23,934 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity and selection bias, respectively.

TABLE A-9. INSTRUMENTED EFFECT OF DIABETES ON EMPLOYMENT FOR NON-BLACK NON-HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	3.0844*	0.0178	0.4916	5.1793*	0.2514	0.4856	2.5083	-0.0376	0.2110
	(1.7951)	(0.1048)	(0.7333)	(2.9776)	(0.2297)	(3.1824)	(2.2949)	(0.1271)	(0.8681)
Diabetes^2			-0.0251			-0.0182			-0.0131
			(0.0391)			(0.1755)			(0.0464)
Log Likelihood	-12960	-12962	-12961	-5947	-5948	-5944	-6755	-6755	-6755
Wald Statistic	$X^2(119)$= 13694***	$X^2(119)$= 13696***	$X^2(121)$= 13696***	$X^2(118)$= 6036***	$X^2(118)$= 6039***	$X^2(120)$= 6035***	$X^2(118)$= 7436***	$X^2(118)$= 7434***	$X^2(120)$= 7434***
Joint Tests									
Diabetes			0.48			0.64			0.14
p-value			(0.789)			(0.726)			(0.932)
Fixed-effect	708***	707***	709***	451***	445***	453***	367***	367***	367***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Tests									
Endogeneity	4.86**	4.50**	10.78***	20.32***	17.78***	24.46***	2.42	0.31	0.06
p-value	(0.028)	(0.034)	(0.001)	(0.000)	(0.000)	(0.000)	(0.120)	(0.577)	(0.804)
Diabetes	5.62**	4.46**	63.54***	0.99	1.37	9.84***	5.07**	3.60*	46.46***
p-value	(0.018)	(0.035)	(0.000)	(0.320)	(0.242)	(0.000)	(0.024)	(0.058)	(0.000)
Employment	14.51***	14.76***	15.21***	8.84***	8.72***	9.70***	3.49*	3.66*	3.76
p-value	(0.000)	(0.000)	(0.001)	(0.003)	(0.003)	(0.008)	(0.062)	(0.056)	(0.153)

Notes: Whole sample estimations contain 50,327 observations; male estimations contain 26,393 observations; and female estimations contain 23,934 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for *Fixed-effect* is no indication of unobserved heterogeneity. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's wages or the error term from estimation of wages. The null of the Instrument Test (Employment) is the instruments are not correlated with a person's employment or the error term from estimation of employment.

TABLE A-10. EFFECT OF DIABETES ON WAGES FOR BLACK SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.0599	-0.0055	-0.0077	-0.0525	0.0016	-0.0248	-0.0077	-0.0069	0.0138
	(0.0725)	(0.0088)	(0.0203)	(0.0744)	(0.0121)	(0.0220)	(0.1068)	(0.0148)	(0.0261)
Diabetes^2			0.0002			0.0019			-0.0013
			(0.0012)			(0.0013)			(0.0020)
R^2	0.7389	0.7392	0.7391	0.7078	0.7081	0.7078	0.7832	0.7833	0.7835
Wald Statistic	$X^2(140)$= 9790***	$X^2(140)$= 9810***	$X^2(142)$= 9870***	$X^2(139)$= 5824***	$X^2(139)$= 5778***	$X^2(141)$= 6327***	$X^2(138)$= 59943***	$X^2(138)$= 60615***	$X^2(140)$= 60533***
Joint Tests									
Diabetes			0.32			2.51			0.42
p-value			(0.852)			(0.285)			(0.813)
IMR	105***	87***	103***	70***	67***	74***	72***	88***	52***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	1197***	1175***	813***	693***	595***	560***	1665***	10584***	3900000***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 12,072 observations; male estimations contain 7,102 observations; and female estimations contain 4,970 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for both *Fixed-effect* is no indication of unobserved heterogeneity

Table A-11. Instrumented Effect of Diabetes on Wages for Black Sample

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	0.3995	0.0804	-0.0154	-1.4922	0.0265	-0.5347	1.1839	0.1018	0.0553
	(1.2635)	(0.0558)	(0.2840)	(1.7620)	(0.0984)	(0.5834)	(1.3813)	(0.0885)	(0.2624)
Diabetes^2			0.0046			0.0293			0.0019
			0.0127			(0.0283)			(0.0122)
R^2	0.7391	0.7391	0.7391	0.7079	0.7080	0.7079	0.7838	0.7834	0.7842
Joint Tests									
Diabetes			2.23			1.30			1.34
p-value			(0.328)			(0.523)			(0.513)
IMR	157***	87***	94***	62***	53***	67***	54***	72***	62***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)
Fixed-effect	1095***	948***	978***	748***	512***	611***	12006***	9933***	83122***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Test									
Endogeneity	2.28	3.86**	7.48***	0.05	0.87	0.97	1.68	3.87**	5.70**
p-value	(0.131)	(0.049)	(0.006)	(0.819)	(0.350)	(0.326)	(0.195)	(0.049)	(0.017)
Diabetes	2.45	7.31***	952.40***	3.93**	11.33***	4636.79***	0.14	1.84	215.66***
p-value	(0.117)	(0.007)	(0.000)	(0.047)	(0.001)	(0.000)	(0.710)	(0.175)	(0.000)
Wages	0.18	2.37	4.77*	0.48	0.12	0.79	0.66	1.74	4.00
p-value	(0.673)	(0.124)	(0.092)	(0.488)	(0.731)	(0.672)	(0.417)	(0.187)	(0.135)

Notes: Whole sample estimations contain 12,072 observations; male estimations contain 7,102 observations; and female estimations contain 4,970 observations. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Wages) is the instruments are not correlated with a person's wages or the error term from estimation of wages.

TABLE A-12. EFFECT OF DIABETES ON WAGES FOR HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.1195	0.0003	-0.0117	-0.1447	-0.0031	-0.0090	-0.0992	0.0026	-0.0261
	(0.0809)	(0.0081)	(0.0167)	(0.1360)	(0.0085)	(0.0257)	(0.1199)	(0.0110)	(0.0351)
Diabetes^2			0.0007			0.0003			0.0020
			(0.0010)			(0.0013)			(0.0023)
R^2	0.6858	0.6856	0.6856	0.6927	0.6924	0.6920	0.7092	0.7093	0.7094
Wald Statistic	$X^2(140)=$ 14032***	$X^2(140)=$ 14026***	$X^2(142)=$ 14086***	$X^2(139)=$ 9458***	$X^2(139)=$ 9467***	$X^2(141)=$ 10528***	$X^2(139)=$ 16586***	$X^2(139)=$ 16461***	$X^2(141)=$ 16643***
Joint Tests									
Diabetes			0.51			0.15			0.84
p-value			(0.776)			(0.927)			(0.656)
IMR	191***	213***	184***	78***	102***	118***	111***	170***	127***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	1707***	1337***	2398***	937***	847***	1140***	1134***	907***	1468***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 19,420observations; male estimations contain 10,705 observations; and female estimations contain 8,715 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively.

TABLE A-13. INSTRUMENTED EFFECT OF DIABETES ON WAGES FOR HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	0.3688	-0.1563	-1.2357	-0.4789	-0.6056**	-3.3610	1.9594**	0.4654***	1.5458
	(0.8285)	(0.2045)	(1.4073)	(1.2248)	(0.2957)	(2.1445)	(0.8702)	(0.1541)	(1.0382)
Diabetes^2			0.1482			0.3780			-0.1475
			(0.1600)			(0.2423)			(0.1217)
R^2	0.6858	0.6856	0.6856	0.6923	0.6925	0.6928	0.7099	0.7097	0.7097
Joint Tests									
Diabetes			1.15			2.46			6.67**
p-value			(0.563)			(0.293)			(0.036)
IMR	245***	189***	204***	126***	148***	87***	145***	156***	173***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	1180***	1700***	1398***	903***	1046***	743***	1624***	2076***	1150***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Test									
Endogeneity	0.57	0.01	0.00	0.42	0.00	0.05	0.21	0.00	0.14
p-value	(0.451)	(0.939)	(0.990)	(0.518)	(0.974)	(0.823)	(0.649)	(0.985)	(0.712)
Diabetes	3.62*	1.99	21.69***	3.73*	2.05	37.03***	0.27	0.16	0.98
p-value	(0.057)	(0.159)	(0.000)	(0.053)	(0.153)	(0.000)	(0.604)	(0.691)	(0.419)
Wages	0.33	0.57	1.74	0.03	4.57**	13.64***	4.53**	10.58***	19.88***
p-value	(0.563)	(0.450)	(0.419)	(0.860)	(0.033)	(0.001)	(0.033)	(0.001)	(0.000)

Notes: Whole sample estimations contain 19,420observations; male estimations contain 10,705 observations; and female estimations contain 8,715 observations. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Wages) is the instruments are not correlated with a person's wages or the error term from estimation of wages.

TABLE A-14. EFFECT OF DIABETES ON WAGES FOR NON-BLACK NON-HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.1906**	-0.0201***	-0.0139	-0.2076*	-0.0151	-0.0310	-0.1174	-0.0181***	-0.0047
	(0.0905)	(0.0060)	(0.0125)	(0.1226)	(0.0139)	(0.0315)	(0.1167)	(0.0067)	(0.0132)
Diabetes^2			-0.0003			0.0009			-0.0005
			(0.0004)			(0.0015)			(0.0006)
R^2	0.7062	0.7062	0.7062	0.6679	0.6680	0.6680	0.7384	0.7384	0.7384
Wald Statistic	$X^2(140)$= 25062***	$X^2(140)$= 25084***	$X^2(142)$= 25141***	$X^2(139)$= 14011***	$X^2(139)$= 14027***	$X^2(141)$= 14077***	$X^2(139)$= 14165***	$X^2(139)$= 14164***	$X^2(141)$= 14249***
Joint Tests									
Diabetes			9.28***			1.11			5.66*
p-value			(0.010)			(0.573)			(0.059)
IMR	147***	150***	166***	57***	65***	61***	98***	96***	118***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	3527***	2872***	3414***	1810***	1956***	2380***	1737***	1481***	1373***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Whole sample estimations contain 39,274observations; male estimations contain 21,848 observations; and female estimations contain 17,426 observations. Standard errors are presented in parenthesis, except where otherwise noted. *, **, and *** represent significance at the 10, 5, and 1 percent level, respectively. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively.

TABLE A-15. INSTRUMENTED EFFECT OF DIABETES ON WAGES FOR NON-BLACK NON-HISPANIC SAMPLE

	All			Male			Female		
	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)	Model 1 (0/1)	Model 2 (Linear)	Model 3 (Quadratic)
Diabetes	-0.4425	-0.0163	2.1614	-2.2005	-0.0194	1.8858	1.0800	-0.0018	1.9546
	(1.2732)	(0.0622)	(2.4790)	(1.9448)	(0.0738)	(4.9570)	(1.5305)	(0.0792)	(6.2900)
Diabetes^2			-0.1390			-0.1229			-0.1238
			(0.1560)			(0.3109)			(0.3928)
R^2	0.7063	0.7063	0.7064	0.6683	0.6683	0.6688	0.7385	0.7385	0.7386
Joint Tests									
Diabetes			1.42			0.81			0.25
p-value			(0.492)			(0.668)			(0.882)
IMR	122***	141***	164***	76***	74***	70***	109***	133***	113***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed-effect	3482***	2640***	3628***	2306***	2377***	1906***	1087***	1301***	1283***
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Instrument Test									
Endogeneity	1.98	8.04***	9.70***	2.16	1.98	2.01	1.20	10.03***	8.28***
p-value	(0.159)	(0.005)	(0.002)	(0.141)	(0.159)	(0.156)	(0.273)	(0.002)	(0.004)
Diabetes	5.62**	4.46**	63.54***	0.99	1.37	9.84***	5.07**	3.60*	46.46***
p-value	(0.018)	(0.035)	(0.000)	(0.320)	(0.242)	(0.000)	(0.024)	(0.058)	(0.000)
Wages	0.04	0.00	13.52***	1.26	0.06	4.22	0.81	0.08	1.18
p-value	(0.834)	(0.981)	(0.001)	(0.262)	(0.811)	(0.121)	(0.368)	(0.781)	(0.554)

Notes: Whole sample estimations contain 39,274observations; male estimations contain 21,848 observations; and female estimations contain 17,426 observations.. The null hypothesis for both *Fixed-effect* and *IMR* is no indication of unobserved heterogeneity and selection bias, respectively. The null of the Durbin–Wu–Hausman Test is no indication of an endogeneity bias, or diabetes is exogenous with respect to wages. The null of the Instrument Test (Diabetes) is the instruments do not significantly explain own diabetes. Model 3 presents the results of a multivariate test of significance of all instruments on all diabetes specifications. The null of the Instrument Test (Wages) is the instruments are not correlated with a person's wages or the error term from estimation of wages.

CHAPTER 2
THE EFFECT OF DIABETES ON LABOR FORCE DECISIONS: NEW EVIDENCE FROM THE NATIONAL HEALTH INTERVIEW SURVEY

2.1 INTRODUCTION

The prevalence of diabetes among Americans has become a major concern, with an estimated 23.6 million people suffering from the disease today. This represents an annual increase of approximately five percent since 1990 (CDC 2007). Diabetes was listed as the seventh leading cause of death, and the fifth deadliest disease in 2006 by the Center for Disease Control (CDC). This may be underreported, as most analysts suggest that complicating factors such as stroke, hypertension, or old age may be confounding the effect of diabetes. Since 1987, the death rates of heart disease, stroke, and cancer have all declined. In contrast, the death rate attributable to diabetes has increased by 45 percent (CDC 2007). The growth of reported diabetes cases imposes substantial direct and indirect medical costs on individuals. Estimates from the American Diabetes Association suggest that diabetes accounts for $92 billion in direct medical costs and approximately an additional $40 billion in indirect costs.[1] The CDC estimates that about 20.6 million Americans over the age of 20 suffer from some form of diabetes. As this population also accounts for the majority of the American workforce, diabetes could have a significant impact on the U.S. labor market. Due to their illnesses, individuals diagnosed with diabetes may be less productive, miss more days of work, and even earn less than those free from the disease.

---

[1] The American Diabetes Association defines indirect costs as disability, work loss, and premature mortality.

Much of the previous literature on this topic makes no clear distinction between type I and type II diabetes despite basic fundamental differences in the two types of the disease. This could be problematic as these differences could cause type I and type II diabetes to have drastically different effects on workers. In this paper, a distinction between type I and type II diabetes is made for two important reasons. First, type I diabetes is typically diagnosed early on in life, and it is a genetic disorder. This provides a clear, exogenous source of variation. Type II diabetes, on the other hand, can occur at any point during a person's lifetime, and its onset may be linked to the individual's weight, diet, or numerous other health factors that change over a person's lifetime. The causes of type II diabetes are highly contested and it is not clear whether diet and exercise are deterrents to the disease. However, research indicates that proper diet and exercise greatly reduce the risk of developing type II diabetes, suggesting that the disease is at least partially a result of lifestyle choices rather than predetermined.[2] Second, type II diabetes accounts for approximately 90 to 95 percent of all reported diabetes cases.

This paper extends the diabetes literature by initially examining the impact of type I diabetes on labor market behavior. Type I diabetes is examined to show the effect of an exogenously determined case of diabetes on the labor market. Then type II diabetes is included in estimation to see if the two diseases affect the labor market differently. I use data from the 2006 National Health Interview Survey (NHIS) to estimate the impact of diabetes on labor-force participation, days out of work, average hours worked, and earnings. Because type II diabetes may be subject to simultaneity bias with respect to labor market decisions, it may be necessary implement instrumental variable estimation. I

---

[2] See CDC 1999. This distinction will be discussed in further detail in Section II.

use information on whether an individual takes medication to reduce their blood sugar levels as an instrument for diabetes. This instrument should be highly correlated with whether an individual has diabetes, as high blood sugar is a direct result of the disease, if not properly managed. However, high blood sugar itself, whether a result of diabetes or arising on its own in non-diabetics, should not affect labor force decisions on the same scale as type II diabetes.[3] Proper instrumentation will show the direct, causal relationship of type II diabetes on labor force decisions, rather than statistical correlations obtained through ordinary least squares estimates. Results show that type I diabetes negatively impacts numerous work outcomes, including earnings, with the average male type I diabetic losing nearly 17 percent of his annual income and the average type II diabetic losing about eight percent. Interestingly, type I diabetes generates wage and productivity losses that are only slightly larger than those from type II diabetes even though type I diabetes is generally regarded as a more debilitating condition.

The remainder of the paper is arranged as follows. Section II provides some background information on diabetes. Section III summarizes the existing literature associated with diabetes. Section IV presents data sources, and Section V describes the empirical methodology. Section VI presents results, and section VII concludes.

2.2 Diabetes Background

The CDC (2005) defines diabetes as "a group of diseases marked by high levels of blood glucose resulting from defects in insulin production, insulin action, or both." The broad term "diabetes" includes a number of different diseases, all of which are

[3] Theoretical and statistical justification for this instrument is presented in Section V and VI.

related to the body's production on insulin.[4] Type I diabetes occurs when the body's immune system destroys all insulin producing cells. This leaves the individual with no natural means of insulin production. Therefore, they must be administered injections of the hormone daily. Type I diabetes accounts for approximately five to ten percent of all reported diabetes cases, and it typically is diagnosed early in childhood or adolescence, although it may be diagnosed later in life. Other than genetics, there are no known causes and no known cures for type I diabetes. Examination of type I diabetes will be straightforward as it is not a result of any lifestyle choices.

Type II diabetics still produce insulin; however, their cells do not properly process the hormone. As the condition persists, the individual's pancreas may cease to produce insulin in the most severe cases. The treatment for type II diabetes differs from that of type I, with most type II diabetics being able to control their disease through a healthy diet, exercise, weight loss, or oral medication. Only in the advanced stages of type II diabetes, when the pancreas stops producing insulin altogether, is insulin prescribed. In fact, estimates suggest that about 27 percent of type II diabetics take insulin injections on a daily basis (Mayfield 2004). Clinical reports indicate that diagnosis of type II diabetes in children is still rare, with the majority of cases occurring during adulthood. To date, the direct causes and complications of type II diabetes are unknown; however, research suggests that type II diabetes is specifically associated with old age, obesity, genetics, impaired glucose metabolism, and race (CDC 2007). Estimates of type II diabetes effect have potential policy implications, as these effects could be managed or even prevented through proper care and monitoring of the disease. This,

---

[4] This, and much of diabetes background information, is taken from the CDC (2005) Diabetes Fact Sheet.

however, also complicates estimation, as the same factors influencing the disease may be highly correlated with work characteristics leading to biased estimates.

Initial estimation considers only type I diabetics because this disease is exogenously determined. Type I diabetics have a condition that should have no relationship with personal behaviors such as diet, exercise, or weight. Eliminating type II diabetics from estimation is beneficial because it shows the casual effect of an exogenous disease on the labor market, without any potential bias type II diabetics may introduce. Also it may be assumed that in the most severe cases of type II diabetes, when the body stops producing insulin, the effects of type I diabetes on labor market outcomes are comparable to the effects of severe cases of type II diabetes.

2.3 Literature Review

Diabetes has recently become a topic of interest within the economics literature. This is likely due to an increase in both the prevalence of the disease and an improvement in available data. Pango (1999) conducts a survey of the literature and suggests that there is no clear consensus on either the labor market or personal effects of diabetes. He attributes the problem to inconsistent data sources and a disagreement among scholars concerning the appropriate estimation methodology.

Following this, there are two strands of literature on diabetes: those that estimate the direct medical costs to health care providers and employers, and those that estimate the indirect costs to an individual who suffers from the disease. The direct costs of diabetes are fairly well-established. Gilmer et al. (2005) and Oliva et al. (2004) both find significantly increased medical expenditure due to diabetes. Duggan (2006) estimates that the relaxation of eligibility requirements in federally-funded programs increases both expenditures and enrollment of diabetes patients, suggesting that the per capita cost of treating this disease could be rising in the U.S.

Papers attempting to estimate the implicit cost of diabetes have yet to reach agreement. Kahn (1998) estimates that the negative effects of diabetes on labor market participation and earnings are decreasing as a result of technological innovations. These results could be due, in part, to the sample period utilized, which consists of data from 1976, 1989, and 1992. Estimates over these years show an overwhelming increase in productivity and a much slower rise in the prevalence of diabetes. Therefore, the productivity increase may be overshadowing the effect diabetes actually has on job market outcomes. Ramsey et al. (2002) finds that employers face high medical costs stemming from diabetic workers, and he shows that the associated costs are higher for the younger work force.[5] Vijan et al. (2004) and Tuceli et al. (2005) both find that diabetics reduce their weeks and hours worked and experience earnings losses. However, these papers are limited by the data from the Health and Retirement Survey (HRS), which samples only people between 51 and 61 years of age. As diabetes is no longer a disease experienced only by the elderly, these results may not accurately represent the effects on the entire population. Lavigne et al. (2003) estimate diminished productivity due to type II diabetes. However, due to the design of their survey,[6] their analysis may not be nationally representative. Brown et al. (2005) find that diabetes has a negative impact on employment, and they suggest that there may be endogeneity issues associated with the

[5] They estimate a per employee cost of about $4,671 for employees aged 18–35 and $4,369 for those aged 56–64 years, suggesting that all ages should be included in estimating the effects of diabetes.

[6] Their telephone survey respondents consisted of 472 New York state residents who had all reported some type of health claim.

impact of diabetes on labor market outcomes. Utilizing mother and father's ethnicity as an IV for diabetes[7] they find some indication of endogeneity bias, specifically with regard to older females.

This study extends the literature in three ways. First, I examine not only the loss of productivity to the employer, measured as days out of work and hours worked per week, but also the effects of diabetes on individual wage rates and labor-force participation. Second, by using the 2006 NHIS, I am able to obtain the most current and nationally representative estimates. Third, I attempt to address the potential endogeneity bias associated with labor market outcomes and diabetes by instrumenting diabetes with whether a person takes medicine to control their high blood sugar.

2.4 DATA

I use data from the 2006 NHIS. This survey is conduced annually by the National Center for Health Statistics (NCHS), a division of the CDC. The NCHS is widely considered one of the principal sources for civilian health information in the U.S. The survey has been ongoing since 1957, but it received a major revision in 1997. Since then, much more detailed personal health, demographic, and health care information has been collected. Approximately 35,000 households with 87,500 persons are interviewed every year. However, this is not a panel data set that allows researchers to track individuals across multiple years. For this reason, I examine only the most recent year of data available.

[7] In an attempt to replicate the results of this paper, I find that Hispanic country of origin (a proxy for parent's ethnicity) is a weak instrument, not adequately explaining the variance of diabetes in a national sample.

The NHIS collects detailed data on numerous diseases and health-related problems. I examine whether a person has ever been diagnosed with *Type I* or *Type II* diabetes. The NHIS also collects numerous variables on work history within the past year. The variable *Working Last Year* is recorded as a zero-one indicator variable for whether the respondent worked for pay at some point during the last year;[8] *Days Missed* is the number of total days of work missed due to an illness in the last year; *Work Hours* records the average number of hours worked each week during the previous year; *Earnings* is recorded on a discrete scale ranging from zero to eleven.[9]

The NHIS records Body Mass Index (BMI) in the survey as weight divided by height squared, and this measure is used as a proxy for body size.[10] Information on numerous medical conditions including heart disease, kidney disease, and blood pressure are all recorded by the NHIS, as is information on age, sex, marital status, self-reported health,[11] education, and region of residence. Detailed information on personal exercise habits, weight loss, and diet are also recorded for every respondent. The NHIS also gathers information on industry and occupation, reported according to the 2002 North American Industry Classification System specifications.[12]

[8] If you were not in the labor force last year, respondents were not asked to list work characteristics.

[9] One is from $1 to $4,999, and with each successive level increasing in increments of five thousand dollars for the first five levels. After this, increments are ten thousand dollar increases until top-coded at level 11. Level 11 captures all incomes above $75,000.

[10] According to the Center for Disease Control, a BMI of over 30 is obese and over 25 is overweight. This measure is not perfect, as it does not account for fat versus muscle tissue. However, BMI has been used as a proxy to control for issues related to being overweight and obese in the economics literature (see Cutler et. al 2003; Baum and Ford 2004; Ruhm 2007).

[11] The NHIS asks respondents whether they are in better, worse, or the same health as last year. This is used as a proxy for health status, as it tells, taking as given any pre-existing medical conditions, the general well-being of an individual.

[12] Question numbers for industry and occupation are ASD.080_00.000 and ASD.090_00.000, respectively. Unemployed individuals receive a zero for all industry and occupation indicator variables.

The NHIS's measure of diabetes is an indicator for both type I and type II diabetes.[13] One of the fundamental differences in the two types of diabetes is the timing of diagnosis. Type I diabetes, sometimes called juvenile onset diabetes, is typically diagnosed early on in life when the individual is still a child. This is because most type I individuals begin to exhibit symptoms early and must then monitor their condition very closely on a daily basis. Type II diabetes, however, can occur at any point in a person's life. That NHIS does not distinguish between the two diabetes types may be problematic if the effect of type II diabetes is different from that of type I. Because type I diabetes is genetically derived, whereas type II diabetes could arise from personal decisions and lifestyle choices, the distinction between the two potentially becomes even more important. Because the NHIS asks when an individual was diagnosed with diabetes, I am able to exclude all respondents who were diagnosed before the age of 20, separating *Type I* diabetics from the cases of *Type II.*[14]

Table 1 presents summary statistics for the variables used in estimation. Because both the work characteristics and incidence of diabetes are different for men and women, the sample is partitioned by gender.[15] About six percent of males and females in the sample have diabetes.[16] *BMI* ranges from underweight to morbidly obese, with the average person being of healthy weight. The sample is restricted to those of working age, 20 to 65, with the average person aged 41. All labor market outcomes appear to be

---

[13] The survey only asks respondents if they have been diagnosed with diabetes. It does not distinguish between *Type-One* and *Type-Two* diabetes.

[14] CDC (2005) reports that cases of *Type II* diabetes being diagnosed before this age are very rare, and this gives the sample about 7 percent *Type I.*

[15] The estimates were also performed as a pooled sample where an indicator variable for gender was included. The estimated coefficients of diabetes were not significantly different.

[16] Sample weights are not included in any summary statistics or estimations. This is primarily due to the nationally representative summary statistics.

nationally representative. About 80 percent of males were employed last year, with the percentage of women working being somewhat smaller at 57 percent. Of those who worked, males and females worked on average about 42 and 40 hours per week, respectively. This number does not include any values for the approximately 30 percent of the sample that are not employed. The mean of *Days Missed* is approximately 3.5 days per year for men and 4.4 days for women. Lastly, *Earnings* for men average about $32,300 per year, and the average female's annual earning is $30,400.

Table 2 provides a comparison of summary statistics for diabetics and non-diabetics. The data show that diabetics, both male and female, are more likely to have heart disease, kidney disease, and high blood pressure. Diabetics also tend to be somewhere between 3 to 6 BMI points heavier, with the average diabetic being obese. The average age of non-diabetics is around forty while the average diabetic is at or above fifty. Diabetics report both feeling worse than they did last year and being told by a physician to change their lifestyle much more frequently than non-diabetics. Only about 58 percent of males with diabetes reported working last year, which is well below the 82 percent of non-diabetics who worked. Similarly, only about 38 percent of female diabetics participated in the work force in the last year, whereas, nearly 60 percent of non-diabetic females worked. This suggests that there is a significant difference in the labor market decisions made by those people who have diabetes.

2.5 Empirical Methodology

The goal of this study is to estimate the effects of having diabetes on various labor market outcomes. Initial estimation takes the form:

$$WorkLastYear_i = \alpha_0 + \alpha_1 TypeI_i + \alpha_2' X_i + \mu_i. \tag{1}$$

The variable $WorkLastYear_i$ is a binary indicator variable equal to one if person i was employed last year; $Type\ I_i$ is an indicator variable equaling one if person i has type I diabetes and zero otherwise; X_i is a vector of person-specific controls; the α_j are parameters to be estimated; and μ_i represents the idiosyncratic error term. In initial estimations all type II diabetics are eliminated from the sample to show the effect of type I diabetes relative to only non-diabetics. Because X_i includes person-specific variables related to health, personal behavior, demographics, and industry/occupation it should absorb any unobserved heterogeneity related to the individual that otherwise may have been attributed to the impact of diabetes. [17] All variables enter the regression equations as controls, but are not reported along with the main results.[18] Equation (1) is estimated first to test the impact of having diabetes on the probability of working and to correct for any selection bias in subsequent models. Because estimates of all other labor-force outcomes will only include those people who worked last year, I employ Heckman's (1976) correction for selection bias. This takes the form:

$$Y_i = \beta_0 + \beta_1 Type\ I_i + \beta_2' X_i + \beta_3 \lambda_i + \varepsilon_i. \tag{2}$$

The variable Y_i represents the various outcome variables (i.e., *Days Missed*, *Work Hours*, and *Earnings*), and λ_i represents the inverse of Mills ratio, which controls for the

[17] Estimates are also performed with all suspected endogenous controls excluded. Estimates change very slightly with all signs and significance remaining intact. Those dropped are heart disease, kidney disease, high blood pressure, drink, smoke, exercise, change to lifestyle, and whether you were told to change your lifestyle.

[18] A 'full' set of results are presented in Appendix-A, Table A-1.

probability that a person was employed in the last year.[19] All remaining outcomes are estimated using a censored regression estimation technique and the inverse-Mills correction for selection bias.[20] The coefficient of interest is β_1.

Secondary estimation includes type II diabetics in the sample with no specific controls for the effect of their disease. Because type I diabetes is viewed as the more severe case of diabetes, the coefficient on β_1 is expected to fall in absolute value when type II diabetics are included in the estimation. This is because type II diabetics may have the similar labor market penalties, but they may be less pronounced. The third estimation model includes a type II diabetes indicator, so that they are removed from the comparison group. Estimation takes the form:

$$WorkLastYear_i = \alpha_0 + \alpha_1 Type\ I_i + \alpha_2 Type\ II_i + \alpha_3' X_i + \mu_i, \tag{3}$$

$$Y_i = \beta_0 + \beta_1 Type\ I_i + \beta_2 Type\ II_i + \beta_3' X_i + \beta_4 \lambda_i + \varepsilon_i. \tag{4}$$

This allows a comparison of *Type I* to not only non-diabetics, but a direct interpretation may be made for *Type II* diabetics as well. The two coefficients of interest are β_1 and β_2.

The concern with *Type II* diabetes is it may be subject to endogeneity bias with respect to the different labor outcomes. To illustrate this, consider the following example. Type II diabetes could affect work decisions. Simultaneously, work decisions may influence a person's diet and exercise, which have a direct impact on their probability of contracting diabetes. For instance, it could be you work more hours and have less time to exercise or eat fast food more often, thereby increasing your risk of diabetes. Also, someone diagnosed with diabetes may decide to get a job because they need health

[19] This value is obtained from the maximum likelihood estimation of *Work Last Year,* and it is used in the estimation of the other labor force outcome variables.

[20] All outcomes are treated as continuous variables; even though, *Earnings* is recorded on a discrete scale.

insurance. In either case, it is likely that *Type II* is an endogenous variable. Therefore, it is necessary to instrument the variable in order to estimate the causal effect of type II diabetes. Instrumentation should also account for any unobserved variables that are correlated with having type II diabetes and labor outcomes. For example, a person could be living a sedentary lifestyle which causes her to be less productive at work, but this also contributes to the probability of contracting diabetes. The two-stage procedure takes the form:

$$\widehat{Type\ II}_i = \gamma_0 + \gamma_1' Z_i + \gamma_2' X_i + \upsilon_i, \tag{5}$$

$$Y_i = \beta_0 + \beta_1 Type\ I_i + \beta_2 \widehat{Type\ II}_i + \beta_3' X_i + \beta_4 \lambda_i + \varepsilon_i. \tag{6}$$

Equation (6) is equivalent to Equation (4), except now $\widehat{Type\ II}_i$ is the predicted value from Equation (5) for each individual *i*. This equation contains all independent variables from previous equations, but it adds the vector Z_i, which contains variables correlated with the probability of having diabetes. To identify the causal effect of diabetes one must find a valid instrument: one that is correlated with the suspected endogenous regressor but uncorrelated with the outcome variable of interest. However, finding a rationally and statistically sound instrument is potentially difficult. In many cases, a theoretical relationship exists with little statistical relevance, or vice versa (see Angrist & Krueger 2001; Altonji et. al 2005; Murray 2006).

For the purposes of this paper, whether a person takes medication for high blood sugar (*HBS*) is used as the instrumental variable for type II diabetes. High blood sugar is the first sign of diabetes and if not treated immediately could lead to a more serious medical condition. Certainly there is a correlation between people who take medication for high blood sugar and those with diabetes, as this is one of the primary methods of controlling a diabetic's blood sugar. Also, a prescription of medication to control blood sugar levels does not necessarily mean the person will contract diabetes. Hyperglycemia, or high blood sugar, arises in non-diabetics for reasons ranging from poor diet and exercise to stress and infection. In the sample utilized, 164 people are taking medication for their blood sugar and do not have diabetes, as opposed to 860 diabetics who currently take an oral agent.

Type I and type II diabetes cases are considered separately, using the same estimation technique. The three-stage estimation technique, described above, is used to deal with any selection or endogeneity bias and is designed to give results that accurately represent the true effect of diabetes on the labor market outcomes considered.

2.6 Results

Table 3 presents the effect of diabetes on employment. Results indicate *Type I* diabetes significantly reduces employment for men, but it has no significant impact on women's labor force participation. Probit estimates, presented as Model I, indicate that a male with type I diabetes is 17 percent less likely to be employed than a non-diabetic holding all health, demographic, and personal variables constant. When type II diabetics are included in the control group in Model II the effect of type I diabetes decreases to a 15.5 percent reduction in the probability of employment. In Model III results suggest that *Type I* diabetes decreases a male's probability of working by about 17.6 percent. Similarly, *Type II* diabetes reduces the probability a male will enter the work force by about 8.1 percent. Where *Type I* has no effect for females across all specifications, *Type*

II has a detrimental effect the probability of a woman working, lowering her percentage probability by 17.2 percent.

Table 4 presents the effect of diabetes on the number of days missed at work. Findings indicate that *Type I* diabetes influences both male and female days missed. Model I indicates that contracting *Type I* diabetes causes males to miss about 2.7 fewer days of work. This becomes larger when type II diabetics are included, with males now missing 3.6 fewer days and women now missing a statistically significant 2.5 fewer days. Lastly, in Model III, findings show that *Type I* diabetics actually miss about 3.2 and 2.4 days less than the general population for males and females respectively. This result seems odd, considering it indicates people with a serious medical condition actually miss fewer days of work than the general population. This could be due to the fact that while type I diabetics do have a condition that they must monitor daily, because their disease was diagnosed early on in life (before the age of 20 in this study) they are more familiar with the effects the disease has on their daily lives, and thus are able to adjust more flexibly than those presented with a relatively new medical condition. For evidence of this, consider the coefficient of *Type II*. Results suggest that *Type II* diabetic males actually miss about 4 more days than the general population, and about 7 more than *Type I* diabetics.

The effect of diabetes on average hours worked per week is presented in Table 5. Here, male and female work hours are both significantly affected by diabetes. Model I indicates that male *Type I* diabetics work an average of 3.8 hours less than those people without diabetes. Similarly, *Type I* diabetic females work about 4.8 hours less per week. Model II shows that these effects change only slightly when other diabetics are included in the sample. Males now reduce their hours about 4.0 hours per week, and females reduce their work hours about 4.7 hours per week. Model III estimates show that *Type I* males reduce their average work hours by about 4.0 hours per week, and *Type II* diabetes has no significant effect on male work hours. For females, *Type I* diabetes causes a reduction in the work week of about 4.8 hours, and *Type II* diabetes causes a smaller reduction of 1.7 hours per week.

Table 6 presents the impact of diabetes on earnings. Initial estimates imply a loss of about $5,155 per year for men from having any sort of diabetes.[21] Diabetes does not significantly affect wages for females. For men, Model II indicates a smaller wage penalty of about $5,005. Model III shows that *Type I* diabetic males actually lose about $5,245 relative to non-diabetics, and *Type II* diabetics lose $2,355. This translates to a loss of about 17 percent of the average male's salary for *Type I* diabetics and a loss of eight percent for *Type II.*

Taken together, results suggest that diabetes will reduce the probability of employment and average hours worked for males and females. Additionally, any type of diabetes is detrimental to male wages.[22] And the effect of *Type I* diabetes is different than that of *Type II* with respect to days out of work. The differing magnitudes, and in one case opposite signs, indicate that type I and type II diabetes do in fact have different impacts on the labor market, making their separate analysis crucial to the understanding of this diseases' economic impact.

[21] Due to the data available each integer of earnings corresponds to a $5000 increment, to interpret the coefficients multiply by 5000 (i.e. -1.031*5000 = -5155).

[22] It is possible that such a low number of working female diabetics are the primary cause for numerous statistically insignificant estimates.

The estimates of specifically *Type II* diabetes may be subject to endogeneity bias and therefore are re-estimated using an instrumental variable. Results are presented in Table 7. Across all outcome variables the coefficient of Type I diabetes does not change significantly. A test of the first stage predictive power of *HBS* with respect to *type II* indicates that the instrument provides substantial explanatory power for whether a person has been diagnosed with type II diabetes.[23] With a valid instrument we are able to test for endogeneity of *type II* diabetes. The Davidson-McKinnon test of *Type II* suggests there is endogeneity for females with respect to the decision to work.[24] This means un-instrumented results were biased, [25] and *Type II* diabetes is estimated to cause a reduction in the female's probability of employment by about 10 percent, suggesting endogeneity bias had previously caused the effect to be understated.[26] Testing for endogeneity of these results indicates no bias for male outcomes, meaning the un-instrumented results are consistent and preferred over the IV estimates. Endogeneity tests also indicate that previous estimates are not consistent with respect to a female's days missed, and so we must examine the instrumented results. These reveal that *Type II* diabetes will result in 3.6 additional days missed. Again, in the presence of endogeneity bias, the effect of *Type II* was previously understated. For average hours worked, instrumented results suggest an average decrease of 3.3 for *Type II* diabetes, up from previous estimates of only about 1.7

---

[23] The t-statistic for all types of diabetes is 67.61 for males and 32.80 for females with corresponding p-values of < 0.001 for both sub-sets. Results of this estimation are present in Appendix Table A-2.

[24] The null hypothesis for the Davidson- McKinnon exogeneity test is no endogeneity. This test is preformed by regressing the error term from a first stage regression of type II diabetes on work outcomes. A significant error term would represent the presence of endogeneity.

[25] This correlates to the findings in Brown, et.al. (2005) who find endogeneity bias with respect to older female diabetics.

[26] The null of the instrument test is that it has no predictive power with respect to the relevant outcome variables. Results indicate *HBS* has no correlation with the decision to work for males; however, there is a statistical correlation for females, suggesting that *HBS* may not be a valid instrument for the female's decision to work.

hours per week reduction. The estimated effect of *Type II* diabetes is now closer to the effect of *Type I*, a reduction of about 4.9 hours. Lastly, earnings show no sign of bias for men or women; therefore, previous estimates are preferred for efficiency.

Across all specifications, the effect of type II diabetes is estimated to be endogenous with respect to the numerous labor force decisions for females. Results change once endogeneity bias is taken into account. This suggests that many previous studies of the economic impact of the disease may have misinterpreted the true effect diabetes has on the labor market. Also, by controlling for related illnesses, such as heart disease, high blood pressure, kidney disease, and obesity, the estimated effects more likely represent the true effect of merely contracting diabetes and not the significant other medical complications that may arise as a result of the disease. Comparing the differences between diabetes types indicates that type II diabetes produces only a slightly smaller and no less significant impact on employment decisions for both men and women. It may seem intuitive that more "manageable" cases have a slightly smaller impact on work decisions, but the effects of these cases, which are primarily attributable to lifestyle choices, definitely have a significant detrimental impact on the labor market that may be easily avoided.

2.7 Conclusion

This paper extends the current literature on the impact of diabetes on the US labor market in several important ways. First, the separation of type I and type II diabetes is shown to be important and relatively new to this branch of the literature. Second, novel instrumental variable estimates are presented that eliminate the endogeneity bias which may have been present in previous studies. A statistically sound instrument also allows a test for endogeneity of diabetes with respect to all the work outcomes. Third, other sources of bias that may have caused some inaccurate results in previous studies, such as the selection problem related to working and the incidence of related health problems, are both accounted for in this paper. Finally, the data set used not only allows for the most up-to-date estimates available, but it is also nationally representative of the US labor force.

In estimating the effects of diabetes on labor force outcomes, I find that both type I and type II diabetes significantly affect numerous work decisions. Primarily, diabetes will reduce employment, the number of hours spent at work, and total earnings. Interestingly, type II diabetes is estimated to be slightly less detrimental to labor market decisions than type I diabetes. This may seem odd, as type I requires constant maintenance and insulin injections, whereas type II may be controlled through diet and exercise. The result could be driven by the differing timing and diagnosis of both types of diabetes. Where type I is typically diagnosed early on in life, probably before the individual enters the work-force, type II diabetes may not be diagnosed until much later in the person's life, or even into the later stages of the disease. This may make the impact much more pronounced, as these are the prime earning years for most of the population. Furthermore, the findings suggest that type II diabetes is subject to endogeneity bias for women when it comes to labor market decisions such as days out of work, average hours worked, and labor force participation.

All of these effects may become even more harmful as the incidence of diabetes continues to rise. A study by the CDC has stated that the number of diabetes sufferers will increase to over 29 million by the year 2050. However, a more recent analysis published in Diabetes Care reported that diabetes cases may reach this number as soon as 2030 (Wild et al. 2004). The problem presented by this ever growing disease will certainly have an impact upon numerous aspects of the economy as a whole.

This paper provides a clear snapshot of the effects diabetes caused in the year 2006, and it illustrates the problem with what many consider to be a highly preventable disease. Further research in this area could prove interesting. A panel approach could provide insight into the effect that diabetes has as it progresses. Testing the length diabetes may provide some insight into the prolonged effect the disease has on a person. As a person becomes familiar with their particular affliction and treatment, the effects may change. Another aspect for future research could involve the analysis of work benefits, such as health insurance and paid leave. With the currently available data these questions, although interesting, are difficult to answer.

REFERENCES

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). "An evaluation of instrumental variable strategies for estimating the effects of catholic schooling." *Journal of Human Resources*, 40(4), 791-821.

Angrist, J. D., & Krueger, A. B. (2001). "Instrumental variables and the search for identification: From supply and demand to natural experiments." *Journal of Economic Perspectives*, 15(4), 69-85.

Baum, Charles L., & Ford, William F. (2004). "The Wage Effects of Obesity: A Longitudinal Study." *Health Economics*, 13(9): 885-899.

Brown, H. S., I.I.I., Pagan, J. A., & Bastida, E. (2005). "The impact of diabetes on employment: Genetic IVs in a bivariate probit." *Health Economics*, 14(5), 537-544.

Centers for Disease Control and Prevention, Atlanta, GA. (1999). "Chronic diseases and their risk factors: The nation's leading causes of death." *Exercise, Health, & Physical Fitness*, PE 03

Centers for Disease Control and Prevention. (2007). "National Diabetes Fact Sheet."

Centers for Disease Control and Prevention. (2005). "National Diabetes Fact Sheet."

Chodick, G., Heymann, A. D., Wood, F., & Kokia, E. (2005). "The direct medical cost of diabetes in israel." *European Journal of Health Economics, 6*(2), 166-171.

Costa-Font, J., & Gil, J. (2005). "Obesity and the incidence of chronic diseases in spain: A seemingly unrelated probit approach." *Economics and Human Biology, 3*(2), 188-214.

Cutler, D. M., Glaeser, E. L., & Shapiro, J. M. (2003). "Why have Americans become more obese?" *Journal of Economic Perspectives*, 17(3), 93-118.

Duggan, M., Rosenheck, R., & Singleton, P. (2006). *Federal policy and the rise in disability enrollment: Evidence for the VA's disability program.* Doctoral dissertation, National Bureau of Economic Research, Inc, NBER Working Papers.

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

Heckman, J. J. (1976). "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic and Social Measurement*, 5(4), 475-492.

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

Lavigne, J. E. (2003). "Reductions in individual work productivity associated with type 2 diabetes mellitus." *PharmacoEconomics*, 21(15), 1123-1134.

Mayfield, Jennifer A, & White, Russell D. (2004) "Insulin Therapy for Type 2 Diabetes: Rescue, Augmentation, and Replacement of Beta-Cell Function." *American Family Physician*, August 1, 2004.

Murray, M. P. (2006). "Avoiding invalid instruments and coping with weak instruments." *Journal of Economic Perspectives*, 20(4), 111-132.

Olivia, J., Lobo, F., Molina, B., & Monero, S. (2004). "Direct health care costs of diabetic patients in Spain." *Diabetes Care*, 27, 2616-2621.

Pagano, E. (1999). "Costs of diabetes: A methodological analysis of the literature." *PharmacoEconomics, 15*(6), 583-595.

Ramsey, Scott Summers, Kent H. Leong, Stephanie A. Birnbaum, Howard G. Kemner, Jason E. & Greenberg, Paul. (2002). "Productivity and Medical Costs of Diabetes in a Large Employer Population." *Diabetes Care*, 25(1). 23-29.

Ruhm, C. J. (2007). *Current and future prevalence of obesity and severe obesity in the united states*. National Bureau of Economic Research, Inc, NBER Working Papers. NBER Working Paper Series, w13181.

Tuncli, K., Bradley, C. J., Nerenz, D., Williams, L., Keoki, P., Manel, L., & Elston J. (2005). "The Impact of Diabetes on Employment and Work Productivity." *Diabetes Care*, 28(11): 2662-2667.

Telford, R. D. (2007). "Low physical activity and obesity: Causes of chronic disease or simply predictors?" *Medicine & Science in Sports & Exercise*, 39(8), 1233-1240.

Vijan, S., Hayward, R. A., & Langa, K. M. (2004). "The impact of diabetes on workforce participation: Results from a national household sample." *Health Services Research, 39*(6), 1653-1669.

Wild, S. Roglic G., Green A., Sicree R., & King H. (2004). "Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030." *Diabetes Care,* 27(5): 1047-1053.

TABLE 1. SUMMARY STATISTICS

	Male				Female			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Key Explanatory Variables								
Type I	0.005	0.07	0	1	0.004	0.07	0	1
Type II	0.068	0.25	0	1	0.046	0.21	0	1
Key Outcome Variables								
Work Last Year	0.80	0.40	0	1	0.57	0.49	0	1
*Days Missed**	3.50	16.49	0	365	4.47	17.82	0	365
*Average Hours**	42.23	12.32	1	95	40.38	11.97	1	95
*Earnings**	6.46	2.75	1	11	6.08	2.72	1	11
Related Health Variables								
Heart Disease	0.07	0.26	0	1	0.06	0.25	0	1
Kidney Disease	0.02	0.12	0	1	0.01	0.12	0	1
High Blood Pressure	0.25	0.43	0	1	0.21	0.40	0	1
BMI	27.67	4.65	17	51	27.14	6.04	15	55
Demographic Variables								
Age	41.76	12.63	20	65	40.24	12.12	20	65
Married	0.58	0.49	0	1	0.58	0.49	0	1
Widowed	0.01	0.10	0	1	0.03	0.17	0	1
Divorced	0.09	0.29	0	1	0.11	0.32	0	1
Better Health	0.18	0.39	0	1	0.20	0.40	0	1
Worse Health	0.07	0.26	0	1	0.09	0.28	0	1
White	0.79	0.41	0	1	0.76	0.43	0	1
Black	0.14	0.34	0	1	0.18	0.38	0	1
Asian	0.06	0.24	0	1	0.05	0.21	0	1
Hispanic	0.19	0.39	0	1	0.22	0.41	0	1
High School Graduate	0.26	0.44	0	1	0.21	0.41	0	1
Higher Education	0.53	0.50	0	1	0.45	0.50	0	1
High School Dropout	0.19	0.39	0	1	0.27	0.44	0	1
Personal Variables								
Exercise	0.55	0.50	0	1	0.64	0.48	0	1
Told to Diet	0.18	0.38	0	1	0.21	0.40	0	1
Told to Lose Weight	0.21	0.41	0	1	0.27	0.44	0	1
Told to Exercise	0.17	0.38	0	1	0.20	0.40	0	1
Actually Dieted	0.37	0.48	0	1	0.48	0.50	0	1
Actually Lost Weight	0.39	0.49	0	1	0.47	0.50	0	1
Actually Exercised	0.39	0.49	0	1	0.53	0.50	0	1
Drink	0.70	0.46	0	1	0.55	0.50	0	1
Smoke	0.26	0.44	0	1	0.20	0.40	0	1

Notes: Number of observations is 8910 males and 12470 females. * indicates that only those who were working last year are included.

TABLE 2. COMPARISON OF SUMMARY STATISTICS FOR DIABETICS AND NON-DIABETICS

	Male				Female			
	Diabetic		Non-Diabetic		Diabetic		Non-Diabetic	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Key Explanatory Variables								
Type I	0.07	0.26	--	--	0.09	0.28	--	--
Type II	0.93	0.26	--	--	0.91	0.28	--	--
Key Outcome Variables								
Work Last Year	0.58	0.49	0.82	0.39	0.38	0.49	0.59	0.49
*Days Missed**	8.87	31.75	3.20	15.14	6.27	16.45	4.39	17.86
*Average Hours**	40.94	12.54	42.30	12.29	38.89	11.67	40.44	11.99
*Earnings**	6.23	2.83	6.47	2.74	5.64	2.52	6.10	2.73
Related Health Variables								
Heart Disease	0.24	0.43	0.06	0.24	0.21	0.41	0.05	0.23
Kidney Disease	0.07	0.25	0.01	0.11	0.05	0.22	0.01	0.11
High Blood Pressure	0.63	0.48	0.22	0.41	0.63	0.48	0.18	0.38
BMI	30.62	5.03	27.43	4.54	32.19	6.67	26.80	5.84
Demographic Variables								
Age	52.36	9.69	40.91	12.46	49.09	11.24	39.65	11.95
Married	0.65	0.48	0.58	0.49	0.54	0.50	0.58	0.49
Widowed	0.02	0.15	0.01	0.09	0.07	0.25	0.03	0.16
Divorced	0.15	0.35	0.09	0.28	0.16	0.37	0.11	0.31
Better Health	0.21	0.41	0.18	0.38	0.21	0.41	0.20	0.40
Worse Health	0.19	0.39	0.06	0.25	0.24	0.43	0.08	0.27
White	0.73	0.45	0.79	0.40	0.70	0.46	0.76	0.43
Black	0.19	0.39	0.13	0.34	0.25	0.43	0.18	0.38
Asian	0.06	0.24	0.06	0.24	0.02	0.15	0.05	0.22
Hispanic	0.19	0.40	0.19	0.39	0.26	0.44	0.22	0.41
High School Graduate	0.31	0.46	0.26	0.44	0.27	0.44	0.21	0.41
Higher Education	0.44	0.50	0.54	0.50	0.32	0.47	0.46	0.50
High School Dropout	0.23	0.42	0.19	0.39	0.36	0.48	0.26	0.44
Personal Variables								
Actually Dieted	0.60	0.49	0.35	0.48	0.71	0.45	0.46	0.50
Actually Lost Weight	0.54	0.50	0.37	0.48	0.62	0.49	0.46	0.50
Actually Exercised	0.63	0.48	0.37	0.48	0.74	0.44	0.52	0.50
Drink	0.54	0.50	0.71	0.45	0.31	0.46	0.57	0.50
Smoke	0.20	0.40	0.27	0.44	0.21	0.41	0.20	0.40

Notes: Number of observations for diabetics is 656 males and 776 females, and non-diabetics include 8254 men and 11847 women. * indicates that only those who were working last year are included.

TABLE 3. THE EFFECT OF DIABETES ON EMPLOYMENT

	Male			Female		
	Model I	Model II	Model III	Model I	Model II	Model III
Type I	-0.170***	-0.155**	-0.176***	0.010	0.028	0.007
	(0.080)	(0.082)	(0.083)	(0.087)	(0.086)	(0.086)
Type II			-0.081***			-0.172***
			(0.019)			(0.028)
R^2	0.3313	0.342	0.345	0.410	0.407	0.409
Log Likelihood	-2651.50	-2937.44	-2926.32	-4768.71	-5049.81	-5030.61
Wald Statistic	$X^2(71)$= 1628.35	$X^2(71)$= 1902.31	$X^2(72)$= 1934.02	$X^2(71)$= 3367.16	$X^2(71)$= 3516.41	$X^2(72)$= 3518.55
N Observations	8300	8910	8910	11900	12470	12470
Type II Included		X	X		X	X
Type II Control			X			X

Notes: Standard errors are presented in parentheses. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively. All reported results are probit estimations, but marginal effects are reported so that the coefficients may be directly interpreted.

TABLE 4. THE EFFECT OF DIABETES ON WORK DAYS MISSED

	Male			Female		
	Model I	Model II	Model III	Model I	Model II	Model III
Type I	-2.771**	-3.636***	-3.218**	-2.357	-2.505*	-2.486*
	(1.193)	(1.389)	(1.322)	(1.465)	(1.466)	(1.461)
Type II			4.073**			0.368
			(1.776)			(1.550)
R^2	0.058	0.064	0.067	0.045	0.046	0.046
F-Statistic	2.61	2.69	2.66	3.43	3.49	3.49
N Observations	6771	7122	7122	6963	7148	7148
Type II Included		X	X		X	X
Type II Control			X			X

Notes: Standard errors are presented in parentheses. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively. All estimations include the Heckman correction for selection bias.

TABLE 5. THE EFFECT OF DIABETES ON AVERAGE HOURS WORKED

	Male			Female		
	Model I	Model II	Model III	Model I	Model II	Model III
Type I	-3.882*	-4.002*	-4.085*	-4.879**	-4.739**	-4.831**
	(2.361)	(2.358)	(2.365)	(2.107)	(2.100)	(2.105)
Type II			-0.800			-1.702**
			(0.729)			(0.863)
R^2	0.037	0.036	0.036	0.032	0.031	0.031
F-Statistic	3.56	3.65	3.61	3.08	3.03	3.02
N Observations	6771	7122	7122	6963	7148	7148
Type II Included		X	X		X	X
Type II Control			X			X

Notes: Standard errors are presented in parentheses. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively. All estimations include the Heckman correction for selection bias.

TABLE 6. THE EFFECT OF DIABETES ON EARNINGS

	Male			Female		
	Model I	Model II	Model III	Model I	Model II	Model III
Type I	-1.031**	-1.001**	-1.049***	0.404	0.421	0.408
	(0.403)	(0.403)	(0.402)	(0.426)	(0.427)	(0.427)
Type II			-0.471***			-0.232
			(0.152)			(0.177)
R^2	0.252	0.247	0.248	0.251	0.250	0.250
F-Statistic	34.64	35.33	35.04	38.26	38.86	38.33
N Observations	6771	7122	7122	6963	7148	7148
Type II Included		X	X		X	X
Type II Control			X			X

Notes: Standard errors are presented in parentheses. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively. All estimations include the Heckman correction for selection bias.

TABLE 7. THE INSTRUMENTED EFFECT OF DIABETES ON LABOR MARKET OUTCOMES

	Work Last Year		Days Missed		Work Hours		Earnings	
	Male	Female	Male	Female	Male	Female	Male	Female
Type I	-0.169**	0.018	-3.370**	-2.706*	-4.122*	-4.919**	-1.044***	0.399
	(0.083)	(0.087)	(1.358)	(1.493)	(2.352)	(2.105)	(0.400)	(0.426)
Type II	-0.059***	-0.101**	2.588*	-3.678*	-1.164	-3.325**	-0.418**	-0.404
	0.018	(0.047)	(1.493)	(1.895)	(0.921)	(1.345)	(0.185)	(0.279)
R^2			0.0665	0.0444	0.036	0.0309	0.248	0.250
F-Statistic			2.66	3.45	3.63	3.06	34.92	38.31
Log Pseudo-likelihood	1611.17	752.65						
Wald Statistic	$X^2(72)=$ 1923.73	$X^2(72)=$ 3512.98						
N Observations	8910	12470	7122	7148	7122	7148	7122	7148
Exogeneity Test	0.78	4.68	0.83	7.63	0.50	2.87	0.26	0.58
	(0.38)	(0.03)	(0.36)	(0.01)	(0.48)	(0.09)	(0.61)	(0.45)
Instrument Test	2.38	3.44	1.44	4.96	0.48	1.60	1.22	1.79
	(0.12)	(0.06)	(0.23)	(0.03)	(0.51)	(0.21)	(0.27)	(0.18)

Notes: Type II diabetes is instrumented in all cases with Blood Sugar Medication. The null of the Davidson-McKinnon test for endogeneity is no endogeneity. The null of the instrument test is Blood Sugar Medication has no significant impact on the outcome variable. Standard errors are presented in parentheses, except where otherwise noted. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively. Days Missed, Work Hours, and Earnings, all include the Heckman correction for selection bias.

APPENDIX

TABLE A-1. "FULL" MODEL OF DIABETES ON EARNINGS

	Male	Female
Type I	-1.049***	0.408
	(0.402)	(0.427)
Type II	-0.471***	-0.232
	(0.152)	(0.177)
Heart Disease	-0.055	-0.264**
	(0.153)	(0.134)
Kidney Disease	-0.347	-0.571*
	(0.338)	(0.338)
High Blood Pressure	-0.159*	-0.248***
	(0.082)	(0.079)
BMI	0.006	-0.027***
	(0.007)	(0.006)
Age	0.024***	0.027***
	(0.003)	(0.003)
Better Health	-0.039	-0.100
	(0.077)	(0.071)
Worse Health	-0.433***	-0.203*
	(0.145)	(0.116)
White	0.809**	0.102
	(0.326)	(0.323)
Black	0.494	-0.078
	(0.334)	(0.326)
Asian	0.633*	0.091
	(0.348)	(0.353)
Hispanic	0.103	-0.360
	(0.330)	(0.325)
Exercise	-0.298***	-0.321***
	(0.064)	(0.064)
Inverse Mills	-0.361***	0.146
	(0.129)	(0.106)
R^2	0.248	0.250
F-Statistic	35.04	38.33
N Observations	7122	7148

Notes: Standard errors are presented in parentheses. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively.

TABLE A-2. LINEAR PROBABILITY MODEL OF TYPE II DIABETES

	Male	Female
HBS	0.858***	0.649***
	(0.013)	(0.020)
Type I	-0.462***	-0.310***
	(0.065)	(0.047)
Heart Disease	0.032***	0.030***
	(0.010)	(0.010)
Kidney Disease	0.037	0.061***
	(0.025)	(0.021)
High Blood Pressure	0.011**	0.028***
	(0.005)	(0.005)
BMI	0.001**	0.001***
	(0.000)	(0.000)
Age	0.001***	0.001***
	(0.000)	(0.000)
Better Health	0.009**	0.010***
	(0.004)	(0.004)
Worse Health	0.017*	0.016**
	(0.009)	(0.007)
White	-0.036	0.009
	(0.025)	(0.023)
Black	-0.024	0.010
	(0.026)	(0.023)
Asian	-0.031	0.015
	(0.025)	(0.024)
Hispanic	-0.032	0.014
	(0.025)	(0.023)
Exercise	0.006*	0.004
	(0.003)	(0.003)
R^2	0.6684	0.4689
F-Statistic	136.05	29.64
N Observations	8910	12470

Notes: Standard errors are presented in parentheses. *, **, and *** indicate significance at the 1, 5, and 10 percent level respectively.

Chapter 3

State Minimum Wage Differences: Economic Factors or Political Inclinations?

(with Mark F. Owens and William F. Ford)

3.1 INTRODUCTION

The passage of the *Fair Minimum Wage Act of* 2007 has once again brought minimum wage laws to the forefront of American politics. The act raised the federal minimum wage from \$5.15 to \$7.25 per hour by July 2009. [1] According to the U.S. Department of Labor (2007), by January 2007 half of the states had established minimum wages greater than the prior federal rate of \$5.15 per hour, which had been in effect since September 1, 1997. Seven states had a minimum wage that exceeded \$7.00 per hour and nearly 150 separate urban areas had either minimum or "living wage rates" above the federal level.[2] This tendency for state and local minimum wages to change between infrequent federal rate changes is not new, and neither is the debate about the merits of such legislation.

Typically, the stated goal of such minimum wage increases is to help low-wage-earning workers. However, whether minimum wages are an effective way to help low-wage workers afford the necessities of modern life, and whether they are they best policy

[1] The new law incorporated three increments, starting with an increase to \$5.85 per hour in July 2007 followed by an increase to \$6.55 per hour in July 2008 before the final step in 2009 to \$7.25 per hour.
[2] In many cases these local rates were substantially higher. Hartford, Connecticut for example had a rate of \$15.39 per hour; nearly triple the federal rate, in July 2007 (see ACORN, 2007).

for doing so, has been widely discussed in the economic literature. To be effective, the minimum wage rate logically would need to be enacted to reflect regional cost-of-living differences, since each cohort of workers, in each state or city, requires different funds to achieve this stated goal, in real terms. Federal legislation, applied uniformly across the entire country, cannot possibly reach this goal. However, as we will see, differences in state minimum wage levels are not attributable to differences in the cost of living.

This study uses state level data from two prior federal minimum wage cycles, spanning from 1991-2007, to assess the extent to which political inclinations and cost of living differences have led to the adoption of various state minimum wage levels in excess of the federal standard. This question has received very little attention compared to the extensive literature debating the impact of minimum wages on the economy.[3] Most of the states that have established minimum wages in excess of the federal level share two common characteristics. First, such states are relatively high cost-of-living areas. Second, voters in those states also tend to reflect more liberal political views on the proper role of government.[4]

On the surface, all state and local minimum wage legislation appears to be driven by both economic and political factors. This paper analyzes the importance of such factors in driving the higher-than-federal minimum wages enacted by various states since 1991. We believe distinguishing between economic and political factors is important because proposals to change the minimum wage at the state and federal levels are almost

---

[3] Exceptions include Waltman, & Pittman (2002) and Levin-Waldman (1998).

[4] For example, at the time of the 2004 Presidential election, of the 31 states that voted Republican, only Alaska and West Virginia had a state minimum wage greater than the federal level. Nineteen states and the District of Columbia voted Democratic in the 2004 election, and 12 of these had a minimum wage higher than the federal level.

universally promoted as responses to regional differences or increases in the cost of living. It is worth investigating whether this is in fact the case as these factors may also serve as a harbinger of coming changes in state minimum wage laws as the new and higher federal levels are enacted.

Supporters of increasing the minimum wage generally have passed legislation at the state and even local level to bring their minimum wage in line with what they believe to be the cost of living factors affecting workers' lives (See for example, Association of Community Organizations for Reform Now (ACORN), 2007; The Center for Policy Alternatives, 2007; and the Ballot Initiative Strategy Center (BISC), 2006). For clarity, we label these "economic" concerns, as they relate to purchasing power and consumption issues. A state minimum wage change that occurs in response to concerns for rising costs of living should introduce few *additional* adverse affects, relative to the prior market conditions. In these instances, the minimum wage is not likely to be a binding wage floor, as the market wage in many of these areas often exceeds the newly-legislated minimum level. Thus, a small increase, (one that is roughly proportional to the increase in cost of living), is not likely to have a large economic impact, or to distort the incentives facing low wage workers *relative to current conditions*. This is not to say that adverse effects from minimum wages do not occur. Rather we take any effects caused by current minimum wage policy as given, since there is no indication that such policies will be discontinued, or scaled back, in the foreseeable future.

While cost-of-living concerns are typically cited as the reason for increases in a state's minimum wage, political issues and beliefs about the proper role of government are also contributing factors. Thus, states that exhibit more liberal political beliefs can be expected to have a greater tendency to enact minimum wages higher than the federal level. We refer to these as "political" inclinations throughout the paper.

Our results indicate that interstate political leanings consistently explain variations in state minimum wages in the federal cycle spanning from April 1991 until August 1997, and in the cycle spanning from September 1997 until 2006. We do not find evidence that cost of living concerns increase the likelihood that a state will raise its minimum wage, and find only weak evidence that cost of living influences the magnitude of a state's minimum wage increase.

This paper is organized as follows. Section 2 briefly examines theoretical concerns relating to the enactment of minimum wage levels. Section 3 presents the data and methods used in our analysis. Section 4 presents estimation results, and Section 5 presents our conclusions.

3.2 Theoretical Background

It is well known that states tend to increase their minimum wage, above the federally-mandated level, as time passes without a federal increase. Figure 1 shows the number of states with minimum wage levels above the federal level by year. Whether this is a beneficial or harmful trend depends on one's reading of the rich literature on the effects of minimum wages.

The numerous criticisms of minimum wage legislation fall into three broad categories. The first concern relates to the inefficiency caused by prohibiting mutually beneficial employment contracts. In this case unemployment is increased as the quantity of labor supplied exceeds the quantity of labor demanded above the market equilibrium level of wages.[5] A second and related point of contention with minimum wage laws is whether they are actually an effective means of helping low-income workers and their families. [6]The third category of minimum wage critiques encompasses philosophical beliefs regarding the proper role of government in the economy.

Supporters of state and local minimum wage legislation (or of increases in their level), such as The Center For Policy Alternatives (2007), maintain that the federal minimum wage is not effective because many workers do not have sufficient earnings to cover the cost of basic needs.[7] Or, they argue that the federal wage floor is too low to be binding for many employers and is thus ineffective. Still others, like the American Federation of Labor and Congress of Industrial Organizations (AFL-CIO, 2003), believe the government should use minimum wage laws to more actively redistribute income. Concerns that full time workers, with families, who are earning the minimum wage are still near or below the poverty line, and normative beliefs about the skewed nature of the U.S. income distribution, drive such movements for higher minimum wages. Some groups argue that minimum wage levels should be directly tied to cost of living measures, effectively creating a "living wage" that will help workers across all industries (see BISC, 2006 and The Center for American Progress, 2007, for a description; and Sander and Williams, 2005, for an assessment of living wages).

---

[5] This idea is so firmly grounded in economic theory that it is presented in principles of microeconomics courses.

[6] See Burkhauser et al. (1996) , Fairchild (2005) and Neumark and Wascher (2007) for comparisons of minimum wages to the Earned Income Tax Credit as one example, and Neumark and Wascher (2002, 2005) for evidence regarding the groups affected by the legislation.

[7] This is similar to the argument for living wages. In some areas the minimum wage has been modified to serve as a living wage that is explicitly tied to the cost of basic needs.

The existing literature on minimum wage legislation is primarily concerned with the effects these laws have on economic efficiency and their distributional consequences. As noted above, standard economic theory predicts that a binding minimum wage will create unemployment and potentially raise prices. However, finding the effect that minimum wage laws have on the economy, empirically, has proven to be somewhat elusive. Card and Krueger (1994, 2000) examine a natural experiment, with variations in minimum wage laws across states. They do not observe negative consequences from an increase in the minimum wage level with no loss in employment, or any significant increase in prices. However, studies since then have looked not only at prices and employment effects, but numerous other economic variables that may be adversely affected (See for example Burkhauser et al., 2007 and Neumark et al., 2005).

Another branch of the literature relates to the altered incentives for non-work activities. When minimum wages exceed the market-determined rate, some low-skilled workers may choose to enter the work force earlier, or to work longer hours, and thus receive less schooling. This decision potentially lowers their human capital acquisition and thus their lifetime earnings. Neumark and Wascher (2003) estimate that exposure to binding minimum wages may lower school enrollment, thus having a negative impact on labor force skill acquisition. Chaplin et al. (2003) also find that a teenagers' school enrollment declines in the presence of a binding minimum wage. More recently, Neumark and Nizalova (2007) examine the longer-run implications of a minimum wage, and they estimate that a prolonged exposure to the minimum wage as a teenager has detrimental effects later in life, which includes less labor force participation and lower long-term wages.[8]

Given the foregoing research on the broader effects of binding minimum wages on human capital acquisition, one concern is whether politically-driven or economically-driven minimum wages are more likely to be binding for employers. This remains an open question as few studies have examined the reasons for minimum wage changes. If for example, politically-driven minimum wage increases are more likely to be binding for employers, they are also more likely to distort the economic incentives that relate to employment and schooling decisions. Neumark and Nizalova (2007) demonstrate that these decisions have important long-term implications for workers. The motivation behind the minimum wage change likely does not matter to the workers, who simply respond to the incentives presented to them. However, the disincentives to human capital formation that are introduced may be more substantial in cases where cost-of-living differentials are not the primary reason for changing the state law. Thus, we might expect lower human capital acquisition in areas covered by the legislation when it is driven primarily by political concerns.[9]

As such, a state minimum wage increase could have a different economic impact, depending not only on the existing conditions of the labor market, but also whether it is the result of "economic" factors relating to the cost of living, or "political" factors. Distinguishing between "economic" and "political" factors is therefore important because

---

[8] Falk, Fehr and Zehnder (2006) find evidence in a laboratory experiment that a minimum wage unambiguously raises an employee's reservation wage, which could adversely affect employment levels.

[9] It is worth noting that many areas with high state minimum wages (i.e. New England states) also have higher education levels and those in states with a minimum wage at or below the federal level (i.e. southern states) have lower educational attainment.

political determinants may have greater long-run detrimental effects on workers and the economy. In addition, increases in the national minimum wage are more likely to be binding in low-income and low cost-of-living areas, and less likely to be binding in high-income and high cost of living areas. This, in turn, may lead to different schooling and long-term employment outcomes in different locations.

In this paper we analyze the extent to which political and economic factors are driving changes in state minimum wage laws, rather than the short-run economic effects caused by state and local minimum wage legislation. This area of research has been somewhat neglected until recently. Levin-Waldman (1998) concludes that the minimum wage is not only an economically motivated law, but is also highly influenced by politics. Waltman and Pittman (2002) also estimate the effects of wealth, politics, and public ideology on the adoption of state level minimum wages. They argue that minimum wages are mainly symbolic since they typically have a small effect on the economy and are determined primarily by public beliefs rather than wealth or political influences. Their measure of political influence ranks a state on a Likert scale of 0 to 5. Instead, we will utilize a percentage scale derived from Congressional voting records that is a combination of the Liberal Quotient scores tabulated for each state by Americans for Democratic Action. We argue that actual votes are a superior measure because they allow for a more accurate description of a state's political climate than a discrete, categorical variable obtained from survey data. The sample size utilized in this paper also greatly exceeds the one utilized by Waltman and Pittman (2002), which allows for more robust results.

3.3 DATA AND ESTIMATION

State minimum wage data were collected individually from each state's department of labor, and federal minimum wage numbers were drawn from the U.S. Department of Labor. Explanatory variables of interest were recorded from a number of other sources. Variable definitions and sources are presented in Table 1.

A measure of the state's political inclination is taken from federal voting records, maintained by the Americans for Democratic Action (ADA) (2007). The ADA records all votes by both U.S. Senate and U.S. House members. It then scores each member on the percentage of the times they voted liberal.[10] State averages for both the U.S. House of Representatives and Senate were calculated, and these were combined to give a single score for each state. ADA records these scores as the Liberal Quotient (*LQ*) of a state, which is scaled from zero to one, with one being the most liberal score a state can receive. Actual state values for *LQ* range from zero to one in the sample with a mean of 0.464 and standard deviation of 0.259. The average suggests a fairly equal division of political beliefs during this sample period with Congress leaning slightly to the conservative side nationwide. This is consistent with what we would expect concerning political inclinations, especially with regard to presidential and Congressional elections during the sample time period analyzed.

Our proxy for the cost of living in a state is an interstate housing price index, collected from the Office of Federal Housing Enterprise Oversight (OFHEO). The

[10] We use United States House and Senate voting records to proxy political views at the state level. Alternatively, we could have constructed our political variables from state government voting records since they directly influence a state's minimum wage. However, there is considerable variation in how state governments operate and we are not aware of a consistent means to characterize state voting patterns between states and over time.

OFHEO collects a quarterly housing price index for each state in the U.S. and records the data with a base year of 1980. Throughout our sample the national mean for the index increased from 163.84 in 1991 to 372.49, in 2006.[11] For the estimations that follow, we use the yearly state level of the home price index divided by 100, *hpi,* and also the state's yearly growth rate of the index, *growhpi.* The Consumer Expenditure Survey (2006) reports that the average person spends about 33.8 percent of their annual income on housing expenditures. As this holds true across income levels, we believe that an indication of housing prices also reflects the relative cost-of-living for an area, at any given time. The housing price index has the distinct advantage over other characterizations of regional cost of living differences, such as the local CPI, or food and lodging cost indices, in that it is exogenous with respect to a state's minimum wage level. It is likely that a minimum wage is actually set in response to some broader measure of the cost of living. However, the idea that differences in housing costs are largely determined by a state's minimum wage law is improbable, whereas food and lodging costs are clearly more sensitive to existing minimum wages.[12]

We control for observable differences in state populations with three variables. The state population divided by one million, *population* and the yearly growth rate of the state's population, *population growth,* are included as controls. We also use the ratio of births to deaths in each state, in each year as a proxy for the *age* of a state's population. States with a higher ratio are more likely to have younger populations which may influence the passage of minimum wage laws. We include *income per capita* and the

---

[11] The minimum state-level value for the index is 94.07 and the largest value is 729.91.

[12] Singell and Terborg (2007) find different employment effects from minimum wage changes in the food sector where it is binding, versus the lodging sector, where it is not binding.

percentage of total population which is employed, *employment*, as controls for other labor market characteristics. States with high employment percentages and high incomes are more likely to have high equilibrium wages in the absence of minimum wage laws.

Lastly, we code zero-one indicator variables for geographic regions according to the U.S. Census Bureau's protocol. These variables enter the estimation to pick up any region-specific unobservables that our other included covariates do not capture. Likewise, yearly indicator variables enter all specifications to control for any macroeconomic factors which political and cost of living variables do not capture. The inclusion of both of these controls will account for any number of unobserved factors in our data.

We use three different model specifications to obtain our estimated coefficients. First, we estimate the model using Cox proportional hazard specifications. Survival analysis is appropriate since states are observed to increase their minimum wage above the federal level over time. This estimation will indicate whether our explanatory variables appear to influence when a state will increase its minimum wage above the federal rate. Our dependent variable is a zero-one indicator equal to one if a state has a legally-mandated minimum wage *above the federal minimum wage* ,*afedmw* ,at time t and zero otherwise. A state is a "survivor" as long as *afedmw* is zero and "fails" when *afedmw* is observed to be equal to one. Two features of the data suggest that the Cox proportional hazard specification is appropriate for analyzing state level responses for a given federal minimum wage level. First, states which choose to increase their minimum wage above the federal level are observed to maintain the higher minimum wage at least until the federal rate increases. Second, once a state increases its minimum wage, it is less likely to do so before the federal rate adjusts. Formally our proportional hazard model is:

$$afedmw_{it} = \beta_0 + \beta_1 LQ_{it} + \beta_2 hpi_{it} + \beta_3 growhpi_{it} + \beta_4 X_{it} + \varepsilon_{it}.$$

The variable LQ_{it} is the political leaning of a particular state at time t, hpi_{it} is the state housing price index, and $growhpi_{it}$, are the preferred measures for cost-of-living differences. X is a vector which includes population and employment characteristics and regional indicator variables, and ε is the residual term. Our null hypothesis states β_1 will be positive, as enactment of minimum wage legislation is typically considered a more liberal policy. We also expect β_2 and β_3 to be positive, which implies that states with high cost of living levels and states with increasing housing costs will be more likely to increase their minimum wage levels.

The nature of hazard analysis does not allow us to pool data from the two minimum wage cycles together, because it cannot account for states which "fail" (i.e. increase their wage above the federal level) and then are observed at a later time to be "survivors," once the federal minimum wage is raised. Hazard estimates also cannot utilize information for states which are observed as "failures" in the initial period of the sample. Further, it appears that different baseline hazard rates are present in the 1991-1997 cycle than in the 1997-2006 cycle.

Our second specification models the influence of our explanatory variables over the entire span of the data. We estimate a panel probit, with state specific random effects, for whether a state's minimum wage is higher than federally-mandated. This specification takes the form:

$$afedmw_{it} = \beta_0 + \beta_1 LQ_{it} + \beta_2 hpi_{it} + \beta_3 growhpi_{it} + \beta_4 X_{it} + \lambda_i + \varepsilon_{it}.$$

All explanatory variables of interest remain unchanged for this estimation. The probit estimations include indicator variables for each year to capture time trends. This specification will allow state-specific attributes, not already accounted for by the independent variables, to be controlled for in our regression, by the term λ_i.

Technically a state can increase or decrease its legislated minimum wage at any time. In reality states that introduce minimum wages higher than the federal level are not observed to decrease their minimum wage. For this reason we restrict our analysis to include state-year observations in which the state either maintains the federal level, or increases its wage for the first time. We drop observations for states that offered a higher than federal minimum wage in the previous year from analysis because the factors present after the time of adoption are irrelevant to maintaining higher than federal minimum wages.[13]

We first estimate this model for the two minimum wage cycles in isolation as a means to verify whether they are similar to the previous hazard estimation. Then, we expand the analysis to cover the span of both minimum wage cycles.

Our third estimation technique utilizes a continuous outcome variable to capture the effect of our explanatory variables on the magnitude of state minimum wage changes. We estimate the effect of our variables on the magnitude of minimum wage increases

[13] We also conducted the same analysis on a sample that includes observations for states that had already increased their minimum wage. Including these observations does not change the sign or significance of the *LQ* coefficient in the probit and Tobit models for the entire sample or for the 1997-2006 time period. The *LQ* variable is no longer significant in the probit and Tobit for 1991-97. The *hpi* and *growhpi* variables become positive and significant in probits and Tobits for the entire sample. Only the *hpi* level is significant in either regression for 1997-2006 and only the *growhpi* is significant for 1991-1997. We do not report these regressions because they include information that is irrelevant at the time of the state's decision.

using a Tobit regression with state-level random effects (RE).[14] These regressions address a slightly different, yet equally relevant, question by indicating whether the size of a minimum wage increase is influenced by our explanatory variables. In many cases a state has either no minimum wage legislation or a state minimum wage that is less than the federal rate. In these instances we use the federal wage rate as the value for the state in year t, as it is the binding level.[15]

The state percentage deviation from the federal level, $dsmw_{it}$, is constructed by taking the difference in state i's minimum wage level from the federal minimum wage, at time t, and then dividing it by the federal minimum wage at time t.[16] Formally,

$$dsmw_{it} = (smw_{it} - federalmw_t) / federalmw_t .$$

The variable smw_{it} is state i's effective minimum wage during year t, $federalmw_t$ is the federal minimum wage for year i.[17]

---

[14] All specifications were also estimated with state-level fixed effects models (without region indicators), and yielded qualitatively similar results. Since the regional indicators are time invariant and cannot be included in fixed effects models, and because Hausman tests of random versus fixed effects and Breusch-Pagan LaGrange Multiplier tests and all favor random effects in each of the regressions, we report only the random effects results.

[15] All regressions were also calculated using an alternative which specified the dependent variable in terms of the state-mandated minimum wage instead of treating lower wages as simply the federal rate. This change in the dependent variable did not affect the sign or significance of any of the coefficients reported. We do not focus on these estimates because this characterization of the state minimum wage may not represent the "true" value either. This is especially problematic because states with lower-than-federal minimum wages are not likely to adjust their state law if the change does not bring the state level above the federal mandate. Thus some state minimum wages are a non-binding artifact remaining from a point in time where the federal limit overtook the state's mandated minimum wage level. We use the federal level for states that have a lower minimum wage level for this reason in addition to the fact that the federal level is binding.

[16] For states that experience a change in minimum wage within the year, we construct a weighted average of the minimum wage and use this value for the state's year observation. For example, the federal minimum wage value for 1997 is recoded as 4.88 because the minimum wage changed from 4.75 to 5.15 on September 1, 1997.

[17] All of the reported analyses were also conducted with several other specifications of the dependent variable. These include state binding minimum wages, state deviations from federal minimum wage, deviations from the national mean minimum wage, and deviations from the national mean minimum wage weighted by the national standard deviation. Each of these was found to have qualitatively similar results, and significance levels to what is reported here. We report the results for the percentage deviations from the

Tobit estimation accounts for the fact that the percent deviation in the state minimum wage dependent variable is censored at zero for all states with minimum wages less than or equal to the federal minimum wage. This censorship is important as values of zero may not accurately reflect the true preference of the state. The state specific random effects account for other unobservable characteristics which may be influencing state minimum wages but are not captured by our other control variables. Estimation takes the form:

$$dsmw_{it}^* = \beta_0 + \beta_1 LQ_{it} + \beta_2 hpi_{it} + \beta_3 growhpi_{it} + \beta_4 X_{it} + \lambda_i + \varepsilon_{it}$$

$$\text{in which } dsmw_{it} = \begin{cases} dsmw_{it}^* & \text{if } dsmw_{it}^* > 0 \\ 0 & \text{if } dsmw_{it}^* \leq 0. \end{cases}$$

All other variables are the same as the probit estimation. We again restrict our analysis to states that remain at the federal level and the first year of a higher than federal minimum wage.

3.4 Results

Our data spans two major federal minimum wage episodes. The first, earlier cycle, goes from April 1991 until August 1997.[18] The second cycle begins in September 1997 and continues until 2006, the last complete year of data for our political variable. Table 2 presents the Cox proportional hazard estimation results for theses two sample periods. Of the three main explanatory variables, *LQ*, *hpi,* and *growhpi*, only *LQ*, our

federal minimum wage as the dependent variable because they are somewhat more intuitive and because they have a lower bound of zero for all censored observations.

[18] There was a change in both 1996 and 1997, but this was very small, and both were part of the same piece of legislation, so it is treated as one large change in 1997 in these regressions. We performed the same set of regressions using 1996 as the last year and the signs and significance levels did not change qualitatively.

proxy for a state's political views significantly affects a state's minimum wage level in both cycles. *LQ* is positive and significant at the 1 percent level in both cases, indicating that the liberal leaning of a state does significantly contribute to a state raising its minimum wage level above that which is federally mandated. Neither the level, nor the growth rate of our cost of living variable are statistically significant in either sample. Our controls for population, population growth and per capita income are significant in the 1991-1997 subsample, but not in the 1997-2006 subsample.

Table 3 presents three sets of estimates for our random effects probit specification. The first two data columns check this estimation versus the previous hazard estimation. The probits for the two sample periods in isolation produce similar estimates to the hazard models in Table 2. The coefficient on the political variable is again positive and significant, and the cost of living variables are not found to be significant. In the 1991 to 1997 sample the probits do not attribute significance to the population growth control, whereas the hazard estimation finds it to be significant at the 10% level. As in the hazard estimation for 1991-1997 *population* is negative and income per capita is positive, with both significant at the 5% level in both models. From 1997-2006, only *LQ* is found to be significant and positive in the latter sample, just as the hazard model predicted. These two sets of results indicate that the probit and hazard models are behaving similarly and closely measuring the same effects.

Column three of Table 3 presents probit estimates for the entire sample spanning both minimum wage cycles. The political variable is positive and significant at the 1% level over the entire sample. None of the other variables are found to be significant for predicting whether a state increases its minimum wage above the federal level.

Table 4 presents the estimates from the Tobit regressions which capture the magnitude of a minimum wage increase in relation to the explanatory variables. These regressions indicate a positive and significant effect from the political variable over each time period. However, the significance of the variable is somewhat lower in the subsamples than in the proportional hazard or probit estimates, achieving the 10% level for 1991-1997 and 5% level for 1997-2006. Neither cost of living variable is significant for the 1991-1997 subsample. However, there is some evidence that the magnitude of a minimum wage increase depends on the level of home prices in the 1997-2006 sample, as the *hpi* variable is positive and significant at the 1% level. For the entire sample from 1991-2006 *LQ* is positive and significant at the 1% level. Also, the *growhpi* variable becomes positive and significant at the 10% level. This is an indication that over the entire sample, conditional on a state increasing its minimum wage, the *magnitude* of the increase appears to be influenced by growth in the cost of living measure.

Taken together, our findings indicate that political factors are the only force which consistently explains whether a state will raise its minimum wage level above the federal standard. Controlling for characteristics of the population, employment rates, and regional characteristics, liberal-leaning states are significantly more likely to choose to increase their minimum wage above the federal rate. It is somewhat surprising that cost of living concerns do not significantly influence a state's decision to adopt a higher than federal minimum wage.[19] There is some evidence, however, that the *magnitude* of a state minimum wage increase is sensitive to cost of living.

---

[19] We measure state level differences in the cost of living using the home price index variables and feel this is appropriate given that more than one third of household income is spent on housing. However, since this finding is somewhat unexpected it is worth investigating whether it does in fact capture enough variation.

3.5 CONCLUSION

Most of the previous literature on minimum wages has looked at the effect they have on short-run labor force participation, unemployment, or other specific economic outcomes. We extend the existing literature by examining how political and economic factors contributed to differences in state minimum wage laws over the two federal minimum wage cycles spanning from 1991 until 2006. Our results indicate that political leanings are the only factor that is significant in explaining differences in minimum wage laws within each of the last two minimum wage cycles and also over our *entire* observed sample. It is not surprising that states with liberal voting records are significantly more likely to have a higher than federal minimum wage. However, we find little evidence in the data linking cost of living considerations to state minimum wage legislation. The level of our cost of living variable appears to influence the magnitude of increases since 1997, but cost of living factors do not have any statistically significant influence on a state's decision to increase its minimum wage above the federal level. This result is interesting since proponents of raising the minimum wage usually cite the rising the cost of living as the main justification.

These findings could have predictive value if the latest federal legislation turns out to trigger a new round of state and local minimum wage changes driven by the same influences that we have analyzed. Not only could the new federal wage rate possibly be

---

Additional regressions estimated without the *LQ* variable indicate a positive and significant effect from the *hpi* and *growhpi* which suggests they have some impact. Also, other cost of living measures which could potentially be used are likely to be endogenously determined with minimum wages.

binding in some low-income areas of the country, but many states may now be more motivated than in the past to increase their wage rate above the new federal level.

Research by Neumark and Nizalova (2007) has shown that binding minimum wages distort incentives for work and schooling decisions among workers which leads to negative long-run consequences. Whether politically or economically driven minimum wages are more likely to be binding is an open question for future research. Our findings suggest minimum wages are more closely related to political leanings than economic conditions, and this could prove economically detrimental in the long run.

REFERENCES

American Federation of Labor and Congress of Industrial Organizations (AFL-CIO). "Agenda to Create Jobs and Lift the Economy." 2003 http://www.aflcio.org/mediacenter/prsptm/pr01062003.cfm.

Americans for Democratic Action (ADA). "ADA Voting Records." http://www.adaction.org/votingrecords.htm, 2007.

Association of Community Organizations for Reform Now (ACORN). "ACORN helped raise the federal minimum wage." http://www.acorn.org/index.php?id=2668, 2007.

Ballot Initiative Strategy Center (BISC). "State Minimum Wage Campaigns." http://www.ballot.org/index.asp?Type=B_DIR&SEC= {4A31D1AF-7D75-413B-847C-41F5AC3B2751}, 2006.

Burkhauser, R. V.; K. A. Couch, and A. J. Glenn. "Public Policies for the Working Poor: The Earned Income Tax Credit versus Minimum Wage Legislation." *Research in Labor Economics*. 15, 1996, 65–109.

Burkhauser, R. V., and J. J. Sabia. "The Effectiveness of Minimum-wage Increases in Reducing Poverty: Past, Present, and Future." *Contemporary Economic Policy*, 25(2), 2007, 262–81.

Center for American Progress. "Life at Minimum Wage." http://www.americanprogress.org/issues/2007/07/min_wage.html, 2007.

Center for Policy Alternatives. "Minimum Wage Policy Brief." http://www.stateaction.org/issues/issue.cfm/issue/MinimumWage.xml, 2007.

Chaplin, D. D., M. D. Turner, and A. D. Pape. "Minimum Wages and School Enrollment of Teenagers: A Look at the 1990's." *Economics of Education Review*, 22(1), 2003, 11–21.

Consumer Expenditure Survey. "Table 46. Income Before Taxes." ftp://ftp.bls.gov/pub/special.requests/ce/standard/2006/income.txt

Fairchild, D. J. "Does the Minimum Wage Help the Poor?" *Forum for Social Economics* 34(1–2), 2005, 31–42.

Falk, A., E. Fehr, and C. Zehnder. 2006. "Fairness Perceptions and Reservation Wages: The Behavioral Effects of Minimum Wage Laws." *Quarterly Journal of Economics*, 121(4), 2006, 1347–81.

Levin-Waldman, O. M. "Exploring the Politics of the Minimum Wage." *Journal of Economic Issues*, 32(3), 1998, 773–803.

Levin-Waldman, O. M. "The Minimum Wage and Regional Wage Structure: Implications for Income Distribution." *Journal of Economic Issues*, 36(3), 2002, 635–57.

Neumark, D., and O. Nizalova. "Minimum Wage Effects in the Longer Run." *Journal of Human Resources*, 42(2), 2007, 435–52.

Neumark, D., M. Schweitzer, and W. Wascher. "The Effects of Minimum Wages on the Distribution of Family Incomes: A Nonparametric Analysis." *Journal of Human Resources*, 40(4), 2005, 867–94.

Neumark, D., W. Wascher. "Do Minimum Wages Fight Poverty?" *Economic Inquiry*, 40(3), 2002, 315–33.

Neumark, D. and W. Wascher. "Minimum Wages and Skill Acquisition: Another Look at Schooling Effects." *Economics of Education Review*, 22(1), 2003, 1–10.

Neumark, D., and W. Wascher. "Minimum Wages and Employment: A Review of Evidence from the New Minimum Wage Research." Ph.D. dissertation, University of California-Irvine, 2006.

Neumark, D., and W. Wascher. "Minimum Wages, the Earned Income Tax Credit, and Employment: Evidence from the Post-Welfare Reform Era." National Bureau of Economic Research Working Paper No. 12915, 2007.

Sander, R. H., and E. D. Williams. 2005. "Santa Monica's Minimum Wage: Assessing the Living Wage Movement's New Frontier." *Economic Development Quarterly*, 19(1), 2005, 25–44.

Singell, L. D., and J. R. Terborg. "Employment Effects of Two Northwest Minimum Wage Initiatives." *Economic Inquiry*, 45(1), 2007, 40–55.

US Department of Labor. "Minimum Wage in America." http://www.dol.gov/esa/minwage/america.htm, 2007.

Waltman, J. and S. Pittman. "The Determinants of State Minimum Wage Rates: A Public Policy Approach." *Journal of Labor Research*, 23(1), 2002, 51–56.

FIGURE 1. NUMBER OF STATES WITH HIGHER THAN FEDERAL MINIMUM WAGES BY YEAR

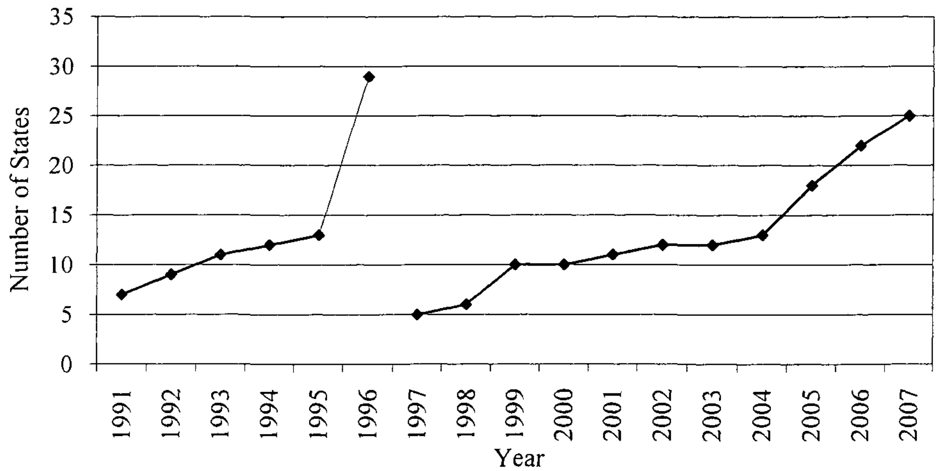


TABLE 1. VARIABLE DEFINITIONS

Variable	Definition	Source
afedmw	Equal to one if state's minimum wage level is greater than the federal level	Created from US and State Departments of Labor
dsmw	Percent deviation of a state's minimum wage from the federal minimum wage	State's Department of Labor
LQ	Liberal voting percentage	Americans for Democratic Action
hpi	Interstate housing price index divided by 100	Office of Federal Housing Enterprise Oversight
growhpi	Percentage growth of housing price index	Office of Federal Housing Enterprise Oversight
age	Population age measured as births to deaths ratio	US Census Bureau
population growth	Percentage change in state population	US Census Bureau
population	State population estimate divided by 1 million	US Census Bureau
income per capita	Total state income divided by population	Regional Economic Information System
employment	Ratio of employed persons to the entire population	Regional Economic Information System

TABLE 2. COX PROPORTIONAL HAZARD ESTIMATES OF STATE MINIMUM WAGES

	1991-1997	1997-2006
LQ	1.994***	3.490***
	(0.698)	(1.181)
hpi	-0.837	0.398
	(0.556)	(0.389)
growhpi	-3.739	4.106
	(6.710)	(6.044)
age	0.618	0.097
	(0.384)	(1.342)
population growth	-23.578*	-11.677
	(14.013)	(28.446)
population	-0.086**	-0.016
	(0.034)	(0.074)
income per capita	0.167**	0.014
	(0.083)	(0.040)
employment	-2.502	0.920
	(4.802)	(5.920)
Log Likelihood	-103.39	-66.63
Wald Statistic	$X^2(11)$=	$X^2(11)$=
	44.47***	66.26***
N Observations	269	405

Notes: Standard errors are presented in parentheses. *** indicates significance at the 1 percent level, ** 5 percent, and * 10 percent. These also include region indicator variables which are not reported.

TABLE 3. RANDOM EFFECTS PROBIT ESTIMATES OF STATE MINIMUM WAGES

	1991-1997	1997-2006	1991-2006
LQ	1.804**	2.256***	1.811***
	(0.731)	(0.645)	(0.465)
hpi	-0.874	0.544	0.192
	(0.600)	(0.332)	(0.266)
growhpi	-0.103	2.365	2.873
	(5.909)	(3.784)	(2.762)
age	0.532	0.298	0.259
	(0.379)	(0.337)	(0.244)
population growth	-17.871	-18.114	-7.965
	(18.484)	(16.206)	(12.651)
population	-0.068**	0.003	-0.025
	(0.031)	(0.027)	(0.019)
income per capita	0.163**	-0.007	0.043
	(0.080)	(0.046)	(0.039)
employment	-1.887	0.430	-0.764
	(3.972)	(3.513)	(2.547)
Log Likelihood	-62.29	-57.30	-118.25
Wald Statistic	$X^2(17)=$	$X^2(20)=$	$X^2(26)=$
	44.55***	31.28**	70.29***
N Observations	269	405	624

Notes: Standard errors are presented in parentheses. *** indicates significance at the 1 percent level, ** 5 percent, and * 10 percent. All regressions include year and region indicator variables.

TABLE 4. RANDOM EFFECTS TOBIT ESTIMATES OF STATE MINIMUM WAGES

	1991-1997	1997-2006	1991-2006
LQ	0.104*	0.128**	0.147***
	(0.062)	(0.040)	(0.044)
hpi	-0.037	0.056***	0.029
	(0.050)	(0.022)	(0.025)
growhpi	0.285	0.183	0.435*
	(0.422)	(0.262)	(0.260)
age	0.035	0.002	0.020
	(0.026)	(0.022)	(0.023)
population growth	-0.894	-1.573	-0.870
	(1.444)	(1.032)	(1.086)
population	-0.004	0.001	-0.001
	(0.003)	(0.002)	(0.002)
income per capita	0.009	-0.001	0.003
	(0.007)	(0.003)	(0.004)
employment	-0.130	0.129	-0.065
	(0.387)	(0.232)	(0.245)
Log Likelihood	4.623	3.812	40.030
Wald Statistic	$X^2(17)=$	$X^2(20)=$	$X^2(26)=$
	38.270***	49.430***	56.240***
N Observations	269	405	624

Notes: Standard errors are presented in parentheses. *** indicates significance at the 1 percent level, ** 5 percent, and * 10 percent. All regressions include year and region indicator variables.