Undergraduate Economic Review

Volume 14 | Issue 1 Article 12

2017

Examining the Impact of Education on Diabetes Rates

Mackenzie Bogiages
University of North Carolina at Chapel Hill, bogiages@live.unc.edu

Examining the Impact of Education on Diabetes Rates

Abstract

Diabetes rates are a problem of growing concern in the U.S. Over 30 million people in the U.S. have been diagnosed with diabetes, and over 90 percent of that group has the largely preventable type 2 diabetes. In order to thwart the disease that costs the U.S. $245 billion a year, the populations most susceptible to diabetes need to be identified and educated. Using data from MEPS, education and several other factors have been identified and their relationships with diabetes have been analyzed to target at risk populations and outline tactics to educate those in dire need of preventative care.

Examining the Impact of Education on Diabetes Rates

Introduction

As healthcare costs continue to rise, the costs are often broken into subcategories based on diseases, with diabetes near the top of the list. A whopping $245 billion was spent on diabetes care in 2012 [3]. This number is huge considering that type 2 diabetes is largely preventable with non-medicinal precautions such as a healthy diet and exercise. In spite of this, Americans continue to lead dangerously unhealthy lives that lead to their premature deaths while producing large excess costs along the way. How much influence education has on diabetes needs to be estimated so that insurers and healthcare providers can identify their target audience before spending more money on diabetes prevention. A 2013 report by the State of Illinois shows that in Illinois, adults with less than a high school education have the highest diabetes rates at 14.6 percent and those who graduated college have the lowest at 5.5 percent [2]. Another study done at Virginia Commonwealth University shows that it is not just that those with less education are more prone to diseases, it is specifically diseases with relatively easy prevention such as diabetes and heart disease [7]. Of adults age 20 and older, 35.8 percent are overweight, 25.1 percent are obese, and 4.5 percent are extremely obese according to data published by the Medical Expenditure Panel Survey in 2012 [4]. This means that less than 34 percent of American adults are considered to be a healthy weight.

Data analysis and regression techniques were used in this project to determine the relationship between the dependent variable (diabetes diagnosis) and the independent variable (highest level of education obtained) while controlling for age, race, sex, ethnicity, level of English spoken, and other variables. Data from the Medical Expenditure Panel Survey were used in this project to gather information on adult diabetes patients, their backgrounds, and their highest level of education obtained to determine the relationship among the variables. Before analyzing the data, I hypothesized that there would be a negative correlation between diabetes diagnosis and highest degree of education obtained. From data gathered in the Medical Expenditure Panel Survey, it can be seen that obesity and extreme obesity both decrease as level of education increases, so I expect diabetes diagnoses to follow the same trend [4]. The purpose of this project is to determine if the relationship between diabetes rates and education is statistically significant. This information could be used to determine if the populations with no or lower degrees of education should be

targeted for preventative diabetes education and information on how to take care of your diabetes to reduce national diabetes expenditures.

Health disparities are everywhere. While analyzing the data presented in this project, we must keep in mind that there is a disparity in overall health based on differences in education [1]. As age increases, particularly above 69, the probability of survival for a college graduate decreases less than the probability of survival of someone with less than a high school degree [1]. There are many theories that could explain this discrepancy. The efficient producer hypothesis suggests that better educated individuals are more efficient at taking care of themselves and producing better health because lessons or schooling may teach students the importance of taking better care of themselves (education influences health). The direct income hypothesis, allostatic load hypothesis, income inequality hypothesis, and access to care hypothesis also suggest that education level impacts health level [1]. On the other hand, the productive time hypothesis suggests that health level influences education level, because bad health leads to lower productivity and less time to spend on education [1]. Regardless of which factor is independent and which is dependent, we can see that a positively correlated relationship exists between education level and health level, and diabetes diagnosis is considered in a patient's level of health. I expected that this study would support the efficient producer hypothesis, and demonstrate that higher levels of education positively impact levels of health, and in this case, that higher levels of education are negatively correlated with diabetes rates.

Measuring the data

Gathering data

Data for this project were gathered from the Medical Expenditure Panel Survey's file "MEPS Panel 19 Longitudinal Data File." Due to the brevity of this project, cuts were made to condense the size of this massive data set which originally contained 3,489 variables and spanned five time periods. This study will focus on only data gathered in the year 2015, rather than making it a longitudinal study and utilizing all five of the different times that data were gathered for the original MEPS data file. Numerous variables were dropped from the data set; the variables used in this study will be diabetes diagnosis, region of the United States where the patient resides, age at the time of the last survey (only patients 18 and older are considered to focus on adult diabetes alone, rather than with childhood diabetes), sex, race, ethnicity, student status, how well English is spoken by the patient, high blood pressure diagnosis, heart disease diagnosis, heart attacks, strokes, and high cholesterol diagnosis. All of the utilized variables can be seen in the Appendix in Figure 1. The choices of which independent variables were to be used in this model were based on which of the variables were included in this MEPS data set, which of the variables had the least missing

values from the last time period surveyed, and which of the independent variables most frequently appeared as statistically significant in previously published studies about diabetes rates and education levels in the United States. In the end, 18 variables were chosen in hopes that they could help determine the impact of different levels of education on diabetes rates. The levels of highest degree of education obtained include: no degree (did not finish high school or obtain GED), GED, high school diploma, bachelor's degree, master's degree, doctorate degree, and other degree.

Cleaning the data

The first step in creating this Stata output was to drop the statistically insignificant variables, keeping only the variables that I was most interested in. Next, we changed the highest degree of education obtained variable to eliminate all values that were not an actual level of education (such as did not respond, not applicable, etc.), because the goal of this study is to determine the relationship between diabetes and education. The next step was to eliminate all patients under the age of 18, so that the study focuses on adult diabetes rather than childhood diabetes, which would affect the results, since no child has a high degree of education. The diabetes diagnosis was changed to be a binary variable rather than just a categorical variable, as it was originally coded to be in the MEPS data file. Finally, the summary statistics of the data can be computed and analyzed while specifying indicators at each category for the discrete variables: region, highest degree of education obtained, race, and student status, as can be seen in Figure 2. After tabulating both diabetes diagnosis and highest degree of education obtained, we can move forward while keeping in mind that 10.73 percent of our sample patients have diabetes (Figure 3), 20.01 percent of them have no educational degree, 48.34 percent have a high school degree or its equivalent, and 31.66 percent have a bachelor's degree or higher degree of education (Figure 4).

The dependent variable in this model, diabetes diagnosis, is a binary value with zero meaning a patient does not have diabetes and one meaning that they have diabetes. Due to the fact that the dependent variable is binary, a nonlinear regression model must be used, which is why logit is implemented in Figure 5. From these results, we can see that having a bachelor's degree or higher appears to have a strong and statistically significant negative impact on diabetes diagnosis, meaning those with higher degrees of education are less likely to have diabetes according to this model. Having a doctorate degree makes someone .9578 times less likely to have diabetes than someone who has no degree, and this is statistically significant at a confidence level of .01. Having a master's degree makes someone only.4558 times less likely to have diabetes than someone who has no degree. Having a bachelor's degree makes someone .4827 times less likely to have diabetes than someone without a degree, but having a high school

diploma only makes someone .07564 times less likely to have diabetes than someone without a degree: a dramatic difference for a few more years of education. The variables how well English is spoken by the patient and coronary heart disease diagnosis are not statistically significant at the .05 level.

Examining the impact of education

From running the margins analysis sas een in Figure 6, we can tell if all patients had no degree, the average rate of diabetes would be .118677. If all patients had a doctorate degree, the average rate of diabetes for this sample would be .05633. The biggest difference in margin among the different levels of education is between high school diploma recipients and bachelor's degree recipients at .02964. Another large difference comes between those with master's degrees and those with doctorate degrees, but this may be influenced by the fact that only one percent of the sample holds a doctorate degree, so information may be insufficient and biased. Another reason it could be biased is that those with doctorate degrees include doctors, who have more information and, most likely, more incentive to take care of themselves and prevent diabetes and other health maladies. In Figure 7, I ran a joint test of constituent contrasts; the chi value is large, so we fail to reject the null. While excluding the other degree category, which may contain different kinds of degrees that require varying numbers of years of schooling, we can see a decrease in diabetes rates with every increase in education level from no degree to doctorate degree (Figure 13).

Examining the impact of race

Another important factor to consider in this model is race. In this data file, all races are represented at every level of education. The race categories are not mutually exclusive, but because there is no specific breakdown of what other races someone is. If they are listed as Asian with other races, the other races are not listed. Running the model combining mixed race data with those who identified with one race did not yield significantly different coefficients. For this reason, I chose to execute the next portion of the test as if race was mutually exclusive. In Figure 8 it can be seen from A, B, and C respectively that the rate of diabetes is .1092 for the black portion of the sample, .1473 for the Asian portion of the sample, and .0968 for the white portion of the sample. Typically, African Americans have a slightly higher rate of diabetes than Asians, but these results may be biased seeing as I did not have enough information and tested them like mutually exclusive variables [5]. The rate of diabetes for blacks and whites that I found is similar to the rates published by the Center for Disease Control, but the rate of diabetes among Asians that I found is much higher than the rate of 9 percent that the CDC published [5].

In Figure 9, logit was run again, but this time with an interaction variable between highest degree of education obtained and high blood pressure diagnosis. Then, the margins were run, and in Figure 10 we can see that 20.32 percent of the part of the sample with no degree and high blood pressure have diabetes, while only 13.81 percent of those with a bachelor's degree and high blood pressure have diabetes. This may be because those with higher education know how to take better care of themselves, to better manage their high blood pressure, and to prevent diabetes more successfully. In Figure 11 and Figure 12, the same tests were done, but this time using the interaction variable between highest degree of education obtained and high cholesterol diagnosis. The results of these tests are lesser than the previous ones, but they point in the same direction: patients with higher levels of education who know the importance of taking care of themselves are able to better manage their high cholesterol and to prevent diabetes more successfully.

Conclusions

From running these tests and analyzing the results, we can see that highest degree of education obtained has a strong and statistically significant impact when comparing those who got at least a bachelor's degree with those who did not. While the rate of diabetes decreases with every jump in education, the "make or break point" is the separation between high school graduates and those with a bachelor's degree. To further analyze this separation of those who went to college and those who did not, I temporarily disregard those with other degrees as there is no way to further separate it. Highest degree of education obtained is then changed to be a binary variable (patients did obtain a minimum of a bachelor's degree as one and patients did not obtain a minimum of a bachelor's degree as zero). From Figure 13, it can be seen that getting a bachelor's degree has a huge and statistically significant impact on getting diabetes when compared to those without a bachelor's degree which shows us that college dramatically influences the rate of diabetes. On average, those with bachelor's degrees are -.4578 times less likely to get diabetes when compared to the population without bachelor's degrees.

A major issue to consider when studying the health of adults is the categorization of weight categories. The Medical Expenditure Panel Survey uses the Body Mass Index (BMI) to categorize adults into underweight, health weigh, overweigh, obese, and extremely obese. While these categorizations do not impact diabetes diagnosis, it is important to note, that BMI is often viewed as an outdated method of determining health status. The BMI is a formula developed two centuries ago in order to measure obesity in a large population [7]. It is based on body measurements of the average man and tends to incorrectly classify those who are not average men. More women are categorized as overweight and obese,

because they are not men. BMI also tends to classify those with significant muscle masses (such as professional athletes) as overweight or obese, because they are not average (and not always men) [7]. Keeping in mind that the categorization of body types is flawed, we are still able to rely on diabetes diagnosis as an accurate, but not comprehensive, measure of one's health.

Recommendations

While far from perfect, this study does show results that agree with published data. If this study were to be recreated, a few changes could yield more information. It would be beneficial to break the other degree category of education down so that it could be further analyzed. Race should be more thoroughly recorded in the data so that mixes of different races can be analyzed. Health factors that cause diabetes such as poor diet and lack of exercise are often attributed to before someone goes to college, which suggests that there may be more a relationship between diabetes rates and whether or not someone is projected to go to college from a certain age, but this would require further analysis on data not available in this data file. After thorough analysis, it can be concluded that the initial proposed hypothesis is correct, and highest level of education obtained does have an impact on diabetes diagnosis. More specifically, whether or not someone obtains a bachelor's degree (at the minimum) appears to have a very large and statistically significant on whether or not someone has diabetes.

Grade schools require numerous health and physical education courses throughout the years, so it makes sense that the populations that complete more school have lower diabetes rates. Most universities require a health or personal fitness course to be completed prior to graduation. The impact of these courses on the student body's health should be studied prior to determining the true impact of a bachelor's degree on diabetes rates. There may be a confounding variable that makes young students on the "college bound track" less likely to have health issues such as diabetes.

One possible solution to the diabetes epidemic would be to offer diabetes preventative education in public schools at a young age, before students are legally old enough to drop out and miss this critical information, and before they are old enough to be responsible for their own health via diet and exercise regime. To reduce diabetes rates (as well as the national expenditure on diabetes), we should invest more in the health education of those without bachelor's degrees, as well as those not projected to go to college and teach them about the importance of diabetes prevention and care.

Bibliography

1. Bhattacharya, Jay, Timothy Hyde, and Peter Tu. Health Economics. Palgrave Macmillan, 2014.

2. “Chronic Disease Burden Update.” *Illinois Department of Public Health Report*, vol. 2, no. 18, Nov. 2013.

3. Institute for Health Metrics and Evaluation. "Diabetes, heart disease, and back pain dominate US health care spending”. ScienceDaily. ScienceDaily, 27 December 2016.

4. Medical Expenditure Panel Survey. “MEPS HC-183: MEPS Panel 19 Longitudinal Data File.” Sept. 2017.

5. “National Diabetes Statistics Report, 2017.” *Center for Disease Control*, 2017.

6. O’Neil, Cathy. *Weapons of Math Destruction.* Penguin Books, 2017.

7. Seuring, Till, Olga Archangelidi, and Marc Suhrcke. “The Economic Costs of Type 2 Diabetes: A Global Systematic Review.” *Pharmacoeconomics* 33.8 (2015): 811–831. *PMC*. Web. 29 Oct. 2017.

Appendix

```
. de

Contains data from /Users/Bogiages/Desktop/Diabetes Project/C:\MEPS\PROG/H181.dta
  obs:        13,658
 vars:            18                          30 Nov 2017 10:35
 size:       245,844
--------------------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
--------------------------------------------------------------------------------------------
REGION15        byte    %15.0g     H1810966X
                                              CENSUS REGION AS OF 12/31/15
AGE15X          byte    %8.0g                 AGE AS OF 12/31/15(EDITED/IMPUTED)
SEX             byte    %18.0g     H1811008X
                                              SEX
RACEAX          byte    %32.0g     H1810953X
                                              ASIAN AMONG RACES RPTD (EDITED/IMPUTED)
RACEBX          byte    %32.0g     H1810954X
                                              BLACK AMONG RACES RPTD (EDITED/IMPUTED)
RACEWX          byte    %32.0g     H1810958X
                                              WHITE AMONG RACES RPTD (EDITED/IMPUTED)
HISPANX         byte    %14.0g     H1810335X
                                              HISPANIC ETHNICITY (EDITED/IMPUTED)
EDUYRDG         byte    %41.0g     H1810268X
                                              YEAR OF EDUCATION OR HIGHEST DEGREE
HIDEG           byte    %23.0g     H1810334X
                                              HIGHEST DEGREE WHEN FIRST ENTERED MEPS
FTSTU15X        byte    %18.0g     H1810312X
                                              STUDENT STATUS IF AGES 17-23 - 12/31/15
HWELLSPE        byte    %41.0g     H1810443X
                                              HOW WELL PERSON SPEAKS ENGLISH
HIBPDX          byte    %18.0g     H1810333X
                                              HIGH BLOOD PRESSURE DIAG (>17)
BPMLDX          byte    %18.0g     H1810105X
                                              MULT DIAG HIGH BLOOD PRESS (>17)
CHDDX           byte    %18.0g     H1810136X
                                              CORONARY HRT DISEASE DIAG (>17)
MIDX            byte    %18.0g     H1810612X
                                              HEART ATTACK (MI) DIAG (>17)
STRKDX          byte    %18.0g     H1811067X
                                              STROKE DIAGNOSIS (>17)
CHOLDX          byte    %18.0g     H1810158X
                                              HIGH CHOLESTEROL DIAGNOSIS (>17)
DIABDX          byte    %18.0g     H1810206X
                                              DIABETES DIAGNOSIS (>17)
--------------------------------------------------------------------------------------------
Sorted by:
     Note: Dataset has changed since last saved.
```

Figure 1

```
. summ DIABDX i.REGION15 AGE15X SEX i.RACEAX i.RACEBX i.RACEWX HISPANX i.HIDEG i.FTSTU15X i.HWELLSPE HIBPDX BPMLDX CHDDX MIDX ST
> RKDX CHOLDX
```

Variable	Obs	Mean	Std. Dev.	Min	Max
DIABDX	13,658	.1072631	.3094588	0	1
REGION15					
2 MIDWEST	13,658	.1930004	.3946678	0	1
3 SOUTH	13,658	.376922	.4846328	0	1
4 WEST	13,658	.2686338	.4432652	0	1
AGE15X	13,658	45.7969	17.84607	18	85
SEX	13,658	.4657344	.4988427	0	1
RACEAX					
2 ASIAN -..	13,658	.006077	.0777209	0	1
3 ALL OTH..	13,658	.9158003	.2776973	0	1
RACEBX					
2 BLACK -..	13,658	.0120808	.1092509	0	1
3 ALL OTH..	13,658	.8031923	.3976003	0	1
RACEWX					
2 WHITE -..	13,658	.0219651	.1465751	0	1
3 ALL OTH..	13,658	.2803485	.4491859	0	1
HISPANX	13,658	.2866452	.4522109	0	1
HIDEG					
2 GED	13,658	.0451018	.207535	0	1
3 HIGH SC..	13,658	.4381315	.4961757	0	1
4 BACHELO..	13,658	.1534632	.3604466	0	1
5 MASTER'..	13,658	.0682384	.2521638	0	1
6 DOCTORA..	13,658	.0181579	.133527	0	1
7 OTHER D..	13,658	.0767316	.2661748	0	1
FTSTU15X					
1 FULL-TIME	13,658	.0472251	.2121277	0	1
2 PART-TIME	13,658	.0120808	.1092509	0	1
HWELLSPE					
1 VERY WELL	13,658	.1798946	.3841137	0	1
2 WELL	13,658	.060697	.2387825	0	1
3 NOT WELL	13,658	.0706546	.2562563	0	1
4 NOT AT ..	13,658	.0453214	.2080158	0	1
HIBPDX	13,658	.3246449	.4682592	0	1
BPMLDX	13,658	.3793381	.5940597	0	2
CHDDX	13,658	.0452482	.2078556	0	1
MIDX	13,658	.0353639	.1847046	0	1
STRKDX	13,658	.039098	.1938351	0	1
CHOLDX	13,658	.28152	.4497569	0	1

Figure 2

```
. tabulate DIABDX
```

DIABETES DIAGNOSIS (>17)	Freq.	Percent	Cum.
0	12,193	89.27	89.27
1 YES	1,465	10.73	100.00
Total	13,658	100.00	

Figure 3

```
. tabulate HIDEG
```

HIGHEST DEGREE WHEN FIRST ENTERED MEPS	Freq.	Percent	Cum.
1 NO DEGREE	2,734	20.02	20.02
2 GED	616	4.51	24.53
3 HIGH SCHOOL DIPLOMA	5,984	43.81	68.34
4 BACHELOR'S DEGREE	2,096	15.35	83.69
5 MASTER'S DEGREE	932	6.82	90.51
6 DOCTORATE DEGREE	248	1.82	92.33
7 OTHER DEGREE	1,048	7.67	100.00
Total	13,658	100.00	

Figure 4

```
. logit DIABDX REGION15 AGE15X SEX RACEAX RACEBX RACEWX HISPANX i.HIDEG FTSTU15X HWELLSPE HIBPDX BPMLDX CHDDX MIDX STRKDX CHOLDX
>

Iteration 0:   log likelihood = -4654.0282
Iteration 1:   log likelihood = -3926.2989
Iteration 2:   log likelihood = -3615.6711
Iteration 3:   log likelihood = -3605.6363
Iteration 4:   log likelihood =   -3605.24
Iteration 5:   log likelihood = -3605.2378
Iteration 6:   log likelihood = -3605.2378

Logistic regression                             Number of obs     =     13,658
                                                LR chi2(21)       =    2097.58
                                                Prob > chi2       =     0.0000
Log likelihood = -3605.2378                     Pseudo R2         =     0.2254
```

DIABDX	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
REGION15	.015833	.0312994	0.51	0.613	-.0455128	.0771787
AGE15X	.0255081	.0022763	11.21	0.000	.0210465	.0299696
SEX	.0903842	.0625352	1.45	0.148	-.0321826	.212951
RACEAX	.0146007	.1367209	0.11	0.915	-.2533674	.2825687
RACEBX	-.2008183	.1225681	-1.64	0.101	-.4410472	.0394107
RACEWX	.0966742	.1192835	0.81	0.418	-.1371172	.3304657
HISPANX	.520955	.1074076	4.85	0.000	.3104399	.7314701
HIDEG						
2 GED	.0274676	.1459302	0.19	0.851	-.2585504	.3134856
3 HIGH SCHOOL DIPLOMA	-.0756538	.0845175	-0.90	0.371	-.2413051	.0899974
4 BACHELOR'S DEGREE	-.4826827	.1173266	-4.11	0.000	-.7126386	-.2527268
5 MASTER'S DEGREE	-.4557508	.1566461	-2.91	0.004	-.7627714	-.1487301
6 DOCTORATE DEGREE	-.9575485	.2925963	-3.27	0.001	-1.531027	-.3840704
7 OTHER DEGREE	-.1146969	.1334959	-0.86	0.390	-.3763442	.1469503
FTSTU15X	-1.020703	.3855698	-2.65	0.008	-1.776406	-.2650006
HWELLSPE	.0311724	.0379445	0.82	0.411	-.0431975	.1055424
HIBPDX	1.591172	.1406433	11.31	0.000	1.315516	1.866827
BPMLDX	-.4040496	.1034303	-3.91	0.000	-.6067693	-.20133
CHDDX	.1423223	.1158403	1.23	0.219	-.0847206	.3693651
MIDX	.2597604	.1269764	2.05	0.041	.0108912	.5086297
STRKDX	.2295882	.1101262	2.08	0.037	.0137448	.4454315
CHOLDX	1.166295	.069479	16.79	0.000	1.030119	1.302472
_cons	-4.392157	.879675	-4.99	0.000	-6.116288	-2.668025

Figure 5

```
. margins HIDEG

Predictive margins                              Number of obs     =     13,658
Model VCE    : OIM

Expression   : Pr(DIABDX), predict()
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
HIDEG						
1 NO DEGREE	.118677	.0057141	20.77	0.000	.1074776	.1298765
2 GED	.1210181	.0109617	11.04	0.000	.0995336	.1425026
3 HIGH SCHOOL DIPLOMA	.1123979	.0038238	29.39	0.000	.1049034	.1198925
4 BACHELOR'S DEGREE	.082754	.0060024	13.79	0.000	.0709896	.0945184
5 MASTER'S DEGREE	.0845064	.0090727	9.31	0.000	.0667242	.1022885
6 DOCTORATE DEGREE	.0563383	.0133657	4.22	0.000	.0301421	.0825346
7 OTHER DEGREE	.1092538	.0089692	12.18	0.000	.0916744	.1268332

Figure 6

```
. margins HIDEG, contrast

Contrasts of predictive margins
Model VCE    : OIM

Expression   : Pr(DIABDX), predict()
```

	df	chi2	P>chi2
HIDEG	6	42.31	0.0000

Figure 7

```
. margins, at (RACEAX=1 RACEBX=3 RACEWX=3)

Predictive margins                              Number of obs     =     13,658
Model VCE    : OIM

Expression   : Pr(DIABDX), predict()
at           : RACEAX          =           1
               RACEBX          =           3
               RACEWX          =           3

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .1092172   .0108373    10.08   0.000     .0879764    .1304579
------------------------------------------------------------------------------

. margins, at (RACEAX=3 RACEBX=1 RACEWX=3)

Predictive margins                              Number of obs     =     13,658
Model VCE    : OIM

Expression   : Pr(DIABDX), predict()
at           : RACEAX          =           3
               RACEBX          =           1
               RACEWX          =           3

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .1472752   .0067099    21.95   0.000      .134124    .1604263
------------------------------------------------------------------------------

. margins, at (RACEAX=3 RACEBX=3 RACEWX=1)

Predictive margins                              Number of obs     =     13,658
Model VCE    : OIM

Expression   : Pr(DIABDX), predict()
at           : RACEAX          =           3
               RACEBX          =           3
               RACEWX          =           1

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .0967929    .002769    34.96   0.000     .0913658    .1022199
------------------------------------------------------------------------------
```

Figures 8A, 8B, 8C

```
. logit DIABDX REGION15 AGE15X SEX RACEAX RACEBX RACEWX HISPANX FTSTU15X HWELLSPE BPMLDX CHDDX MIDX CHOLDX STRKDX HIDEG##HIBPDX

Iteration 0:   log likelihood = -4654.0282
Iteration 1:   log likelihood = -3947.2983
Iteration 2:   log likelihood = -3614.9855
Iteration 3:   log likelihood = -3603.4636
Iteration 4:   log likelihood =   -3603.03
Iteration 5:   log likelihood = -3603.0272
Iteration 6:   log likelihood = -3603.0272

Logistic regression                             Number of obs     =     13,658
                                                LR chi2(27)       =    2102.00
                                                Prob > chi2       =     0.0000
Log likelihood = -3603.0272                     Pseudo R2         =     0.2258
```

DIABDX	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
REGION15	.0156509	.0313316	0.50	0.617	-.045758	.0770598
AGE15X	.0255183	.0022781	11.20	0.000	.0210533	.0299832
SEX	.0931052	.0626248	1.49	0.137	-.029637	.2158475
RACEAX	.0174	.1368441	0.13	0.899	-.2508095	.2856096
RACEBX	-.1992937	.1227234	-1.62	0.104	-.4398271	.0412397
RACEWX	.0992974	.1194311	0.83	0.406	-.1347832	.3333781
HISPANX	.523184	.1075015	4.87	0.000	.312485	.733883
FTSTU15X	-1.01975	.3857167	-2.64	0.008	-1.77574	-.2637589
HWELLSPE	.0311446	.0379768	0.82	0.412	-.0432887	.1055778
BPMLDX	-.4048676	.1034869	-3.91	0.000	-.6076981	-.2020371
CHDDX	.1449296	.1160478	1.25	0.212	-.0825199	.3723791
MIDX	.253592	.1272158	1.99	0.046	.0042535	.5029305
CHOLDX	1.166406	.0694924	16.78	0.000	1.030204	1.302609
STRKDX	.2316503	.1103003	2.10	0.036	.0154656	.4478349
HIDEG						
2 GED	-.3452829	.320347	-1.08	0.281	-.9731515	.2825857
3 HIGH SCHOOL DIPLOMA	-.1292237	.1429172	-0.90	0.366	-.4093363	.1508888
4 BACHELOR'S DEGREE	-.41354	.1936191	-2.14	0.033	-.7930265	-.0340534
5 MASTER'S DEGREE	-.5631919	.2605699	-2.16	0.031	-1.0739	-.0524843
6 DOCTORATE DEGREE	-.6085837	.4753001	-1.28	0.200	-1.540155	.3229875
7 OTHER DEGREE	-.0203965	.22123	-0.09	0.927	-.4539994	.4132063
HIBPDX						
1 YES	1.557138	.1830864	8.50	0.000	1.198295	1.91598
HIDEG#HIBPDX						
2 GED#1 YES	.4802039	.3561945	1.35	0.178	-.2179245	1.178332
3 HIGH SCHOOL DIPLOMA#1 YES	.0784362	.1655475	0.47	0.636	-.2460309	.4029033
4 BACHELOR'S DEGREE#1 YES	-.1073748	.2319233	-0.46	0.643	-.5619361	.3471866
5 MASTER'S DEGREE#1 YES	.168299	.3164859	0.53	0.595	-.452002	.7886
6 DOCTORATE DEGREE#1 YES	-.509412	.5911711	-0.86	0.389	-1.668086	.649262
7 OTHER DEGREE#1 YES	-.1432627	.2674485	-0.54	0.592	-.6674523	.3809268
_cons	-4.386427	.8851756	-4.96	0.000	-6.121339	-2.651514

Figure 9

. **margins HIDEG##HIBPDX**

Predictive margins   Number of obs = **13,658**
Model VCE : **OIM**

Expression : **Pr(DIABDX), predict()**

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
HIDEG						
1 NO DEGREE	.1187301	.0057438	20.67	0.000	.1074725	.1299878
2 GED	.117936	.0106412	11.08	0.000	.0970797	.1387922
3 HIGH SCHOOL DIPLOMA	.1123877	.0038145	29.46	0.000	.1049114	.1198639
4 BACHELOR'S DEGREE	.0823052	.0060415	13.62	0.000	.0704641	.0941462
5 MASTER'S DEGREE	.0853363	.009227	9.25	0.000	.0672517	.103421
6 DOCTORATE DEGREE	.0559517	.0134863	4.15	0.000	.029519	.0823844
7 OTHER DEGREE	.109206	.0090524	12.06	0.000	.0914637	.1269483
HIBPDX						
0	.0496043	.0035355	14.03	0.000	.0426747	.0565338
1 YES	.1839685	.0111558	16.49	0.000	.1621035	.2058335
HIDEG#HIBPDX						
1 NO DEGREE#0	.0576167	.0065866	8.75	0.000	.0447071	.0705262
1 NO DEGREE#1 YES	.2032123	.0158575	12.81	0.000	.1721321	.2342925
2 GED#0	.0420846	.0117786	3.57	0.000	.0189991	.0651702
2 GED#1 YES	.2230738	.0246845	9.04	0.000	.174693	.2714547
3 HIGH SCHOOL DIPLOMA#0	.05128	.0046735	10.97	0.000	.0421201	.0604399
3 HIGH SCHOOL DIPLOMA#1 YES	.1960537	.0133643	14.67	0.000	.1698602	.2222472
4 BACHELOR'S DEGREE#0	.0395113	.0059567	6.63	0.000	.0278364	.0511862
4 BACHELOR'S DEGREE#1 YES	.1380527	.0145866	9.46	0.000	.1094635	.1666418
5 MASTER'S DEGREE#0	.0343719	.0076278	4.51	0.000	.0194218	.0493221
5 MASTER'S DEGREE#1 YES	.1521503	.0216696	7.02	0.000	.1096787	.1946219
6 DOCTORATE DEGREE#0	.032941	.0142986	2.30	0.021	.0049163	.0609657
6 DOCTORATE DEGREE#1 YES	.084658	.025687	3.30	0.001	.0343125	.1350036
7 OTHER DEGREE#0	.0565718	.0099236	5.70	0.000	.0371219	.0760218
7 OTHER DEGREE#1 YES	.1807703	.0206456	8.76	0.000	.1403057	.2212349

Figure 10

```
. logit DIABDX REGION15 AGE15X SEX RACEAX RACEBX RACEWX HISPANX FTSTU15X HWELLSPE BPMLDX CHDDX MIDX HIBPDX STRKDX HIDEG##CHOLDX

Iteration 0:   log likelihood = -4654.0282
Iteration 1:   log likelihood = -3955.5289
Iteration 2:   log likelihood = -3613.8752
Iteration 3:   log likelihood = -3601.8788
Iteration 4:   log likelihood = -3601.4402
Iteration 5:   log likelihood = -3601.4374
Iteration 6:   log likelihood = -3601.4374

Logistic regression                             Number of obs     =     13,658
                                                LR chi2(27)       =    2105.18
                                                Prob > chi2       =     0.0000
Log likelihood = -3601.4374                     Pseudo R2         =     0.2262
```

DIABDX	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
REGION15	.0164383	.0313244	0.52	0.600	-.0449563	.077833
AGE15X	.0254595	.0022783	11.17	0.000	.0209941	.029925
SEX	.0907195	.0626354	1.45	0.148	-.0320437	.2134827
RACEAX	.0132509	.1367993	0.10	0.923	-.2548708	.2813726
RACEBX	-.2050308	.1226894	-1.67	0.095	-.4454976	.0354359
RACEWX	.0951892	.1193705	0.80	0.425	-.1387727	.329151
HISPANX	.525496	.1075267	4.89	0.000	.3147475	.7362444
FTSTU15X	-1.022948	.3862995	-2.65	0.008	-1.780081	-.2658152
HWELLSPE	.0299319	.0378845	0.79	0.429	-.0443203	.1041841
BPMLDX	-.4105111	.1036166	-3.96	0.000	-.6135958	-.2074263
CHDDX	.1490349	.1162207	1.28	0.200	-.0787536	.3768233
MIDX	.2545394	.1273303	2.00	0.046	.0049766	.5041023
HIBPDX	1.599163	.1408662	11.35	0.000	1.323071	1.875256
STRKDX	.2285776	.110382	2.07	0.038	.0122329	.4449223
HIDEG						
2 GED	-.3405312	.2683958	-1.27	0.205	-.8665772	.1855149
3 HIGH SCHOOL DIPLOMA	-.1817109	.1278834	-1.42	0.155	-.4323577	.0689359
4 BACHELOR'S DEGREE	-.4253252	.1791842	-2.37	0.018	-.7765198	-.0741305
5 MASTER'S DEGREE	-.6158751	.2543232	-2.42	0.015	-1.114339	-.1174109
6 DOCTORATE DEGREE	-.3985868	.4378757	-0.91	0.363	-1.256807	.4596338
7 OTHER DEGREE	-.1665443	.2075589	-0.80	0.422	-.5733523	.2402638
CHOLDX						
1 YES	1.06305	.1262959	8.42	0.000	.8155144	1.310585
HIDEG#CHOLDX						
2 GED#1 YES	.542153	.3158786	1.72	0.086	-.0769577	1.161264
3 HIGH SCHOOL DIPLOMA#1 YES	.1731492	.1566343	1.11	0.269	-.1338483	.4801467
4 BACHELOR'S DEGREE#1 YES	-.0892903	.2238326	-0.40	0.690	-.5279942	.3494136
5 MASTER'S DEGREE#1 YES	.2580365	.313244	0.82	0.410	-.3559105	.8719836
6 DOCTORATE DEGREE#1 YES	-.8313523	.5714412	-1.45	0.146	-1.951356	.2886518
7 OTHER DEGREE#1 YES	.0848479	.2606463	0.33	0.745	-.4260094	.5957052
_cons	-4.312232	.8837244	-4.88	0.000	-6.0443	-2.580164

Figure 11

```
. margins HIDEG##CHOLDX

Predictive margins                  Number of obs     =     13,658
Model VCE    : OIM

Expression   : Pr(DIABDX), predict()
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
HIDEG						
1 NO DEGREE	.1187809	.0057572	20.63	0.000	.107497	.1300648
2 GED	.118871	.0105872	11.23	0.000	.0981204	.1396216
3 HIGH SCHOOL DIPLOMA	.1124175	.0038084	29.52	0.000	.1049532	.1198817
4 BACHELOR'S DEGREE	.0825889	.00605	13.65	0.000	.0707311	.0944466
5 MASTER'S DEGREE	.084884	.0090489	9.38	0.000	.0671484	.1026196
6 DOCTORATE DEGREE	.0592615	.0144767	4.09	0.000	.0308876	.0876353
7 OTHER DEGREE	.1092374	.0089769	12.17	0.000	.0916429	.1268318
CHOLDX						
0	.0640557	.002902	22.07	0.000	.0583679	.0697435
1 YES	.1628093	.0052072	31.27	0.000	.1526034	.1730152
HIDEG#CHOLDX						
1 NO DEGREE#0	.0759001	.0065307	11.62	0.000	.0631002	.0887001
1 NO DEGREE#1 YES	.1752618	.0108131	16.21	0.000	.1540684	.1964552
2 GED#0	.0564215	.0124158	4.54	0.000	.0320869	.080756
2 GED#1 YES	.2016992	.0213371	9.45	0.000	.1598794	.2435191
3 HIGH SCHOOL DIPLOMA#0	.0648826	.0044229	14.67	0.000	.0562139	.0735512
3 HIGH SCHOOL DIPLOMA#1 YES	.1741949	.0078662	22.14	0.000	.1587774	.1896125
4 BACHELOR'S DEGREE#0	.0523167	.00688	7.60	0.000	.0388322	.0658012
4 BACHELOR'S DEGREE#1 YES	.1191076	.0111523	10.68	0.000	.0972494	.1409657
5 MASTER'S DEGREE#0	.0440536	.0093058	4.73	0.000	.0258146	.0622925
5 MASTER'S DEGREE#1 YES	.1345228	.0178514	7.54	0.000	.0995346	.169511
6 DOCTORATE DEGREE#0	.0535813	.0203219	2.64	0.008	.0137511	.0934116
6 DOCTORATE DEGREE#1 YES	.0657262	.0208431	3.15	0.002	.0248743	.106578
7 OTHER DEGREE#0	.0657459	.010339	6.36	0.000	.0454818	.08601
7 OTHER DEGREE#1 YES	.1652673	.0173487	9.53	0.000	.1312645	.1992701

Figure 12

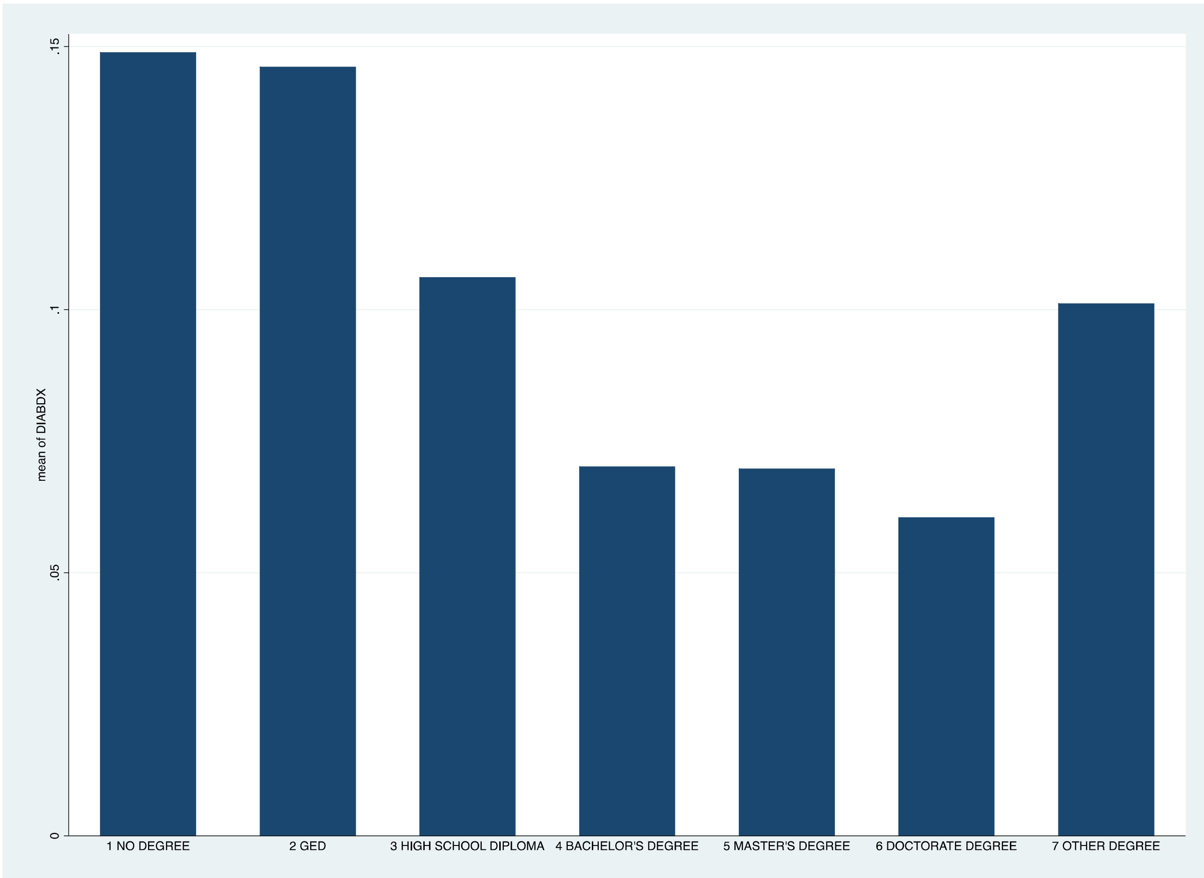
```
. logit DIABDX REGION15 AGE15X SEX RACEAX RACEBX RACEWX HISPANX HIDEG FTSTU15X HWELLSPE HIBPDX BPMLDX CHDDX MIDX STRKDX CHOLDX

Iteration 0:   log likelihood = -4310.4869
Iteration 1:   log likelihood = -3628.9544
Iteration 2:   log likelihood = -3338.6625
Iteration 3:   log likelihood =  -3329.165
Iteration 4:   log likelihood = -3328.8264
Iteration 5:   log likelihood =  -3328.825
Iteration 6:   log likelihood =  -3328.825

Logistic regression                             Number of obs     =     12,610
                                                LR chi2(16)       =    1963.32
                                                Prob > chi2       =     0.0000
Log likelihood =  -3328.825                     Pseudo R2         =     0.2277
```

DIABDX	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
REGION15	.0178419	.0325332	0.55	0.583	-.0459219	.0816058
AGE15X	.025022	.0023461	10.67	0.000	.0204237	.0296203
SEX	.0645958	.0650232	0.99	0.321	-.0628473	.1920388
RACEAX	-.0092368	.1443616	-0.06	0.949	-.2921804	.2737068
RACEBX	-.2407851	.1298869	-1.85	0.064	-.4953587	.0137885
RACEWX	.0635859	.1265782	0.50	0.615	-.1845027	.3116745
HISPANX	.5362938	.1118127	4.80	0.000	.3171449	.7554428
HIDEG	-.457839	.0843793	-5.43	0.000	-.6232195	-.2924586
FTSTU15X	-.9721893	.3854274	-2.52	0.012	-1.727613	-.2167655
HWELLSPE	.0440172	.0372776	1.18	0.238	-.0290456	.11708
HIBPDX	1.646154	.1461103	11.27	0.000	1.359783	1.932524
BPMLDX	-.4287176	.1073332	-3.99	0.000	-.6390869	-.2183483
CHDDX	.1394808	.1201189	1.16	0.246	-.0959479	.3749094
MIDX	.2580339	.1322544	1.95	0.051	-.0011799	.5172478
STRKDX	.2307694	.1150419	2.01	0.045	.0052916	.4562473
CHOLDX	1.168589	.0722132	16.18	0.000	1.027053	1.310124
_cons	-4.213106	.9244478	-4.56	0.000	-6.024991	-2.401222

Figure 13



mean of DIABDX
0
.05
.1
.15
1 NO DEGREE
2 GED
3 HIGH SCHOOL DIPLOMA
4 BACHELOR'S DEGREE
5 MASTER'S DEGREE
6 DOCTORATE DEGREE
7 OTHER DEGREE

Figure 14