NBER WORKING PAPER SERIES

DIABETES AND THE RISE OF THE SES HEALTH GRADIENT

James Smith

Working Paper 12905
http://www.nber.org/papers/w12905

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2007

Diabetes and the Rise of the SES Health Gradient
James Smith
NBER Working Paper No. 12905
February 2007
JEL No. I1

ABSTRACT

This paper investigates the salient diabetes prevalence patterns across key SES indicators, and how they changed over time. The investigation spans both the conventional concept of diagnosed diabetes and a more comprehensive measure including those whose diabetes is undiagnosed. By doing so, I separate the distinct impact of covariates on disease onset, better self-management, and the probability of disease diagnosis. Emphasis is given to SES correlates of undiagnosed diabetes and how these changed as those with undiagnosed diabetes plummeted over the last 25 years. I estimate the differential ability by education to successful self-manage diabetes, especially when disease self-management became more complicated.

James Smith
Labor and Pop Studies Program The RAND Corporation
1776 Main Street
Santa Monica, CA 90406
smith@rand.org

Diabetes is a serious illness that deals with the body's inability to produce (Type 1) or regulate (Type 2) insulin which controls the level of glucose in the blood. Diabetes prevalence rises rapidly with age, is believed to be increasing rapidly over time,[1] and prevalence is apparently particularly high among Americans.[2] The consequences of diabetes can be quite severe, including heart and kidney disease, poor circulation occasionally resulting in amputation of limbs, vision problems with blindness a possibility, a diminished quality of life, and premature death.[1]

Diabetes may not be an equal opportunity disease. Research has indicated that the incidence and prevalence of diabetes is much more common among those at the bottom of several alternative prominent SES markers, such as education and income, as well as for America's principal ethnic and racial minorities- African-Americans and Hispanics.

In this paper, I investigate several dimensions of the SES gradient with diabetes. These include the salient prevalence patterns across these key indicators of SES and especially the way those patterns have been changing over time. My investigation spans both the conventional concept of diagnosed diabetes, a more comprehensive measure that includes within it those whose diabetes is as yet undiagnosed, and a measure that excludes those whose diabetes is under control. By doing so, I am able to separate the distinct impact of covariates on disease onset, better self-management, and the probability of disease diagnosis. Special emphasis is given in the research to SES correlates of undiagnosed diabetes and how these changed as the fraction of the American population with undiagnosed diabetes plummeted over the last twenty-five years. I also summarize results about the differential ability of those of different education groups to successful self-manage their diabetes, especially during a time span when disease self-management became more complicated.

METHODS

Data

In this research, I will use various waves of the National Health and Nutrition Examination Surveys (NHANES). These waves include what I will label NHANES IV, which took place during the years 1999-2002, NHANES III, fielded during 1988-1994, and NHANES II, implemented in all years in the 1976-1980 interval. These series of NHANES contain data obtained through personal interviews, physical exams, and lab tests. All data are available for

adults ages 25-70, the age span studied in this analysis. Details of the specific survey and sampling procedures used can be found in the references cited for each of the NHANES.[3,4,5,6] However, it is important to note that NHANES II over-sampled low income households while the latter two NHANES over-sampled African-Americans and Hispanics. All tabular data presented in this paper are weighted.

In all waves, information is available on the self-reported prevalence of a wide variety of illnesses including diabetes. Moreover, individual attributes including age, gender, race, marital status, family income, education, and parental prevalence of diabetes are obtained from individual interviews conducted during each annual wave. In particular, all variants of NHANES collect data on individual self-reports of diabetes of the general form 'Did a doctor ever tell you that you had diabetes…' Unfortunately, gestational diabetes, a significant component of diabetes among women, is neither consistently included nor excluded in the NHANES over time and for that reason this research will focus on men only.

In addition to personal interviews, the key advantage of NHANES for this research is that all waves contain data obtained both through physical tests and lab exams (blood, urine, and swabs). Particularly relevant for research with a focus on diabetes, physical tests were performed on height and weight so that BMI can be computed and with it an objective indicator of whether or not the respondent is obese (BMI $\geq$ equal to 30) or overweight (BMI $\geq$ 25 and $<$ 30). In addition, I separate the obese population into three groups- class 1 ($\geq$ 30 and $<$35), class 2 ($\geq$ 35 and $<$ 40) and morbid ($\geq$ 40).[7]

The lab exam for diabetes in NHANES III and IV is a glycosylated hemoglobin A1c (HbA1c) test, a measure of the percent of hemoglobin molecules that are bound to glucose. Although not usually a screener for diabetes, HbA1c is highly correlated with fasting plasma glucose levels[8,9]. While there is no strict diagnosis threshold value, I follow the standard convention by using values greater than or equal to 6.5% as indicating clinical diabetes. My principal results are not sensitive to the specific threshold chosen.

NHANES II relied instead on fasting glucose. For NHANES II, I defined clinical diabetes using a classification of the Oral Glucose Challenge/Tolerance Test (OGTT) results developed by NHANES. The OGTT involves measurement of plasma glucose concentrations in the fasting state, and cut by whether the fasting level is 140 milligrams per deciliter or more.

Lab tests in HHANES II were not given to all sample participants, but only those randomly selected for the blood tests. Additional respondents in NHANES II had to be excluded from the clinical analyses since their tests results were not useable for a variety of reasons. Consequently, sample sizes for analyses which rely on the lab test results are much smaller in NHANES II compared to the other two surveys, limiting the analysis that can be conducted by on lab results with this data.

Finally, all versions of NHANES collect several health related behaviors, some of which are thought to be significant 'risk factors' for diabetes. These risk factors include whether the respondent ever smoked or is a current smoker, and the amount of physical exercise in which one normally engages.

The definition of exercise is not identical in the three waves. In NHANES IV respondents are asked "Over the last 30 days, did you do any vigorous activities for at least 10 minutes that caused heavy sweating or large increases in breathing or heart rates? Some examples are running, lap swimming, aerobics classes or fast bicycling." In NHANES II respondents are asked "In the things you do for recreation, for example, ports, hiking, dancing and so forth, do you get much exercise, moderate exercise, or no exercise" – only much is counted as vigorous. In NHANES III, you are asked how often over the last month you did a set of activities (walk a mile, jog or run, bicycle, swim, aerobics, other dancing, calisthenics or other exercise, yard work, lift weights. You were also asked what other (exercise, sports, or physically active hobbies) activities you did in last month. An intensity rated scale is given to each activity. We examined the intensity scales of activities counted as vigorous in NHANES III, and only activities in NHANES III receiving that score or above are counted as vigorous exercise.

These revisions create imprecision in the amount of real change in vigorous exercise over time. However, these exercise variables provide the same threshold at a point in time so there impact on the gradient can be ascertained. There is also relevant information about diet in NHANES, which will be dealt with a latter version of this paper.

Measures of SES

The principal measure of SES used is years of education. In all waves, I divide men into three education groups- less than a high school education, a high school education, and more than a high school education. This is the only measure of schooling available from NHANES IV

and for reasons of comparability education is defined similarly in the other waves. For some of the analyses using NHANES II and III, I separate college graduates from those with 13-15 years of schooling in the highest schooling group. This division is not yet possible in NHANES IV.

Since income is crudely measured in NHANES with annual family income categorized into relatively few brackets at the top, NHANES is certainly not the data source to conduct a detailed investigation of the role of income. Instead, I divide total family income roughly three equally sized groups or terciles. That goal produced income tercile cutoffs points of $35,000, and $65,000 in NHANES IV, $25,000, and $50,000 in NHANES III, and $10,000, and $20,000 in NHANES II. Race and Hispanic ethnicity is ascertained from a self-report of respondents.

Statistical methods

Combined, these series of NHANES allow one to monitor over a twenty-five year period (both overall and by various SES measures) trends in diagnosed and actual diabetes. The combination of biological measures alongside self-reports of disease prevalence also permits one to track differential trends in undiagnosed diabetes, a central focus of this research. Age specific diabetes prevalence is modeled as a function of indicators for race and ethnicity, a quadratic in age, ever and current smoking, vigorous physical activity, overweight (BMI $\geq$ 25 and <30), the three stages of obesity class 1 ($\geq$ 30 and <35), class 2 ($\geq$ 35 and < 40) and morbid ($\geq$ 40), height (in inches), and parental prevalence of diabetes (either the father or mother was a diabetic). Probit models were used in estimation, but the main conclusions do not depend on the specific statistical model chosen.

As mentioned above, the measures of obesity, overweight, and height used are based on physical exams of respondents and hence are free of the well-known types of measurement errors associated with self-reports of these concepts. For each of the series of NHANES, comparable models are estimated both for self-reports of diabetes prevalence as well as the more comprehensive prevalence measure based on both self-reports and biological measurement. I also estimate models estimating the probability of having undiagnosed diabetes and the probability of successful diabetes management. Both of these analyses are conditioned on being a diabetic.

RESULTS

Table 1 documents trends in alternative measures of diabetes prevalence for all men 25-70 years old as well as for major ethnic and racial groups separately. If we only examine self-reports of diagnosed diabetes, the increase in prevalence among all men in this age group is dramatic. For example, the fraction of men of all races who self-report being diabetic has well more than doubled from 3.1% to 7.1%. Indeed, the absolute increase in the fraction of men reporting that a doctor told them they were diabetic was larger than the initial proportion of diagnosed male diabetes in the late 1970s. In the most recent NHANES used in this study, self-reported male diagnosed diabetes is about one-third larger among African-Americans (8.4%) and Hispanics (8.5%) compared to non-Hispanic White men (6.3%).

If we examine instead the more comprehensive measure of prevalence obtained by combining clinically evaluated with self-reported diagnosed diabetes, levels of male diabetes are not surprisingly higher making this serious disease even more common with about one in every eleven men in this age group afflicted between 1999 and 2002. At the same time, the observed secular trends- while still real and significant- are also far less dramatic. For example, among men of all races, overall diabetes prevalence rose over this period from 6% to 9%, a fifty percent increase compared to more than a doubling when respondent reported diagnosed diabetes was used earlier. The absolute percentage point increase in the comprehensive measure of diabetes prevalence is 75% as large as the absolute percentage point increase in diagnosed diabetes. The discrepancy in these two measures of diabetes prevalence is particularly large among African-American and Hispanic men.

The reason for the discrepancy is trends in the two alternative prevalence measures results from the quite large decline in the fraction of male diabetics whose diabetes is undiagnosed. Over these twenty-five years, rates of male undiagnosed diabetes fell sharply from almost half to a little more than one in five. Especially during the 1990s, these declines in undiagnosed diabetes were particularly large among Hispanic and African-American men.

Table 1 also documents levels and trends some well established and prominent risk factors associated with diabetes. Male heights increased by less than half an inch while mean age was also relatively constant over time within this broad age group. The main exception to this stagnant portrait was the very large expansion in the fraction of men who were obese. Male obesity more than doubled over this period, growing from 13% to 28%. The increase in obesity

was particularly large among White Non-Hispanics as male rates of obesity converged over time amongst the different ethnic and racial groups.

Finally, secular trends in education are strongly positive. Between the late 1970s and early 2000s, the fraction of men without a high school diploma fell from a third to a fifth. Similarly, the percent of men who went beyond high school increased from 35% to 55% over this period. There was a steady secular advance in the education accomplishments of all male ethnic groups over time, which was only partially offset by the increased weight given to Latinos as the numbers of Hispanic immigrants increased dramatically over this period.

A central concern of this research involves understanding reasons for gradient of diabetes with education and the manner in which that gradient has evolved over time. To set the background, Table 2 provides a list of key outcomes stratified by my division of years of schooling in each of the three sets of NHANES.

Especially in recent years, there exists a pronounced negative gradient of diagnosed diabetes prevalence with years of schooling. To illustrate, in NHANES IV 9.8% of men who went behind high school had diagnosed diabetes compared to 6% of men who failed to obtain a high school degree. In contrast, the education gradient in diagnosed diabetes in NHANES II is more muted with the only real diagnosed prevalence difference taking place after high school graduation.

In fact, once we control for age, there essentially is no education gradient in diagnosed diabetes in NHANES II. For example, among men ages 55-70, the fraction with diagnosed diabetes are 6.4%, 6%, and 7.2%, among those with less than a high school degree, only a high school degree, and more than a high school degree respectively. In all three sets of NHANES, the education gradients appear to be much sharper in the more comprehensive and presumably more accurate total prevalence measure.

The other rows in Table 2 are suggestive about some possible reasons for the education gradient and the manner in which it has evolved over time. Those in the lowest education group are much more likely to be Latino or African-American, are less likely to engage in vigorous physical exercise, are much more likely to smoke, and are more than an inch and a half shorter in stature. All these gradients are reasonably steep. While the education gradient in % overweight is essentially flat, there is a negative education gradient in the far more critical obesity, one that becomes steeper as we move up the obesity categories toward the morbidly obese. Although the

percent with a parent who was a diagnosed diabetic does decline with education, the strength of this relationship is very weak. This may suggest that the fundamental reasons for any education gradient in male diabetes may not be largely genetic.

Turn next to the risk factors associated with schooling that have changed the most over time. First on that list would have to be the increasing fraction of Latinos in the lower education groups. In NHANES IV, among men who were not high school graduates, one third were Hispanics- in NHANES II the comparable fraction was ten percent. Today, more than half of male non-high school graduates are either Latino or African-American- twice the rate in the late 1970s.

Second prize goes to age. The cross-sectional negative relation between age and schooling basically reflects the rate of cross cohort improvement in education. With the slowdown in educational advancements over time, the age-education relation has basically disappeared. For example, there was more than a nine year age difference between college graduates and non-high school grads in NHANES II with less than a two year difference now. To some extent, the initial education gradient with diabetes prevalence was simply picking up the correlation with age, a point made above. The negative smoking gradient with education also became much larger over time, but as we will see in the next section smoking behavior apparently has little to do with the onset of diabetes.

While rates of obesity have clearly risen over the last 25 years, there does not appear to be any significant steepening of that gradient as obesity rates in all categories of obesity appear to have risen at roughly similar rates. Education gradients in height and parental diabetes also seem not to have been altered sharply over time.

The data in Table 3 list by education, income, and ethnicity rates of undiagnosed diabetes in each of the three sets of NHANES. Especially in NHANES III and IV, there exists a pronounced negative gradient in those with undiagnosed diabetes across education and income groups. For example, using NHANES III to illustrate the point, 38% of diabetics who have less than a high school degree are not aware of their condition. The comparable fraction for those who went past high school is only 21%. Overtime there has taken place a very pronounced decline in the percent of diabetes that is undiagnosed diabetes falling from almost half to about one fifth. These improvements in eliminating undiagnosed diabetes appear to be larger among the more educated and to a lesser extent those with the most income. In sharp contrast, the huge

differentials in undiagnosed diabetes across racial and ethnic groups have for all practical purposes been eliminated.

Model Estimates

Table 4 contains estimates of diabetes prevalence models estimated or all three NHANES waves. There are three alternative models of prevalence estimated in each wave, each of which illuminates a different dimension of behavior related to prevalence. By including both diagnosed and undiagnosed diabetes, the most comprehensive model of prevalence in the final two columns estimates the relationship of covariates to the actual presence of diabetes in this population. Because the two middle columns (labeled clinical) models the presence of diabetes above the clinical threshold, the difference between the comprehensive and clinical models is that clinical models exclude those respondents whose diabetes is under control. Thus the difference between these two models is an indication of how co-variates affect good disease management. Since some diagnosed diabetes were excluded from the NHANES lab tests, modeling clinical diabetes only is not possible in NHANES II.

Finally, the estimates in the first set of columns tell us about the relation of covariates to diagnosed diabetes only. The difference between the estimated effects in the diagnosed and the comprehensive prevalence models are indicative of the impact of covariates on the probability of having undiagnosed diabetes. I will present more direct models of undiagnosed diabetes below. Table 4 lists for each of the three types of prevalence models estimated partial derivatives alongside the associated 'z' statistic for the estimated effect being different from zero. Robust standard errors are used in all models.

First examine estimated effects for the comprehensive measure of total diabetes. Even after controlling for this set of personal attributes, diabetes prevalence is significantly higher among both Latino and African-American men, but the estimated disparity is somewhat smaller in the most recent NHANES especially for African-Americans. Not surprisingly, diabetes prevalence increases with age, albeit at a decreasing rate. The probability of being a diabetic is considerably higher if one of the respondent's parents was also a diabetic. The extent to which this strong generational transmission reflects common genetic influences or a shared family social and environmental background is not knowable from these estimates alone.

Being either overweight or obese raises the likelihood that one is a diabetic with the estimated impact of obesity considerably larger than that of being overweight. Within the obesity classification, the estimated effects increase as we move across the three stages of obesity with very larges effects estimated for those who have a BMI ≥ 40. These estimated effects of excessive weight are much stronger in the combined clinical-self report prevalence estimates, an issue to which I will return below. Engaging in vigorous exercise is negatively associated with diabetes, although the impact of this risk factor does not vary much depending on the definition of diabetes prevalence. In none of the models does additional height (measured in inches) have any systematic association with the probability of being a diabetic. Only in NHANES III is the positive estimated impact of ever smoking statistically significant, but even this effect is negated by past smoking cessation.

Next examine some differential patterns that emerge across the three measures of diabetes prevalence and consider first the difference between the diagnosed prevalence and total prevalence models, which are indicative of the impact of co-variates on the probability of being undiagnosed. Perhaps, the most systematic pattern is that estimated effects of excessive weight are much larger in the total prevalence models, indicating that obesity is strongly negatively correlated with the probability of being diagnosed. As mentioned above, these results also suggest that race was associated with not being diagnosed but that its effect has been eliminated over time.

In addition to excessive weight, the variables whose estimated impacts differ the most between the diagnosed and total diabetes prevalence are the two SES markers- education and income. For both education and income, I estimate a more consistent and larger negative impact on prevalence when the most inclusive definition of prevalence is used that includes undiagnosed diabetes detected through the biological exams. Indeed, for years of schooling there exists no statistically significant association with self-reported diagnosed diabetes prevalence conditional on the other personal attributes included in the model. In contrast, the estimated effects of education on prevalence are large and statistically significant when undiagnosed diabetes is included in the prevalence computation. Especially for the middle income group, a similar difference is found for the relationship with income between the self-report and inclusive measure of prevalence.

These differences in the estimated effects of an important sub-set of variables that include the SES markers between the narrow and broad definitions of prevalence suggest the alternative model that is contained in Table 5. This model represents the probability of being an undiagnosed diabetic conditional on being a diabetic. This conditioning lowers sample sizes considerably and thus makes obtaining statistical significance more difficult. Therefore these models were estimated combing men and women- tests for differences by gender did not indicate any significant differences between the male and female samples outside of an intercept shift. Even after combining the sexes, samples sizes in NHANES II were too small for any meaningful analysis so these models were only estimated for NHANES III and NHANES IV.

A few additional co-variates were added to these models in addition to those contained in the prevalence models. These include two measures meant to capture the extent of contact with the medical system-whether one has health insurance and the last time one saw a doctor-one to 3 years and more than three years with one year or less the omitted category. To capture the possible impact of misplaced self-perceptions, the difference between clinically measured and self-reports of obesity is also included, a variable that is also interacted with an indicator variable that one is clinical obese. Finally, an indicator for marriage was added, an effect that was allowed to differ by gender.

The covariates that do not appear to be related to conditional non-diagnosis are smoking exercise, Hispanic ethnicity, marriage, mistaken BMI perceptions, and height. In NHANES III, female diabetics were more likely to be diagnosed and the probability of not being diagnosed increased with age- while still present, both patterns were not statistically significant in NHANES IV.

Given that doctors are trained to have a standard checklist to use to query patients for the likelihood of a disease and that these checklists typically include familial disease histories, it is not all surprising that in both waves of NHANES having a parent who is a diabetic reduces the likelihood that diabetes is not diagnosed. Parental diabetes is the best predictor in these models for diabetes prevalence and it is reassuring that it is taken into account in detection. However, the second best predictor of diabetes prevalence is obesity and it appears not to be sufficiently taken into account. In al three stages of obesity and in both NHANES waves, the obese are more likely to be undiagnosed. Why this would be so is a bit of a mystery. One possibility is that the evidence relating obesity to diabetes is more recent and there are (most surely unnecessary) lags

in implementation. Doctors may also discount signals that they see as the patients fault and obesity may fall within that group.

Duration of time since last contact with a physician is positively related to being undiagnosed although the interpretation of this effect is somewhat problematic as diagnosis induces additional physician visits. Health insurance appears not to be related to the probability of diagnosis.

These results suggest a declining significance of race in being diagnosed over this time period. Race and ethnicity have been highlighted in NIH campaigns to reduce health disparities in health outcomes including disease detection, and- at least based on these results- with some success. In NHANES III, African-Americans were more likely to not have been diagnosed, a result that is statistically significant. By NHANES IV, this disparity had disappeared.

However, at the same time that race and ethnic disparities in diabetes diagnosis were eliminated, disparities by education appeared. In NHANES IV, diabetics in the highest education group are much more likely to be diagnosed- a statistically significant difference that was not present in NHANES III. Health disparities appear in all sorts of ways, with race and ethnic differences more easy to monitor in terms of patient attributes. Disparities across markers such as education level are more difficult to monitor and perhaps easier to ignore or dismiss as an underachieving patient problem. They are no less real. Finally, in neither of the two NHANES waves was income a marker for undiagnosed disease.

Self- Management of Diabetes and the SES Health Gradient

As explained above, a comparison of estimates of the models for total prevalence with those for clinical prevalence in Table 4 provides some evidence of the determinants of successful self-management of diabetes since the clinical measures excludes those diabetics whose lab readings are below the clinical threshold. (These may also be false positives, a point to which I return below). The clinical models are much closer to the total prevalence models than the diagnosed diabetes models were so I turn to an alternative specification to identify the determinants of successful disease management.

Those with diagnosed diabetes whose clinical tests indicate levels that are below the diagnosis threshold may be considered very successful in their disease management. Table 6 presents models of this form of successful disease management, once again estimated conditional

on being a diabetic. These models include the same set of covariates as in the models for undiagnosed diabetes in Table 5. The threshold for success is high in these models requiring management that brings one below the diagnostic threshold. Thus, success will be a function not only how one manages the disease but how far one is from the diagnostic threshold.

These estimates indicate a significant shift in the determinants of successful self-management between the NHANES III and NHANES IV. Besides race and age, the only variables that appear to be statistically significantly related to successful disease management in NHANES III are the length of time since the respondent saw a doctor (lowering the probability of good self-management). By NHANES IV, however, two types of variables are associated with successful self-management of diabetes. The first component reflects variables that measure good personal behaviors in other ways. For example, better self-management is associated with not currently smoking, not being overweight or obese as well as seeing a physician more often.

In addition, both markers of SES-education and income- are positively associated with good self-management of diabetes in NHANES IV. I interpret the emergence of these SES gradients with successful self-management of diabetes as indicating an enhanced premium on education as well as income in adopting and adhering to the new but complex treatments for diabetes. New recommended treatment regimens for diabetes are quite complex and often require high quality and persistent patient self-management on a daily basis, and not all patients will be equally adept at complying. Extensive self-management is important, including frequent monitoring of blood glucose levels, balancing dosages with food intake and physical activity, prevention and treatment of hypoglycemia, monitoring timing and dosage of insulin injections, and regular consultation with health care providers. While a greater tendency to have false positives among the more educated, non-smokers and the non-obese could also explain these patterns that seem a more unlikely mechanism.

New treatments for diabetes are known to be efficacious but the treatment is complicated and places great demands on a patient's ability to self-monitor his condition. In an earlier study with Goldman[10], I used data from an important clinical trial—the Diabetes Control and Complications Trial (DCCT). In the DCCT patients with type 1 diabetes were randomized into treatment and control groups. The treatment arm involved a quite intensive regimen where there was very close self and external monitoring of blood glucose levels and encouragement of strict adherence. In particular, patients in the treatment arm were seen weekly until a stable treatment

program was achieved. While not insignificant, the treatment in the control arm consisted of a more standard regimen and far less intrusive external monitoring of patients.

Table 7 shows that prior to the intervention there were large differences across education groups in several measures of good behavior at the study's baseline. Whether it involved checking blood, following insulin regimens, exercise, or smoking, those with less education were not doing as well. Given these initial but unsurprising baseline differences by education in adherence to good practice, we hypothesized that imposing a good behavior regimen—which is essentially what the rigorous treatment regimen did—would impart more benefits to the less educated who were having more problems with treatment to begin with.

We used an objective health outcome measure in the DCCT—glycosolated hemoglobin, which measures the amount of sugar binding to the blood. Higher levels indicated worse control. The impact of enforcing a common treatment regime can be obtained by subtracting what normally would occur (the control sample) with what took place under an enforced treatment regimen (the treatment sample). The data in Table 8 demonstrate that while those in all education groups benefited from being in the treatment arm, the benefits from enforced better adherence relative to the control group were largest for the least educated (see the final row in Table 8). Thus, a differential ability to adhere to beneficial albeit complicated medical regimens appears to be one reason for the association between education and health outcomes for those with diabetes.

In our study, Goldman and I were also able to provide some evidence on why education might matter for adherence. Two factors that did not matter in promoting better adherence were household income or having a better memory. Alternatively, it does appear to be related to higher-level aspects of abstract reasoning, part of which included the ability to internalize the future consequences of current decisions.

Additional research on why education matters so much in effective self-management should receive high priority. One possibility is that the education experience itself has little to do with it, but it is simply a marker for personal traits (reasoning ability, rates of time preference, etc.) that may lead people to acquire more education and to be healthier. But education may not be that passive. It may help train people in decision-making, problem solving, and adaptive skills, forward looking behavior, all of which have pretty direct applications to a healthier life. Education may well have biological effects on the brain, which result in improved cognitive

function and problem-solving ability some of which may impart benefits to choices made regarding one's health. This is similar to the argument that more active brain functioning when younger pushes off the onset of dementia.

Explaining trends in Diabetes Prevalence

My goal in this section is to isolate those factors most responsible for increasing diabetes prevalence over time.[11] Consider the impact of a single risk factor *j*. Let *P(A)* and *P(B)* be the (predicted) diabetes prevalence rates in years A and B and let $P(A)^{-j}$ and $P(B)^{-j}$ the predicted prevalence in years A and B for the "counterfactual" situation that nobody suffers from this risk factor *j*. $P(A)-P(A)^{-j}$ can be interpreted as the diabetes rate in year *A* due to that risk factor and similarly for year B. This assignment of importance depends both on the risk factor prevalence and on the sensitivity of the probability of diabetes to that risk factor (the corresponding coefficients). I will separate them below.

The difference in diabetes prevalence in the two years is:

$$P(B)-P(A)=[P(B)^{-j}-P(A)^{-j}]+[P(B)-P(B)^{-j}]-[P(A)-P(A)^{-j}]$$

The first term on the right hand side can be interpreted as the difference between diabetes prevalence in the two years *not* due to the chosen risk factor. The sum of the second and third term is the part due to the chosen risk factor. The latter two terms can each be further separated in a 'prevalence' effect (the percentage with the risk factor) and an 'impact 'effect (the impact of the risk factor on diabetes). We can write:

$$P(A)-P(A)^{-j}=1/N_A\sum_{i\in A}\{g(x_i,b_A)-g(x_i^{-j},b_A)\}=$$

$$[\sum_{i\in A}x_{ij}/N_A][\sum_{i\in A,x_{ij}=1}\Delta g(x_i,b_A)/\sum_{i\in A}x_{ij}]$$

where $g(x_i,b_A)$ is the probability of having the risk factor for an individual with characteristics x_i and parameter vector b_A.

The first factor is the fraction in year *A* that suffers from the chosen risk factor (the "quantity effect" for year *A*). In the second term, $\Delta g(x_i,b_A)$ is the marginal effect ("partial derivative") for a dummy variable, the difference if it is set to 1 or 0, with other variables set to their values for observation *i*. Thus the second term can be seen as the average marginal effect

for those who have the risk factor. The same decomposition can be used for all co-variates in the model allowing one to compare the importance of each to the prevalence rates of diabetes in each year and the difference between the years.

Table 9 presents my accounting for the increase in total diabetes prevalence between the three NHANES waves. Some attributes included in the model either did not change over time (age, height) or had relatively small estimated effects on diabetes prevalence (smoking and exercise) and are therefore excluded from this accounting. Those that remain include demographic factors (race and ethnicity), SES variables (income and education), and high levels of BMI (overweight and the three stages of obesity).

Demographic forces (including age) had a relatively small impact as the small increase in prevalence predicted mostly as the increasing numbers of Hispanics was partly offset by the diminished importance of race as a predictor of diabetes. Combined race and ethnicity predicts that diabetes prevalence would have risen by 0.19 percentage points.

Increasing numbers of men who had a parent who was a diabetic and the increase impact of parental diabetic inheritance combined to predict that male diabetes prevalence would have increased by 1.39 percentage points between the three NHANES data sets.

Excessive weight by height was the most important factor leading to rising levels of obesity over time. Being overweight itself was not critical (except that it makes you more likely to be obese), but all three stages of obesity are. The three stages combined predict an increase of obesity of 2.02 percentage points between NHANES II and NHANES IV- adding in the small contribution of overweight implies that excessive BMI lead to a 2.15 percentage point rise in male diabetes prevalence. The increase in obesity did not come out of thin air. A number of papers have argued that the recent growth in obesity can at least partially be explained by declines in the relative price of food, reinforced by even steeper declines in the relative price of foods dense with calories.[7,12] The decline in physical activity associated with work (especially in lower SES more manually intensive occupations) may also have played a contributing role.

The principal factor operating in the opposite direction was the improving levels of SES in the population and most importantly higher levels of education. SES related factors predict a decline in diabetes prevalence of 1.21 percentage points.

Using data from Table 1, total diabetes prevalence increased by three percentage points between the late 1970s and the beginning of this new century. Three factors loom largest in

attempting to explain the increase in diabetes prevalence over time. The biggest impact flows from the increasing obesity in all three stages of the population, following by the increasing fractions of Americans with a parent who was a diabetic, with a relatively small effect due to the changing racial and ethnic demographics of the country. The important offsetting force was the improving SES levels of the population over the last 25 years which negated some of the deleterious effects of the other factors. Combined all these factors predict that diabetes prevalence would increase by 2.48 percentage points, 83% of the increase in male diabetes prevalence that actually took place.

If the objective instead was to explain the secular increase in total diagnosed diabetes prevalence then the large decline in undiagnosed diabetes would be added towards the top of the list of factors accounting for trends. Twenty-five percent of the increase in diagnosed diabetes since the late 1970s actually represents improved detection.

Explaining the SES education gradient in diabetes

My second objective was to isolate the prominent reasons for the SES health gradient in diabetes as well as the manner in which it has changed over time. To do so, I re-estimate the total prevalence models in Table 5 starting with only controlling for education- that is the unadjusted education gradient. I then add variables to this model in the following order always keeping in the previous variables in the list- (a) age quadratic, (b) race and ethnicity, (c) smoking and exercise, (d) parental diabetes, (e) excessive weight, (6) income groups. The (b) model in this grouping represents the education gradient without controlling for any behavioral factors related to schooling and the (e) models the adjusted gradient with only a single SES marker- schooling- without trying to parcel out the distinct effects of schooling and incomes. This would be my preferred model for understanding the nature of the schooling gradient with diabetes. Estimates are provided for all three waves of NHANES using the same education stratifications for all waves in Table 10.a and using the additional separation of higher education possible only in NHANES 2 and NHANES 3 in Table 10.b.

The unadjusted gradients in prevalence in (1) are large, generally statistically significant and have increased only slightly over these twenty five years. Even thought the average age is about the same in all three education groups (see Table 2), controlling for age still significantly

diminishes the schooling gradient in diabetes due to the non-linear effect of age on prevalence. Approximately half of the schooling gradient with diabetes prevalence is accounted for by age, race, and ethnicity controls but now it appears that the demographically adjusted schooling gradient with diabetes prevalence may have risen over time.

Smoking and exercise contribute to an additional diminishing of the gradient but this is largely offset by including past parental diabetes. Adding in the effects of being overweight or being obese further attenuates the estimated gradient- combined all of these behavioral and demographic controls explain somewhere between 60% and 75% of the schooling gradient. Once again, the estimated schooling gradient appears to increase slightly over these twenty five years.

Even when schooling and income battle it out as competing SES markers in the final row of Table 10, a schooling gradient remains (albeit smaller). It should be mentioned that the controls for income and education in such models are of very different standing. Education has been showed elsewhere to be related to the onset of diabetes while income is not related to diabetes onset.[13] Instead, reduced income appears to be the consequence of diabetes onset and at least based on that reasoning income does not belong in these models. For this reason, the education effects in row (6) of Table 10 represent my preferred summary of the net effect of education on diabetes prevalence.

Conclusions

While the increase in diabetes prevalence over time is less than that indicated by the commonly used measured of diagnosed diabetes, it remains a significant public health concern. The prevalence models estimated in this paper suggest that the most important forces leading to higher diabetes prevalence are excessive weight and obesity, an increase inheritance of diabetes through parents, and to a smaller extent the changing ethnic and racial demographics of the country. Some but not all of this predicted increase in diabetes due to these factors was offset by improvements in the education of the population over time.

Undiagnosed diabetes remains an important health problem with a little over one in five male diabetes undiagnosed by the years 1999-2002. However, this is far less of a problem than it was twenty-five years ago when almost half of male diabetes were undiagnosed. Although race and ethnic differentials in undiagnosed diabetes were eliminated over the last twenty five years,

the disparities became larger across other measures of disadvantage such as education. Moreover, undiagnosed diabetes is a particularly severe problem among the obese, a group at much higher risk on diabetes onset.

Successful diabetes management is also related to education. With the introduction of new efficacious but highly complex therapies for treating diabetes, the more educated have another advantage over the least educated in that they appear to be more able to adapt and to adhere to these new therapies.

Those in lower education groups face a triple threat with diabetes. First, at least in more recent years, they are of slightly higher risk in contracting the disease. Second, they remain at considerably greater risk of having their diabetes undiagnosed and presumably untreated. Third, even after diagnosis, they have considerably more difficulty in successful self-management of the disease using the complex but effective treatments necessary to diminish the negative health consequences associated with diabetes.

References

1. *Diabetes: A National Plan for Actions. Steps to a Healthier US*. US. Department of Health and Human Services, December 2004.

2. Banks, J., M. Marmot, Z. Oldfield, and J. P. Smith. "Disease and Disadvantage in the USA and in England," *JAMA* 295(16):2037-2045, May 3, 2006.

3. *Plan and Operation of the Second National Health and Nutrition Examination Survey 1976-1980.* Programs and Collection Procedures Series 1, No. 15, U.S. Department of Health and Human Services, July 1981.

4. *Vital and Health Statistics. Plan and Operation of the Third National Health and Nutritional Health and Nutrition Examination Survey, 1988-1994*. Series 1, Program and Collection Procedures No 32, U.S. Department of Health and Human Services, DHHS Publication No 94-1308.

5. Centers for Disease Control and Prevention, National Center for Health Statistics. National Health and Nutrition Examination Survey Data Sets and Related Documentation (Survey Questionnaire, Examination and Laboratory Protocols, 1988-1994 and 1999-2002). Available at http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm. Accessed 9 January 2006.

6. *NHANES Analytical Guidelines*, National Center for Health Statistics, June 2004.

7. Rashid, I., M. Grossman, and S. Chou. "The Super Size of America—An Economic Estimation of Body Mass Index and Obesity in Adults," NBER Working Paper 11584, National Bureau of Economic Research, August 2005.

8. Goldman, N., I Lin, M. Weinstein, and Y. Lin. "Evaluating the Quality of Self-Reports of Hypertension and Diabetes," *Journal of Clinical Epidemiology* 56:148-154, 2003.

9. Davidson M., D. Schriger, A. Peters et al. "Relationship Between Fasting Plasma Glucose and Glycosylated Hemoglobin. Potential for False Positive Diagnoses of Type 2 Diabetes Using New Diagnostic Criteria," *JAMA* 281:1203-1210, 1999.

10. Goldman, D., and J.P. Smith. "Can Patient Self-Management Help Explain the SES Health Gradient?" *Proceedings of the National Academy of Sciences* 99(16):10929-10934, August 6, 2002.

11. Kapteyn, A., J. P. Smith, and A. Van Soest. "Vignettes and Self-Reported Work Disability in the US and The Netherlands," July 2006. *American Economic Review* forthcoming.

12. Lakdawalla, D., and T. Philipson. "The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination," NBER Working Paper 8946, National Bureau of Economic Research, May 2002.

13. Smith J.P. "Unraveling the SES-Health Connection," *Population and Development Review Supplement: Aging, Health and Public Policy: Demographic and Economic Perspectives*, Volume 30, 2004. pp. 108-132. The Population Council.

Table 1
Diabetes Prevalence Rates by Race and Ethnicity
Men 25-70 years old

	All Races			White Non-Hispanics			African-Americans			Hispanics		
	1999-2002	1988-1994	1976-1980	1999-2002	1988-1994	1976-1980	1999-2002	1988-1994	1976-1980	1999-2002	1988-1994	1976-1980
Diagnosed	7.0	4.6	3.1	6.3	4.7	3.0	8.4	5.7	5.0	8.5	3.9	1.4
Clinical	7.1	5.2	NA	6.4	4.8	NA	8.5	8.8	NA	7.6	5.3	NA
Total Prevalence	8.9	6.8	6.0	8.0	6.3	5.6	11.1	10.3	8.4	10.8	7.0	4.0
% Undiagnosed	21.6	32.5	48.2	21.2	26.4	46.0	24.3	45.1	40.3	21.4	44.0	65.4
% Obese (clinical)	28.2	21.0	13.1	28.8	21.3	12.8	29.2	21.5	17.2	24.3	23.0	15.1
Height (clinical)	69.4	69.3	69.0	69.9	69.6	69.2	69.7	69.5	69.0	67.1	66.9	67.0
Age	44.6	43.1	44.4	45.7	43.8	44.6	43.9	41.8	43.5	40.8	40.1	42.0
Low Ed	20.9	23.1	33.2	12.8	18.0	29.4	36.4	32.3	53.0	46.7	56.5	69.8
Middle Ed	24.9	31.1	31.3	26.4	32.0	33.1	22.7	36.9	24.9	19.2	21.6	17.3
High Ed	55.1	45.7	35.4	60.8	50.0	37.5	40.9	30.7	22.1	33.9	21.9	12.9

Source- NHANES II, III, IV (all data are weighted). All data are percents except for age (measured in years) and heights (measured in inches).

Table 2
Patterns by Education Group

NHANES II—MEN Ages 25-70
Education Level

	low	middle	high	all	Ed 13-15	Ed 16+
Diagnosed prevalence	3.9	3.4	2.0	3.1	2.0	1.9
Total prevalence	8.1	6.6	3.1	5.8	1.9	3.9
% undiagnosed	48.2	43.7	49.6	47.9	51.2	49.1
Hispanic	9.9	2.6	2.5	4.7	2.3	2.6
Black	14.6	7.3	5.7	9.3	8.3	3.8
Now Smoker	49.1	44.8	35.2	42.8	42.8	29.7
Ever Smoker	82.6	76.3	64.9	74.3	72.6	59.5
Vig-exercise	20.9	23.4	23.9	22.8	22.6	24.9
% overweight	43.1	41.9	41.0	41.9	42.6	39.9
% obese	15.4	15.5	9.1	13.1	13.1	7.1
Obese—1	12.6	12.2	7.4	10.6	9.4	6.1
Obese—2	2.1	3.1	1.3	2.1	2.0	0.8
Obese—3	0.8	0.2	0.4	0.4	0.6	0.2
Height	68.3	69.1	69.7	69.0	69.6	69.7
Parent diabetic	17.8	16.8	15.2	16.5	15.7	15.1
Age	49.7	42.8	41.9	44.4	41.1	40.3

NHANES III—MEN Ages 25-70

	low	middle	high	all	Ed 13-15	Ed 16+
Diagnosed prevalence	7.2	4.4	3.4	4.6	4.4	2.6
Total prevalence	12.1	6.4	4.3	6.8	5.8	03.3
% undiagnosed	40.5	30.7	22.9	32.5	24.6	20.7
Hispanic	23.0	6.4	4.4	9.5	6.8	3.3
Black	13.6	11.6	6.6	9.9	9.6	4.3
Now Smoker	44.9	42.8	21.0	33.7	27.6	15.9
Ever Smoker	78.1	71.6	54.2	65.2	60.2	49.6
Vigorous-ex	20.5	29.3	50.8	37.0	43.3	56.4
% overweight	39.4	45.0	41.2	42.0	41.5	40.9
% obese	24.7	22.2	18.6	21.1	21.7	16.1
Obese—1	17.0	15.8	14.3	15.4	15.9	13.1
Obese—2	5.1	4.0	2.7	6.3	3.7	1.9
Obese—3	2.6	2.4	1.6	2.1	2.0	1.2
Height	68.1	69.4	69.8	69.3	69.8	69.8
Parent diabetic	21.9	20.4	18.6	19.9	19.7	17.7
Age	46.3	42.2	42.1	43.1	40.7	43.1

NHANES IV—MEN Ages 25-70
Education Level

	low	middle	high	all
Diagnosed prevalence	9.8	7.0	6.0	7.0
Total prevalence	14.2	8.7	7.2	8.9
% undiagnosed	31.1	19.3	16.4	21.8
Hispanic	32.4	10.7	7.2	13.9
Black	17.6	8.8	8.5	9.6
Now Smoker	43.2	36.6	18.9	28.2
Ever Smoker	71.7	65.9	49.2	57.9
Vig-exercise	23.5	33.5	51.4	41.4
% overweight	43.0	41.9	42.2	42.3
% obese	28.3	31.1	26.9	28.2
Obese- 1	17.9	20.5	17.6	18.3
Obese- 2	5.9	6.3	6.4	6.3
Obese-3	4.5	4.2	2.9	3.5
Height	68.3	69.3	69.9	69.4
Parent diabetic	26.1	25.4	23.4	24.4
Age	45.3	43.6	44.9	44.6

All data are weighted.

Table 3
% of Men 25-70 Who Have Undiagnosed Diabetes

	1999-2002	1988-1994	1976-1980
Ed 0-11	31.1	40.5	48.2
Ed 12	19.3	30.7	43.7
Ed>12	16.4	22.9	49.6
Income			
Lowest	27.5	37.8	55.1
Middle	13.8	26.3	44.3
Highest	19.4	25.7	45.5
Ethnicity			
White Non-Hispanic	21.2	26.4	46.0
Hispanic	21.4	44.0	65.4
African-American	24.3	45.1	41.7
All	21.6	32.5	48.2

All data are weighted and are from NHANES IV, III, and II.

Table 4
Probit Models for Prevalence of Diabetes—Men Ages 25-70
NHANES wave IV

	Diagnosed dF/dx	Diagnosed z	Clinical dF/dx	Clinical z	Total Prevalence dF/dx	Total Prevalence z
Hispanic	0.0146	1.37	0.0051	0.49	0.0191	1.52
Black	0.0171	1.57	0.0143	1.34	0.0283	2.19
Age	0.0124	3.78	0.0107	3.37	0.0155	4.12
Age2	-0.0008	2.63	-0.0001	2.19	-0.0001	2.77
Ed Mid	-0.0037	0.33	-0.0188	1.87	-0.0155	1.26
Ed High	-0.0052	0.50	-0.0226	2.26	-0.0202	1.69
Income Mid	-0.0100	1.10	-0.0212	2.42	-0.0156	1.47
Income High	-0.0379	3.54	-0.0346	3.26	-0.0467	3.73
Ever smoked	0.0044	0.50	0.0049	0.55	0.0110	1.07
Current Smoker	-0.0142	1.44	-0.0015	0.15	-0.0160	1.37
Vig Exercise	-0.0133	1.52	-0.0104	1.20	-0.0162	1.58
Overweight*	0.0157	1.47	0.0450	3.77	0.0390	2.96
Obesity* stage 1	0.0236	1.75	0.0801	4.77	0.0692	3.94
Obesity* stage 2	0.0955	4.45	0.1996	7.29	0.1984	6.88
Obesity* stage 3	0.1520	4.73	0.2562	6.58	0.2875	7.02
Height*	-0.0019	1.38	-0.0022	1.58	-0.0024	1.43
A Parent Diabetic	0.0713	7.36	0.0633	6.67	0.0797	7.18
N	3106	3106	3109	3109	3109	3109

All models estimate robust standard errors. * measured during physical exams

NHANES wave III

	Diagnosed dF/dx	Diagnosed z	Clinical dF/dx	Clinical z	Total Prevalence dF/dx	Total Prevalence z
Hispanic	0.0148	2.35	0.0227	3.07	0.0262	3.06
Black	0.0197	3.09	0.0455	5.89	0.0498	5.70
Age	0.0068	3.74	0.0096	4.70	0.0117	4.95
Age2	-0.00004	2.42	-0.0001	3.27	-0.0001	3.36
Ed Mid	-0.0045	0.81	-0.0091	1.46	-0.0108	1.47
Ed 13-15	-0.0016	0.23	-0.0156	2.05	-0.0124	1.37
Ed 16+	-0.0064	0.79	-0.0238	2.76	-0.0249	2.45
Income Mid	-0.0055	1.04	0.0015	0.24	-0.0029	0.39
Income High	-0.0131	1.98	-0.0103	1.25	-0.0156	1.69
Ever Smoked	0.0162	3.18	0.0164	2.68	0.0182	2.58
Current Smoker	-0.0188	3.64	-0.0142	2.36	-0.0210	2.98
Vig Exercise	-0.0105	1.79	-0.0109	1.60	-0.0174	2.23
Overweight*	0.0145	2.48	0.0213	3.03	0.0306	3.81
Obesity* stage 1	0.0338	4.12	0.0588	5.93	0.0806	7.00
Obesity* stage 2	0.0656	4.19	0.1284	6.50	0.1663	7.25
Obesity* stage 3	0.1386	5.20	0.2100	6.55	0.2790	7.63
Height*	-0.0005	0.65	0.0008	0.82	0.0006	0.50
A Parent Diabetic	0.0566	8.91	0.0629	8.87	0.0784	9.62
N	5419	5419	5426	5426	5426	5426

All models estimate robust standard errors. * measured during physical exams

NHANES wave II

	Diagnosed dF/dx	Diagnosed z	Total Prevalence dF/dx	Total Prevalence z
Hispanic	-0.0096	0.97	0.0122	0.41
Black	0.0274	3.66	0.0457	1.99
Age	0.0033	2.07	0.0004	0.10
Age2	-0.00001	0.87	0.0003	0.79
Ed Mid	0.0104	1.95	0.0233	1.60
Ed 13-15	0.0048	0.63	-0.0175	0.84
Ed 16+	0.0084	1.05	0.0111	0.51
Income Mid	-0.0099	2.14	-0.0207	1.53
Income High	-0.0112	1.98	-0.0317	1.98
Ever Smoked	0.0007	0.14	-0.0319	2.02
Current Smoker	-0.0086	1.89	0.0047	0.34
Vig Exercise	-0.0127	2.46	-0.0044	0.30
Overweight*	-0.0004	0.08	0.0295	2.31
Obesity* stage 1	0.0071	0.95	0.0295	1.31
Obesity* stage 2	0.0946	4.00	0.0279	3.03
Obesity* stage 3	0.1573	3.18	0.6641	3.58
Height*	0.0009	1.11	-0.0041	1.99
A Parent Diabetic	0.0399	6.34	0.0336	2.21
Intercept	5708	5708	1562	1562

All models estimate robust standard errors. * measured during physical exams

Table.5
Probability of Undiagnosed Diabetes Given that One is a Diabetic

	NHANES IV dF/dx	NHANES IV z	NHANES III dF/dx	NHANES III z
Hispanic	-0.0060	0.16	0.0247	0.63
Black	0.0395	0.98	0.1051	2.89
Female	-0.0480	0.81	-0.1075	1.93
Age	0.0077	0.59	0.0298	2.86
Age2	-0.00001	0.59	-0.0003	2.90
Married	-0.0684	1.34	-0.0519	1.09
Married Female	0.0182	0.28	0.0185	0.32
Ed Mid	-0.0533	1.33	0.0289	0.87
Ed High	-0.0836	2.36	0.0001	0.02
Income Mid	0.0263	0.65	0.0632	1.75
Income High	0.0918	1.62	0.0130	0.24
Ever smoked	0.0375	1.06	-0.0189	0.59
Current Smoker	-0.0148	0.35	0.0480	1.30
Vig Exercise	-0.0226	0.57	0.0784	1.70
Overweight*	0.1021	1.69	0.0369	0.85
Obesity* stage 1	0.1335	2.05	0.1520	3.24
Obesity stage 2	0.1792	2.42	0.2178	3.81
Obesity stage 3	0.1531	1.95	0.1749	2.80
BMI*-BMI (self)	-0.0128	0.86	-0.0029	0.25
BMI*-BMI (self) X obese	0.0217	1.30	0.0032	0.72
Height*	0.0013	0.23	0.0079	1.48
A Parent Diabetic	-0.0684	2.24	-0.1380	5.26
Have health insurance	0.0298	0.75	-0.0265	0.63
Last saw doctor-1-2 yrs	0.4464	4.54	0.4695	7.12
Last saw doctor 3 or more yrs	0.4354	4.47	0.3822	6.09
N	746	746	1289	1289

All models estimate robust standard errors. * measured during physical exams

Table 6
Probability of Successful Management of Diabetes Given that One is a Diabetic

	NHANES IV dF/dx	NHANES IV z	NHANES III dF/dx	NHANES III z
Hispanic	-0.0417	1.08	-0.0316	0.90
Black	-0.0258	0.64	-0.1038	3.34
Female	0.0250	0.39	-0.0695	1.29
Age	0.0005	0.04	-0.0476	5.33
Age2	-0.0000	0.05	0.0004	4.64
Married	0.0034	0.07	-0.0264	0.54
Married X Female	0.0222	0.33	0.0226	0.39
Ed Mid	0.0794	1.79	0.0070	0.23
Ed High	0.0628	1.50	0.0514	1.32
Income Mid	0.0792	1.95	-0.0011	0.03
Income High	0.0142	0.23	0.0269	0.56
Ever Smoked	0.0451	1.23	-0.0002	0.01
Current Smoker	-0.0931	2.20	-0.0113	0.32
Vig Exercise	-0.0134	0.32	-0.0243	0.62
Overweight*	-0.0956	1.99	0.0406	1.10
Obesity* stage 1	-0.1159	2.44	-0.0028	0.07
Obesity stage 2	-0.1429	2.84	-0.0546	1.18
Obesity stage 3	-0.1482	2.83	-0.0659	1.36
BMI*-BMI (self)	-0.0069	0.19	-0.0029	0.29
BMI*-BMI (self) X obese	0.0210	1.28	-0.0092	0.79
Height*	-0.0017	0.30	-0.0005	0.10
A Parent Diabetic	-0.0187	0.59	0.0049	0.19
Have health insurance	-0.0939	2.04	0.0460	1.24
Last saw doctor 1-2rs	-0.1283	1.61	-0.1049	2.14
Last saw doctor 3 or more yrs	-0.1477	1.86	-0.0735	1.33
N	746	746	1289	1289

All models estimate robust standard errors. * measured during physical exams

Table 7
Educational Differences in Treatment Adherence at DCCT Baseline

Measure of Adherence	College grad/HS degree		
	Postgrad degree	Some college	Some secondary
Number of times self-monitored blood glucose per week	8.8	7.7	6.7
Missed insulin injection at least once in past month (%)	4.3	6.0	9.2
Did not follow insulin regimen at least once in past month (%)	15.7	25.2	26.6
Did not self-test blood or urine at least one day in past month (%)	66.1	74.1	77.2
Minutes of very hard exercise per week	58.1	49.6	19.7
Currently smoking cigarettes (%)	16.5	19.2	40.8

Source: Goldman and Smith (2002).

Table 8
Educational Differences in Treatment Impact for Diabetics

Group	Glycosolated Hemoglobin: Postgraduate Degree	College grad/ Some college	HS degree/ Some secondary
Conventional Therapy Only (n=495)			
Baseline	8.42	8.76	8.96
End-of-study	8.88	9.08	9.59
Difference	0.46	0.32	0.63
Intensive Treatment Only (n=490)			
Baseline	8.04	8.86	8.93
End-of-study	7.18	7.30	7.43
Difference	-0.85	-1.56	1.51
Treatment Effect[#]	-1.31	-1.88*	2.14**

*p<.10; **p<.05

[#]Treatment effect is the improvement in glycemic control among the intensive treatment group relative to conventional therapy. Significance levels are for a test of equivalence with the postgraduate category and control for duration in study, gender, marital status, and age.

Table 9
Factors Explaining Time Series Increase in Male Total Prevalence of Diabetes

	Prevalence			Impact Effect on Probability of Diabetes			Contribution to Diabetes Prevalence			Contribution to Changes Over Time
	NHANES			NHANES			NHANES			NHANES
	IV	III	II	IV	III	II	IV	III	II	IV-II
Hispanic	13.9	9.4	4.4	1.8	2.4	1.0	0.24	0.23	0.05	0.19
Black	9.5	9.7	7.6	2.8	4.6	4.1	0.27	0.45	0.31	-0.04
Total demographics										0.15
Ed mid	24.9	31.1	32.5	-1.6	-1.1	1.9	-0.40	-0.33	0.62	-1.02
Ed-high	54.9	45.6	38.2	-1.9	-1.5	-0.2	-1.02	-0.69	-0.07	-0.95
Income mid	32.1	36.6	38.0	-1.8	-0.3	-2.0	0.57	-0.10	-0.77	1.20
Income high	30.0	26.3	35.3	-4.7	-1.6	-2.7	-1.40	-0.42	-0.96	-0.44
Total SES										-1.21
parent diabetic	23.8	19.5	17.7	8.1	7.9	3.1	1.93	1.13	0.54	1.39
overweight	41.5	41.9	42.3	2.9	2.3	2.5	1.19	0.97	1.06	0.13
obesity 1	18.0	15.4	10.7	5.4	7.4	2.3	0.96	0.49	0.24	0.72
obesity 2	6.2	3.6	2.3	16.1	13.6	9.9	0.99	0.49	0.23	0.76
obesity 3	3.5	2.1	0.3	21.1	18.2	58.1	0.73	0.27	0.19	0.54
total obesity										2.02
total obesity and overweight										2.15
All factors										2.48

Table 10.A
Measuring the Schooling Gradient With Diabetes Prevalence

	NHANES IV		NHANES III		NHANES II	
	ed med	ed high	ed med	ed high	ed med	ed high
Other Controls						
(1) None	-.058	-.078	-.045	-.071	-.019	-.066
	(4.06)	(6.04)	(5.06)	(2.81)	(1.24)	(4.08)
(2)=(1)+age	-.033	-.057	-.020	-.043	-.010	-.027
	(2.62)	(5.12)	(2.51)	(5.39)	(0.55)	(1.72)
(3)=(2)+race	-.021	-.041	-.012	-.029	.017	-.021
and ethnicity	(1.55)	(3.31)	(1.43)	(3.39)	(1.08)	(1.27)
(4)=(3)+smoking +	-.049	-.036	-.010	-.026	.013	-.025
exercise	(1.38)	(2.79)	(1.21)	(2.93)	(0.86)	(1.56)
(5)=(4)+parental	-.025	-.041	-.012	-.026	.012	-.023
diabetes	(1.96)	(3.43)	(1.56)	(3.19)	(0.03)	(1.40)
(6)=(5)+excessive	-.020	-.033	-.012	-.023	.015	-.014
weight	(1.65)	(2.89)	(1.68)	(2.95)	(1.02)	(0.91)
(7)=(6)+income	-.015	-.019	-.010	-.018	.020	-.005
groups	(1.23)	(1.62)	(1.30)	(2.14)	(1.44)	(-0.30)

Estimated DF/dx for schooling coefficients with robust standard errors in parenthesis.

Table 10.B
Measuring the Schooling Gradient with Diabetes Prevalence

	NHANES IV		NHANES III			NHANES II		
	ed med	ed high	ed med	some college	college or more	ed med	some college	college or more
Other Controls								
(1) None	-.058	-.078	-.046	-.055	-.075	-.017	-.062	-.055
	(4.06)	(6.04)	(5.18)	(5.31)	(7.00)	(1.13)	(3.10)	(3.06)
(2)=(1)+age	-.033	-.057	-.021	-.025	-.054	.011	-.034	-.017
	(2.62)	(5.12)	(2.61)	(2.64)	(5.88)	(0.77)	(1.67)	(0.88)
(3)=(2)+race and ethnicity	-.021	-.041	-.013	-.016	-.042	.019	-.030	-.008
	(1.55)	(3.31)	(1.58)	(1.62)	(4.09)	(1.23)	(1.43)	(0.43)
(4)=(3)+smoking+ exercise	-.049	-.036	-.011	-.015	-.038	.016	-.031	-.015
	(1.38)	(2.79)	(1.37)	(1.52)	(3.58)	(1.02)	(1.52)	(0.77)
(5)=(4)+parental diabetes	-.025	-.041	-.013	-.017	-.036	.015	-.030	-.013
	(1.96)	(3.43)	(1.72)	(1.81)	(3.66)	(0.99)	(1.46)	(0.66)
(6)=(5)+excessive weight	-.020	-.033	-.013	-.016	-.030	.017	-.024	-.002
	(1.65)	(2.89)	(1.81)	(1.79)	(3.21)	(1.16)	(1.19)	(0.08)
(7)=(6)+income groups	-.015	-.019	-.011	-.012	-.025	.023	-.018	.011
	(1.23)	(1.62)	(1.47)	(1.37)	(2.45)	(1.60)	(0.84)	(0.51)

Estimated DF/dx for schooling coefficients with robust standard errors in parenthesis.