The Relationship between Education and Type 2 Diabetes Morbidity: Decomposing Change for Cohorts Born between 1935-1954

Erik L. Hernandez

The primary question motivating this paper is: What has been the relationship between education and diabetes morbidity in the United States during the Epidemiological Transition, specifically during the last decades of the 20th Century and the first decades of the 21st Century? We know that during the Epidemiological Transition (ET) countries, experience increasing prevalence of chronic diseases, including diabetes, which is associated with increasing obesity, tobacco use, sedentary occupations, and other factors (Flegal et al. 2010; McKeown 2009; Lopez and Mathers 2006). We also know that the educational gradient related to overall mortality increased between 1986 and 2006 for selected ages between 54 and 84, with differences in the magnitude and timing of the changes depending on race and gender (Montez at al. 2011). Additional analyses find that among earlier generations in the United States, the most highly educated were more likely to engage in risky health behavior, such as smoking, which contribute to chronic disease (Baker et al. 2011), but that, as such behaviors became more prevalent, and knowledge of negative effects reached the most educated, the relationship between education and risky health behavior reversed. It is not known, however, how the educational gradient has changed across generations with regard to the prevalence of and mortality from specific chronic diseases. Additionally, it is not known whether overall increases in education have acted as a social vaccine against greater prevalence of and mortality from specific chronic diseases in general and diabetes in particular. This paper seeks to fill these gaps in knowledge specifically with regard to diabetes morbidity.

Using cohort analysis, I plan to assess the changing effect of education with regard to type 2 diabetes prevalence. The main hypothesis is that education is more negatively associated with chronic disease for later cohorts. Data for this research were collected by the 1997-2016 National Health Interview Surveys (NHIS). The paper examines the relationship between education and type 2 diabetes prevalence among four cohorts at ages 60-64 (born 1935-39, 1940-44, 1945-49, and 1950-54). The analysis involves logistic regressions of type 2 disease prevalence by educational attainment. Regression results are then incorporated into a Fairlie decomposition analysis, to measure the effect changing educational composition across these cohorts. The analysis reveals the contributions of changing educational composition and of changing educational gradients for overall diabetes prevalence across cohorts.

PRELIMINARY ANALYSIS

Weighted descriptive statistics (not shown) reveal growing diabetes prevalence and education expansion with each successive cohort. Between 1935-39 and 1950-54 cohort, diabetes prevalence increase by approximately 23 percent. Meanwhile, the proportion of persons in the 1950-54 cohort with at least a bachelor's degree (compared to the 1935-39 cohort) increased by nearly 62 percent. Later cohorts experienced increases in these two variables, but the expansion of education was greater than the increase in diabetes prevalence. Additional descriptive results reveal differences in the growth of diabetes by educational groups, with the least educated (less than high school) experiencing more extreme levels of diabetes prevalence growth than more educated. These estimates suggest a need to analyze whether

education's effect on diabetes prevalence has changed for later cohorts, and whether the educational composition of later cohorts has worked as a social vaccine against diabetes prevalence.

Table 1 provides four pooled logistic regression models of the association between education and type 2 diabetes with controls for race (omitted group: Non-Hispanic White), marital status (omitted group: Married), body mass index (BMI), and region (omitted group: South). The four pooled logistic regression models include odds ratios using two cohorts in each model; 1) persons born in 1935-39 and 1940-44, 2) 1940-44 and 1945-49, 3) 1945-49 and 1950-54, and 4) 1935-39 and 1950-54. Across all models, the effect of educational attainment on being diagnosed with diabetes is consistently negative and statistically significant. Across cohorts controlling for population characteristics, BMI, and region, the odds of being diagnosed with diabetes for persons with at least a bachelor's degree are between 49 and 53 percent less than persons without a high school degree. The relationship is strongly significant ($p<.001$) and is not statistically different between models. Logistic regressions were also calculated for each individual cohort (not shown) with similar results.

Table 2 provides Fairlie decomposition models for the same four pooled logistic regressions in Table 1. Time 1 is the earlier cohort in each model, comparing to time 2 (later cohort). In all models, BMI has the largest positive coefficient, associated with the rising diabetes prevalence. Due to limited space of this extended abstract, results focus on the composition of educational attainment (main independent variable). In the 1935-39 and 1940-44 cohort analysis, the later cohort (1940-44) has higher diabetes prevalence than the earlier cohort (approximately 1.2 percentage points greater). Decomposition reveals endowments (changes in composition of independent variables) explains approximately 55 percent of the difference between the two cohorts. The education coefficient is negative and strongly statistically significant ($p<.001$), and in the opposite direction of the difference between the two cohorts. This indicates that if there had been no change in educational composition between the 1935-39 cohort and the 1940-44 cohort, that is, if education had not increased across cohorts, then diabetes prevalence for the 1940-1944 cohort would have been (substantially) higher than it actually was.

The 1940-44/1945-49 cohort analysis indicates the 1945-49 cohort has greater diabetes prevalence (approximately 2.58 percentage points greater). Decomposition reveals endowments explain approximately 19 percent of the difference between cohorts. The education coefficient is negative and strongly statistically significant ($p<.001$), and in the opposite direction of the difference between the two cohorts. This indicates that cohort if there had been no change in educational composition between the 1940-1944 cohort and the 1945-49 cohort, that is, if education had not increased across cohorts, then diabetes prevalence for the 1945-1949 cohort would have been higher than it actually was. This result also indicates that the expansion of educational attainment played an important role in minimizing the increase in diabetes prevalence.

The 1945-49/1950-54 cohort analysis presents a different picture. The prevalence of diabetes increased, rather than decreased, across these cohorts, that is, the later 1950-1954 cohort has slightly lower diabetes prevalence than the earlier 1945-1949 cohort, negative difference of approximately 0.25 percentage points. This difference, however, is small. Additionally, the total explained by endowments is in the opposite direction, explaining a -21 percent of the difference. The interpretation that follows from negative values for both the "difference" (change through time) and the "total explained endowments" is that the endowments tend to lead to a decrease in diabetes prevalence across the two cohorts. The small decline in prevalence of 0.25 percentage points is not statistically significant, suggesting that the effect of endowments roughly counter-balanced the (something, but I'm not certain what, that is, what word to use here…perhaps just drop this whole sentence, and focus on the next which is what you really care about

here, that is, the impact of the education "endowment") Due to the "difference" and "total explained" having opposite values, diabetes prevalence between the 1945-49 and 1950-54 cohorts are not significantly different. Nonetheless, education remains strongly negatively statistically significant ($p<.001$), that is, in the analysis of change for these two cohorts, the expansion of educational attainment played contributed to the measured (but not statistically significant) decline in diabetes prevalence. In other words, without the expansion in education, the overall prevalence of diabetes would have been lower than the measured prevalence for the 1950-54 cohort, compared to the 1945-1949 cohort

The final model compares the 1935-39 and 1950-54 cohorts. The 1950-54 cohort has higher diabetes prevalence (approximately 3.53 percent points greater). Decomposition reveals endowments to explain approximately 39 percent of the difference between the cohorts. Again, education has the largest negative coefficient and is strongly statistically significant ($p<.001$), and, it is opposite in direction (negative) to the difference in prevalence (positive) across cohorts. This indicates that if there had been no change in educational composition between the 1935-39 cohort and the 1950-54 cohort, that is, if education had not increased across these two cohorts, then diabetes prevalence for the 1950-54 cohort would have been higher than it actually was. This result indicates that, overall, across the cohorts studied here, the expansion of educational attainment played an important role in minimizing the increase in diabetes prevalence.

DISCUSSION

There are two main findings to the preliminary analysis. First, results in Table 1 indicate a stable effect of education on diabetes prevalence across cohorts, that is, the education gradient did not change. The effect of having at least a bachelor's degree was consistently negative, and in the narrow range of .47-.51. Thus, persons with higher levels of education (a bachelor degree) were less likely to be diagnosed with diabetes than persons with lower levels of education, for each cohort. Second, results in Table 2 indicate changing educational composition, that is, increases from one cohort to the next, affected the level of diabetes prevalence. Later cohorts experienced greater levels of educational attainment, and these higher attainments resulted in lower diabetes prevalence than would have been the case if there had been no educational expansion.

Table 1: Pooled Logistic Regressions of Education's Association to Type 2 Diabetes, 5-year Cohorts

	1935-39 and 1940-44		1940-44 and 1945-49		1945-49 and 1950-54		1935-39 and 1950-54	
	O.R.	**S.E.**	**O.R.**	**S.E.**	**O.R.**	**S.E.**	**O.R.**	**S.E.**
HS Only	0.79 **	0.06	0.78 **	0.06	0.75 ***	0.06	0.78 **	0.06
Some College	0.75 ***	0.07	0.77 **	0.06	0.69 ***	0.05	0.69 ***	0.06
BA+	0.49 ***	0.05	0.51 ***	0.04	0.47 ***	0.04	0.50 ***	0.05
Female	0.70 ***	0.04	0.71 ***	0.04	0.77 ***	0.04	0.79 ***	0.05
NH Black	1.93 ***	0.17	1.67 ***	0.13	1.60 ***	0.12	1.79 ***	0.15
Hispanic	1.57 ***	0.16	1.65 ***	0.15	1.79 ***	0.15	1.87 ***	0.17
Asian/Pacific Islander	2.30 ***	0.50	2.21 ***	0.38	2.03 ***	0.32	2.08 ***	0.39
Native American	1.65	0.52	1.37	0.37	1.50	0.37	1.82 *	0.50
Other	1.16	0.72	1.16	0.57	1.37	0.60	1.49	0.80
Single/Never Married	0.88	0.12	1.02	0.11	1.19	0.11	1.22	0.13
Cohabiting	0.96	0.19	0.81	0.14	0.98	0.15	1.23	0.23
Separated	1.11	0.21	1.89 ***	0.29	1.61 ***	0.22	1.02	0.16
Divorced	1.12	0.09	1.13	0.08	1.07	0.07	1.07	0.08
Widowed	1.13	0.10	1.14	0.10	0.98	0.09	0.91	0.09
BMI	1.13 ***	0.01	1.14 ***	0.01	1.13 ***	0.01	1.13 ***	0.01
Northeast	0.81 *	0.07	0.76 ***	0.06	0.82 *	0.06	0.87	0.07
Midwest	0.92	0.07	0.98	0.07	0.95	0.07	0.88	0.07
West	0.93	0.08	0.95	0.07	0.81 **	0.06	0.74 ***	0.06
N	14590		16885		19583		17288	
Psuedo R-Squared	0.0821		0.0995		0.0991		0.0878	

***p<.001 **p<.01 *p<.05

Table 2: Fairlie Decomposition of Education's Association to Type 2 Diabetes, 5-year Cohorts

	1935-39 and 1940-44	1940-44 and 1945-49	1945-49 and 1950-54	1935-39 and 1950-54
N (Time 2)	7557	9328	10255	10255
N (Time 1)	7033	7557	9328	7033
Diabetes Prev T2	0.1382	0.1640	0.1616	0.1616
Diabetes Prev T1	0.1263	0.1382	0.1640	0.1263
Difference	0.0120	0.0258	-0.0025	0.0353
Total explained	0.0066	0.0049	0.0030	0.0137
% explained by endowments	55.34	18.92	-21.00	38.72

diabetes	1935-39 and 1940-44		1940-44 and 1945-49		1945-49 and 1950-54		1935-39 and 1950-54	
	Coef.	**S.E.**	**Coef.**	**S.E.**	**Coef.**	**S.E.**	**Coef.**	**S.E.**
Education	-0.006 ***	0.001	-0.007 ***	0.001	-0.006 ***	0.001	-0.014 ***	0.002
Gender	-0.002 ***	0.000	-0.001 **	0.000	-0.001 ***	0.000	-0.001 ***	0.000
Race	-0.002 ***	0.000	-0.002 ***	0.000	0.000	0.000	0.003 ***	0.000
Marital Status	0.000	0.000	0.000	0.000	0.000	0.000	0.002 *	0.001
BMI	0.016 ***	0.001	0.014 ***	0.001	0.009 ***	0.000	0.024 ***	0.001
Region	0.001	0.000	0.001 **	0.000	0.001 **	0.000	0.000	0.000

***p<.001 **p<.01 *p<.05

REFERENCES

Baker, D.P., Leon, J., Smith-Greenaway, E.G., Collins, J, and M. Movit. 2011. "The education effect on population health: a reassessment." Population and Development Review 37:307-332.

Flegal, Katherine M., Carroll, Margaret D., Ogden, Cynthia L., and Lester R. Curtin. 2010 "Prevalence and Trends in Obesity Among US Adults, 1999-2008." *Journal of the American Medical Association* 303(3), 235-241.

Lopez, A.D., Mathers, C.D., 2006. Measuring the global burden of disease and epidemiological transitions: 2002e2030. Ann. Trop. Med. Parasitol. 100 (5e6), 481-499.

McKewon, Robert E. 2009. "The Epidemiologic Transition: Changing Patterns of Mortality and Population Dynamics." *American Journal of Lifestyle Medicine.* 3 (1 Suppl), 19S-26S.

Montez, Jennifer Karas, Robert A. Hummer, Mark D. Hayward, Hyeyoung Woo, and Richard G. Rogers. 2011. "Trends in the Educational Gradient of U.S. Adult Mortality From 1986 Through 2006 by Race, Gender, and Age Group." *Research on Aging* 33 (2): 145–71.