12-9-2022

Effect of Education on Self-reported Health

Mai Le '24
DePauw University

Effect of Education on Self-reported Health

Mai Le

12/09/2022

Submitted in partial fulfillment of the requirements for ECON 385: Regression and Simulation

ABSTRACT

Human Capital Theory pointed out health as a possible return to education. The question at the center of this research is if education can improve health. Replicating the work of Goesling (2007) on new data from the 2000-2022 Current Population Survey (CPS), a cross-sectional probit analysis shows a positive relationship between educational level and self-reported health. This relationship is robust and significant across age groups.

INTRODUCTION

The US marked the 20th century with unprecedented growth for both its economy and citizens, allowing for a century-long decline in mortality rate and increasing life expectancy. We can attribute this improvement to higher income, which allows individuals access to healthcare and gives countries the resources needed to spend on publicly financed education and healthcare. This paper will instead highlight the other main contributor to better health - education. Higher levels of education is not only positively associated with longer life, but also improved overall health during that span. So in trying to predict the health of an individual, it would be beneficial to consider their access to education. After all, while education and health are undeniable fundamental human rights, substantive equality for all is not yet achieved. Stepping into the 21st century, income and social inequality in the US is only rising more. Given the positive relationship between health and education, along with the trend of rising equality in recent decades, it's reasonable to expect college graduates to report even better health than people without college degrees. Or in other words, the health outcomes for education groups are diverging over time.

This has led to the research question of this paper: *What is the effect of educational level attainment on self-reported health?* This topic is highly relevant on the national level, since population wellbeing is an important measure of a nation directly by itself and through its influence on productivity and thus, wealth. But it also resonates on a personal level. Education and income are both main aspects of personal well-being, yet we know much less about health differences between education groups than we do about earnings or income disparities (Goesling, 2007).

LITERATURE REVIEW

The theoretical basis behind this work is Human Capital Theory (Becker, 1964) . In this model, education is a “black box” (Oreopoulos & Salvanes, 2011) where inputs in terms of spending and education go in, something happens, and increased income is yielded. By neatly capturing returns to schooling in a single equation, Becker’s foundational work popularized Human Capital research. While labor market outcomes to education are more often studied, it’s also reasonable to also try to look into nonpecuniary outcomes. After all, schooling supposedly results in gain to human capital such as skill and knowledge. Then this acquired stock is assumed to give access to jobs that pay more than what someone with, for example, illiteracy can get. This increased spending power can also allow higher healthcare spending, and knowledge can lead to adoption of preventative measures. In short, health is a possible return to education (Oreopoulos & Salvanes, 2011).

The empirical study of how education predicts longevity was first established by Kitagawa and Hauser (1973). Using the 1960 Matched Records Study, they computed a “mortality index” by finding the ratios of actual to expected death for 4 education categories: less than elementary, elementary, high school and college. Information on age, race and sex for each cohort was gathered as the basis for the expected number of deaths from each causes. They found that white males aged 25-64 are around 29% less likely to die from diabetes (mortality index of .71) than expected when they have a college degree; while females of the same demographic are 58% less likely to die (mortality index of .42). But these ratios change in the case of those without an elementary degree to 3% more likely (mortality index of 1.03) for white males aged 25-64 and 116% more likely (mortality index of 2.16) for white females. Therefore, lower educated groups are expected to have a higher mortality rate to diabetes mellitus. The trend also holds true for another mortality factors cirrhosis of the liver (for which the dataset was large enough to separate by sex) and cardiovascular-renal diseases, influenza and pneumonia, etc (for which, the dataset was only large enough to distinguish between the different education categories but not sex). Other studies have since replicated the positive effect of education on both health itself and health behaviors. Their results are shown in **Table 1**.

Table 1. *Found association between education and health*

Health outcome(s) that is correlated with education	Population of interest	Study referenced
Mortality rate	US adult men & women	Zajacova, 2006
Self-reported health	US adult	Zajacova, 2012
Smoking	US adult	Walque 2007
Self-reported health Smoking Obesity Work limitation Exercise Binge Drinking	US Foreign-born and native-born, with racial/ethnic subgroup	Kimbro et al., 2008
Depression	US adolescent	Fletcher, 2008

The source of this association can stem from i) health affecting education, or education affecting health - a concept known as reverse causality, or ii) omitted variables. Utilizing compulsory schooling laws (which is related to education but not to health), the reverse causality link between health and education can be broken, thus enabling a casualty effect of education on health to be examined. However, results from studies following this method from both within and outside the US have been inconsistent and often null (Fletcher, 2015). The heterogeneity in estimated effect can be attributable to common factors that affect both education and health, such as individual ability, socioeconomic background, and time preference.

Therefore, an established pattern of correlation alone is not the full picture. One of the most pressing concerns is how contemporary social and political context can change this association. It's reasonable to think that the amount of capital gain isn't constant across the population. For example, children who attend overcrowded and underfunded schools would acquire less skill and knowledge compared to their richer counterparts. Accordingly, Lynch (2006) found that the variability in health-education across cohorts is strongly influenced by the "indirect influence" of income both on education and on health. Given the rising income inequality in the US in recent decades, differentials between education categories have diverged since the 1980s. Goesling (2007) studied this trend by performing a binary logistic regression on National Health Interview Survey data from 1982 to 2004. Results confirmed the positive association between health and education, controlled for basic demographic factors and survey year. But the difference in self-reported health between the college graduates and lower education levels have been widening with time for adults aged 70 and up. Interestingly, this disparity is relatively stable across younger age groups during this 23-year span.

In summary, these literatures provide a consistent result showing that education is an important predictor for health outcomes, regardless of demographics. However, they also pointed out the need for future research that places health and education in its current contemporary social contexts. This paper seeks to do so by replicating the work of Goesling (2007) with updated data from Current Population Survey (CPS) 2000-2022.

DATA AND METHODS

1. Data

The dependent variable is *h* - a dummy variable on respondent self-rated health. The CPS records HEALTH on a five-point scale, as excellent, very good, good, fair, or poor. This data is recoded into 0 (signifying good health) when the respondent answered “excellent”, “very good”, or “good”; and into 1 (diminished health) for the response of “fair” or “poor”.

The first predictor variable is *e*, a categorical variable that measures the respondent’s highest educational level obtained. CPS EDUC is recoded into 3 categories: those with less than a high school degree (less than 12 years of education), those with high school degree but less than college degree (12-15 years of education), and those with a college degree and above (more than 15 years of education). The base is college graduates. The second predictor variable is *y* - a linear index of the survey year. This is a continuous variable included to test the claim that health outcomes are diverging for different educational attainment levels (Goesling 2007). Using a 23-year time frame similar to Goesling, *y* will start at 0 for the year 2000 and end at 22 for the year 2022.

To limit the influence of confounding variables, controls such as age, gender, race, and geographically location are included. Of which, age is the most notable for its interaction with the two main predictor variables. A control *age* variable is included, but restricted from 30 and up so that education is mostly completed. CPS top-coded *age* at value 90. Besides that, the data is grouped into 3 age ranges: 30-49, 50-69, and 70 and older. I performed the analysis separately by age groups. This way is done instead of one pooled analysis of all ages in order to see how the trend differs between the young, middle and old (Goesling 2007). Any variable on income or wealth isn’t included, since rising economic inequality is often pointed to as the source of the increasing gap in educational differences. Including income as a control (and thus holding it constant) can make this discrepancy non-significant (Lynch 2006). The full list of control variables is included in **Table 1**, and the summary statistics of all the variables are in **Table 2**.

Table 1: *Information on 6 control variables*

Variable name	**CPS source**	**Type**	**Recode**
age	AGE	Continuous	None.
women	SEX	Dummy	0* *(men)* - if SEX = 1 1 *(women)* - if SEX = 2
race	RACE	Dummy	1* (*white)* - if RACE = 100 2 (*black)* - if RACE = 200 3 (*others)* - otherwise
hispan	HISPAN	Dummy	0* (*not hispanic)* - if HISPAN = 000 1 (*hispanic)* - otherwise
marst	MARST	Dummy	1* (*married)* - if MARST = 1,2 2 (*separated or divorced)* - if MARST= 3,4 3 (*widowed)* - if MARST= 5 4 (*never married)* - if MARST = 6
region	REGION	Dummy	1* (*northeast)* - if REGION = 11,12 2 (*midwest)* - if REGION = 21,22 3 (*south)* - if REGION = 31,32,33 4 (*west)* - if REGION = 41,42
(*) indicates base case			

Table 2: *Summary statistics on 1 dependent variable and 8 independent variables*

Variable	Mean	Std. Dev.	Min	Max
h				
good health	(base)			
dimininshed health	0.16161	0.368093	0	1
e				
edu < 12	(base)			
edu 12-15	0.562127	0.496125	0	1
edu > 15	0.307371	0.461405	0	1
y	10.82584	6.449776	0	22
age	51.43117	14.42365	30	90
women	0.528405	0.499193	0	1
race				
white	(base)			
black	0.115529	0.319659	0	1
other races	0.087315	0.282296	0	1
marst				
married	(base)			
separated or divorced	0.144363	0.351457	0	1
widowed	0.071216	0.257186	0	1
never married	0.126574	0.332496	0	1
hispan	0.145314	0.352417	0	1
region				
northeast	(base)			
midwest	0.213107	0.409503	0	1
south	0.331463	0.470739	0	1
west	0.264523	0.441079	0	1

Obs: 2,538,388

Women contribute to 52.9% of the dataset, so our data is balanced between male and female. Age ranges from 30 to 90 (the median is 60) but it has a mean of 51. This means that the data is left-skewed, most concentrated in the 30-49 age group. The lower number of observations for older ages affects the precision of our estimations in the upcoming regression. For main variables *h* and *e*, further summary statistics are presented in **Table 3**. More than 50% of the data is clustered in the 12-15 years of education category. Looking at all observations that reported good health, 33.9% of them have more than 15 years of education, while people with less than 12 years of education make up only 10.43%. Overall, the table is consistent with the claim that self-reported health improves along with educational attainment. But we can’t prove the claim just from this table because the result hasn’t been controlled for omitted variable bias.

Table 3: *Self-reported Health by Education Level, in percentage*

diminished health status	education level edu < 12	edu 12-15	edu > 15	Total
good health	10.43	55.68	33.90	100.00
dimininshed health	26.65	59.00	14.35	100.00
Total	13.05	56.21	30.74	100.00

2. Methods

This research uses data from the US Current Population Survey (CPS). Since I’m trying to predict a dichotomous outcome (i.e diminished health or not), a probit model regression analysis is used. The functional form is based on Goesling (2007):

$$P(\mathit{Health}) = \Phi^{-1}\,(\beta_0 + \beta_1 \mathit{Educ} + \beta_2 \mathit{Year} + \beta_3[\mathit{Educ} \times \mathit{Year}] + \Sigma\, \alpha_i X_i + \varepsilon)$$

Where *h* is the predicted probability of reporting diminished health; *e* is the categorical variable for educational attainment; *y* is a continuous index of survey year; X_i represents the set of 6 control variables. The coefficient on the probit function represents shifts in the standard cumulative normal and therefore cannot be easily interpreted. So, I applied a marginal transformation of the coefficients. After which, the coefficients represent the average marginal effects of the regressors, that is, how much the conditional probability of the outcome variable changes when you change the value of a regressor, holding all other variables constant at their means. In my model, for example, the marginal effects of estimated β_1 shows the change in predicted probability of reporting diminished health at different values of *e*.

In this study, education is discrete rather than continuous since the meaning of each additional year of education is not the same. For example, finishing 15 years is still classified as “some college”, but finishing 16 years means obtaining a bachelor’s degree. Beyond the yearly accumulation of skill and knowledge, gaining a degree sends outward signals about an individual's skills that give economic returns, known as the “sheepskin effect”. This effect is also present in health returns (Liu et al, 2011).

Time, age, and education are all interconnected with each other. Age doesn’t just affect education, but also the effect of education on health: the association between education and health is stronger in older people than younger people (Hammond, 2002). This is why an interaction term between time and education is included. But instead of also having interaction terms for age, data is instead separated into three age groups: 30-49, 50-69, 70 and older. For each subset of the data, a separate regression is run. I did this instead of one single analysis to avoid a three-way interaction between time, age and education.

Lastly, because the data from IPUMS CPS are acquired through the complex survey, a weighting variable can improve the reliability of the result. This study uses ASECWT, a person-level weight provided by IPUMS to be used in analysis of the March supplement data.

EMPIRICAL ANALYSIS

The main result of the study is in **Table 3**, which shows coefficients from three separate probit models for the young, middle and old age groups. The coefficients for the variables of interest and their interaction term are included along with their robust SE. The original finding in Goesling (2007) is presented in **Table 4**. To make the magnitudes of probit and logit estimates roughly comparable, I applied a “rule-of-thumb” correction by multiplying the logit estimates by 0.625 (Woolridge, 2015).

The results from mine and the original are closely related: the signs of the coefficients have similar magnitude, and their signs largely similar. The original study covered 1982 to 2004, while mine covered 2000-2020, so there’s a 4 year overlap in our data. But the other 19 different years produced some minor differences. For the education dummy variables in the first 2 rows: the oldest age group’s *12-15 years of education* coefficient has become 1.7 times larger, the middle age group stayed relatively unchanged, while the *<12 years of education* coefficient for youngest age group has gotten 1.7 times smaller. While not yet clear, this is directly in line with Goesling’s claim of that since the 1980s, educational disparity in self-rated health has grown larger for oldest age group, held stable for the middle age group, and narrowed for the youngest age group.

The interaction terms between 12-15 years of education and year for model (2) and (3) are both negative in my result, yet both positive in Goesling. But considering the relatively large robust SEs in my result for these two coefficients, the sign difference could just be the result of chance rather than any meaningful change in the parameter value.

Table 3: *Probit Regression Result*

	Age Group (1) 70+	(2) 50-69	(3) 30-49
diminished health ~s			
edu < 12	0.606***	1.037***	0.986***
	(0.0161)	(0.0119)	(0.0126)
edu 12-15	0.251***	0.459***	0.478***
	(0.0149)	(0.00963)	(0.00990)
year	-0.0132***	-0.00319***	0.00455***
	(0.000907)	(0.000619)	(0.000669)
edu < 12 # year	0.00321**	-0.00698***	-0.0130***
	(0.00121)	(0.000926)	(0.00102)
edu 12-15 # year	0.00300**	-0.000247	-0.00138
	(0.00105)	(0.000717)	(0.000774)
Observations	350588	905934	1281866

Standard errors in parentheses
Source: Current Population Survey
* p<0.05, ** p<0.01, *** p<0.001

Table 4: *Goesling (2007) Unadjusted Logistic Regression Result, scaled for comparison*

	(1) 70+	(2) 50-69	(3) 30-49
Edu: <12	**0.544****	**1.148*****	**1.501*****
Edu: 12-15	**0.148*****	**0.451*****	**0.659*****
year	**-0.013*****	**-0.011*****	**0.009*****
<12#year	**0.0113*****	**0.008*****	**-0.012*****
12-15#year	**0.009*****	**0.008*****	**0.001**
*p < .05 **p < .01 ***p < .001			

Table 5: *Average Marginal Effect of the Result in Table 3*

	Age Group (1) 70+	(2) 50-69	(3) 30-49
edu < 12	0.230***	0.270***	0.139***
	(0.00306)	(0.00208)	(0.00152)
edu 12-15	0.0945***	0.104***	0.0574***
	(0.00228)	(0.000997)	(0.000580)
year	-0.00390***	-0.00109***	0.000348***
	(0.000144)	(0.0000781)	(0.0000476)
age	0.00937***	0.00533***	0.00403***
	(0.000212)	(0.0000898)	(0.0000531)
women	-0.0100***	-0.00176	0.0130***
	(0.00209)	(0.00102)	(0.000589)
midwest	-0.000388	0.0115***	0.00156
	(0.00297)	(0.00155)	(0.000924)
south	0.0452***	0.0386***	0.0109***
	(0.00273)	(0.00143)	(0.000857)
west	0.0130***	0.0111***	0.00286**
	(0.00301)	(0.00154)	(0.000911)
black	0.0964***	0.0635***	0.0202***
	(0.00320)	(0.00165)	(0.000991)
other races	0.0605***	0.0308***	0.00815***
	(0.00413)	(0.00201)	(0.00113)
hispanic	0.0520***	0.000468	-0.0122***
	(0.00349)	(0.00159)	(0.000790)
separated or divorced	0.0400***	0.0917***	0.0523***
	(0.00346)	(0.00145)	(0.000980)
widowed	0.0246***	0.0793***	0.0764***
	(0.00237)	(0.00230)	(0.00379)
never married	0.0233***	0.0947***	0.0562***
	(0.00486)	(0.00205)	(0.000935)
Observations	350588	905934	1281866

Standard errors in parentheses
Source: Current Population Survey
* p<0.05, ** p<0.01, *** p<0.001

1. Coefficient Interpretation

1.1. Education

The marginal transformation of my probit regression result is presented in **Table 5**, along with their robust SE. Looking at column 1, row 1 and 2 shows the marginal effect of education on health is for adults aged 70 and up. *Less than 12 years of education* has a coefficient of .23 ± .003. This means that holding constant year, age, gender, region, race, ethnicity and marital status, an average person aged 70 and up without a high school degree has a 23 percentage point higher chance of reporting poor or fair health compared to a college graduate. Whereas *12-15 years of education* has a coefficient of .094 ± .002. This means that an average person aged 70 and up with a high school degree has a 9.4 percentage point higher chance of reporting poor or fair health compared to a college graduate, holding other independent variables constant.

Looking horizontally across all age groups, the average marginal effect of education on health has the same economic meaning. Overall, these 6 numbers are positive, and the numbers on the 1st row are higher than the 2nd row. So, people with less than 12 years of education are more likely on average than 12-15 years of education to report poor or fair health, compared to the same base case of more than 15 years of education. In other words, there is a positive correlation between educational level and self-rated health.

1.2. Year and Year×Education

The probit coefficients on the 6 interaction terms between age and education are reported in Table 1. However, the magnitude and statistical significance of the interaction effect varies by observation (Ai & Norton, 2003). In other words, the effect of *Year×Education* on health varies for each of the 23 years included in the sample, and no singular number can capture this. Finding the marginal effect of these interaction terms comes with additional difficulties, because "the magnitude of the interaction effect in nonlinear models does not equal the marginal effect of the interaction term" (Ai & Norton, 2003).

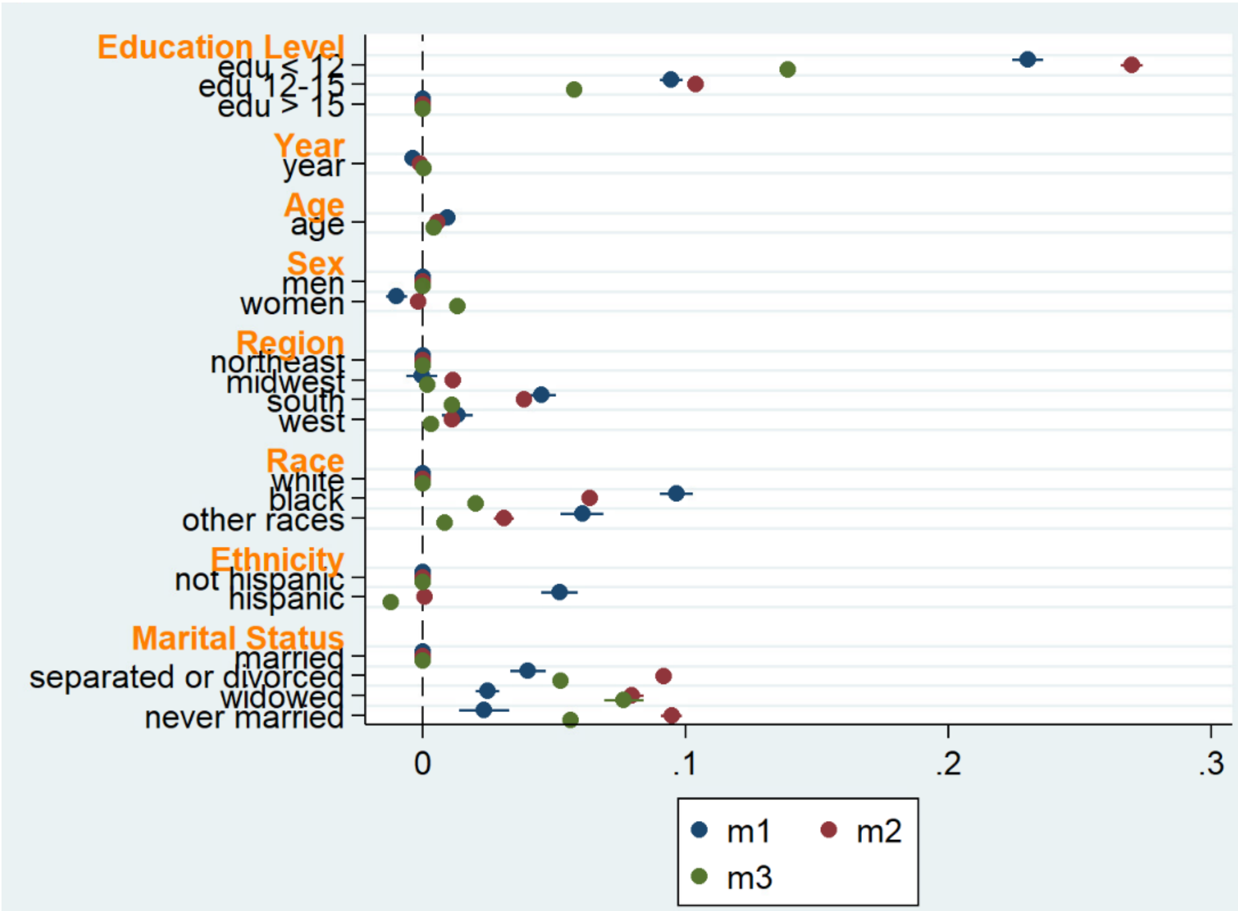
Goesling (2007) was interested in interaction terms since they were trying to study how the relationship between health and education changes over time. This is beyond the scope of my study, which is only about if education has an effect on health. Although CPS data from 2000-2022 does seem to be in support of Goesling's finding (see Appendix).

2. Confidence Interval

If it was possible to run a regression on the entire US population of interest, we could find what the actual effect of education (and other factors) on health is. Because we're working on a sample, the estimations of the true, unbiased effects each have a bounce to them. This is why the coefficients in **Table 5** are reported with their robust SE. Looking at *y* for age group 70+, we can construct a 95% confidence interval as -0.0039 ± 2× (0.00014) = (-0.0042; -0.0036). This means that 95% of the intervals constructed in this way will cover the true, unbiased parameters. **Figure 2** presents all the estimates as dots and their 95% CI as whiskers.

The base cases of dummy variables all have coefficients equal to zero. The interval lengths for model 1's coefficients (the blue bars) are longer overall than that of the other 2 models. This reflects the fact that the oldest age group has the smallest sample size, which is almost 4 times smaller than that of the youngest age group (333,578 versus 1,240,966 observations).

Fig 2: *Average Marginal Effects Estimates with 95% Confidence Interval*



From **Table 5** and **Figure 2**, the dots for the coefficients of *midwest* for model 1 (70+) and model 3 (30-49) along with *women* and *hispan* for model 2 (50-69) have bars that cross the zero line. This means that the true value of these corresponding parameters could be zero, and they might not have an effect on health. Other than these four, no other 95% CI includes zero. The coefficients for *y* and both levels of *e* are all highly statistically significant. The value of *y* estimates across all 3 models look close to zero, but none of their 95% CI crosses the dashed line since their robust SEs are tiny. To further illustrate this point, we can perform hypothesis testing on *y* for age group 70 and older, as shown in **Table 7**.

Table 6: *Hypothesis Testing for y variable of age group 70+*

	Delta-method dy/dx	Std. Err.	z	P>|z|
y	-.0039759	.0001548	-25.68	0.000

Its p-value is miniscule and rounded to 0 by Stata. This means that if *y* truly didn't have an effect on health, the probability of getting our estimate of -0.004 or more extreme is close to 0%. So we can reject the null and accept that *y* does impact health.

CONCLUSION

Using data from IPUMS CPS 2000-2022, this study found a robust positive relationship between education level and self-rated health. Across three age groups, an average college degree holder is less likely to report poor or fair health than an average high school degree holder, and an average high school degree holder is in turn less likely to report poor or fair health than an average person without a high school degree, holding constant time, age, gender, race, ethnicity, marital status and region. This finding is in line with Human Capital Theory and the literature review. Rates of return to education, as traditionally calculated, only account for labor market earnings. If the effect of education on health is causal, it makes sense to attribute additional returns to education. Unfortunately, a clear causal relationship cannot be established from this paper. Asides from income, many variables correlate to education and health, such as ability, parent's education level, health insurance coverage. Consequently, this study suffers from omitted variable bias. The CPS variable *Health* may also be affected by self-report bias. Despite any limitations, this paper reminds us once again of the importance of education. Further work done on providing equal access to education has the opportunity to improve health and wellness for all.

WORKS CITED

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, *80*(1), 123-129.

Becker, G. S. (1964). Human capital: a theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research.

Fletcher J. M. (2008). Adolescent depression: diagnosis, treatment, and educational attainment. *Health economics*, *17*(11), 1215–1235. https://doi.org/10.1002/hec.1319

Fletcher J. M. (2015). New evidence of the effects of education on health in the US: compulsory schooling laws revisited. *Social science & medicine (1982)*, *127*, 101–107.

Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J. R., & Westberry, M (2022). Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. *IPUMS*. Minneapolis, MN

Goesling, B. (2007). The rising significance of education for health?. *Social Forces*, *85*(4), 1621-1644.

Jann, B. (2014). Plotting regression coefficients and other estimates. The Stata Journal 14(4): 708‐737.

Kimbro, R. T., Bzostek S., Goldman N. & Rodríguez G. (2008). Race, ethnicity, and the education gradient in health. *Health Affairs*, 27(2), 361–372.

Kitagawa, E. & Hauser, P. (1973). Socioeconomic Differentials by Cause of Death. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology* (pp.73-92). Cambridge, MA and London, England: Harvard University Press.

Liu, S. Y., Buka, S. L., Linkletter, C. D., Kawachi, I., Kubzansky, L., & Loucks, E. B. (2011). The association between blood pressure and years of schooling versus educational credentials: test of the sheepskin effect. *Annals of Epidemiology*, *21*(2), 128-138.

Lynch S. M. (2006). Explaining life course and cohort variation in the relationship between education and health: the role of income. *Journal of health and social behavior*, *47*(4), 324–338.

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic perspectives*, *25*(1), 159-84.

Walque, D. (2007). Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of health economics*, *26*(5), 877-895.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

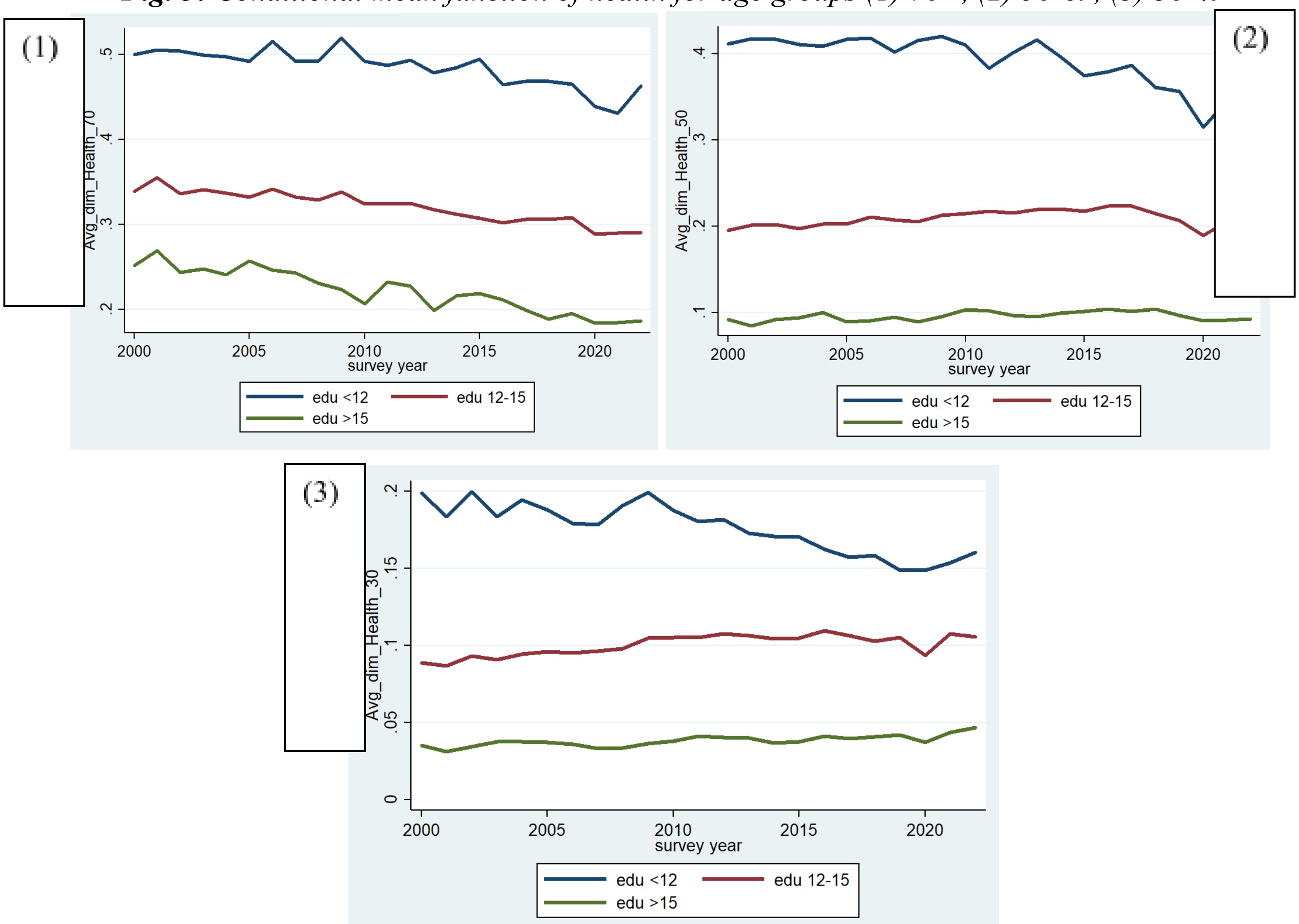
Zajacova, A. (2006). Education, gender, and mortality: Does schooling have the same effect on mortality for men and women in the US?. *Social science & medicine*, *63*(8), 2176-2190.

Zajacova, A., Hummer, R. A., & Rogers, R. G. (2012). Education and health among U.S. working-age adults: a detailed portrait across the full educational attainment spectrum. *Biodemography and social biology*, 58(1), 40–61.

APPENDIX

While interpreting the interaction term for a non-linear model is tricky, we can still get a sense of how education affects health over time. **Figure 3** shows the conditional mean of percentage of people reporting diminished health across education level from 2000-2020, broken down into 3 age groups.

Fig. 3: *Conditional mean function of health for age groups (1) 70+, (2) 50-69, (3) 30-49*



Given the interaction between year and education, the coefficient on variable *year* by itself can be interpreted as the likelihood of our base case reporting poor or fair health. Here the base case is more than 15 years of education. In **Table 3**, *year* has the coefficient -0.0132, -0.00319, 0.005 across three models. We can see this reflected in the slope of the green lines, which represents college graduates: model (1) shows a decreasing percentage of adults aged 70 and up reporting fair or poor health, model (2) shows a stable growth for adults 50-69 years old, and model (3) shows an increasing percentage of adults aged 30-40 reporting fair or poor health.

Furthermore, the red and green lines have similar slopes that make them almost parallel in model (2) and (3). This implies that the health trends of two higher education levels experienced similar growth over the last two decades. This may explain why the coefficients of the interaction between 12-15 years of education and years (see **Table 3**) aren't statistically significant.

Looking at each model as a whole, the three lines are getting further apart in model (1), stay relatively horizontal in model (2), and are becoming closer to each other in model (3). From 2000 to 2020 there seems to be an increasing health disparity for the oldest age group, while the younger age group sees stable or narrowing health disparity. This is in line with Goesling's finding and suggests the continuing existence of diverging health disparity in education levels has been reported since the 1980s. However, this figure only shows the mean of health on education and year without controlling for any confounding variables. While my paper can't confirm Goesling's claim, it does corroborate it.