Abstract

Income, education, occupation, age, sex, marital status, and ethnicity are all correlated with health in one context or another. This paper reflects on the difficulties encountered in deriving robust scientific conclusions from these correlations or drawing reliable policy applications. Interactions among the variables, nonlinearities, casual inference, and possible mechanisms of action are discussed. Strategies for future work are suggested, and researchers are urged to pay special attention to possible interactions among health, genes, and socio-economic variables.

© 2004 Elsevier B.V. All rights reserved.

JEL classification: I12

Keywords: Health; Income; Education; Occupation; Age; Sex; Marital status; Ethnicity; Correlations and causality

Income, education, occupation, age, sex, marital status, and ethnicity are all correlated with health in one context or another. Effective health policy requires an understanding of these correlations; such understanding is also relevant for broader economic issues such as productivity, income distribution, labor force participation, and intergenerational transmission of poverty. No one has done more to advance understanding of these correlations than Michael Grossman. His seminal work on the demand for health (Grossman, 1972) quickly had a major impact on health economics and has continued to inspire streams of research ever since. In addition to his own exceptional productivity, Grossman has been the foremost mentor of Ph.D.s in health economics. He and his students have investigated almost every aspect of health from infant mortality to substance abuse, with special attention to the roles of education and income in determining health outcomes.

* Tel.: +1 650 858 1527; fax: +1 650 328 4163. E-mail address: fuchs@newage3.stanford.edu (V.R. Fuchs).
Despite the contributions of Grossman and others, it seems to me that considerable uncertainty remains concerning the socio-economic correlates of health, the extent to which they reflect causal chains, and their implications for policy. There are numerous reasons for this uncertainty. Many of the socio-economic variables are correlated with each other; sometimes it is difficult to estimate the independent relationship of each one with health. To skirt the problem of multi-collinearity, some researchers refer simply to “socio-economic status” (SES), but this blending of income, education, and occupation limits the scientific value and policy applicability of the research. Interactions among the variables are numerous and varied, as are nonlinearities. One of the biggest problems is establishing causality. Moreover, even when a causal connection appears to be particularly robust, the mechanism of action is usually unknown. All of these problems are exacerbated by the fact that health is multi-dimensional. Frequently used measures such as mortality, morbidity, disability, and self-reported health status are usually positively correlated, but sometimes the correlations are weak and occasionally are even negative. For example, holding age constant, women report worse health and greater disability than men but have higher life expectancy. Additional complications arise from possible interactions among genes, changing medical technology, and the socio-economic variables.

This brief overview of my reflections on these matters cannot possibly cover them in depth; it summarizes my views about the state of knowledge in this area and concludes with some suggestions for future research.

1. Income

Of all the socio-economic variables, the relationship between income and health is probably the most complicated. The correlation can vary from highly positive to weakly negative, depending on context, covariates, and level of aggregation. Even when the positive correlation is strong and stable, the interpretations can include causality running from income to health, from health to income, and/or “third variables” that affect health and income in the same direction.

In high-income countries, researchers usually conclude that the correlation is positive and that the causality runs from higher income to better health. The strength of this effect, however, varies greatly by age, disease, level of income, and other variables. The mechanism(s) through which income might affect health are usually not clearly established. Income can facilitate access to medical care (through insurance or direct purchase), but some of the most carefully controlled studies such as the RAND health insurance experiment do not show significant health effects from better insurance coverage, and universal coverage in European countries does not eliminate the income–health correlation. Income can support acquisition of food, shelter, and other goods and services that contribute to better health, or it can be an indicator of greater efficiency in home production of health (if productivity in the market and in the household are positively correlated). On the other hand, higher income (holding education and other variables constant) may signal longer hours of work, more stress, or participation in dangerous occupations, thus offsetting possible favorable effects of higher income on health.
Some researchers make a distinction between income and wealth. This distinction would be redundant if the two variables were measured comprehensively (including human capital and imputed income), if capital markets worked perfectly, if there were no transaction costs, and if individuals did not engage in peculiar forms of mental accounting. When empirical research produces different results for income and wealth, or for different forms of income, it would be useful to determine why.

The relationship between health and income is particularly strong across countries at below average levels of income. When 149 countries are sorted into quintiles according to Gross Domestic Product per capita in the late 1990s, those in the middle quintile ($3860 average GDP per capita) have 20 years more life expectancy than countries in the lowest quintile ($800 GDP per capita), but only 10 years less than countries in the highest quintile ($20,910 GDP per capita).

The stronger relation across poorer countries may reflect a greater effect of income on health or a greater impact of health on income. Both explanations are probably valid. Higher income in poor countries can increase access to health-enhancing necessities and can support investments in sanitation. Better health means much higher survival rates for infants and children, which raises GDP per capita by increasing the ratio of workers to dependents. Better health among adults increases labor force participation and improves the productivity of those who are at work. Even in high-income countries, a causal chain from health to income is observed when sicker older workers withdraw from work prior to normal retirement, thus reducing their current income as well as their subsequent public and private retirement income.

To see how health and income can be negatively correlated, consider two high-income countries with equal endowments of human, physical, and natural resources but sharply different attitudes toward work. In one country the work week is 35 h, there are many holidays, and four to eight weeks of annual vacation. In the other country, everyone works nearly all the time. The gross domestic product per person would almost surely be higher in the second country, but it would not be surprising if life expectancy were lower. Could this help explain why life expectancy in Italy is higher in the South than the North while income is higher in the North?

Many different “third variables” could be responsible for the positive correlation between income and health, but the obvious ones, such as education, are usually controlled for. Other possibilities, such as the level of serotonin in the brain or the strength of the immune system, are never controlled for. These variables are, no doubt, partly endogenous, but also partly reflect differences in genes, intrauterine developments, and environmental factors in infancy and early childhood. Several additional questions concerning the income–health relationship must be mentioned, albeit all too briefly.

1.1. Nominal or real income?

Most health studies use nominal income measures, even when the individuals or populations live in different locations within the US. Significant differences in cost-of-living across these locations suggest that nominal income is an imperfect and probably biased measure of real income. Because the latter is presumably the variable expected to affect health, the former should be deflated by a cross-location cost-of-living index. The presumption of bias
arises from the likelihood that nominal income and cost-of-living are positively correlated. If so, the elasticity of health with respect to real income would be greater than for nominal income.

1.2. Permanent or transitory income?

Some researchers find that it is permanent income that affects health; other models assume that transitory changes are very important. No one has produced a definitive answer to this question, probably because there isn’t one. The answer may vary depending upon the health measure under study, the level of income, the age of the subjects, and other factors.

1.3. Absolute or relative income?

For a long time most researchers assumed that the positive correlation between income and health was driven primarily by the poor health of those living in poverty, defined as a very low absolute level of income. As income rose over time, however, the income–health gradient showed little evidence of diminishing. Thus, considerable attention is now given to relative income, i.e., to the fact that those at the bottom of the income distribution have worse health than their more affluent peers regardless of the average level of income. This change in focus requires a reexamination of the mechanisms that were invoked for the income–health gradient when average income levels were much lower. When those in poverty went hungry, the connection between health and income was obvious. Now, when obesity is widespread in the US, especially among low-income individuals, different explanations are required. It may be that having much less income than most of society has adverse effects on health in part through psychological mechanisms. More than 2 centuries ago Adam Smith in *The Wealth of Nations* wrote about the importance of relative income for well-being when he defined “necessaries” as “not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it indecent for creditable people, even of the lowest order, to be without.”

2. Education

Compared with income, the education–health relationship appears to be much less complicated. Causality running from health to education is possible, but there is less evidence of this effect than in the case of income. A negative correlation between health and education is very unlikely. Moreover, measurement is usually less problematic for education than for income. Michael Grossman, who has led the way in empirical as well as theoretical investigation of the education–health connection concludes that “years of formal schooling completed is the most important correlate of good health” (Grossman, 2003, p. 32). Proponents of income can argue that education is serving as a proxy for long-term income, but proponents of education can counter that education’s favorable effect on health works, in part, through higher income. The correlation is undoubtedly high in most contexts, but the case for a causal interpretation would be strengthened if certain questions could be resolved.
For example, in cross-section studies, income has sharply diminishing effects on health as income rises, but the apparent effect of additional years of schooling is undiminished at all levels. It is plausible that high school graduates learn something in school that makes them healthier than their peers who dropped out after tenth grade. But is it plausible that adding 2 years of schooling beyond college should have as large a causal effect on health?

The analogy between education’s contribution to health and its contribution to higher earnings is also problematic. With respect to earnings, we know that college graduates who majored in science or engineering earn much more than humanities majors, and MBAs make much more than teachers with masters in education. Do similar differences exist in the education–health relationship? Are college graduates who majored in biology healthier than French literature majors? Does a masters degree in computer science confer more health benefits than one in art history? Does the content of schooling matter at all? If not, what is it that schooling does to improve health?

Education and intelligence are usually correlated, but inclusion of a control for intelligence in earnings regressions does not have much effect on the education coefficient. Is the same true for the education–health relationship? One study of health differences among elderly with chronic conditions found that inclusion of intelligence test scores eliminated the significance of the education variable.

Finally, there are large differences among high school students with respect to cigarette smoking, binge drinking, unprotected sex, and other unhealthy behaviors. Participation in these behaviors tends to be negatively correlated with the amount of schooling the students will eventually complete and negatively correlated with adult health. Thus, when we observe a positive correlation between completed education and adult health, we cannot assume that the additional education caused the better health. Some economists claim that teenagers know how much schooling they will complete (and how high their earnings will be) and that those who expect to have high incomes will take better care of their health even as teenagers so they can enjoy the higher income over a longer period. This may be theoretically correct, but probably weak in importance. One could also imagine a theoretical model in which teenagers who know they will have higher incomes in the future engage in more unhealthy behaviors because they will be better able to afford the medical care to deal with the consequences.

An alternative to inferring that schooling is the cause of better health is to posit the existence of one or more “third variables” that explain the correlation. One possibility is time preference. Individuals who have lower rates of time discount (more willing to delay gratification) are more likely to stay in school longer and do (or not do) the things that contribute to better (worse) health. Another possibility is self-efficacy, a concept developed by psychologist Albert Bandura to describe an individual’s effective control over his (her) behavior. It is, of course, possible that additional schooling lowers time preference, or increases self-efficacy, thus bringing schooling back as a cause of better health.

To explain the education–health connection, some researchers have proposed that those with more schooling are quicker to act on new health information or take advantage of improvements in medical technology. This seems reasonable, but is it important? The persistence of the negative gradient between education and cigarette smoking many decades after information about the harmful effects of smoking became widespread raises questions about the robustness of this explanation.
In comparisons across low-income countries, women’s education is found to be highly correlated with the health of infants and children. Because the level of schooling is usually quite low, it is certainly possible that learning to read and write and receiving some elementary instruction about sanitation and hygiene can have direct health effects. But again one must raise the possibility of omitted variables: those societies that afford greater educational opportunities to women probably also have other customs, traditions, and policies more favorable to women. Such differences would have favorable effects on the health of women and their children independently of the additional schooling.

3. Occupation

Some occupations are much less healthy than others. For example, work related death rates in construction and transportation are triple the rate for all workers. Death rates in agriculture and mining are even higher—five to six times the national average. In addition to fatal injuries, there are large differences across occupations in non-fatal injuries, exposure to toxic materials, and long-term damage to health through physical and psychological stress. Information about past occupation(s) could be as important as current occupation.

Extensive research on earnings differentials across occupations concludes that (ceteris paribus) riskier occupations pay higher wages. But this fact poses a dilemma for investigators. If occupation is not included in a regression of health on income and education, the coefficients may be biased. But inclusion of occupation as an additional RHS variable is also problematic without more understanding of occupational “choice.” The constraints facing a young man born amid the coal mines of West Virginia are undoubtedly different from those confronting a native of Manhattan.

4. Age

That health decreases with age (on average) is a fundamental fact of biology. Because age is frequently correlated with income, education, and other variables, it is essential to control for age through age-adjustment, or better still, analysis of age-specific data. Interactions between age and other variables such as income and education are often quite important. In particular, the correlation between health and income varies greatly over the life cycle, and not monotonically.

5. Sex

At the close of the 20th century, the life expectancy of women exceeded that of men in every country. This is attributed to biological differences, but interactions between these differences and the physical and social environment are of major importance. In low-income countries ($ 1000 Gross Domestic Product per capita or less), the female–male differential in life expectancy is only about 4%. For countries with GDP per capita around $ 3000, it is 8%, and at $ 9000 it averages 11%. But the differential does not continue to increase with
GDP per capita. In the 30 highest income countries (GDP per capita averaging $21,000), the sex differential in life expectancy is only 8%. A similar shift can be observed over time in the female–male differential in life expectancy in high-income countries. Over the past 30 years the differential has tended to fall, reversing the previous long-term trend toward a higher ratio in the course of economic development. In the US this downward trend occurred at the same time that the female–male education ratio was rising.

6. Marital status

In every country, married men and women are healthier than their unmarried peers. Moreover, marital status is usually correlated with income and education. Thus, researchers usually include marital status as a RHS variable when it is available. The presence of a spouse is assumed to make a positive contribution to the household production of health or to increase the demand for health. But there is another possible explanation for the correlation that has the causality running from health to marital status: healthier men and women tend to do better in the marriage market. If that is the case, the typical equation is misspecified.

Both views come sharply into play when considering the fact that the marital status–health correlation is much stronger for men than women. The household production story would say that wives are much more important for the maintenance of their husband’s health than the reverse. The marriage market selection story would say that health is a more important determinant of marital status for men than for women. Panel data might help shed light on this puzzle.

7. Ethnicity

Ethnic differences in health are extensively discussed but convincing explanations are rare. Because ethnicity is frequently highly correlated with other socio-economic variables, a deeper understanding of ethnic differences could contribute substantially to understanding the relation between socio-economic variables and health.

This understanding will not come easily. Measurement problems abound because classification of an individual in a socially constructed ethnic group can depend on self-identification or on the perception of others. The categories are inherently imprecise, and often include heterogeneous sub-categories. Ethnicity may be a marker for genetic differences, cultural differences, or differential treatment by others. Understanding how the legacy of centuries of slavery, segregation and discrimination may affect the health of blacks in the US today is particularly problematic. Researchers must try to avoid stigmatization or the reinforcement of stereotypes, but must also avoid providing incorrect or incomplete explanations because they are politically correct.

The most challenging ethnic health problem in the US is to explain why black life expectancy is 7 years less than white. The most frequently proposed answers – lower income, less education, and poorer access to medical care – probably have some validity;
however, Hispanic-Americans rank lower than blacks on all three measures but do not experience such adverse health outcomes.

8. Concluding observations and suggestions for research

The focus of most empirical research is on hypothesis testing and determining whether the data are “consistent with the model.” Policy makers would benefit from more focus on quantitative estimates of the relation between health and the independent variables.

Most quantitative estimates (and some qualitative inferences) of the relation between socio-economic variables and health are context specific. Policy decisions, therefore, should be based as much as possible on empirical research that covers the relevant contexts.

Even when researchers are confident that they have established a causal connection, they rarely are explicit about the mechanism(s) of action. The elucidation of mechanisms can be enhanced by inclusion of previously omitted variables. For example, if the investigator believes that income increases health through access to medical care, inclusion of one or more measures of access should provide support or refutation of this belief.

Another way of probing for mechanisms is through disaggregation. If the investigator believes that income (or education) affects health through mechanism X, it will usually be possible to draw on medical and epidemiological knowledge to say, “If it is X, the effect will be greater at certain ages than at others, or greater for mortality from certain causes of death than for others.” Repeating the analysis with disaggregated data will lend support to or cast doubt on the proposed mechanism.

Most researchers look for regularities in the data, but there may be significant pay-offs to pursuing anomalies. For example, why do Hispanic-Americans have higher life expectancy than blacks despite less education and lower income? Why is the black–white ratio of heart disease mortality at ages 45–74 much higher for females than males even though black–white differentials in hourly earnings and education are smaller for females than males?

Another question deserving careful attention is why the correlation between income and health is so much stronger in cross-section individual data than in cross-area mean data? Are there differences in income distribution across areas (plus nonlinear relationship between income and health) that are concealed in the area means? Or is there more genetic variation (third variable explanations) within than between areas? Or is there more causality from health to income within areas because the between area differences in income are affected by inherited physical capital, industry mix, location, and other income determinants not controlled for in the regressions? Or are all three explanations relevant?

My final suggestion for research is to pay more attention to interactions between genes and socio-economic variables. For example, social scientists have often identified stress as a major cause of bad health, but sceptics have noted that the effects of apparently the same stress vary greatly across individuals. Recent genetic research helps explain why. Among individuals exposed to considerable stress, the probability of depression is very high for those with two short versions of a particular gene, moderately high for those with one short and one long version, and no higher than that of unstressed individuals for those with two long versions. It seems inevitable that similar interactions will be found for many genes.
and many diseases. Economists seeking to broaden understanding of the socio-economic correlates of health will have to incorporate these interactions into their analyses.

Acknowledgements

Helpful comments from Jay Bhattacharya, Angus Deaton, Jim Poterba, and Richard Zeckhauser are gratefully acknowledged. Thanks are also due to the Robert Wood Johnson Foundation and the California Health Care Foundation for their grants to the NBER to support Fuchs’ research in health economics.

References