Subjective Health and Income Since 1972

Michael Hout
University of California, Berkeley

Scott M. Lynch
Princeton University

**Prepared for the Annual Meetings of the
Population Association of America
17-19 April 2008**

Draft
Comments welcome

Acknowledgments

We acknowledge the support of both the Berkeley Population Center, funded by NIH grant number R21HD056581-01, and the Office of Population Research, funded by NIH grant number R24HDxxx. The contents of this paper are the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or the National Institute of Child Health and Human Development. For the 2007-08 academic year, both authors are at the Office of Population Research, Princeton University, Wallace Hall, Princeton NJ 08544.
Contact us by email: mikehout@berkeley.edu and slynch@princeton.edu.

Introduction

Affluent American adults are healthier than poor ones. Dozens of studies document differences by income and other socioeconomic indicators in a wide variety of health outcomes (e.g., Williams & Collins 1995; Deaton 2001; Mullahy, Robert, & Wolfe 2004; Schnittker 2004; Warren & Hernandez 2007). Gaps differ in size depending on the outcome. But the existence of health disparities is not in doubt.

Although social scientists have been documenting health inequalities at least since the 1950s, they have produced few estimates of long-term trends. Challenges include finding data sets that measure health and income consistently over a long time span, separating the trends in the existence of health problems from medical science's ability to detect those problems, and the complex interactions between income and diagnosis.

Subjective health measures can be a useful tool for assessing changes in disparities over time (e.g., Lynch 2003; Warren & Hernandez 2007). Many surveys include subjective health questions. The subject diagnoses herself; subjective health requires no diagnosis. Taking data from untrained observers may seem unreliable but, as we document below, the measures are surprisingly robust and predictive. Precisely because they do not depend on professional assessment, subjective health measures do not select for the kinds of people who seek out medical exams. And, because the measure is so simple and straightforward, interpreting trends is correspondingly simpler and more straightforward.

The General Social Survey (GSS) and the National Health Interview Survey (NHIS) offer the potential to track the association between family income and the subjective health of adults from 1972 to 2006. The GSS has measured subjective health and family income more consistently over time; the NHIS interviews far more people. As the health disparities literature would predict, the trends in subjective health differ sharply by family income in both surveys. We begin by sketching the overall trends. Then we turn to statistical models that allow us to infer change in the population in both the gross relationship between income and health over time and the net relationship after adjusting for other important factors.

We find that the gap between the subjective health of low-income and affluent adults narrowed significantly during the 1970s, mostly because the fraction of the lowest-income adults who reported good or excellent health rose from about 40 percent in 1972 to 55 percent by 1982. The gap subsequently widened again as adults in affluent families increasingly described their health as "excellent" while, after 1982, adults in low-income families experienced no further improvements in their health.

Health Trends by Income

Low-income Americans reported improved health throughout the 1970s while higher-income Americans reported no change in their health (Figure 1).[1] In the early 1970s, a distinct minority – two-fifths – of low-income Americans had good or excellent health. Middle-income and affluent Americans had substantially better health at that time; three-fourths of middle-income and over 80 percent of affluent Americans felt their health was good or excellent. That sorry situation improved quickly through the 1970s, as more low-income American adults had good or excellent health; the percentage increased from 40 to 53 percent between 1972 and 1980. Relatively few low-income adults described their health as excellent in any year (not shown). Meanwhile the health of middle and upper Americans improved slowly but steadily after the mid-1980s. Middle-income Americans have recently felt their health decline while affluent Americans are more likely than ever to say they have good or excellent health.[2]

Figure 1 is appealing for its simple and straightforward depiction of U.S. health disparities, but we cannot trust our inferences from these data without accounting for the uncertainty of samples and the confounding influence of other variables. The GSS are very high quality – face-to-face interviews, a very high response rate of 77 percent until 2000 and 70 percent since, and the added option of completing the interview in Spanish in 2006. But survey estimates are intrinsically un-

[1]This is an illustrative result from the GSS. As our project evolves, we plan to use the much larger NHIS and CPS to get more precise estimates of these trends. We also plan to investigate if the improved health among low-income adults coincided with public policy. In particular many states improved their Medicaid coverage through the early 1970s. States funded Medicaid more intensively in different years. If improving health among adults in low-income families was due to Medicaid, then the trend should appear sooner in states that enrolled in Medicaid right away and later in states that waited to join Medicaid.

[2]The data for four other income groups – $15,000–$24,999, $25,000–$34,999, $50,000–$64,999, and $65,000–$84,999 – fall between the data shown. The smoothed trend lines are monotonic, though the point estimates occasionally depart from the rank order of income categories.

certain due to the nature of sampling, so tests are appropritate. Confounding factors that influence health include education, age, gender, racial ancestry, ethnicity, marital status, and region. Education and age are most immediately relevant. Income differences are bigger for less-educated adults (Schnittker 2004), and the relationship between education and health increased across cohorts (Lynch 2003). Furthermore, the composition of the low-income population changed dramatically between 1970 and 1990, becoming more feminine, younger, and more urban (Bianchi 1981; Iceland 2003). Marriage also emerged as a major factor in poverty (Ellwood & Jencks 2004; Fischer & Hout 2006). As most of these variables correlate with subjective health as well as income, we will have to make statistical adjustments to the trends in Figure 1 before we can infer that the income-health relationship changed.

Criteria for a Statistical Model

A useful statistical model for trends in the relationship between family income and subjective health would have these properties:

1. Take account of the order among the categories of the dependent variable. The subjective health responses – "excellent, good, fair, or poor" – are ordered almost certainly not cardinal.

2. Express the relationship between family income and subjective health in way that can detect the difference between the trend for people in low-income families and the trend for people at higher income levels.

3. Specify all the other factors that might confound the income-health relationship. The literature contains examples of interactions between age and income, gender and income, and cohort and education. We will want all those interactions in our model as well as terms for changes in important differences among racial ancestry groups, marital statuses, and the like.

Ordered logit and probit models are, by now, well-known and widely used, but our initial attempts to apply them to subjective health failed. Preliminary tests unfortunately turned up strong evidence against the "proportional odds" assumption of the ordered logit model and the equivalent "parallel slopes" assumption of the ordered probit model. Stereotyped ordered regression (what we will refer to as "slogit") is a parsimonious model for ordered dependent variables that does not require the parallel logits assumption (Anderson 1984; DiPrete 1990).[3]

Spliced line ("spline") functions are a useful tool for situations where the researcher is interested in hypotheses about particular segments of the distribution of an independent variable. A spline specifies one slope for the relationship up to some point (called the "knot") in an independent variable's range and another slope beyond the knot. We experimented with several exploratory techniques to pick a good knot. We looked at scatterplots and fit models with many dummy variables for narrow income categories to help us choose appropriate knots. Then, using a random one-third of the non-missing cases, we tried knots at 6, 8, 10, 12, 16, 20, 24, 32, and 64 thousand dollars. Spline functions are very flexible and can have more than one knot. We tried three pairs of knots: 10, 12, and 16 thousand dollars paired (one at a time) with 64 thousand dollars. A single knot at $12,000 produced the best results as gauged by Bic' (Raftery 1995). The relationship below the knot refers to the low-income population – our main focus – while the relationship above the knot refers to the people whose family income exceeded $12,000 per year.

The slogit model with a spline function for the relationship between income and health is also a multivariate model well-suited for incorporating any number of other independent variables and covariates.

Details of the slogit model

The GSS subjective health question is "Would you say your own health, in general, is excellent, good, fair, or poor?" To have health scores rise with better health, we score the answers 4, 3, 2, and 1 for excellent, good, fair, and poor, respectively. The general form of the slogit model (Anderson 1984; DiPrete 1990) is:

$$y_{ij} = \ln\left(\frac{p_{ij}}{p_{i1}}\right) = (\theta_j - \theta_1) + \sum_{k=1}^{K} (\phi_j - \phi_1)\beta_k X_{ki} \tag{1}$$

where $i = 1, \ldots, N$ indexes people, $j = 1, \ldots, J$ indexes responses, $k = 1, \ldots, K$ indexes the independent variables, the θs are intercepts, the βs are slopes, and the ϕs alter the slopes as j

[3] Anderson (1984) discusses the prospect of ordering the dependent variable along more than one dimension. Our data provide very weak evidence of more than one dimension to the subjective health responses.

changes.[4] If the ϕs are monotonic with respect to *j*, i.e.,: $\phi_1 > \phi_2 > \ldots > \phi_J$ or $\phi_1 < \phi_2 < \ldots < \phi_J$, the model is appropriate for an ordinal dependent variable. The model is under-identified without restrictions on the θs and ϕs. We use $\theta_1 = \phi_1 = 0$ and $\phi_J = 1$ as our identifying restrictions because they make interpretation easier for us. Those identifying restrictions simplify equation (1) to:

$$\ln\left(\frac{p_{ij}}{p_{i1}}\right) = \theta_j + \phi_j \sum \beta_k X_{ki} \tag{2}$$

for $j = 2, \ldots, J$.[5] Inserting the spline function into the slogit model we get this expression:

$$\ln\left(\frac{p_{ij}}{p_{i1}}\right) = \theta_j + \phi_j \left(\beta_1 lnIncome_i + \beta_2 lnSpline_i + \sum_{t=1}^{22} \beta_{t+2} T_{ti}\right) \tag{3}$$

where j = 2, 3, 4, and T_t =1 for year t and 0 otherwise, for t = 1972, 1973, ..., 2004, skipping 1978, 1983, and 1986 because the subjective health question was not asked in those years, and 1979, 1981, 1992, 1995, 1997, 1999, 2001, 2003, and 2005 because no survey was done in those years. We use 2006 as the reference year.

The year dummies in equation (3) shift subjective health for everyone. But our primary interest is in changes over time for low-income persons but not middle- and upper-income ones. That is where the spline function comes in. If β_1 became smaller over time while β_2 did not, or if β_1 became smaller over time while β_2 got larger, then we would have statistical evidence to support our reading of the trends in Figure 1. Formally, this calls for interaction effects between the income variables and year dummies:

$$\begin{aligned} \ln\left(\frac{pij}{pi1}\right) = \quad & \theta_j + \phi_j(\beta_1 lnIncome_i + \beta_2 lnSpline_i + \textstyle\sum_{t=1}^{22} \beta_{t+2} T_{ti} \\ & + \textstyle\sum_{t=1}^{22} \beta_{t+24} lnIncome_i T_{ti} + \textstyle\sum_{t=1}^{22} \beta_{t+46} lnSpline_i T_{ti}) \ . \end{aligned} \tag{4}$$

[4] In the original article, Anderson (1984, eq. 7) wrote the model as $f_s(\mathbf{x})/f_k(\mathbf{x}) = \exp\left(\beta_{0s} - \phi_s \beta^{\mathbf{T}}\mathbf{x}\right), (s = 1, \ldots, K; \beta_{0k} = \phi_k = 0)$. The minus sign in front of ϕ_s makes interpretation unnecessarily confusing, so we write the model with a plus sign in front of ϕ_s. Stata estimates the model as Anderson wrote it, so we reverse the signs of the β coefficients produced by Stata to align the results with our equations.

[5] By controlling the ϕs researchers can use the slogit model to specify a wide variety of interesting models. For example, if $\phi_j = (J - j)/(J - 1)$, the model generalizes Duncan's (1979) "uniform association" model (also see Goodman 1979; DiPrete 1990). And with the specification $\phi_j = z_j$ where z_j is the score for category j on some scale of interest, we get a multivariate version of the linear-by-linear interaction model (Haberman 1979; Hout 1984; DiPrete 1990).

In equation 4, β_1 and the interaction terms that include $lnIncome$ apply to the low-income population; β_2 and the interaction terms that include $lnSpline$ apply to people with family incomes above \$12,000. See:

$$
\begin{aligned}
\frac{dy}{dIncome} &= \beta_1 + \sum_{t=1}^{22} \beta_{t+24} && \text{if} \quad Income \leq \$12,000 \\
&= (\beta_1 + \beta_2) + \sum_{t=1}^{22} \beta_{t+24} + \beta_{t+46} && \text{if} \quad Income > \$12,000
\end{aligned}
$$

Therefore β_{25} to β_{30} – or equivalent coefficients – are the coefficients of greatest interest. If they are statistically significant and greater than zero while the rest of the coefficients of that type (β_{31} to β_{46} are not, then the slogit model will have confirmed our interpretation of the trends in Figure 1.

Using 44 parameters to assess changes in the income-health relationship is pretty inefficient. Preliminary analyses showed we would lose little information if we grouped years into seven time periods – 1972–1975, 1976–1980, 1982–1985, 1987–1990, 1991–1994, 1996–2000, 2002–2006, reducing the 44 parameters to 12 (with 2002–2006 as reference). That left us with:

$$
\begin{aligned}
\ln\left(\frac{pij}{pi1}\right) = \quad & \theta_j + \phi_j(\beta_1 lnIncome_i + \beta_2 lnSpline_i + \textstyle\sum_{t=1}^{6} \beta_{t+2} T_{ti} \\
& + \textstyle\sum_{t=1}^{6} \beta_{t+9} lnIncome_i T_{ti} + \sum_{t=1}^{22} \beta_{t+46} lnSpline_i T_{ti}) \ . \qquad (5)
\end{aligned}
$$

A Wald test failed to reject the null hypothesis that the interactions involving $lnSpline$ were zero (i.e., $\beta_9 = \beta_{10} = \ldots = \beta_{14}$ = 0).

We begin with this basic model, then add demographic variables (an additional 38 parameters) and dummy variables for specific places (76 parameters) for the full model. The income coefficients for the basic, demographic, and geographic models are in Table 1. The means and standard deviations of all variables in the analysis are appended (Table A1), as are the coefficients for all the substantive variables in the model (Table A2).

The Changing income-health Relationship

American adults in families with incomes below \$12,000 per annum had significantly worse health than Americans with more family income throughout the 1972-2006 period. Differences among low-income people, up to \$12,000 a year, are not significant in most years. Beyond the \$12,000 a year threshold, health improves proportionally with family income. The slope is steepest for the most extreme contrast – excellent vs. poor health. Its precise value is $\phi_4(\beta_1 + \beta_2) = 1.41$ for the 2002-2006 time period (the default). To see how income-health relationship varies among adjacent categories, we compare the $\hat{\phi}$s.

$$\begin{aligned}
\text{Excellent : Good} \quad & \hat{\phi}_4 - \hat{\phi}_3 && = 1 - .771 = .229 \\
\text{Good : Fair} \quad & \hat{\phi}_3 - \hat{\phi}_2 && = .771 - .377 = .394 \\
\text{Fair : Poor} \quad & \hat{\phi}_2 - \hat{\phi}_1 && = .377 - 0 = .377
\end{aligned}$$

Thus income affects the odds on good versus fair health and fair versus poor health more than it affects the odds on excellent versus good health. And these differences apply proportionately to the two parts of the income-health relationship under the slogit model.

The first time period – 1972-1975 – was exceptional, as indicated by the significant interaction effect for that period one. The magnitude of $\hat{\beta}_4$ = .397 relative its standard error of .089 indicates that the evidence for a different pattern in the first period is quite strong. The second period was less distinctive $\left(\hat{\beta}_5 = .224\right)$ but still significantly different from the last period. Subsequent periods are less distinct from the last one (and one another), further supporting our interpretation of Figure 1. Just as in Figure 1, we see in these coefficients evidence of a reduction in the income-health relationship in the 1970s with no significant change after that.

Figure 2 helps us visualize these changes by displaying each period's regression line in a simple graph of the odds on good versus fair health by family income. Both axes are logged to highlight the linearities in the model. We clearly see, emphasized by its broad line and red color, the steeper slope of the income relationship for the first period compared to later ones. The second period is just above it – also red but thinner line. The most recent period is at the top in bold blue. This figure shows the improved health of low-income people. In 1972-1975, more people with incomes below $12,000 had fair health than good health (the odds of good versus fair were less than one). Over time the health of the poorest of the poor caught up with that of other low-income Americans. It also shows the comparatively trivial change in how income affected the health of persons from families with incomes over $12,000. Indeed in every year the steep slope is above the $12,000 threshold.

The model fits the observed data quite well as we see in Figure 3. Observed data points scatter around all three lines from the model, but the spline function captures the key features of the underlying data.

In the conclusion we will discuss what happened in the early 1970s that might have influenced poor peoples' health and reduced the income gradient at the lowest end. But first we turn to the multivariable results in order to minimize the possibility that we are making a spurious inference.

Multivariate Results

We now develop a full multivariate model for subjective health, building on our results in the previous sections, of course, but also on the vast health disparities literature. In addition to the income terms in the basic model, we include gender, racial ancestry (African Americans and Latinos compared with all other groups), education, age, marital status, principal activity, region, location at age 16 (U.S. vs. foreign country), and – in the geographic model – specific places within the United States.[6] We include interaction effects between income and ancestry, gender, and age, between age and gender, between education and cohort, and between Latino ancestry and foreign residence. The income coefficients are in the second and third columns of Table 1; results for the other substantive variables are shown in Appendix Table A2.[7]

The multivariate results confirm the previous finding that the income-health relationship was substantially stronger in the early 1970s than it has been since. The coefficient gauging the income--time interaction for 1972-1975 in the demographic model is statistically significant but only half as large as in the baseline model. Adding the place dummies does not reduce the coefficient any more, but it does raise the standard error, leaving us uncertain about its significance within places. The coefficient for the second period – 1976-1980 – is significant in baseline model only.

The demographic controls also substantially reduce our estimate of the income-health relationship in families with over $12,000 income. In the baseline model $\hat{\beta}_2$ is 1.505; in the demographic model it is .760. However, several of the demographic interact with income in producing health

[6] We used dummy variables for the places represented by the GSS variable "sampcode," recoded to recognize when two code numbers represent the same particular place.

[7] Coefficients for the year dummies are available on request; coefficients for the place dummies cannot be released because they depend on access to a confidential data set. Releasing them would violate the agreement we signed to gain access.

disparities so that $\hat{\beta}_2$ of .760 in the demographic model actually only pertains to men who are neither African American nor Latino, and are 25-34 years old. Income above $12,000 influences women's health about 30 percent more than men's. African Americans and Latinos are less affected by increases in income – the income effect is approximately half as large for these two groups as it is in the rest of the population. The income effect is significantly bigger for 45-54 year olds and nearly zero among people 75 years old and over.

Education, like income, is a major source of health disparity (Mirowsky & Ross 2003; Lynch 2003). That shows up clearly in these results. The coefficient for education in the range from 8 years to 16 is .296; thus a 2 year increase in education – whether from 10 to 12, 12 to 14, or 14 to 16 – boosted underlying health by just about the same amount as doubling annual income – whether from $12,000 to $24,000 or $48,000 to $96,000. Anyway you read that, it is a large effect. Importantly, though, these very large educational disparities did not change in the 1970s or later. The change was concentrated in the income disparities.

We also note the extremely large health disparities among people with different kinds of principal activities. Employed people are healthiest by far. Unemployed, retired, and other-status people are substantially less healthy than the employed. Even people keeping house are significantly less healthy than employed people. These activities are as likely consequences as causes of health. Many people retire early precisely because their health prevents them from working anymore. Most respondents coded "other" are disabled.

Conclusion

Poor peoples' health improved in the early 1970s, significantly reducing health disparities. Looking back we see that the relationship between income and health was significantly stronger in the 1970s than it has been since. The change is visible in simple, descriptive trend lines. Multivariate analysis shows that ignoring other variables that contribute to health disparities leads us to exaggerate income differences in health but not the direction or significance of the trend. In short, the evidence for the change is both broad and fairly robust.

What changed in the early 1970s that could have resulted in improvements in health for low-income Americans but not others? First, we ought not limit our attention to the early 1970s. The changes we recorded all took place in the 1970s, mostly in the early 1970s. But we lack data for the period before 1972. For all we know, the 1972-1977 trend is the tail end of something that had been going on longer.

Medicare and, especially, Medicaid are our candidates for reducing health disparities by income. Introduced in 1965, Medicare and Medicaid spending varied more from state to state in 1972 than it did later. Congress originally mandated comprehensive coverage in every state by 1975 but pushed the deadline back to 1977 in 1969. By 1978, the poor-peoples' health provisions were as close to completely implemented as they ever would be. So the timing works right; events in Medicaid's history correspond very nicely to the chronology of changing disparities. The particular pattern of income differences we find in the GSS also accords well with the Medicaid story (though other explanations may be consistent with it as well). By the late 1990s, the GSS shows that increases in income from very low levels to $12,000 per year have no affect on health. Disparities kick in above that $1,000 a month threshold. In the early 1970s, by contrast, the disparities were significant below the $12,000 a year threshold, though they were, even then, smaller than the disparities above it.[8]

Coincident timing and differences in the pattern of disparities are suggestive but weak evidence. It is impossible to say with any certainty that there was no other change underway in 1972 and spent by 1977 that could have produced the patterns we see in the GSS data. For example, the 9-1-1 emergency dispatch system was introduced in 1968 and spread.[9] 9-1-1 seems less plausible to us because most health issues are not emergencies and people of all income brackets use emergency services. But it has the timing right. And we cannot rule it out with evidence any stronger than the preceding sentence.

Indeed the quandary we face is probably the reason why we could not find an academic article in social science or public health journals or a government report that proved the success of Med-

[8]An income that gets expressed as $12,000 at 2005 prices was reported in the 1972-1975 GSS as falling in the $3,000-$3,999 per year income bracket. It is a meaningful amount for the senior author because he was supposedly supporting two kids on his $3,600 per year NIMH fellowship in 1974 and 1975. That proved insufficient, even in subsidized married student housing that cost just $108 per month, so he supplemented his fellowship working as a research assistant to his dissertation chair.

[9]We got this suggestion from a colleague at Columbia when we presented a preliminary version at the department colloquium there.

icaid. Our methodology of assessing treatment effects is well-suited to discrete treatments. We need to be able to assess the efficacy of the treatment on the treated as well as average treatment effects. What is the treatment in a broad program like Medicaid. It is insurance not treatment. It steps in to reimburse patients and healthcare providers for costs and services. It is diffuse. Following the money, its value is contingent on the incidence of trouble. The "treatment" for a seriously ill cancer patient may be far more extensive than that for a person who sprains her ankle while exercising. But the program improves health and removes some health disparities if it removes income as a factor in seeking treatment when either of those things occurs.

Our plans for future work involve finding larger data sets that contain data on separate states. Our idea is that we might be able to correlate changes in the income-health association within states with the timing of full implementation of Medicaid, levels of enrollment and spending at the state level. At first we thought we could use the National Health Interview Study. It is an order of magnitude bigger than the GSS so it gives us substantially more statistical power and many observations in most states. Unfortunately for us, the public use files do not reveal the respondent's state of residence. The March CPS has included health since 1996. There are lots of cases and useful geographic detail. But the CPS time series starts too late; the key years are in the 1970s not the last 12 years.

References

Bianchi, Suzanne. 1981. *Household Composition and Racial Inequality.* New Brunswick: Rutgers University Press.

Deaton, Angus. 2001. "Inequalities in Income and Inequalities in Health." Pp. 285-363 in *The Causes and Consequences of Increasing Inequality*, edited by Finis Welch. Chicago: University of Chicago Press.

Ellwood, David, and Christopher Jencks. 2004. "The Uneven Spread of Single-Parent Families: What Do We Know? Where Do We Look for Answers?" Pp. xxx-yyy in *Social Inequality*, edited by Kathryn Neckerman. New York: Russell Sage Foundation.

Fischer, Claude S., and Michael Hout. 2006. *Century of Difference: How America Changed Over the Last One Hundred Years.* New York: Russell Sage Foundation.

Iceland, John. 2003. *Poverty in America*. Berkeley: University of California Press.

Lynch, Scott M. 2003. "Cohort and Life Course Patterns in the Relationship Between Education and Health: A Hierarchical Approach." *Demography* 40: 309-331.

Mirowsky, John, and Catherine E. Ross. 2003. *Education, Social Status, and Health*. New York: Aldine de Gruyter.

Mullahy, John, Stephanie Robert, and Barbara Wolfe. 2004. "Health, Income, and Inequality." Pp. zz-bb in *Social Inequality*, edited by Kathryn Neckerman. New York: Russell Sage Foundation.

Schnittker, Jason. 2004. "Education and the Changing Shape of the Income Gradient in Health." *Journal of Health and Social Behavior* 45: 286-305.

Warren, John Robert, and Elaine M. Hernandez. 2007. "Did Socioeconomic Inequalities in Morbidity and Mortality Change in the United States over the Course of the Twentieth Century?" *Journal of Health and Social Behavior* 48: 335-351.

Williams, David R., and Chiquita Collins. 1995. "U.S. Socioeconomic and Racial Differences in Health: Patterns and Explanations." *Annual Review of Sociology* 21: 349-386.

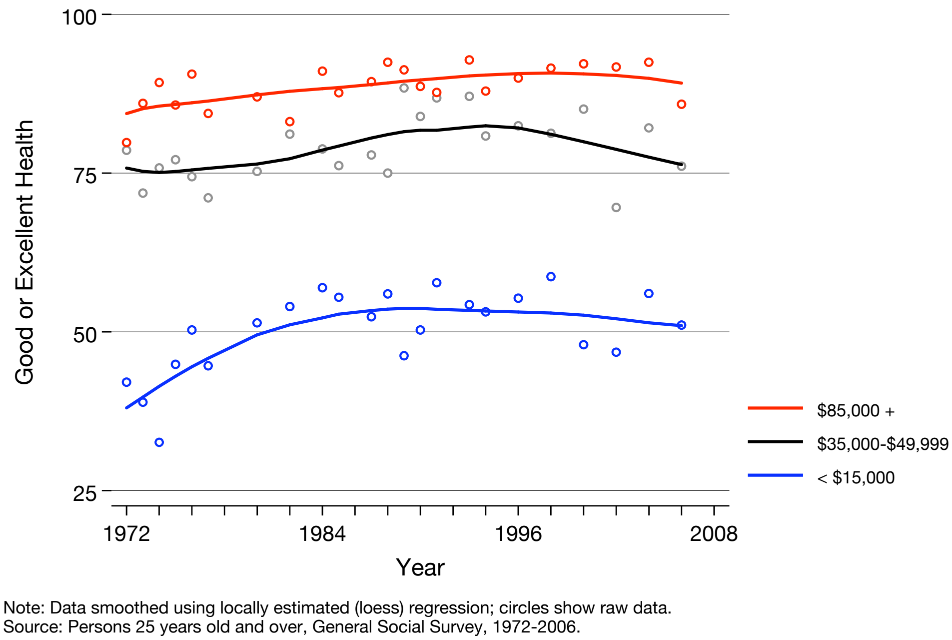

Figure 1. Good or Excellent Subjective Health by Year for Three Income Groups: 1972-2006.

Table 1
Income Coefficients from Slogit Regressions of Subjective Health on Family Income (with and without controls)

		Model	
Variable	Baseline	Demographic	Geographic
Family income[a]	.038	.099	.109
	(.083)	(.095)	(.104)
Family income $\geq$ \$12,000[a]	1.505*	.760*	.767*
	(.097)	(.142)	(.148)
Interactions: Family income $\times$ year			
1972-1975	.397*	.197*	.204*
	(.089)	(.097)	(.112)
1976-1980	.224*	.069	.067
	(.100)	(.107)	(.112)
1982-1985	.058	-.129	-.160
	(.096)	(.105)	(.110)
1986-1990	.185	.022	.002
	(.098)	(.103)	(.108)
1991-1994	.122	.038	.016
	(.098)	(.108)	(.114)
1996-2000	.134	.086	.059
	(.081)	(.087)	(.093)
Slope shifts			
ϕ_1	.000	.000	.000
	—	—	—
ϕ_2	.377*	.371*	.376*
	(.020)	(.014)	(.014)
ϕ_3	.771*	.761*	.762*
	(.013)	(.010)	(.010)
ϕ_4	1.000	1.000	1.000
	—	—	—
Fit statistics			
Observations	30,813	30,813	28,378
Initial log-likelihood	-36,636.74	-36,635.74	-33,744.16
Model log-likelihood	-33,877.57	-32,335.65	-29,802.22
Degrees of freedom used	35	83	151
BIC’			

$*p < .05$.

[a]Transformed to ratio scale by taking the natural logarithm.

Source: Persons 25 years and over, General Social Survey, 1972-2006.

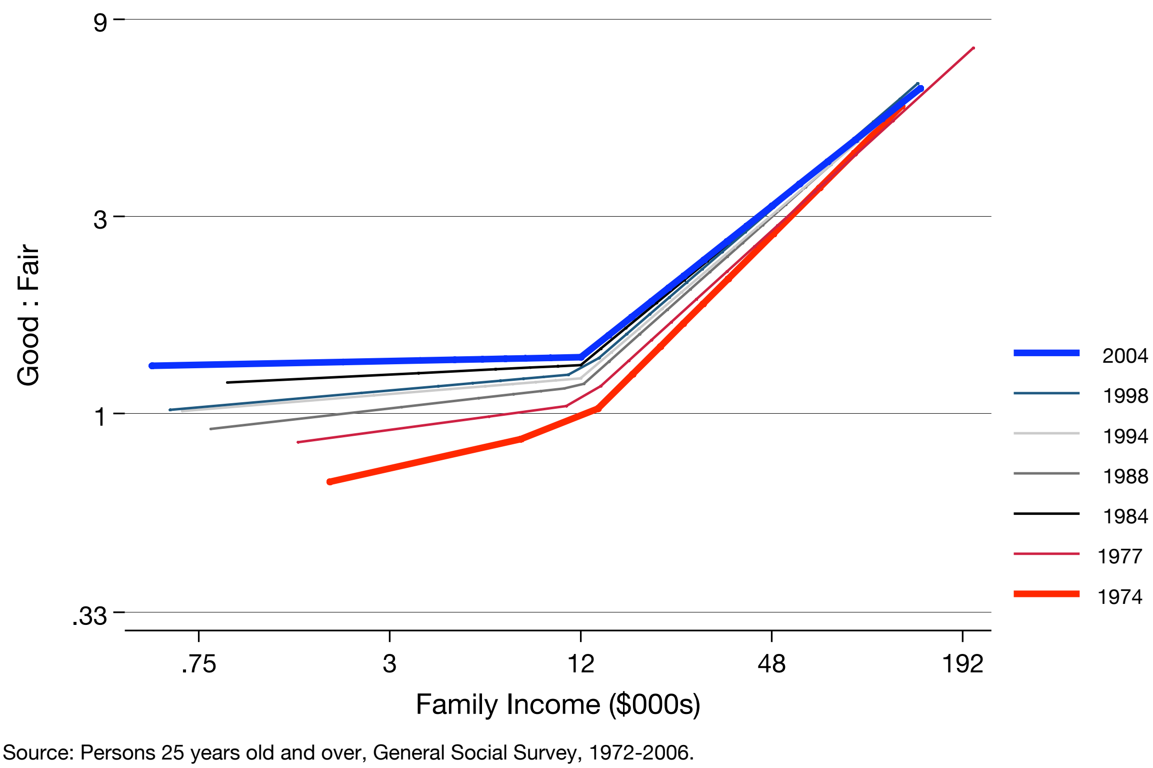

Figure 2. Good versus Fair Health by Income and Time Periods.
Source: Coefficients in Table 1.

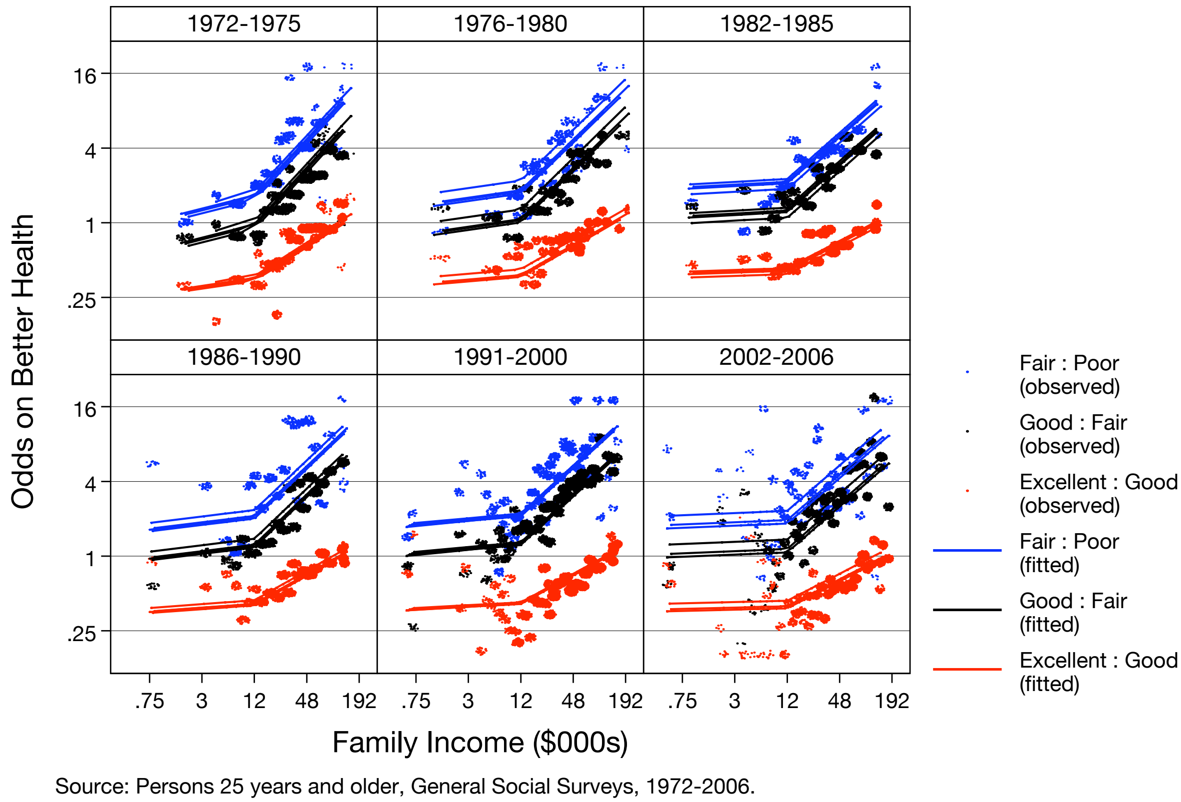

Figure 3. Subjective Health by Income in Six Time Periods.
Source: Basic model in Table 1 overlaid on the observed data.

Appendix Tables

Appendix Table A1
Means and Standard Deviations for the Slogit Analysis

Variable	Mean	Standard deviation
Family income	$40,928	$37,812
Family income (log)	3.712	.868
Family income $\geq$ $12,000 (log)	1.285	.722
Woman	.535	.499
Black	.110	.312
Latino	.058	.234
Age		
25-34	.258	.438
35-44	.238	.426
45-54	.204	.403
55-64	.148	.355
65-74	.100	.301
75 and over	.052	.222
Marital status		
Married once	.551	.497
Remarried	.141	.348
Divorced or separated	.124	.330
Widowed	.068	.252
Never married	.115	.319
Principal activity		
Employed	.653	.476
Unemployed	.027	.161
Retired	.118	.323
Student	.012	.110
Keephouse	.174	.379
Othstatus	.016	.125
Education level	1.502	1.190
Advanced degree	.071	.256
Living in a foreign country at age 16	.058	.233
Region		
Northeast	.199	.399
Midwest	.263	.441
South	.343	.475
Mountain	.060	.237
Pacific	.135	.342

Appendix Table A2
Slogit Coefficients for Three Models of Subjective Health

		Model	
Variable	Baseline	Demographic	Geographic
Family income[a]	.038	.099	.109
	(.083)	(.095)	(.104)
Family income $\geq$ \$12,000[a]	1.505*	.760*	.767*
	(.097)	(.142)	(.148)
Interactions: Family income $\times$ year			
1972-1975	.397*	.197*	.204
	(.089)	(.097)	(.112)
1976-1980	.224*	.069	.067
	(.100)	(.107)	(.112)
1982-1985	.058	-.129	-.160
	(.096)	(.105)	(.110)
1986-1990	.185	.022	.002
	(.098)	(.103)	(.108)
1991-1994	.122	.038	.016
	(.098)	(.108)	(.114)
1996-2000	.134	.086	.059
	(.081)	(.087)	(.093)
Years of Education			
Total		.296*	.293*
		(.015)	(.015)
Elementary		-.169*	-.164*
		(.037)	(.040)
Postgraduate		-.174*	-.172*
		(.055)	(.056)
Born before 1930		-.036*	-.037*
		(.008)	(.008)
Foreign		-.141*	-.150*
		(.029)	(.032)
Woman		-.158	-.062
		(.145)	(.150)
Black		-.043	-.067
		(.111)	(.120)
Latino		.165	.141
		(.207)	(.219)
Interactions: Family income $\geq$ \$12,000[a] $\times$ social group			
Woman		.224*	.191*
		(.079)	(.082)
Black		-.427*	-.362*
		(.108)	(.114)
Latino		-.321*	-.295

continued on next page

Appendix Table A2 continued

Variable	Baseline	Demographic	Geographic
		(.163)	(.176)
Age			
25-29		.000	.000
		—	—
30-34		-.228*	.000
		(.114)	—
35-39		-.483*	-.497
		(.199)	(.192)
40-44		-.807*	-.497*
		(.203)	(.192)
45-49		-1.822*	-1.769*
		(.212)	(.205)
50-54		-1.960*	-1.769*
		(.214)	(.205)
55-59		-1.911*	-1.613*
		(.222)	(.216)
60-64		-1.706*	-1.613*
		(.224)	(.216)
65-69		-1.065*	-.837*
		(.233)	(.230)
70-74		-1.089*	-.837*
		(.240)	(.230)
75-79		-.694*	-.301
		(.263)	(.259)
80-84		-.425	-.301
		(.282)	(.259)
85 and over		-.402	-.301
		(.332)	(.259)
Interactions: Woman $\times$ Age			
35-44		-.388*	-.410*
		(.161)	(.168)
45-54		.390*	.372*
		(.167)	(.175)
55-64		.110	.050
		(.173)	(.182)
65-74		.079	.005
		(.185)	(.192)
75 and over		-.072	-.208
		(.216)	(.226)
Interactions: Family income $\geq$ \$12,000^a $\times$ Age			
35-44		.047	.057
		(.119)	(.123)

continued on next page

Appendix Table A2 continued

Variable	Baseline	Demographic	Geographic
45-54		.256*	.296*
		(.122)	(.128)
55-64		.192	.217
		(.129)	(.135)
65-74		-.055	-.057
		(.149)	(.155)
75 and over		-.639*	-.660*
		(.172)	(.181)
Marital Status			
Married once		.220*	.173
		(.096)	(.101)
Remarried		-.216	-.265*
		(.112)	(.118)
Divorced or separated		-.151	-.184
		(.102)	(.106)
Widowed		.096	.006
		(.122)	(.127)
Never married		.000	.000
		—	—
Principal activity			
Employed		.000	.000
		—	—
Unemployed		-.696*	-.684*
		(.159)	(.168)
Retired		-1.276*	-1.299*
		(.105)	(.109)
Student		-.554*	-.500
		(.287)	(.306)
Keeping house		-1.023*	-1.020*
		(.080)	(.084)
Other status		-3.206*	-3.180*
		(.189)	(.197)
Foreign		1.591*	1.791*
		(.401)	(.442)
Foreign × Latino		-.141	-.254
		(.282)	(.307)
Region			
Northeast		.000	.000
		—	—
Midwest		-.086	-.173
		(.078)	(.116)
South		-.263*	-.518*

continued on next page

Appendix Table A2 continued

Variable	Baseline	Demographic	Geographic
		(.074)	(.110)
Mount		.114	-.244
		(.125)	(.165)
Pacific		-.034	.009
		(.095)	(.161)
Slope shifts			
ϕ_1	.000	.000	.000
	—	—	—
ϕ_2	.377*	.371*	.376*
	(.020)	(.014)	(.014)
ϕ_3	.771*	.761*	.762*
	(.013)	(.010)	(.010)
ϕ_4	1.000	1.000	1.000
	—	—	—
Intercepts			
θ_1	.000	.000	.000
	—	—	—
θ_2	.653*	.360*	.286
	(.092)	(.140)	(.154)
θ_3	.736*	-.126	-.241
	(.179)	(.284)	(.309)
θ_4	-.220	-1.575*	-1.714*
	(.234)	(.378)	(.411)
Year and place dummies			
Year dummies	Yes	Yes	Yes
Place dummies	No	No	Yes
Fit statistics			
Observations	30,813	30,813	28,378
Initial log-likelihood	-36,636.74	-36.635.74	-33,744.16
Model log-likelihood	-33,877.57	-32,335.65	-29,802.22
Degrees of freedom used	35	83	151
BIC’			

*$p < .05$.

[a]Transformed to ratio scale by taking the natural logarithm.

Source: Persons 25 years and over, 1972-2006 GSS.