Health, Mental Health and Labor Force Participation: The Role of Self-Reporting Bias

April 2008

Abstract

This paper relates physical and mental health status to labor force participation and compares these relationships among self-report and proxy respondents. Previous research on male subjects has conjectured that subjective self-reports of health status may lead to an upward bias in the estimated effect of health on labor force participation because subjects who are out of the labor force may be more likely to understate their health status so as to justify their lack of employment. We initially find some evidence of systematic differences between proxy and self-reporters in the labor force participation effects of health, raising the possibility that self-reporters may be biased in their health assessments. Once we control for objective baseline indices of physical and mental health status, however, differences between subjective health assessments and labor force participation become smaller and statistically insignificant. These results suggest that self-reports do not lead to overestimates of the importance of good physical or mental health on labor force participation, once one conditions on more objective measures validated for their psychometric properties.

Keywords: health economics, propensity scoring, logistic regression, self-reported health, labor force participation, mental health.
JEL classification numbers: C10, C14, I18, J24.

1. Introduction

The impact of health on labor force participation is a topic of considerable interest in the labor and health economics literatures. Growing recognition of the indirect labor productivity costs associated with poor health and chronic disease has raised the interest of policymakers as well. Previous research (e.g., Bound, 1991) has conjectured that self-reports of health status may lead to an upward bias in the estimated effect of health on labor force participation. The reason is that subjects who are out of the labor force may be more likely to understate their health status so as to justify their lack of employment. This effect is believed to be more prevalent among males than females.

Consistent with this view, research has found that men "perceive the state of unemployment as more stigmatic" (Kulik 2000a, p. 487).[1] However, the dearth of information relating mental health status to labor force participation[2] and the difficulties in quantifying the relative importance of physical health versus mental health in effecting labor force participation suggest that linkages between health and labor force participation may not be firmly established. Empirical evidence on these issues would be quite useful to self-insuring employers who are designing health insurance benefits packages, to employer-based insurance in general, and to policymakers concerned with understanding the full social benefits of alternative health insurance plans.

In this paper we apply propensity scoring methods to analyze treatment effects among labor force participation, physical, and mental health status. We utilize a unique set of data, the Medical Expenditure Panel Survey (MEPS). We use the MEPS to relate physical and mental health status to labor force participation and we compare these relationships among self-report and proxy respondents. The MEPS is well-suited to this purpose because it includes both

[1]See also Kulik (2000b) for further evidence on gender differences in attitudes toward unemployment.

measures of physical and mental status as well as self-reports and proxy responses to questions regarding physical and mental health. In the MEPS survey, members of a subject's household responded to questions regarding the subject's physical and mental health status when the subject was unavailable to answer the surveyor's questions. Separate analyses are conducted for males and females.

Methodologically, treatment effects have been a topic of interest in the literature for some time although recent interest appears to center on issues of robustness. Robins (1999) has suggested a class of new class of non-nested marginal structural models that may be used to estimate the causal effects of time dependent treatment on a binary outcome while Hernan et al. (2001) has applied this method in a model of joint outcomes. Since it was first introduced by Rosenbaum and Rubin (1983), propensity scoring, which examines the conditional probability of receiving treatment given pre-treatment variables, has been motivated by its non-parametric methodology. It eliminates bias due to an assumed (and incorrect) parametric relationship between the outcome and the observed covariates whose parametric relationship with the outcome variable is unknown. Imbens (1999) extended this methodology to allow for estimation of average causal effects with multi-valued treatments.

The paper is structured as follows. In the next section we discuss in more detail the empirical evidence that links labor force participation and health. Section 3 provides theoretical treatments and the rationale behind propensity scoring methods and propensity score binning in particular. Section 4 discusses the results of our propensity scoring analyses and their implications. Section 5 concludes.

2. Evidence from the Medical Expenditure Panel Survey

This study uses data from the years 2000-2005 Household Component Consolidated

[2]We use "health" to refer to physical health and "mental health" otherwise.

Data file from the Medical Expenditure Panel Survey (MEPS). This database, cosponsored by the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS), provides nationally representative estimates of medical treatments, health care expenditures, health status, labor force participation, and sociodemographic characteristics for the civilian, non-institutionalized population in the United States. Our sample includes subjects aged 18 to 65. Students were excluded from the sample[3]. A total of 26,757 males and 11,817 female subjects had physical and mental health status reported by a proxy (someone in that person's household). Another 16,774 males and 38,978 females self-reported their physical and mental health status.

Table 1 shows respondents' answers to questions about their physical and mental health according to whether they are in or out of the work force. Separate summaries of responses are reported by gender and respondent status (e.g., self-reporter or proxy). Turning first to the males, Table 1 indicates that among subjects in the labor force, there is close agreement between proxies and self-reporters in terms of health status. Thus, 92 (respectively 91) percent of male (respectively female) self-reporters in the work force indicate that they are in good physical health. "Good "physical health is defined as a binary variable equal to 1 if the subject indicates that he or she is in good, very good, or excellent health and 0 if in poor or fair health. The same procedure is used to define good mental health. Proxy respondents indicate that 93 percent of both males and females in the workforce are in good physical health. Proxy and self-reporters' assessments of mental health status are nearly identical for males and females in the work force. MEPS has a complex survey design involving stratification and clustering. We have used the weights in MEPS to deal with these issues.

(INSERT TABLE 1)

[3]The MEPS sample was chosen as a nationally representative subsample of the ongoing National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics, and may be linked to the NHIS database as well. The MEPS survey respondents were interviewed in person. The survey achieved a response rate of 77.7 percent (see Cohen, Monheit, Beauregard, et al., 1996 for further details).

In contrast, among subjects out of the work force, substantial differences emerge between proxy and self-reporters for physical health. In particular, while 61 percent of out-of-work males indicate that they are in good physical health, the corresponding figure for proxies is 71 percent. A similar though less pronounced pattern occurs for females. Proxy and self reports of mental health are similar for out-of-work subjects.

Collectively, this pattern seems to suggest that out-of-work male and female self-reporters may be understating their physical health status and that this tendency is stronger among the males. These discrepancies are all the more striking since proxy and self-reports are found to be remarkably similar among subjects who were in the labor force.

While the patterns observed in Table 1 are suggestive, differences in the characteristics of subjects whose health status is self or proxy reported may confound these results. In the next section we attempt to isolate the linkages between health conditions and labor force participation utilizing robust nonparametric procedures. Since there is no evidence that proxy and self-reports have been randomly assigned, the problem of sample selection of subjects may be an issue. Propensity scoring methods were employed to deal with the selection bias that arises in such a situation. The methods are introduced below.

3. Propensity Scoring

Propensity scoring methods were first introduced by Rosenbaum and Rubin (1983) as a way to significantly reduce bias in observational studies. Early applications found in the biometrics literature analyze medical treatment effects. Propensity scoring techniques have been found to be an efficient alternative to most common econometric bias-reducing techniques (e.g. Heckman's two-step estimation procedure). The objective is to statistically evaluate the effect of a particular treatment on a population (e.g. the effect of smoking on mortality). A randomized experiment with human subjects would be not only impractical but also unethical. Observational data are typically used in such cases to study causal effects.

The main problem in observational studies is that selection for treatment is not randomized. Therefore, the treated and the non-treated may differ in characteristics other than treatment intake. In the present case, subjects are not randomly assigned to be in or out of the labor force. These cohorts may differ from each other in a variety of respects besides subjective health status. Hence, to isolate the effects of self-reported health status on labor force participation it is necessary to condition on these factors.

The propensity scoring approach requires one to estimate the probability that each subject receives the treatment, whether or not the subject actually does. At first blush, this may not seem reasonable since we already know which subjects were assigned to a particular treatment. However, if we use the probability that a subject would have been treated given his values of the covariates (the propensity score) in order to adjust our estimates of the treatment effect, we can create a ”quasi-randomized” experiment. In other words, two subjects (regardless of their treatment intake) who have similar propensity scores can be thought of having the same *ex ante* probability of being assigned the treatment or among subjects with similar propensity scores. One can imagine that subjects were ”randomly” assigned the treatment in the sense of being equally likely to be treated or control.

To implement the propensity score, we matched subjects on the following characteristics: age, education, race, marital status, health insurance status, and geographic location. We also conditioned on two indices of physical and mental health provided in MEPS. The physical health index is known as the Physical Component Summary Scale Score (PCS) score and the mental health index is the Mental Component Summary Scale Score (MCS). These indices are designed to provide comprehensive summary measures of mental and physical health status based on the ability to perform a number of specific daily activities as well as feelings and vitality. Because they are based on the ability to perform specific activities and measure specific mental-health related attitudes, these provide more objective measures of

physical and mental health the subjective health assessments. These scales have been described in more detail elsewhere (Jenkinson et al. 1997; Ware et al. 1994). To gauge the importance of conditioning on objective physical and mental health status, we estimate treatment effects using propensity score models with and without these health indices.

4. Propensity Scoring and the Relationship between Health and Labor Force Participation

Table 2 provides descriptive statistics for the variables used in this study, by respondent status (proxy or self-reporter) and gender. Self-reporters and proxy respondents are quite similar in terms of most observable characteristics. The main exception is marital status. For both males and females, substantially lower shares of self-reporters are married. This likely reflects that fact that unmarried subjects may have more difficulty obtaining a proxy respondent from their household. Self-reporters have slightly lower physical and mental health status measures.

(INSERT TABLE 2)

In this section we utilize propensity scoring techniques to analyze the relationship between health and labor force participation. Propensity score binning is applied to subjects whose health conditions have been proxy-reported and to self-reporters, separately.

4.1. No Matching on Health Indices

We first analyze the effect of physical and mental health (the treatments) on labor force participation *without matching on the physical and mental health indices described above.* We apply propensity score binning to the proxy-reported subjects and to the self-reporters separately. We estimate the effect of physical and mental health on labor force participation using the sample frequencies within each bin as weights. This is a robust approach that mitigates sensitivity to outliers.

The main results are presented in Table 3. Turning first to the subgroup of male subjects whose health status was reported by a proxy, the marginal increment in the probability of being in the labor force due to good health condition is 18%. In other words, being in good health increases one's chances of being in the labor force by 18%. This marginal increment is greater for self-reporting males – 24.6%. Moreover, this difference is significant at the 1% level. The association between mental health status and labor force participation is also higher for self-reporting males than for proxies, but this differences is not statistically significant.

(INSERT TABLE 3)

For females, we find no difference in the relationship between physical health status and labor force participation for self-reporters and proxy respondents. In contrast, we find a larger effect of mental health on labor force participation among proxy responders (29.4%) than among self-reporters (22.2%), a difference that is statistically significant at the 1% level.

4.2. Matching on Health Indices

We repeat the propensity score approach, now adding the health and mental health indices described above to our matching procedure. The results are summarized in Table 4. As the Table indicates, the magnitudes of the treatment effects decline and in most case the differences between self-reporters and proxies decline as well. Most importantly, we now find that none of the differences between self-reporters and proxy respondents is statistically significant.

(INSERT TABLE 4)

The results suggest that any differences between proxies and respondents in the effects of perceived physical and mental health on labor force participation become insignificant once subjects are matched on more objective health and mental health indices. This, in turn, implies that the apparent reporting biases in the estimates that did not control for these health indices in

fact reflect actual differences in health.

4.3 Limitations

A limitation of this study is that our subjective health and mental health measures may be endogenous. In particular, our estimates which did not control for objective measures of health and physical health status may suffer form endogeneity bias. While we have more confidence in the estimates that adjust for the objective health and mental health indices, it is possible that this matching strategy, too, has missed an important yet unobservable factor affecting perceived health. Our findings must be viewed with this limitation in mind.

5. Conclusion

This paper relates subjective measures of physical and mental health status to labor force participation and compares these relationships among self-reporter and proxy respondents. We used propensity scoring methods to help control for confounding factors.

We do not find significant differences in the associations between these measures of health and labor force participation, once treatments and controls are matched on more objective measures of health and mental health. Our results suggest that the labor force participation costs of poor physical health are not seriously overstated when based on estimates using self-reports of health. Such measures, which are much more readily available in health economic databases than detailed objective health indices, thus may be useful in understanding the labor market consequences of health and mental health. From a policy perspective, estimates of health and labor force participation based on subjective health measures do not appear to significantly overstate employer labor force participation costs associated with poor health in the workforce due to response bias.

References

[1] Abadie A., Observational studies I: Matching and Regression, Lecture notes of Spring 2001, handout #5. (Unpublished).

[2] Behrman, J., Sickles, R. C., Taubman, P. (1998), *The Causes, Correlates and Consequences of Death Among Older Adults: Some Methodological Approaches and Substantive Analysis*, Boston: Kluwer Academic, 1-186.

[3] Behrman, J., Sickles, R. C., Taubman, P., and Yazbeck, A. (1992), "Black-White Differentials in Mortality: Results from the Retirement History Survey," *Journal of Econometrics*, 50, 183-204.

[4] Bound, J. (1991), "Self-reported versus Objective Measures of Health in Retirement Models," *Journal of Human Resources* 26(1), 106-138.

[5] Cochran, W. G. (1968), " The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies," *Biometrics*, 26, 295-313.

[6] Cohen J, Monheit A, Beauregard K, et al. (1996). The Medical Expenditure Panel Survey: a national health information resource. Inquiry 33: 373-389.

[7] D.Agostino, R. B. Jr. (1998), "Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group," *Statistics in Medicine*, 17, 2265-2281.

[8] Hernan, M. A., Brumback, B., and Robins, J. (2001), "Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments," *Journal of American Statistical Association*, 96, 440-447.

[9] Imbens, G. W. (1999), " The Role of Propensity Score in Estimating Dose Response Functions," NBER Working Paper.

[10] Jenkinson, C., Layte, R., Jenkinson, D., et al. (1997). "A Shorter Form Health survey: can the SF-12 Replicate Results from the SF-36 in Longitudinal Studies?" *Journal of Public Health*

Medicine, 19, 179-286.

[11] Kulik, L. (2000a), "Jobless Men and Women: A Comparative Analysis of Job Search Intensity, Attitudes Toward Unemployment, and Related Responses," *Journal of Occupational and Organizational Psychology*, 73, 487-500.

[12] Kulik, L. (2000b), "The Impact of Gender and Age on Attitudes and Reactions to Unemployment: The Israeli Case," *Sex Roles*: *A Journal of Research*, 43, 85-104

[13] Medical Expenditure Panel Survey (MEPS); Household Component Full Year Files (2000-2005). Agency for Healthcare Research and Quality, Rockville, MD.

[14] Robins, J.M. (1999), "Marginal Structural Models Versus Structural Nested Models as Tools for Causal Inference," *Statistical Models in Epidemiology: The Environment and Clinical Trials*, eds. M. E. Halloran and D. Berry, New York: Springer-Verlag, pp. 95-134.

[15] Rosenbaum, P. R., and Rubin, D. B. (1983), " The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41-55.

[16] Sickles, R. C., and Taubman, P. (1986), "An Analysis of the Health and Retirement Status of the Elderly," *Econometrica*, 54, 1339-1356.

[17] Sickles, R. C., and Taubman, P. (1997), "Mortality and Morbidity Among Adults and the Elderly," Chapter 16 of the *Handbook of Population and Family Economics*, edited by Mark R. Rosenzweig and Oded Stark, Amsterdam: North-Holland, 559-643.

[19] Ware, J., Kosinski, M., Keller, S.D. (1994). SF-36 Physical and Mental Health Summary Scales: A User Manual. Boston, MA: New England Medical Center.

Table 1: Subjective Health and Mental Health, by Respondent and Labor Force Participation Status

MALES

Health Measure	**In Labor Force Proxy N=22,943**	**In Labor Force Self N=13,647**	**Not In Labor Force Proxy N=3,814**	**Not In Labor Force Self N=3,127**
Good Physical Health[1]	0.93	0.92	0.71	0.61
Good Mental Health[1]	0.97	0.96	0.82	0.80

FEMALES

Health Measure	**In Labor Force Proxy N=8,104**	**In Labor Force Self N=26,409**	**Not In Labor Force Proxy N=3,713**	**Not In Labor Force Self N=12,570**
Good Physical Health[1]	0.93	0.91	0.79	0.74
Good Mental Health[1]	0.97	0.96	0.87	0.86

1. "Good "physical health is defined as a binary variable equal to 1 if the subject indicates that he or she is in good, very good, or excellent health and 0 if in poor or fair health. The same procedure is used to define good mental health.

Table 2. Descriptive Statistics for Matching Variables, by Reporting Status and Gender

VARIABLE NAME	MALE SELF N=16,774	MALE PROXY N=26,757	FEMALE SELF N=38,979	FEMALE PROXY N=11,817
DEPENDENT VARIABLES				
Labor force participation rate	0.84 (0.36)	0.87 (0.33)	0.74 (0.44)	0.73 (0.44)
Subject is in good physical health	0.87 (0.34)	0.90 (0.30)	0.86 (0.34)	0.89 (0.31)
Subject is in good mental health	0.94 (0.24)	0.95 (0.21)	0.93 (0.25)	0.94 (0.23)
MATCHING VARIABLES				
Age in years	42.13 (12.08)	41.81 (11.96)	42.34 (11.65)	41.90 (12.83)
Education in years	13.45 (2.79)	12.82 (2.93)	13.24 (2.73)	12.95 (3.02)
Race				
Hispanic	0.12 (0.32)	0.15 (0.35)	0.11 (0.32)	0.14 (0.35)
Black	0.11 (0.32)	0.09 (0.29)	0.13 (0.33)	0.11 (0.31)
Other non white	0.06 (0.24)	0.05 90.21)	0.05 (0.21)	0.10 (0.30)
Caucasian	0.71 (XXX)	0.70 (XXX)	0.71 (XXX)	0.65 (XXX)
Subject is married	0.41 (0.49)	0.77 (0.42)	0.58 (0.49)	0.72 (0.45)
Subject lacks health insurance	0.18 (0.39)	0.18 (0.38)	0.13 (0.33)	0.15 (0.35)
PCS index: physical health	50.99 (9.62)	51.39 (8.69)	50.01 (10.10)	50.29 (9.38)
MCS index: mental health	51.22 (9.32)	52.19 (8.80)	49.60 (10.05)	50.48 (9.66)
Census Region				
Northeast	0.18 (0.38)	0.19 (0.39)	0.19 (0.39)	0.19 (0.39)
Midwest	0.23 (0.42)	0.23 (0.42)	0.24 (0.42)	0.21 (0.40)
South	0.35 (0.48)	0.36 (0.48)	0.36 (0.48)	0.37 (0.48)
West	(0.25 (0.43)	0.22 (0.41)	0.22 (0.41)	0.24 (0.43)
Subject lives in urban location	0.85 (0.36)	0.80 (0.40)	0.82 (0.39)	0.84 (0.36)

Table 3. Average treatment effect of Physical and Mental Health on Labor Force Participation
MEPS Years 2000-2005

Not using health indices for PS matching			ATT	Std. Err.
Males	Physical H	Self	**0.246**	0.012
		Proxy	**0.180**	0.009
	Mental H	Self	**0.323**	0.020
		Proxy	**0.273**	0.017
Females	Physical H	Self	**0.188**	0.007
		Proxy	**0.188**	0.014
	Mental H	Self	**0.222**	0.011
		Proxy	**0.294**	0.022

Table 4. Average treatment effect of Physical and Mental Health on Labor Force Participation
MEPS Years 2000-2005

Using health indices for PS matching			ATT	Std. Err.
Males	Physical H	Self	**0.070**	0.018
		Proxy	**0.054**	0.011
	Mental H	Self	**0.160**	0.034
		Proxy	**0.133**	0.018
Females	Physical H	Self	**0.056**	0.018
		Proxy	**0.051**	0.022
	Mental H	Self	**0.100**	0.019
		Proxy	**0.163**	0.042