PERSPECTIVES

PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES

Revisiting Internet Use and Depression Among the Elderly

George S. Ford, PhD*

June 7, 2013

Introduction

Four years ago in a paper entitled *Internet Use and Depression Among the Elderly*, my colleague and I presented evidence indicating that Internet use reduces symptoms of depression among older Americans (55 or older).[1] This showing was based on the sample of over 7,000 retired persons taken from the Health and Retirement Study ("HRS") conducted at the University of Michigan. Applying a variety of advanced empirical techniques, we found that Internet use is associated with a 20% reduction in depression classification based on a commonly used measure of depressive symptoms (the eight-item CESD score).

Given the importance of this issue, in this PERSPECTIVE I present the results of an update (done with Shelia Cotton, Sherry Ford, and Tim Hale) to the earlier analysis that uses additional data and more sophisticated econometric techniques.[2] Specifically, while the data we used in our first paper was large, the HRS is nonetheless a longitudinal survey. Since depression, at least in some forms, is often a recurrent condition that may persist over a number of years, we decided to exploit the time-series component of the HRS to study the role of past depression and Internet use on current depression in older adults. To do so, we use a dynamic probit model, which in some cases is enhanced by matching algorithms.

This updated analysis of the data again reveals that Internet use reduces depression categorization and does so by 34%, a reduction similar to, but slightly larger than, that found in our earlier research (20-28%). Thus, consistent with our earlier findings, as depression is estimated to cost the United States about $100 billion annually, the data indicate that policies which expand Internet use among the elderly may have significant economic payoffs.

... Internet use reduces depression categorization and does so by 34%, a reduction similar to, but slightly larger than, that found in our earlier research (20-28%). Thus, consistent with our earlier findings, as depression is estimated to cost the United States about $100 billion annually, the data indicate that policies which expand Internet use among the elderly may have significant economic payoffs.

Data

As before, the analysis of Internet use and depression employs the HRS, which is a longitudinal household survey of more than 22,000 persons over the age of 50 every two

years. For this study, we employed four waves of biennial data covering the years 2002 through 2008. After limiting the analysis to a balanced sample of retired persons and those not living in a nursing home, the final sample included 3,075 observations in each year, or 12,300 observations. As before, depression is measured by the eight-item version of the commonly-used Center for Epidemiologic Studies ("CES-D") scale calculated as the sum of dysphoric (yes/no) responses to eight questions, and we again followed standard practice by creating a dummy dependent variable with a value of 1 when the CES-D ≥ 4. Depression categorization across the sample periods was found to be relatively stable (2002, 12.4%; 2004, 11.9%; 2006, 13.0%; 2008, 13.3%). The sample mean of depression was 12.6%, with Internet users having a mean of 8.2% and non-users of 14.6%. Notably, roughly half (47%) of those categorized as depressed in the current wave were also depressed in the preceding wave.

Internet use is based on a direct question asking participants, "Do you regularly use the World Wide Web, or the Internet, for sending and receiving e-mail or for any other purpose…?" The response is dichotomous (1=Yes, 0=No). Internet use was relatively stable and persistent over the four waves (30.5%, 32.0%, 32.2%, 31.5%), with 86% of users in a current wave also being users in the preceding wave.

Other explanatory variables include age (in years); gender (1 = male); race (1 = White); and education (1 = more than a high school education); marital status (1 = married and living with spouse); physical activity (1 = active at least once weekly); health (1 = has a debilitating physical health condition); household size (1 = four or more persons in the home); family size (1 = four or more immediate family members living); and seasonal affective disorder (1 = survey taken in months of November, December, or January). Dummy variables for each wave (except the final wave) are also included.

Results

We use a dynamic probit model to estimate the effect of Internet use on depression categorization, permitting current depression categorization to depend on a past depression categorization. More specifically, we use the Conditional Maximum Likelihood Estimator ("CMLE") crafted by Wooldridge (2005).[3]

Table 1. Summary

	Coeff.
Internet Use	-0.314**
Depression Lagged	0.311**
Depression First Period	1.080**
Age	-0.011**
Married	-0.360**
Educ. Years	-0.212**
Active	-0.214**
Health	0.548**
White	-0.020
Male	-0.150**
November	0.437**
December	0.161
January	0.248
Year 2006	0.070
Year 2008	0.063
Constant	-0.550*
ρ	0.34

Significance: * 5%, ** 10%

Given the use of lagged values and a balanced sample, the CMLE is estimated with 9,225 observations. Estimated coefficients are summarized in Table 1. This summary shows that depression categorization is less common among married persons, males, active adults, and more educated persons, and also declines with age. Having poor health has a strong positive effect on the likelihood of being depressed. Some evidence is found to support Seasonal Affective Disorder. Race, household

size, and living relatives do not appear to be important determinants of depression categorization.

The coefficient on the lag of the dependent variable (depression in the previous wave) was estimated to be 0.316 with a t-statistic of 3.50 ($p < 0.01$), indicating the presence of state dependence—i.e., current depression depends on past depression. An average partial effect was estimated as the change in the predicted probability of depression categorization due to presence or absence of depression in the prior period. Those categorized as depressed in the prior period had an average probability of depression in the current period of 0.127, whereas those not categorized as depressed in the prior period have a predicted depression rate of only 0.084. Thus, depression in the previous period increases the probability of depression in the current period by 50% (consistent with the sample means). This marginal effect has a bootstrapped t-statistic of $t_b = 8.34$ ($p < 0.01$). The initial depressive state is likewise a significant determinant of current depression (1.08, $t = 10.19$, $p < 0.01$); a person categorized as depressed in the initial wave of the sample is nearly five times more likely to be depressed in later waves.

Turning to Internet use, the coefficient on Internet use is -0.33 with a t-statistic of -2.61 ($p < 0.01$). Internet users are less likely to be categorized as depressed than are non users. Other things constant, Internet users have an average predicted probability of depression categorization of 0.068 whereas that probability for non-users is 0.104. Based on the difference, Internet use leads to a 35% reduction in the probability of depression categorization. This marginal effect has a bootstrapped t-statistic (50 simulations) of $t_b = 3.14$ ($p < 0.01$).

Coarsened Exact Matching

Internet use is a choice and is not randomly assigned. Not surprisingly, there are some significant differences between users and non-users, particularly with respect to age and education level, with Internet use being much higher for younger and more educated respondents. To address such differences, we employed Coarsened Exact Matching ("CEM") to ensure comparable covariate distributions across Internet users and non-users.[4] The matching algorithm is applied to the 2008 cross section and the weights then distributed across all waves. The balanced, weighted estimation sample was 7,278 observations (2,426 for each wave).

Matching had little effect. The coefficient on Internet use is -0.311 ($t = -2.37$, $p < 0.05$), for a percentage reduction in depression categorization of 32%. The effect of lagged depression is also similar (0.33, $t = 3.06$, $p < 0.01$), indicating a 52% increase in the probability of current depression categorization given depression in the previous state ($t = 3.06$, $p < 0.01$).

Unconfoundedness

A core assumption required for the estimation of causal effects is a concept sometimes referred to as unconfoundedness.[5] The unconfoundedness assumption states that there are no unobserved covariates affecting both Internet use and depression that would lead to spurious correlation. For example, say that Internet use has no effect on depression, but both Internet use and depression are influenced by a third factor unaccounted for in the model. A statistical model may detect a relationship between Internet use and depression, but this finding mistakenly attributes to Internet use what is really caused by some other and unmeasured factor. While there is no direct test of unconfoundedness, there are indirect means by which to assess its presence. Obviously, we could include the missing variable, but exactly

what is "missing" is not always known or measurable. Alternately, another method is to construct a type of test which evaluates the effect of a future treatment on a current outcome (i.e., determined prior to the treatment).[6]

Our approach is outlined as follows. First, we took the data from 2004 and excluded all Internet users from that period. Then, for these non-users, we assigned a pseudo-treatment for Internet use from actual use in 2006. Presumably, those non-users in 2004 that become users in 2006 are those most likely to have high values of the potential unobserved factor. With this data, we ran a single-period probit model to test the null hypothesis that Internet use has no effect; a null hypothesis we know to be true. We conducted the same analysis across a number of different year pairings. In none of the scenarios did we find this pseudo-Internet use variable to influence depression. Notably, for each year, actual Internet use reduced depression.

... consistent with our earlier findings, as depression is estimated to cost the United States about $100 billion annually, the data indicate that policies which expand Internet use among the elderly may have significant economic payoffs. These benefits can be part of a full cost-benefit analysis assessing the wisdom of committing public resources to such efforts.

While we cannot claim that these findings represent a formal statistical test of the unconfoundedness assumption, the results add credibility to the causal nature of our findings on Internet use and depression.

Conclusion

In my original paper, *Internet Use and Depression Among the Elderly*, we found that Internet use reduced depressive symptoms in older Americans. More recently, we exploited the longitudinal nature of the HRS data to quantify the relationship in a more sophisticated econometric model. Our dynamic probit model indicates that current depression depends on past depression. About half of those depressed in a current wave of the data were also depressed in the preceding wave. Importantly, Internet use was found to reduce depression categorization by 34%. This estimated reduction in depression categorization is comparable to our earlier findings. Thus, consistent with our earlier findings, as depression is estimated to cost the United States about $100 billion annually, the data indicate that policies which expand Internet use among the elderly may have significant economic payoffs. These benefits can be part of a full cost-benefit analysis assessing the wisdom of committing public resources to such efforts.

NOTES:

[*] **Dr. George Ford** ***is Chief Economist of the Phoenix Center for Advanced Legal and Economic Public Policy Studies. The views expressed in this*** **PERSPECTIVE** ***do not represent the views of the Phoenix Center or its staff.***

[1] G.S. Ford and S.G. Ford, *Internet Use and Depression Among the Elderly*, PHOENIX CENTER POLICY PAPER NO. 38 (October 2009) (available at: http://www.phoenix-center.org/pcpp/PCPP38Final.pdf) *and reprinted as* 28 COMPUTERS IN HUMAN BEHAVIOR 496 (2012) (available at: http://www.sciencedirect.com/science/article/pii/S074756321100238X).

[2] S. Cotton, G. Ford, S. Ford, and T. Hale, *Internet Use and Depression among Older Adults: A Longitudinal Analysis*, Unpublished Manuscript (2012) (draft available upon request).

[3] J. Wooldridge, *Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity*, 20 JOURNAL OF APPLIED ECONOMETRICS 39-54 (2005).

[4] Iacus, King & Porro (2012) describe several advantages of CEM over more traditional propensity score matching algorithms. S. M. Iacus, G. King, and G. Porro, *Causal Inference without Balance Checking: Coarsened Exact Matching*, 20 POLITICAL ANALYSIS 1-24 (2012).

[5] J. Pearl, *Causal Inference in Statistics: An Overview*, 30 STATISTICS SURVEY 396-146 (2009); G. W. Imbens and J. M. Wooldridge, *Recent Developments in the Econometrics of Program Evaluation*, 47 JOURNAL OF ECONOMIC LITERATURE 5-86 (2009).

[6] Imbens & Wooldridge, *id*. at p. 48