PHOENIX CENTER POLICY PAPER SERIES

Phoenix Center Policy Paper Number 38:

Internet Use and Depression Among the Elderly

George S. Ford, PhD
Sherry G. Ford, PhD

(October 2009)

© Phoenix Center for Advanced Legal and Economic Public Policy Studies, Sherry G. Ford and George S. Ford (2009).

Phoenix Center Policy Paper No. 38
Internet Use and Depression Among the Elderly

George S. Ford, PhD†
Sherry G. Ford, PhD‡

(© Phoenix Center for Advanced Legal & Economic Public Policy Studies, Sherry G. Ford and George S. Ford (2009).)

Abstract: The American Recovery and Reinvestment Act of 2009 directs over $7 billion to expand broadband Internet availability and adoption in the United States. One target of such funding is the elderly population, a group of Americans for which broadband adoption is relatively low. An interesting question is what benefits do such efforts afford? We employ a dataset of over 7,000 elderly retired persons to evaluate the role of Internet use on mental well-being. Well-being is measured using the eight-point depression scale developed by the Center for Epidemiologic Studies (CES-D). Empirical techniques include single equation regression, instrumental variables and propensity score methods. All procedures indicate a positive contribution of Internet use to mental well-being of elderly Americans, and estimates indicate that Internet use leads to about a 20% reduction in depression classification. As depression is estimated to cost the United States about $100 billion annually, expanding Internet use among the elderly may have significant economic payoffs.

---

† Chief Economist, Phoenix Center for Advanced Legal & Economic Public Policy Studies. The views expressed in this paper are the authors' alone and do not represent the views of the Phoenix Center, its Adjunct Fellows, or any of its individual Editorial Advisory Board members. We are indebted to Professor Randy Beard, Phoenix Center Senior Fellow, and Professor John Jackson for their comments and suggestions.

‡ Visiting Scholar, Phoenix Center for Advanced Legal & Economic Public Policy Studies; Associate Professor, Director of Honors Program, University of Montevallo, Alabama.

TABLE OF CONTENTS:

I. Introduction

With the enactment of the American Recovery and Reinvestment Act of 2009 ("ARRA"),[1] the United States has made a renewed commitment to ensuring that all Americans have access to affordable broadband services. While there are certainly detractors who argue that these efforts are all folly, we present in this paper one empirical example of the societal benefits of broadband: improving the mental health of the elderly. This is not a subject to be taken lightly, as late-life depression affects about six million Americans age 65 and older.[2] Depression is estimated to cost the country about $100 billion annually in direct medical

[1] American Recovery and Reinvestment Act of 2009 at Title VI—Broadband Technology Opportunities Program, (Public Law 111-5, 2009)(available at: http://www.gpo.gov/fdsys/pkg/PLAW-111publ5/pdf/PLAW-111publ5.pdf).

[2] *See* WebMd, *Depression in the Elderly* (available at: http://www.webmd.com/depression/guide/depression-elderly, visited August 18, 2009).

costs (31%), increased suicide mortality (7%), and workplace costs (62%).[3] Curbing depression offers sizeable economic returns to society.

With an eye on measuring causal relationships, we show below that applying a variety of econometric and statistical analyses to a large sample of retired Americans (age 55 or older) reveals that Internet use contributes positively to mental well-being, and estimates indicate that Internet use leads to about a 20% reduction in depression classification. Our findings therefore suggest that the development of demand-side broadband access and education programs for the elderly may produce significant societal benefits. These benefits can then be measured against the projected costs of such programs to determine a part of the expected net payoff of public policies aimed at improving Internet use by the elderly.

The paper is outlines as follows. First, we provide a literature review on the role of Internet use on mental health. Second, we describe the data. Third, we conduct a variety of empirical tests in an effort to quantify the causal relationship between Internet use and the mental health of the elderly. Fourth, we provide a brief summary of the results of our analysis. Conclusions are provided last.

II. Background

The seemingly ubiquitous nature of the Internet suggests that access and use are widespread; however, studies continue to indicate that certain segments of the U.S. population lack availability or have yet to adopt. One such population is the elderly, which is consistently shown to have lower subscription rates for Internet services, both domestically and internationally. A recent survey of U.S. households indicates that, while 79% of the general population reports using the Internet, only 42% of those 65 and older indicate use.[4] Empirical studies consistently show that Internet subscription rates are lower in states and countries with older populations.[5] In some respects, a lower level of subscription

---

[3] *See, e.g.*, P. E. Greenberg, S. Leongand, and H. Birnbaum, *The Economic Burden of Depression in the United States: How Did It Change Between 1990 and 2000?*, 64 JOURNAL OF CLINICAL PSYCHIATRY 1465-1475 (2003). We adjust their 2000 estimate to 2009 dollars using the Consumer Price Index (June 2000 to June 2009).

[4] Pew Internet and American Life Project 2009. *Demographics of Internet users, July 15, 2009*, http://www.pewinternet.org/Static-Pages/Trend-Data/Whos-Online.aspx (accessed on July 31, 2009).

[5] *See, e.g.*, G. Ford, T. Koutsky and L. Spiwak, *The Demographic and Economic Drivers of Broadband Adoption in the United States*, PHOENIX CENTER POLICY PAPER NO. 31 (February 2007)(available at: http://www.phoenix-center.org/pcpp/PCPP31Final.pdf); G. Ford, T. Koutsky and L. Spiwak *The Broadband Efficiency Index: What Really Drives Broadband Adoption Across the*

(Footnote Continued. . . .)

is not that surprising since the computer and the Internet are relatively modern technologies. Moreover, many factors associated with aging, such as changes in physical abilities like eyesight and hand movement, can prove prohibitive for elderly access.[6] Still, expanding Internet adoption by the elderly is a policy-relevant topic. The ARRA included a $7.2 billion provision for programs aimed at the expansion of broadband availability, adoption, and use.[7] Expanding Internet use among the elderly is a key target of such funds.

What are the expected benefits from expanded Internet use for the elderly or other social groups presently under-represented in the online community? Improvements in health and healthcare are a commonly referenced benefit of Internet use. One frequently studied area of health impact focuses on the effects of Internet use on mental well-being. Depression is a significant health and economic concern. Estimates of its costs to society are about $100 billion annually in healthcare costs and lost wages and productivity.[8] While in many policy debates the Internet is typically viewed as an unqualified good, the scientific research presents a more sobering view on how the Internet promotes (or detracts from) mental well-being. Indeed, there are conflicting hypotheses on the impact of Internet use on mental health.

For example, one common hypothesis is that Internet use facilitates interpersonal communication and, as such, should improve mental well-being, for instance, by reducing loneliness. Alternately, some hypothesize that Internet use leads to social exclusion, thereby reducing well-being. In an early study, Kraut et al. (1998) found that Internet use was associated with (i) declines in participants' communications with family members in the household; (ii) declines in the size of their social circle; and (iii) increases in participant depression and loneliness.[9] A follow-up study (published in 2002) by some of the same authors using the same sample evaluated later in time, in addition to

---

OECD? PHOENIX CENTER POLICY PAPER NO. 33 (May 2008)(available at: http://www.phoenix-center.org/pcpp/PCPP33Final.pdf).

6 K. O'Hara, *"Curb Cuts" on the Information Highway: Older Adults and the Internet,* 13 TECHNICAL COMMUNICATION QUARTERLY 423-45 (2004).

7 *Supra* n. 2.

8 *Supra* n. 4.

9 R. Kraut, M. Patterson, V. Lundmark, S. Kiesler, T. Mukhopadhyay, and W. Scherlis, *Internet Paradox: A Social Technology That Reduces Social Involvement and Psychological Well-Being?,* 53 AMERICAN PSYCHOLOGIST 1017-1032 (1998).

other data, partially negated these findings.[10] However, McKenna and Bargh (2000) found supporting evidence for the isolation-depression outcome from Internet use.[11] In a later study by Morgan et al. (2003), the authors find that for a sample of college students Internet use in the form of communications services, such as Instant Messaging ("IM"), improves mental well-being, whereas surfing and gaming led to greater depression symptoms.[12] A similar result was reported in Selfhout et al. (2009).[13] Yet, one study found that excessive Internet use was a key factor behind academic failure by college students,[14] and Scherer (1997) finds that synchronous communications (such as IM) leads to Internet dependency.[15] The results on the benefits of Internet use are decidedly mixed on this question and many other questions as well.

Most of the research on Internet use and mental health has focused on teens and college students. The effect of Internet use on the elderly is addressed in relatively few studies. Included in these are such topics as health effects (Campbell and Wabby 2003)[16], psychological well-being and social support (Mellor et al. 2008[17]; Wright 2000[18]), as well as loneliness and social isolation (Sum

---

[10] R. Kraut, S. Kiesler, B. Boneva, J. Cummings, V. Helgeson, and A. Crawford, *Internet Paradox Revisited,* 58 JOURNAL OF SOCIAL ISSUES 49-74 (2002). Furthermore, the results of the 1998 study could not be replicated. J. Katz, and R. Rice, SOCIAL CONSEQUENCES OF INTERNET USE: ACCESS, INVOLVEMENT, AND INTERACTION (2002) at 221.

[11] K. McKenna and J. Bargh, *Plan 9 from Cyberspace: The Implications of the Internet for Personality*, 4 PERSONALITY AND SOCIAL PSYCHOLOGY REVIEW 57–75 (2000).

[12] C. Morgan, and S. Cotten, *The Relationship Between Internet Activities and Depressive Symptoms in a Sample of College Freshman*, 6 CYBERPSYCHOLOGY & BEHAVIOR 133-142 (2003).

[13] M. Selfhout, S. Branje, M. Delsing, T. ter Bogt, and W. Meeus, *Different Types of Internet Use, Depression, and Social Anxiety: The Role of Perceived Friendship Quality*, 32 JOURNAL OF ADOLESCENCE 819-833 (2009).

[14] On Line (April 26, 1996), CHRONICLE OF HIGHER EDUCATION, 42(33), A21; *see also* R. Kubey, M. Lavin, and J. Barrows, *Internet Use and Collegiate Academic Performance Decrements: Early findings*, 51 JOURNAL OF COMMUNICATION 366-382 (2001).

[15] K. Scherer, *College Life Online: Healthy and Unhealthy Internet Use*, 38 Journal OF COLLEGE STUDENT DEVELOPMENT 655–665 (1997).

[16] R. Campbell and J. Wabby, *The Elderly and the Internet: A Case Study,* 3 INTERNET JOURNAL OF HEALTH 2-18 (2003).

[17] D. Mellor, L. Firth, and K. Moore, *Can the Internet Improve Well-Being of the Elderly?* 32 AGEING INTERNATIONAL 25-42 (2008).

[18] K. Wright, *Computer-Mediated Social Support, Older Adults, and Coping,* 50 JOURNAL OF COMMUNICATION 100-118 (2000).

et al. 2008[19]; Bradley and Poppen 2003[20]). For the elderly, more so than teens, mobility limitations and other factors may increase the relative importance of the Internet for interpersonal communication and expanding social networks (McMellon and Schiffman 2000[21]; O'Hara 2004[22]). Bradley and Poppen (2003), for example, found that training seniors to use the Internet resulted in increased social contact and greater satisfaction with that contact.[23] Mellor et al. (2008) conducted interviews with elderly participants who reported experiencing a positive impact on social connectedness.[24] Trocchia and Janda (2000) report that elderly users perceive the Internet enhances personal connection to family and friends.[25] Sum et al. (2008) found a negative correlation between Internet use and loneliness, and a positive correlation between loneliness and depression (suggests Internet use reduces depression).[26]

Unfortunately, these conclusions are typically based on small samples, and this fact limits both the sophistication of the statistical methodology and the potency and generality of the findings. In the Bradley and Poppen study, the sample consisted of only 20 individuals. Mellor et al. (2008) began with 20 participants, which, due to attrition, dropped to 12 in the final analysis. They report conflicting results between their own quantitative and qualitative data, with the quantitative data indicating a negative impact of Internet use, noting that the small sample size, among other factors, may have contributed to the discrepancy.[27] Even a study attempting to collect data from a random national sample (Eastman and Iyer, 2004) ended with only 171 survey responses for analysis.[28] Sum et al. (2008) based their findings on the responses to 222 online

---

19 S. Sum, R. Mathews, I. Hughes, and A. Campbell, *Internet Use and Loneliness in Older Adults*, 11 CYBERPSYCHOLOGY & BEHAVIOR 208-211 (2008).

20 B. Bradley and W. Poppen, *Assistive technology, Computers and Internet may Decrease Sense of Isolation for Homebound Elderly and Disabled Persons,* 15 TECHNOLOGY AND DISABILITY 19-25 (2003).

21 C. McMellon and L. Schiffman, *Cybersenior Mobility: Why Some Older Consumers May Be Adopting the Internet*, 27 ADVANCES IN CONSUMER RESEARCH 139-144 (2000).

22 *Supra* n. 7.

23 *Supra* n. 21.

24 *Supra* n. 18.

25 P. Trocchia and S. Janda, *A Phenomenological Investigation of Internet Usage Among Older Individuals*, 17 JOURNAL OF CONSUMER MARKETING 605-161 (2000).

26 *Supra* n. 20.

27 *Supra* n. 18.

28 J. Eastman and R. Iyer, *The Elderly's Uses and Attitudes Towards the Internet,* 21 JOURNAL OF CONSUMER MARKETING 208-220 (2004).

questionnaires by Australian adults age 55 or older.[29] Those who call into question altogether the positive effect of technology use on the elderly (Dickinson and Gregor 2006) indicate that the small sample sizes used in many prior studies may contribute to problematic results.[30]

In this study, we attempt to shed more light on the role of Internet use on the mental well-being of the elderly by applying a number of statistical and econometric techniques to a large sample of elderly persons. The Health and Retirement Study ("HRS") is longitudinal household survey data for the study of retirement and health among the elderly in the United States, surveying more than 22,000 persons over the age of 50 every two years.[31] Even after placing a few restrictions on the full sample, we are left with a sample of over 7,000 individuals, an amount that vastly exceeds any prior study on this topic. Mental well-being is measured using the "workhorse of depression epidemiology"—the depression scale developed by the Center for Epidemiologic Studies (the CES-D scale) (Eaton et al. 2004[32]; Radloff 1977[33]). Using this observational data, we apply a variety of statistical and econometric techniques with an eye toward measuring "causal" effects and not just correlation—including single equation regression, instrumental variables, and propensity score methods. Most of the empirical methods applied in this paper are covered in a recent review of modern techniques for estimating treatment effects by Imbens and Wooldridge (2009).[34] Irrespective of the empirical method, we find a negative relationship between Internet use and depression categorization among the elderly, and the estimated effects are similar across methods.

---

29 *Supra* n. 20.

30 A. Dickinson and P. Gregor, *Computer Use has no Demonstrated Impact on the Well-Being of Older Adults*, 64 INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES 744-753 (2006).

31 http://hrsonline.isr.umich.edu.

32 W.W. Eaton, C. Muntaner, C. Smith, A. Tien, M. Ybarra, *Center for Epidemiologic Studies Depression Scale: Review and Revision (CESD and CESDR)* in THE USE OF PSYCHOLOGICAL TESTING FOR TREATMENT PLANNING AND OUTCOMES ASSESSMENT (3rd edition, volume 3, edited by M. E. Maruish)(Lawrence Erlbaum Associates 2004) at pp. 363-378.

33 L.S. Radloff, *The CES-D Scale: A Self-Report Depression Scale for Research in the General Population*, 1 APPLIED PSYCHOLOGICAL MEASUREMENT 385-401 (1977).

34 G. Imbens and J. Wooldridge, *Recent Developments in the Econometrics of Program Evaluation*, 47 JOURNAL OF ECONOMIC LITERATURE 5-86 (2009).

III. Data

The data used in this study is provided by the Health and Retirement Study ("HRS") conducted at the University of Michigan.[35] The HRS is longitudinal household survey data for the study of retirement and health among the elderly in the United States, surveying more than 22,000 persons over the age of 50 every two years. A user-friendly version of the data is constructed and maintained by the RAND Center for the Study of Aging ("RAND HRS").[36] The RAND HRS processes the data in a number of ways, including organizing the data by individual respondent, assigning household level variables and spouses to each observation, and maintaining comparability across survey waves. The last year for which data is available for both datasets is 2006, and we focus our analysis on this most current release. While there are some proprietary elements of the dataset, most of the data is available online at no charge.

Some filtering of the data is necessary, mainly due to non-responsiveness. Also, to focus the analysis and pare down some variation that cannot be accounted for easily, we limit the respondents in the following way. First, we include in the sample only retired individuals who are currently not working. Second, since we are concerned about self-perceptions of mental well-being, we include only those records obtained from self responses. Third, we exclude all respondents living in a nursing home. Fourth, we limit the sample to respondents 55 years or older, the common cutoff for empirical studies of the elderly. After these sample restrictions, the econometric estimates with the full sample are based on over 7,000 observations, which is substantially larger than all previous efforts (of which we are aware) to determine the impact of Internet use on the psychological health of the elderly. This large sample facilitates a wider variety of modeling techniques. Other sample modifications are made based on the applied technique, and model-specific sample sizes are provided in the table summarizing the output. Even after applying propensity score matching algorithms, sample sizes remain very large.

A. *Key Variables*

The two key variables of interest include Internet use and the measure of depression. From the HRS data, the Internet use indicator is based on a direct question regarding regular use of the Internet for the purpose of "sending or receiving e-mail or for any other purpose." The response is dichotomous (Yes,

---

35 *Supra* n. 32

36 http://hrsonline.isr.umich.edu/data/index.html.

No). We define the variable *INTUSE* to equal 1.0 if the Internet is used regularly for e-mail or other purposes, 0.0 otherwise. Unfortunately, we cannot distinguish between the use of broadband and dialup Internet services.

Our measure of depression is based on the eight-item depression scale developed by the Center for Epidemiologic Studies (the CES-D scale).[37] The CES-D is one of the most common screening tests for helping an individual to determine his or her depression levels, with some describing it as the "workhorse of depression epidemiology."[38] The scale is created by summing the responses to eight (yes/no) questions reflective of the respondents mental well-being.[39] This eight-item CES-D, provided in the HRS data, is a scaled-down version of the larger twenty-item CES-D measure where the final set of eight items was selected based on factor analysis (Radloff and Teri 1986).[40] Six of the eight questions indicate the presence of depression, whereas the remaining two indicate its absence. This 8-item scale has been used in several published studies of mental health based upon HRS data.[41] Studies show that the psychometric properties of the more limited eight-item CES-D are very good in terms of consistency and validity.[42] The CES-D has values ranging from zero to eight,

---

37 *See, e.g.*, Radloff *supra* n. 34; I. McDowell I and C. Newell, MEASURING HEALTH: A GUIDE TO RATING SCALES AND QUESTIONNAIRES (Oxford University Press 1996).

38 Eaton *et al.*, *supra* n. 33; M.J. Naughton and I. Wiklund, *A Critical Review of Dimension-Specific Measures of Health Related Quality of Life in Cross-Cultural Research*, 2 QUALITY OF LIFE RESEARCH 397-432 (1993); P. Snaith, *What Do Depression Rating Scales Measure?* 163 BRITISH JOURNAL OF PSYCHIATRY 293-298 (1993); A.M. Nezu, C.M. Nezu, K.S. McClure, M.L. Zwick, *Assessment of Depression* in I.H. Gotlib and C.L. Hammen eds. HANDBOOK OF DEPRESSION (2002) at 61-85.

39 The questions are (Yes, No) answers and are as follows: Much of the time during the past week: (1) you felt depressed?; (2) you feel that everything you did was an effort?; (3) your sleep restless?; (4) were you happy?; (5) you felt lonely?; (6) you enjoyed life?; (7) you felt sad?; (8) you could not get going? Questions (4) and (6) are reversed coded.

40 L.S. Radloff and L. Teri, *Use of the Center for Epidemiological Studies-Depression Scale with Older Adults*, 5 CLINICAL GERONTOLOGIST 119-136 (1986).

41 W.T. Gallo, E.H. Bradley, M. Siegel, *Health Effects of Involuntary Job Loss Among Older Workers: Findings from the Health and Retirement Survey*, 55 JOURNAL OF GERONTOLOGY SERIES B PSYCHOLOGICAL SCIENCES AND SOCIAL SCIENCES 131-140 (2000); M. Siegel, E.H. Bradley, W.T. Gallo, S.V. Kasl, *Impact of Husbands' Involuntary Job Loss on Wives' Mental Health Among Older Adults,* 58 JOURNAL OF GERONTOLOGY: SOCIAL SCIENCES 30-37 (2003); M. Siegel, E.H. Bradley, W.T. Gallo, S.V. Kasl, *The Effect of Spousal Mental and Physical Health on Husbands' and Wives' Depressive Symptoms Among Older Adults,* 16 JOURNAL OF AGING AND HEALTH 398-425 (2004).

42 Gallo *et al. supra* n. 42; D.E. Steffick, *HRS Documentation of Affective Functioning Measures in the Health and Retirement Study*, DOCUMENTATION REPORT DR-005, Survey Research Center at the Institute for Social Research (Ann Arbor, MI, 2000); R. Wallace, A.R. Herzog, M.B. Ofstedal, D. Steffick, S. Fonda, K. Langa, DOCUMENTATION OF AFFECTIVE FUNCTIONING MEASURES IN THE HEALTH

(Footnote Continued. . . .)

with a score of eight indicating the most depressive symptoms. The mean value for the sample is 1.57, with about 42% of respondents having a CES-D value of zero. The distribution of responses is summarized in Table 1.

Table 1. CES-D Distribution in Sample

Value	Count	Percent	Cumulative Percent
0	3,087	41.87	41.87
1	1,583	21.47	63.35
2	947	12.85	76.19
3	581	7.88	84.87
4	366	4.96	89.07
5	292	3.96	93.00
6	247	3.35	96.35
7	185	2.51	98.86
8	84	1.14	100
Total	7,372	100	100
Average	1.57	St. Dev.	1.98

Importantly, the CES-D is a screening device and not a diagnosis by a trained professional. Nevertheless, the CES-D has been shown in numerous studies to be a reliable indicator of a depression disorder.[43] In practice, the CES-D scale is converted to a dichotomous variable by defining depression as present when the scale exceeds a threshold, or cutoff, level.[44] Given standard practice, we employ the threshold approach and treat depression as a dichotomous variable. (In the future, we intend to use other modeling approaches that retain the CES-D's natural state.) As in the earlier research, the dependent variable we use is *CESD_DUM* with a value of 1.0 when the CES-D ≥ 4, thereby classifying about

AND RETIREMENT STUDY (University of Michigan 2000); Eaton et al. *supra* n. 33; F.J. Kohout, L.F. Berkman, D.A. Evans, *Two Shorter Forms of the CES-D Depression Symptoms Index*, 5 JOURNAL OF AGING AND HEALTH 179-193 (1993); C.L. Turvey, R.B. Wallace, R. Herzog, *A Revised CES-D Measure of Depressive Symptoms and a DSM-Based Measure of Major Depressive Episodes in the Elderly*, 11 INTERNATIONAL PSYCHOGERIATRICS 139-148 (1999).

43 M.M. Weissman, D. Sholomaskas, M. Pottenger, B.A. Prusoff, B.Z. Locke, *Assessing Depressive Symptoms in Five Psychiatric Populations: A Validation Study*, 106 AMERICAN JOURNAL OF EPIDEMIOLOGY 203-214 (1977); R. Pandya, L. Metz, and S.B. Patten, *Predictive Value of the CES-D in Detecting Depression Among Candidates for Disease-Modifying Multiple Sclerosis Treatment*, 46 PSYCHOSOMATICS 131-134 (2005).

44 *See, e.g.*, J. Blustein, S. Chan, and F.C. Guanais, *Elevated Depression Symptoms Among Caregiving Grandparents,* 39 HEALTH SERVICES RESEARCH 1671-1690 (2004); R. Mojtabai and M. Olfson, *Cognitive Deficits and the Course of Major Depression in a Cohort of Middle-Aged and Older Community-Dwelling Adults,* 52 JOURNAL OF THE AMERICAN GERIATRICS SOCIETY 1060-1069 (2004); I. Nygaard, C. Turvey, T.L. Burns, E. Crischilles, R. Wallace, *Urinary Incontinence and Depression in Middle-Aged United States Women*, 101 OBSTETRICS AND GYNECOLOGY 149-156 (2003).

16% of the sample as "depressed" (or having symptoms thereof).[45] Other control variables in the econometric models are discussed in the next section.

B. *Additional Covariates*

In addition to the primary variables of interest (*INTUSE* and *CESD_DUM*), the econometric models include a number of covariates. Since the CES-D is often used in epidemiology research, we include many of the same covariates, as do these earlier studies.[46] Variables used as determinants of depression include the following: the age (in years) of the respondent (*AGE*); a dummy variable indicating the respondent is married and living with a spouse (*MARRIED*); a dummy variable indicating whether the respondent has been married more than once (*MARRIAGES*); the respondent's years of education (*EDUCYEARS*); a dummy variable for gender (*MALE*); and a dummy variable for whether the respondent has a debilitating physical health condition (*HEALTH*). We also have dummy variables for the months November, December, and January, since responses in these months may reflect the recognized problem of "seasonal" depression. All variables are from the HRS and RAND-HRS data.

Additional variables used to estimate the propensity score and an instrumental variable for Internet use include annual household income (*INCOME*) and its square (*INCOME2*); a dummy variable indicating whether an individual is classified as poor (*POOR*); a dummy variable indicating whether there are four or more persons in the home (*MANY*); a dummy variable indicating whether the respondent is African-American (*BLACK*); a dummy variable indicating whether the respondent has four or more family members including grandchildren, siblings, and children (*FAMILY*); and an interaction of the *AGE* and *HEALTH* variables (*AGEHEALTH*). Nine Census-region dummy variables are also included.

---

45 Blustein *et al.*, *id.*; Mojtabai and Olfson, *id*.

46 N. Dragano, Y. He, S. Moebus, K. Jöckel, R. Erbel, J. Siegrist, *Two Models of Job Stress and Depressive Symptoms: Results from a Population-Based Study*, 43 SOCIAL PSYCHIATRY & PSYCHIATRIC EPIDEMIOLOGY 72-78 (2008); K. Conner, M. Pinquart and P. Duberstein, *Meta-Analysis of Depression and Substance Use and Impairment Among Intravenous Drug Users (IDUs)*, 103 ADDICTION 524-534 (2008); E. Kim, S. Jo, J. Hwang, C. Shin, D. Kim, E. Woo, S. Kim, K. Sin, and I. Jo, *A Survey of Depressive Symptoms among South Korean Adults after the Korean Financial Crisis of Late 1997: Prevalence and Correlates*, 15 ANNALS OF EPIDEMIOLOGY 145-152 (2005).

IV. Estimating the Treatment Effect

With observational data, as we have here, the treatment (Internet use) is not randomly assigned, thereby requiring some effort to ensure we distinguish the treatment effect from the selection effect. In econometrics, the Heckman two-stage selection model is frequently employed to resolve the selection bias. In epidemiology and increasingly in the economics literature, propensity score methods are used to create "balance" in the treated and control groups, so the treated and control groups can be viewed as randomly drawn. A propensity score is simply the predicted probability of receiving the treatment taking into account observed covariates. In the case of a dichotomous treatment, the propensity score is the predicted probability of a logistic (or probit) regression of the treatment on a number of covariates. Propensity score methods ("PSM") attempt to mimic the randomness of an experiment by assuming that unobservable heterogeneity does not impact participation in the "treatment," thereby satisfying the conditional independence assumption. PSM is sometimes referred to as a quasi-random experimental technique. The contrast between more traditional econometric techniques and PSM can be summarized simply by noting that PSM is more concerned with selecting a sample with which to measure differences in outcomes using relatively simple statistical tests whereas econometrics is more concerned with estimation techniques (which can be complex) given the sample available.

The propensity score can be used in a variety of ways to measure treatment effects.[47] Matching algorithms are popular, for example, where each treated observation is assigned one or more control observations by reference to the propensity score.[48] Matching on the propensity score rather than the covariates is desirable in that it avoids the curse of dimensionality when there are numerous covariates. Alternately, the propensity score is sometimes used to weight observations to reduce bias.[49] Crump et al. (2009) recommend trimming the data so that propensity score lies between 0.10 and 0.90, thereby ensuring common support for the covariates.[50] Rosenbaum and Rubin (1983), supported by

---

47 Imbens and Wooldridge, *supra* n. 35; A. Cameron and P. Trivedi, MICROECONOMETRICS (2005).

48 P. Rosenbaum and D. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41-55 (1983).

49 Imbens and Wooldrige, *supra* n. 35 at 38-9

50 R. Crump, V. Hotz, G. Imbens and O. Mitnick, *Dealing with Limited Overlap in Estimation of Average Treatment Effects*, 96 BIOMETRIKA 187-199 (2009).

Cochran's results (1968)[51], propose stratifying the sample, perhaps into quintiles, based on the propensity score, then computing an average treatment affect across the groups.[52] Imbens and Wooldridge (2009) recommend stratification with regression and conclude this approach is "one of the more attractive estimators in practice."[53] We employ this latter approach and some matching algorithms to measure the treatment effect.

The goal of propensity score methods is to balance the distributions of measured baseline characteristics across the treated and control groups. In other words, the methods address the problem of selection bias on observables by simulating a random experiment. Propensity score methods cannot, however, resolve the problem caused by an unmeasured source of bias (such as endogeneity).[54] This problem is often dealt with using instrumental variables. These different procedures are not necessarily substitutes, but each addresses different forms of bias that may render unreliable the results from more simple regression analysis or means-difference testing.

A. *Estimating the Propensity Score*

The propensity score is the predicted probability of receiving the treatment, which in this case is Internet use. We estimate the propensity score using logit regression, where the dependent variable is *INTUSE:*

$$INTUSE_i = X\beta + u_i \tag{1}$$

where the covariate vector X contains the variables *AGE*; *INCOME; INCOME2; POOR; MARRIED; MARRIAGES*; *EDUCYEARS*; *MALE*; *HEALTH; MANY*; *BLACK*; *FAMILY*; *AGEHEALTH*; and three month and eight regional dummy variables (with the ninth left out to avoid the dummy trap). The predicted probability from this regression, p(X), is the propensity score. In all, 24 covariates are in the propensity score model and 15 are statistically significant at the 10% level or better.[55] The null hypothesis of the Hosmer-Lemeshow Test (i.e., "the

---

51 W. Cochran, *The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies,* 24 BIOMETRICS 295-313 (1968).

52 *Supra* n. 49.

53 *Supra* n. 35 at 40-1.

54 *See, e.g.,* N. Zohoori, *Does Endogeneity Matter? A Comparison of Empirical Analyses with and without Control for Endogeneity*, 7 ANNALS OF EPIDEMIOLOGY 258-266 (1997).

55 The estimated coefficients and (robust) t-statistics are: *AGE* (-0.07, -16.95); *INCOME* (0.742, 6.21)*; INCOME2* (-0.063, -3.09)*; POOR* (-0.504, -3.47)*; MARRIED* (0.202, 2.12)*; MARRIAGES* (-0.007, -0.114); *EDUCYEARS* (0.263, 20.48); *MALE* (0.012, 0.19); *HEALTH* (-2.08, -2.14)*; MANY* (-

(Footnote Continued. . . .)

model is correctly specified") is not rejected at standard levels [$\chi^2(7192) = 7086$, Prob = 0.75].[56] The area under the Receiver Operator Curve ("ROC") is 0.79, indicating good predictive power.[57]

In most regards, the propensity score regression appears satisfactory. That said, the primary goal of propensity score estimation is to facilitate covariate overlap (Dehejia and Wahba 1999). We evaluate overlap using the normalized means difference between the treated and untreated groups for all the covariates:

$$\Delta x = \left|\overline{x}_1 - \overline{x}_0\right| / \sqrt{s_0^2 + s_1^2} \tag{2}$$

where x_1 and s_1 are a variable's mean and standard deviation for the treated group; the subscript 0 denotes the control group. The standardized difference in Equation (2) is similar to a means-difference t-test, except the value of the difference is not impacted by sample size. Imbens and Wooldridge (2009) recommend modifying the propensity score regression or extending stratification until the normalized differences are less than 0.25, citing evidence that regression methods tend to be very sensitive to model specification at higher differences.[58] In an effort to improve covariate overlap, we drop from the sample observations with very small propensity scores.[59] For the treated sample, the propensity score range is 0.014 to 0.9667, and for the control group the range is 0.0003 to 0.9293. For all estimates, we trim the sample so that the propensity score equals or exceeds 0.014, which reduces the sample size to 7,192 (a loss of 180 observations). As expected, this trimming improves covariate overlap.

In the full sample, the normalized difference for some variables—including *AGE* (0.34), *INCOME* (0.32), *MARRIED* (0.30), *EDUCYEARS* (0.55)—exceeds the

---

0.476, -3.36); *BLACK* (-1.10, -10.05); *FAMILY* (0.242, 2.05); *AGEHEALTH* (0.021, 1.54); *NOVEMBER* (-0.06, -0.31); *DECEMBER* (0.07, 0.25); *JANUARY* (0.501, 1.57); *REG1* (-0.29, -1.76); *REG2* (-0.29, -2.47); *REG3* (-0.135, -1.29); *REG4* (-0.27, -2.21); *REG5* (-0.107, -1.09); *REG6* (-0.518, -3.48); *REG7* (-0.575, -4.38); *REG8* (-0.148, -1.04); *CONSTANT* (0.516, 1.41).

56 S. Lemeshow and D. Hosmer, *A Review of Goodness of Fit Statistics for Use in the Development of Logistic Regression Models*, 115 AMERICAN JOURNAL OF EPIDEMIOLOGY 92-106 (2002); O. Baser, *Too Much Ado About Propensity Score Models? Comparing Methods of Propensity Score Matching*, 9 VALUE IN HEALTH 377-385 (2006).

57 Baser, *id*.

58 *Supra* n. 35 at 24. We note that this evidence is based on linear models, not the non-linear models used here.

59 R. Dehejia and S. Wahba, *Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs*, 94 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 1053-1062 (1999); Imbens and Wooldridge, *supra* n. 35 at 43-4.

0.25 threshold. After dividing the sample into quintiles based on the values of the propensity score,[60] all the normalized differences are well below 0.25.[61] See Table A.1 in the Appendix. Consequently, the propensity score regression performs as intended with quintile stratification. Later, we will use these quintiles to estimate the average treatment effect as recommended by Imbens and Wooldridge (2009: 40-1).[62] Further, we also use the prediction from Equation (2) as an instrumental variable in Madalla's two-step estimation procedure for simultaneous equations with qualitative dependent variables.[63] A number of the additional covariates satisfy the criteria for instruments, in particular the regional variables and income variables.[64] Excluding the covariates in the outcomes regression from Equation (2), the null hypothesis that the coefficients are jointly zero is rejected (χ^2 = 303.2).[65] Similarly, the null hypothesis that the coefficients on these regional dummy variables are jointly zero is rejected (χ^2 = 31.8).

B. *Single Equation Methods*

We begin with some simple models of the role of Internet use in curbing depression among the retired elderly. First, we estimate a binary Logit model with *CESD_DUM* as the dependent variable. Right-hand side variables include *INTUSE* along with *AGE*, *MARRIED*, *MARRIAGES*, *EDUCYEARS*, *MALE*, and *HEALTH* in addition to the dummy variables *NOVEMBER, DECEMBER* and *JANUARY*. As noted above, the data is limited to respondents who are retired and non-working with complete data gathered from self-responses, and we trimmed the data at the minimum propensity score of the treated sample. These restrictions leave us with a sample size of 7,192 respondents, which is substantially larger than any other datasets used to evaluate the role of Internet use on the mental well-being of the elderly. Second, we estimate a Linear Probability Model ("LPM") using this data. While Logit is obviously preferred to LPM given the dichotomous nature of the outcome, we report the results of the

---

60 Cochran, *supra* n. 52.

61 *See* Table 2.

62 *Supra* n. 35.

63 G. Maddala, LIMITED-DEPENDENT AND QUALITATIVE VARIABLES IN ECONOMICS (1983) at Ch. 8.8.

64 M. Posner, A. Ash, K. Freund, M. Moskowitz, M. Shwartz, *Comparing Standard Regression, Propensity Score Matching, and Instrumental Variables Methods for Determining the Influence of Mammography on Stage of Diagnosis,* 2 HEALTH SERVICES AND OUTCOMES RESEARCH METHODOLOGY 279-290 (2001).

65 J. Wooldridge, ECONOMETRIC ANALYSIS OF CROSS SECTION AND PANEL DATA (2002) at 90-4.

LPM for comparability to a few of the PSMs that render estimates more comparable to the LPM. A summary of the estimated coefficients from the single equation methods is provided in the Appendix as Table A.2 and discussed below.

1. *Logit with Full Sample*

Ten of the eleven coefficients are statistically significant at the 5% level or better (using robust t-statistics). Factors that increase the probability of depression (or a high CES-D score) in a statistically-significant manner include multiple marriages (*MARRIAGES*) and a physical disability (*HEALTH*). As expected, higher depression responses are observed for the months November and January. The other factors are found to reduce the probability of depression. Older persons (*AGE*) and males (*MALE*) tend to be less depressed, as do those who are married (*MARRIED*) and have more education (*EDUCYEARS*). Of most interest is the coefficient on *INTUSE* (-0.388), which is negative indicating that Internet use reduces the probability of having symptoms of depression. With a t-statistic of -3.80, the null hypothesis of zero effect is easily rejected (at better than the 1% level). The mean of the dependent variable for the full sample is 0.155, so the marginal effect of *INTUSE* at the sample mean is computed to be -0.045. The percentage change in the probability of depression given Internet use can be computed by comparing the observation-specific outcomes across the two treatment regimes as predicted by the model. Setting *INTUSE* = 0, the average predicted probability for the dependent variable is 0.165; setting *INTUSE* = 1.0, the predicted probability falls to 0.127, a sizeable reduction in the probability of depression. Internet use is estimated to reduce the probability of depression by a point estimate of 25%.

2. *LPM with Full Sample*

In the second column of Table A.2 are the estimates from a linear probability model. The LPM does not account for the dichotomous nature of the dependent variable and the estimates are inefficient. Nevertheless, ignoring the coefficient scaling, the results across LPM and Logit or Probit are typically comparable. As expected, coefficient signs and statistical significance across the models are similar to those from the Logit model. The coefficient on *INTUSE* is -0.031. The percentage reduction in depression classification is 20% at the sample mean.[66]

---

[66] We do not average the differences across all observations since the LPM has predicted values outside the unit interval, a fact that can render unreasonably large percentage differences.

C. *Instrumental Variables (2SE)*

It could be argued that mental state influences Internet use thereby implying that both *INTUSE* and *CESD_DUM* are endogenous variables. If true, we have a simultaneous Probit/Logit model since *INTUSE* and *CESD_DUM* are both dichotomous. There are a few ways of proceeding in such cases, but a popular method is the consistent two-stage estimator (2SE) proposed by Maddala (1983: Ch. 8.8). The 2SE involves first estimating the reduced form for *INTUSE* using a single equation binary Logit or Probit. From this step, the predicted probabilities are used as a proxy for the continuous latent variable and inserted as a regressor into the *CESD_DUM* structural equation.[67] Estimation of the second stage by maximum likelihood provides consistent estimates of all the parameters, but the parameter variance-covariance matrix must be corrected to account for the use of a generated regressor from the first step. For valid estimates of the variance-covariance matrix, we compute the standard errors based on the Murphy-Topel Covariance Matrix (Murphy and Topel 1985).[68]

We replace the *INTUSE* variable in the regression with the predictions from the propensity score regression (*INT_IV*). Estimates are provided as Model 3 in Table A.2. The IV estimates are very similar to single equation logit of Model 1, though the coefficient on the *INT_IV* is about a third smaller (-0.340 versus -0.225). The t-statistics computed using the Murphy-Topel standard errors are provided in parentheses. The sign on *INT_IV* is negative and statistically significant at the 1% level or better. Internet use reduces depression.

The smaller coefficient translates into a slightly smaller reduction in the probability of depression. Averaging across the sample, the reduction in the probability of depression due to Internet use is about 19%. So, this approach renders a smaller but comparable treatment effect than does the single equation method, but the result is still favorable for Internet use and the null hypothesis of "no effect" is easily rejected.

---

67 As described in Maddala (1983: 246), *supra* n. 64, the dichotomous structure of the first-stage identifies parameters only up to a scalar transformation. As is prescribed, we normalize the variances of the prediction to unity in the standard way.

68 K. Murphy and R. Topel, *Estimation and Inference in Two-Step Econometric Models*, 3 JOURNAL OF BUSINESS AND ECONOMIC STATISTICS 370-379 (1985). Perhaps due to the large sample size, the difference in the Murphy-Topel standard errors and those reported by the statistical package are trivial.

D. *Propensity Score Methods*

We now turn to the estimation of the treatment effect using PSM. First, we follow the recommendation of Crump et al. (2009) and estimate both the single equation Logit model and LPM using only observations for which $0.10 \leq p(X) \leq 0.90$. This trimming excludes those highly unlikely and likely to use the Internet, two groups which are most dissimilar. Second, we compute the average treatment effect with subclassification and regression. We use regression to compute the average treatment effect both with and without the additional covariates in the model. Third, we compute the average treatment effect using radius and kernel matching algorithms. For this approach, one or more untreated respondents is assigned to each treated observation based on a defined proximity of propensity scores. Finally, we use the matched sample to estimate the treatment effect using Logit, which should reduce the variance of the estimated effect.

1. *Logit and LPM on the Trimmed Sample*

We estimate both the single equation Logit and LPM using only observations for which $0.10 \leq p(X) \leq 0.90$. Estimates are summarized as Models 4 and 5 in Table A.2. This trimming reduces the sample size to 5,801 observations, still a very large sample. The estimated coefficients and their standard errors are not much different from the full and trimmed sample. For the Logit, the coefficient is -0.290. Averaged across the observations, the probability of a depression outcome for Internet users is about 22% less than those not using the Internet regularly (with means of 0.142 and 0.113). For the LPM, the coefficient on *INTUSE* is -0.027. For the LPM, the percentage reduction in depression classification is 20% at the sample mean.

2. *Subclassification with Regression*

We begin by dividing the data into quintiles based on the propensity score. As discussed above, the normalized differences are all below 0.25 for all variables in all quintiles, so limiting our subclassification to five groups adequately ensures covariate overlap. To compute the average treatment effect, we first estimate a blocking estimator using the following approach.[69] Given dummy variables for each of our quintiles, g_i (i = 1, 2 … 5), we estimate the treatment effect using the logit regression

69 Imbens and Wooldridge, *supra* n. 35 at 32-3, 41.

$$y_i = \sum_{i=1}^{5} \alpha_i g_i + \sum_{i=1}^{5} \lambda_i g_i INTUSE + v_i \,, \tag{3}$$

which Imbens and Wooldridge (2009) describe as the "block" estimator. We can then test the null hypothesis

$$\sum_{i=1}^{5} \lambda_i = 0 \tag{4}$$

as a statistical test of whether *INTUSE* has an effect on depression (Imbens and Wooldridge 2009: 41). We can also add the covariates to Equation (3), re-estimate, and then perform the same joint test. This is referred to as *subclassification with regression*.[70] The treatment effect in either model is computed using

$$ATE = \sum_{i=1}^{5} \lambda_i (n_i / N) \,, \tag{5}$$

where the calculation for Expression (5) can be compared to the coefficient estimates of the *INTUSE* variables from Table A.2. The regression estimates are summarized in Table A.3, including the test statistics for the joint tests. All quintiles are equal-sized (as a practical matter), so (n_i/N) is 0.20. Notably, the test statistic for Expression (4) is no different than that from Expression (5) under the null that Expression (5) equals zero.

For the block estimator, the treatment effect from the logit regression is -0.365, which is slightly larger than the effect from Models 1 and 3 in Table A.2. The effect is statistically significant at better than the 1% level (χ^2 = 11.89). The percentage reduction in the probability of depression from Internet use is about 25% when computed as the average across the sample. Adding in the covariates increases the treatment effect from the logit to -0.402 with a χ^2 of 13.11 (again significant at better than the 1% level). The probability of depression for Internet users is 26% less than for those not using the Internet regularly. These effects are comparable to those computed by the single equation methods, but are larger than those estimated using the IV approach. As with the simple regressions, the effect of Internet use is favorable and the null hypothesis of zero effect is easily rejected.

---

70 Imbens and Wooldrige, *id.* at 41.

3. *Matching*

Another estimation strategy using the propensity score is to match the data on $p(\mathrm{X})$. The goal of this procedure is to mimic the random assignment of the treatment. There are a number of matching algorithms available. We use radius matching (r = 0.001 and r = 0.000083 with common support) and kernel matching (bandwidth = 0.015 with common support).[71] We allow for replacement of selected control observations. With radius matching, control observations are matched to the treated when their propensity scores fall into predefined radius. Multiple matches are possible, and the method we employ uses an average of controls in such cases. Kernel matching uses a weighted average of all controls for each treated observation with weights inversely proportional to the distance between propensity scores. All control observations are weighted, thereby reducing the variance of the estimator. For all matching algorithms, we report the t-statistic proposed by Lechner (2001).[72] While bootstrapping the standard errors is common, Abadie and Imbens (2008) conclude bootstrapping is not valid for matching algorithms.[73] All calculations are based on the *psmatch2* program by Leuven and Sianesi (2003) and performed using STATA 11.[74]

With r = 0.001, successful matches render a sample of 2,081 treated observations and 4,556 control observations. Thus, almost all the treated observations are included in the sample (the total treated is 2,211). The average

---

71 A general discussion of matching is provided in P. Rosenbaum and D. Rubin, *supra* n. 49. On Kernel matching, see J. Heckman, H. Ichimura, and P. Todd, *Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme*, 64 REVIEW OF ECONOMIC STUDIES 604-654 (1997) and J. Heckman, H. Ichimura, and P. Todd, *Matching as an Econometric Evaluation Estimator*, 65 REVIEW OF ECONOMIC STUDIES 261-294 (1997). There is no rule of thumb for choosing the tolerance level for radius matching. We choose 0.001 because it often appears in the literature and 0.000083 because it cuts the treated sample approximately in half. Kernel bandwidth is selected using the rule of thumb: $bw = 1.06 \cdot \mathrm{Min}\{\sigma, R/1.34\} \cdot n^{-0.2}$, where σ is the (estimated) standard deviation of the propensity score, R is the interquartile range (75% - 25%) of the propensity score, and n is sample size. H. Engelhardt, H. Kohler, A. Fürnkranz-Prskawetz, CAUSAL ANALYSIS IN POPULATION STUDIES: CONCEPTS, METHODS, APPLICATIONS (2006) at 190. The Gaussian (normal), Epanechnikov and Tricube kernels produce nearly identical results.

72 The estimated standard error is based on M. Lechner, *Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption*, in ECONOMETRIC EVALUATION OF LABOR MARKET POLICIES (M. Lechner and F. Pfeiffer eds. 2001) at 43-58.

73 A. Abadie and G. Imbens, *On the Failure of the Bootstrap for Matching Estimators,* 76 ECONOMETRICA 1537 (2008).

74 E. Leuven and B. Sianesi, *PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing* (2003) (available at: http://ideas.repec.org/c/boc/bocode/s432001.html).

treatment effect (on the treated) is -0.031 with a t-statistic of -2.68, indicating a 24% reduction in depression categorization.[75] We choose a tighter radius of 0.000083 that shrinks the treated sample by about half, thereby rendering 1,103 treated and 1,551 control observations. The average treatment effect on the treated is -0.026, a 19% reduction in depression classification, and the null hypothesis of zero effect is rejected with a t-statistic of -1.75 (rejecting at the 10% significance level).[76] Kernel matching, with common support and bandwidth 0.015, the estimated treatment effect is -0.022 indicating a 19% reduction in depression classification.[77] The t-statistic is -2.04, so the null is rejected at the 5% level.

Overall, these matching estimates paint a very similar picture to the regression methods. The tighter radius matching rule and kernel matching both indicate an average treatment effect of a 19% reduction in depression classification. In both cases, the null hypothesis of zero effect is rejected at the 10% significance level or better in all cases.

4. *Pseudo-R^2 Analysis of Matching Algorithms*

These matching approaches and those above depend on the performance of the matching algorithms. Sianesi (2004) proposes a test of the matching algorithms based on the psuedo-R^2 of the propensity score regression.[78] The pseudo-R^2 indicates how well the regressors explain the participation probability of Internet use. Sianesi's (2004) proposal is to compare the pseudo-R^2 before and after matching, since after matching the pseudo-R^2 should be low given that there should be no systematic differences in the distribution of covariates between the treated and control groups. Another approach is simply to perform a joint-test of significance of all the regressors both before and after matching. Prior to matching, the null hypothesis of the joint-test of zero coefficients in the propensity score regression should be rejected, but after matching the null should be accepted.

---

75 The mean of the treated is 0.097 and for the controls is 0.128.

76 The mean of the treated is 0.111 and for the controls is 0.137.

77 The mean of the treated is 0.097 and for the controls is 0.119. We also applied the Epanechnikov and Tricube kernels, which rendered the following results: -0.022 (t = -1.95) and -0.022 (*t* = -1.96).

78 B. Sianesi, *An Evaluation of the Active Labour Market Programmes in Sweden*, 86 THE REVIEW OF ECONOMICS AND STATISTICS 133-155 (2004).

For the propensity score model, the pseudo-R^2 is 0.188 and the null hypothesis that all coefficients are jointly zero is easily rejected by the log-likelihood ratio test (χ^2 = 1669.51; Prob < 0.001). For the radius matching algorithm with r = 0.001, the pseudo-R^2 is 0.0014 and the χ^2 of the joint test is 7.83 (Prob = 0.999) after matching. When r = 0.000083, these statistics are 0.004 and 12.39 (Prob = 0.975). With kernel matching, the pseudo-R^2 is 0.0023 and the χ^2 is 13.79 (Prob = 0.9513). By Sinasi's (2004) approach, the matching algorithms are effective. In none of these after-matching regressions are any of the regressors statistically different from zero, suggesting that balance is obtained for all regressors. We likewise evaluated the covariate balance after matching using Equation (2); the normalized differences were all within tolerance.

5. *Matching with Regression*

A final approach involves estimating the treatment effect using Logit regression and the matched sample.[79] This approach is likely to reduce the variance of the estimate and reduces bias due to lingering discrepancies between the covariates of the treated and control groups.[80] To implement the approach, we estimate the Logit regression as a weighted regression where the weights are those estimated from the matching algorithms.[81]

The estimates are reported in Table A.4 and are comparable to the others reported above. As expected, the regression approach reduced the variance of the estimates (causing the t-statistics to rise). Coefficient estimates using the matched sample are similar to those using the entire sample. For r = 0.001, the coefficient on *INTUSE* is -0.347, whereas the coefficient is -0.34 for the entire sample. As the radius tolerance is tightened, the estimated coefficient looks more like that from the IV estimation (-0.26 versus -0.22). The same is true for the kernel matched sample. Again, the estimates indicate approximately a 20% reduction in depression classification due to Internet use, and all are statistically significant.

E. *Summary of Estimates*

In this section, we have evaluated the role of Internet use on depression among the elderly in the United States. A variety of empirical methods has been applied to a large sample. This large sample has facilitated diversity in empirical

79 Imbens and Wooldridge (2009), *supra* n. 35 at 41.

80 Baser (2006), *supra* n. 57; Imbens and Wooldridge (2009), *supra* n. 35 at 41.

81 We use the *iweight* option in STATA 11.

procedures. A summary of the estimated treated effects is provided in Table 2 below. While the results vary across methods, the conclusions are comparable and consistent—Internet use reduces the probability of a depression categorization for elderly persons by about 20%.

A few key findings are as follows. First, Internet use reduces the probability of a depression classification as determined by the CESD (≥ 4). In all cases, the null hypothesis that the treatment effect is zero is rejected at standard significance levels. Internet use reduces depression symptoms in the elderly population. Second, most estimates indicate about a 20% reduction (or more) in the probability of depression classification resulting from Internet use, so the effect is not small.

Table 2. Summary of Results

	ATE	Δ Prob. Depressed	t-stat	χ^2
Single Equation Methods				
Logit, All, Single Eq.	-0.340	-25%	-3.796[a]	…
OLS, All, Single Eq.	-0.031	-20%*	-3.605[a]	…
Instrumental Variables				
Logit, IV	-0.223	-19%	-2.850[a]	…
PSM, Trimming (0.1 ≤ p(X) ≤ 0.9)				
Logit, Trimmed, Single Eq.	-0.290	-22%	-3.058[a]	…
OLS, Trimmed, Single Eq.	-0.027	-20%*	-3.016[a]	…
PSM, Subclassification				
Block Estimator	-0.365	-25%	…	11.889[a]
With Covariates	-0.402	-26%	…	13.113[a]
PSM, Matching				
Radius Matching (0.001)	-0.031	-24%	-2.68[a]	
Radius Matching (0.000083)	-0.026	-19%	-1.75[c]	
Kernel Matching (bw = 0.015)	-0.022	-19%	-2.04[b]	
PSM, Logit Regression				
Radius Matching (0.001)	-0.348	-24%	-3.17[a]	
Radius Matching (0.000083)	-0.256	-17%	-1.87[c]	
Kernel Matching (bw = 0.015)	-0.261	-19%	-2.57[a]	

Statistical Significance: (a) 1%, (b) 5%, (c) 10%.
* Computed at sample mean.

Considering depression costs the U.S. economy about $100 billion annually, Internet use for the elderly may have a significant payoff (though 62% of those costs are employment related and not relevant to our sample). Still, a 20% reduction in depression represents an annual gain of $7.6 billion, of which about $2 billion can be assigned to the elderly (about 30% of the adult population with slightly slower average depression rates). Whether such benefits are sufficient to offset the costs of expanding availability and use among the elderly requires further study. Third, selection bias does not appear to be a significant influence on the measured treatment effects. Overall, the treatment effects are similarly sized.

V. Conclusions

For this analysis, we used a number of methods for testing the hypothesis that Internet use and depression have an inverse relationship. The various methods include logit models, linear probability models, instrumental variables, and propensity score methods. In each case, results favor the hypothesis that Internet use by the elderly reduces the likelihood of depression. Even when statistical techniques required the sample to be trimmed, the results are approximately the same. Thus, this study adds to the literature on the relationship between Internet use and well-being, at least for the elderly. The data set we use is perhaps the largest in the literature to date. However, it is limited to elderly persons, whereas much of the literature has focused on teens and college students for whom Internet use is very intense.

While we have conducted a wide variety of empirical tests, there are a few limitations of the analysis and avenues for future research. First, we limited the sample to non-working retired persons. There are likely to be some interesting questions about Internet use and depression that differ among the working and non-working elderly. Furthermore, by excluding employed persons, we also eliminate from consideration the employment-related costs of depression (about 62% of the total). Second, we followed much of the existing literature and converted the CES-D into a dichotomous variable. In fact, the CES-D is an ordered variable. Future research could estimate similar models using ordered logit/probit or least squares outcome regressions, and we intend to investigate alternative estimation techniques along these lines. Third, we focus on a single year of data, but the Internet use variable is also found in earlier HRS surveys. It is possible, then, to extend the analysis to exploit the longitudinal nature of the data. The time series component is short, however, and the changes across survey waves are small. We suspect the gains from this analysis may also be small.

Finally, as always, this paper is but one piece in a portfolio of evidence. More research is always desirable and encouraged. We hope this paper introduces this rich dataset to both the economics and epidemiology profession for further study of this interesting and policy relevant question.

Appendix. Summary of Statistics and Estimates

Table A.1. Normalized Differences in Covariates

	Quintile					Full
Variable	First	Second	Third	Fourth	Fifth	Sample
EDUCYEARS	0.07	0.07	0.04	0.01	0.13	0.55
AGE	0.04	0.05	0.01	0.08	0.08	0.34
INCOME	0.16	0.08	0.09	0.06	0.09	0.32
MARRIED	0.07	0.05	0.01	0.02	0.03	0.30
BLACK	0.12	0.06	0.00	0.05	0.09	0.24
ALONE	0.01	0.01	0.01	0.01	0.07	0.23
POOR	0.16	0.01	0.02	0.07	0.05	0.20
HEALTH	0.09	0.00	0.04	0.07	0.04	0.16
AGEHEALTH	0.09	0.00	0.03	0.07	0.04	0.16
INCOME2	0.12	0.05	0.06	0.04	0.01	0.10
MANY	0.05	0.01	0.02	0.00	0.06	0.09
REGION7	0.01	0.05	0.03	0.04	0.01	0.09
MALE	0.12	0.06	0.01	0.03	0.03	0.06
FAMILY	0.09	0.01	0.02	0.03	0.03	0.06
REGION2	0.02	0.02	0.02	0.00	0.01	0.05
REGION6	0.07	0.01	0.02	0.03	0.04	0.05
REGION8	0.08	0.08	0.04	0.05	0.06	0.05
MARRIAGES	0.16	0.03	0.01	0.01	0.02	0.03
JANUARY	0.07	0.02	0.03	0.05	0.03	0.03
REGION5	0.19	0.04	0.04	0.07	0.03	0.02
NOVEMBER	0.16	0.00	0.08	0.01	0.02	0.01
DECEMBER	0.00	0.01	0.05	0.01	0.02	0.00

Table A.2. Summary of Single Equation Estimates
(Dep. Variable *CESD-DUM*)

	Model 1 (Logit)	Model 2 (LPM)	Model 3 (IV)	Model 4 (Logit, Trimmed)	Model 5 (LPM, Trimmed)	Mean (St. Dev.)
INTUSE (*INT_IV*)	-0.340 (-3.80)[a]	-0.031 (-3.61)[a]	-0.223 (-2.85)[a]	-0.290 (-3.06)[a]	-0.027 (-3.02)[a]	0.307 (0.362)
AGE	-0.019 (-4.38)[a]	-0.002 (-4.32)[a]	-0.025 (-4.74)[a]	-0.026 (-4.59)[a]	-0.003 (-4.66)[a]	73.59 (8.45)
MARRIED	-0.822 (-10.72)[a]	-0.100 (-10.62)[a]	-0.746 (-8.70)[a]	-0.832 (-9.47)[a]	-0.094 (-9.08)[a]	0.581 (0.493)
MARRIAGES	0.171 (2.28)[b]	0.020 (2.17)[b]	0.176 (2.35)[b]	0.163 (1.86)	0.018 (1.81)	0.291 (0.455)
EDUCYEARS	-0.106 (-8.20)[a]	-0.012 (-7.92)[a]	-0.075 (-3.70)[a]	-0.116 (-5.82)[a]	-0.011 (-5.83)[a]	12.481 (2.79)
HEALTH	1.589 (17.17)[a]	0.295 (15.05)[a]	1.515 (15.32)[a]	1.762 (14.97)[a]	0.308 (12.37)[a]	0.092 (0.289)
MALE	-0.298 (-3.83)[a]	-0.031 (-3.69)[a]	-0.288 (-3.70)[a]	-0.344 (-3.82)[a]	-0.031 (-3.59)[a]	0.436 (0.496)
NOVEMBER	0.542 (2.69)[a]	0.074 (2.41)[b]	0.554 (2.78)[a]	0.445 (1.82)	0.052 (1.63)	0.023 (0.150)
DECEMBER	-0.326 (-1.06)	-0.036 (-1.06)	-0.323 (-1.04)	-0.347 (-0.89)	-0.035 (-0.97)	0.011 (0.105)
JANUARY	1.112 (3.34)[a]	0.179 (2.96)[a]	1.171 (3.58)[a]	1.145 (3.10)[a]	0.183 (2.76)[a]	0.008 (0.086)
Constant	1.319 (3.59)[a]	0.529 (11.42)[a]	1.531 (4.01)[a]	1.904 (4.05)[a]	0.534 (10.45)[a]	…
Obs.	7,192	7,192	7,192	5,801	5,801	7,192
R^2	…	0.11	…	…	0.10	
Mean *INTUSE*	0.155	0.155	0.155	0.132	0.132	

Statistical Significance: (a) 1%, (b) 5%, (c) 10%.

Table A.3. Summary of Single Equation Estimates (Dep. Variable *CESD-DUM*)

	Model 1 (Logit)	Model 2 (Logit)	Model 3 (Logit, Matched)	Mean *INTUSE*
g_1*INTUSE*	-0.880 (-2.43)[b]	-1.034 (-2.66)[a]	…	0.049
g_2*INTUSE*	-0.426 (-2.00)[b]	-0.451 (-2.03)[b]	…	0.144
g_3*INTUSE*	-0.190 (-1.15)	-0.242 (-1.43)	…	0.285
g_4*INTUSE*	-0.187 (-1.04)	-0.181 (-0.98)	…	0.420
g_5*INTUSE*	-0.139 (-0.67)	-0.101 (-0.48)	…	0.640
INTUSE	…	…	-0.234 (-2.01)[a]	
AGE	…	-0.025 (-4.76)[a]	-0.034 (-4.57)[a]	…
MARRIED	…	-0.751 (-8.90)[a]	-0.903 (-7.67)[a]	…
MARRIAGES	…	0.184 (2.44)[b]	0.250 (2.08)[b]	…
EDUCYEARS	…	-0.078 (-3.86)[a]	-0.096 (-3.96)[a]	…
HEALTH	…	1.536 (15.58)[a]	1.663 (11.57)[a]	…
MALE	…	-0.287 (-3.67)[a]	-0.382 (-3.17)[a]	…
NOVEMBER	…	0.558 (2.78)[a]	0.812 (2.22)[a]	…
DECEMBER	…	-0.307 (-1.01)	-0.408 (-0.43)	…
JANUARY	…	1.174 (3.59)[a]	2.200 (3.52)[a]	…
$\Sigma\lambda/5$	-0.365	-0.402	…	
$\Sigma\lambda = 0, \chi^2$	11.89[a]	13.11[a]	…	
Obs	7,192	7,192	6,644	

Statistical Significance: (a) 1%, (b) 5%, (c) 10%.

Table A.4. Summary of Logit Regression on Matched Samples
(Dep. Variable *CESD-DUM*)

	Model 1 (Radius, 0.001)	Model 2 (Radius, 0.000083)	Model 3 (Kernel, 0.015)
INTUSE	-0.348	-0.256	-0.261
	(-3.17)[a]	(-1.87)[c]	(-2.57)[a]
AGE	-0.034	-0.034	-0.034
	(-4.57)[a]	(-3.60)[a]	(-4.63)[a]
MARRIED	-0.903	-0.988	-0.882
	(-7.67)[a]	(-6.90)[a]	(-8.25)[a]
MARRIAGES	0.250	0.106	0.127
	(2.08)[b]	(0.72)	(1.17)
EDUCYEARS	-0.096	-0.095	-0.102
	(-3.96)[a]	(-3.28)[a]	(-4.71)[a]
HEALTH	1.663	1.584	1.655
	(11.57)[a]	(8.03)[a]	(12.07)[a]
MALE	-0.382	-0.261	-0.342
	(-3.17)[a]	(-1.71)[c]	(-3.04)[a]
NOVEMBER	0.812	0.726	0.445
	(2.22)[a]	(1.84)[c]	(1.55)
DECEMBER	0.408	-3.04	-0.244
	(0.43)	(-2.94)[a]	(-0.36)
JANUARY	0.968	1.007	0.878
	(2.44)[b]	(2.00)[b]	(2.13)[b]
Constant	2.200	2.229	2.28
	(3.52)[a]	(2.96)[a]	(4.03)[a]
Obs.	6,644	2,656	7,183
Pseudo-R^2	0.11	0.11	0.09

Statistical Significance: (a) 1%, (b) 5%, (c) 10%.