PHOENIX CENTER POLICY PAPER SERIES

Phoenix Center Policy Paper Number 39:

Internet Use and Job Search

T. Randolph Beard, PhD
George S. Ford PhD
Richard P. Saba, PhD

(January 2010)

Phoenix Center Policy Paper No. 39
Internet Use and Job Search

T. Randolph Beard, PhD*
George S. Ford, PhD†
Richard P. Saba, PhD‡

Abstract: With unemployment levels at record highs, policymakers are struggling to find any means possible to put Americans back to work. In this PAPER, we use the 2007 *Computer and Internet Use Supplement* of the Census Bureau's *Current Population Survey* to estimate the effect of Internet use on job search, and we find this effect to be significant. Our empirical model, which combines multinomial logit and propensity score methods, exploits the distinction between the unemployed and the discouraged, where both desire employment but the latter has ceased active job search due to negative beliefs about the labor market. We find that broadband use at home or at public locations reduces defection from the labor market due to discouragement by over 50 percent (50%). Dialup Internet use also has a statistically significant effect, reducing labor market discouragement by about one-third. These results provide useful insights for policymakers: on the demand-side, our results show that programs to promote Internet use keep the jobless active in job search and may equate to more employment; and, on the supply-side, our results demonstrate that the promotion of shared connections, such as at libraries, in unserved and underserved areas may, in fact, produce substantial societal benefits.

---

* Senior Fellow, Phoenix Center for Advanced Legal & Economic Public Policy Studies and Professor of Economics, Auburn University.

† Chief Economist, Phoenix Center for Advanced Legal & Economic Public Policy Studies.

‡ Adjunct Fellow, Phoenix Center for Advanced Legal & Economic Public Policy Studies and Associate Professor of Economics, Auburn University. 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. The authors would like to thank Phoenix Center President Lawrence J. Spiwak for his helpful input in preparing this paper for publication.

<u>TABLE OF CONTENTS:</u>

I. Introduction

The Internet is widely viewed as one of the most important forces in social, political, and economic development. Because the diffusion of the Internet in society is a relatively recent phenomenon, however, formal research on its impacts and consequences remains limited. Economists have not been able, for the most part, to document its impact in any particularly convincing manner, though not due to any lack of effort.[1] The consequences of the Internet for the labor market, for example, have garnered high interest for some time. Reductions in frictional unemployment, lower wage dispersion, and other phenomena one might associate with increased market efficiency have all been expected, hoped-for, or analyzed in papers including, for example,

[1] The lack of credible evidence is, in our opinion, largely due to the focus on macro-level impacts of broadband. Macro-effects are difficult to quantify, even with large datasets and mature technologies, neither of which characterize Internet service. For a review of some of the literature on the economic effects of the Internet, *see* L. Holt and M. Jamison, *Broadband and Contributions to Economic Growth: Lessons from the US Experience*, 33 TELECOMMUNICATIONS POLICY 575-581 (2009). Micro-level studies are increasingly prevalent and not restricted to issues narrowly construed as economics. *See, e.g.,* Sherry G. Ford and George S. 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 the citations therein.

Krueger (2000), Mortenson (2000) and Autor (2001).[2] Given the present economic crisis and resulting double-digit unemployment and significant underemployment,[3] of obvious and immediate interest is the possibility that the Internet might reduce the costs of job search, leading to lower unemployment through reductions in the typical length of jobless episodes.

Labor market research has not provided much in the way of positive evidence for the economic effects of the Internet. Indeed, the widely discussed findings of Kuhn and Skuterud (2004) are indicative of the typical finding: "[w]e conclude that either Internet job search is ineffective in reducing unemployment durations, or Internet job searchers are negatively selected on unobservables."[4] This lack of strong evidence on the value of the Internet on labor markets, particularly job search, is somewhat surprising. As many, including Autor (2001) have noted, the Internet surely reduces the direct costs of search, both by job seekers and employers.[5] In most plausible circumstances, this will lead to increased job search and more efficient matching of employers and employees.[6] However, the Internet also serves as a source of information about jobs, employers, and relevant economic conditions. In general, Internet search activities may, or may not, generate information leading the searcher to update his or her beliefs about relevant statistics, such as the prevalence of openings in particular industries or trades. Viewed in this way, use of the Internet could discourage or encourage job seekers, depending on the nature of the information they find there.

---

2 A. Krueger, *The Internet is Lowering the Cost of Advertising and Searching for Jobs*, NEW YORK TIMES (July 20, 2000) at C2; D. Mortensen, *Panel: Modeling How Search-Matching Technologies Affect Labor Market,* Presentation to IRPP and CERF Conference on Creating Canada's Advantage in an Information Age, Ottawa, Canada (May 2000); D. Autor, *Wiring the Labor Market*, 15 JOURNAL OF ECONOMIC PERSPECTIVES 25-40 (2001).

3 Unemployment measures only those looking for work but without jobs, whereas underemployment measures include those interested in work but have ceased active job search.

4 P. Kuhn and M. Skuterud, *Internet Job Search and Unemployment Durations*, 94 AMERICAN ECONOMIC REVIEW 218-232 (2004). Contrary findings are found in B. Stevenson, *The Internet and Job Search*, NBER WORKING PAPER NO. 13886 (2008).

5 *Supra* n. 2.

6 For example, recent work by Weber and Mahringer (2008) begins the task of evaluating not whether certain forms of job search lead to employment, but how different modes of job seeking affect the degree of "job fit" obtained. A. Weber and H. Mahringer, *Choice and Success of Job Search Methods*, 35 EMPIRICAL ECONOMICS 153-178 (2008).

Table 1. The Labor Force: December 2008, 2009

	Total Persons ('000)	
	Dec. 2008	Dec. 2009
Civilian Labor Force …………………………………..	154,587	153,059
Employed …………………………………………....	143,188	137,792
Unemployed ………………………………………....	11,400	15,267
Not in the Labor Force ..……………………………...	80,686	84,231
Persons who currently want a job …..……………….	5,180	5,939
Marginally attached to the labor force ……………..	1,908	2,486
Discouragement over job prospects …..……………..	642	929
Reasons other than discouragement ….……………...	1,266	1,588

Source: Bureau of Labor Statistics, *The Employment Situation - December 2009*, USDL0-09-1583 (Jan. 8, 2010) at Table A-2 and A-13.

It is additionally the case that employment status itself is multi-faceted, and it is often important to distinguish between employment, under-employment, unemployment, marginal attachment to the labor force, and so on.[7] Of particular social importance in this regard is the issue of the "marginally attached", defined by the Bureau of Labor Statistics ("BLS") as those persons not in the labor force who want and are available for a job, and who looked for work in the past 12 months, but who are not currently looking (i.e., within the past four weeks).[8] These workers are divided into two classes: discouraged, and marginally attached but not discouraged. A "discouraged" person is no longer seeking work because they believe either that there are no jobs available, or else no jobs for which they are qualified.[9] As shown in Table 1, discouraged workers amounted to around 929,000 individuals as of the final quarter of 2009, or about 37% of the marginally attached and 16% of those who wanted a job.[10] The remainder of the marginally attached includes those persons not looking for work due to reasons such as family responsibilities or transportation problems. In the final quarter of 2009, these marginally attached but not discouraged individuals were 2.5 million strong, or about 26% of those who wanted a job.[11] Clearly, the marginally attached are economically and sociologically significant, and their problems should be an issue of public policy concern.

---

7 *See, e.g.*, Bureau of Labor Statistics, *How the Government Measures Unemployment* (available at: http://www.bls.gov/cps/cps_htgm.htm).

8 *Id*. Notably, casual Internet search for jobs does not count as active job search.

9 *Id*.

10 Bureau of Labor Statistics, *The Employment Situation - December 2009*, USDL0-09-1583 (Jan. 8, 2010) (available at: http://www.bls.gov/cps) at Table A-13.

11 *Id*.

In this POLICY PAPER, we examine the effects of Internet use on worker status by analyzing the pool of workers who are jobless and currently want a job. This group includes the unemployed (jobless but actively seeking a job or awaiting layoff recall) and the marginally attached (available for work, but not actively seeking employment at present). In BLS terms, the marginally attached are not currently searching because they are either discouraged about their job prospects, or else face some other challenge such as caring for an elderly relative or lacking transportation. In our view, Internet use provides three sorts of services relevant to these categories of the jobless, and potentially affects the probabilities with which a person may fall into one or another of them. Since online job searches are inexpensive, Internet use should encourage active search. Additionally, Internet use provides information on jobs, wages, and the like. This information may influence the workers' beliefs about job availability and requirements. Finally, the Internet is widely used by persons in difficult circumstances (such as joblessness) to obtain support and emotional reinforcement. Such encouragement may prevent job seekers from "giving up", i.e., becoming discouraged.

The informational and supportive roles of the Internet can be crudely evaluated by examining the differential impacts of access on the sorting of the marginally attached workers into various categories, where this sorting reflects beliefs about the labor market, versus those categories that reflect largely external circumstances such as childcare duties, poor health, or the lack of reliable transportation. Although we will also examine our results using a modified definition of discouragement, even when using the formal BLS definition, we find evidence that the jobless are more likely to be discouraged when they do not use the Internet. This evidence suggests that support and information obtained from the Internet reduces the likelihood that they feel there are no jobs, or no jobs for which they could qualify. This finding is consistent with, for example, those provided by Stevenson (2008), who reports that numerous jobseekers claim to have found useful job market information on websites.[12] Importantly, we get results that are similar when we use a modification of the BLS definition of discouragement that, in our opinion, better reflects the expected impact of Internet use. For example, the Internet may provide information relevant to solving transportation problems, or may allow the disabled to work from home. These outcomes do not lead to the affected persons being classified as "discouraged" by the BLS, but we suspect that Internet use may alter the

---

12 *Supra* n. 4.

probabilities with which these challenges drive jobless persons out of the labor force.

The problem of inferring the causal effect of various sorts of Internet use on worker unemployment status, as considered here, will be approached using the general framework of Rubin (1974), the modern details of which were reviewed recently in Imbens and Wooldridge (2009).[13] Other excellent treatments of the topic include Cameron and Trivedi (2005) and Angrist and Pischke (2009).[14] In particular, we seek to obtain average treatment effects with a causal interpretation using multivariate regression methods modified to satisfy the requirements of *unconfoundedness* and propensity score methods to address the issue of *covariate overlap*.[15] This approach is outlined in detail in Section III and is based on the idea that one may remove the biases inherent in simple means comparisons of outcomes between groups by accounting for confounding factors and adjusting the groups to reflect observed differences in covariate distributions. Our task is made more interesting and difficult by the presence of multiple treatments—Dialup Internet use at home, Broadband Internet use at home, and Internet use in public settings—and the trichotomous nature of the outcome—unemployed, discouraged, or marginally attached but not discouraged. We therefore combine the approaches of Lechner (2002) and Crump et al. (2009) to estimate the average treatment effects in this relatively complex environment.[16] Specifically, we estimate propensity scores à la Lechner (2002) and then trim the sample to exclude extreme values of these scores in an effort to improve covariate balance à la Crump et al. (2009).[17] This approach

---

[13] D. Rubin, *Estimating Causal Effects of Treatments in Randomized and Non-Randomized Studies*, 66 JOURNAL OF EDUCATIONAL PSYCHOLOGY 688-701 (1974); G. Imbens and J. Wooldridge, *Recent Developments in the Econometrics of Program Evaluation*, 47 JOURNAL OF ECONOMIC LITERATURE 5-86 (2009).

[14] A. Cameron and P. Trivedi, MICROECONOMETRICS: METHODS AND APPLICATIONS (2005); J. Angrist and J. Pischke, MOSTLY HARMLESS ECONOMETRICS (2009).

[15] Imbens and Wooldridge (2009), *supra* n. 13 at 26; Angrist and Pischke (2009), *supra* n. 14 at Ch. 3.

[16] M. Lechner, *Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies*, 84 REVIEW OF ECONOMICS AND STATISTICS 205-220 (2002); R. Crump, V. Hotz, G. Imbens and O. Mitnick, *Dealing with Limited Overlap in Estimation of Average Treatment Effects*, 96 BIOMETRIKA 187-199 (2009); *see also* 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.

[17] *Id.*

allows all the treatment effects to be estimated in a single model (rather than only in a pair-wise fashion) and facilitates hypothesis testing across estimated treatment effects.

Our paper is organized as follows: Section II reviews some literature on three relevant strands in the literature − Internet use for job seeking activities, discouraged workers, and the role of Internet connectivity in isolation and depression. Section III outlines our empirical strategy for estimating causal effects. Section IV summarizes the results. Conclusions are provided in Section V. We also provide an appendix including more detail on the data and econometric results.

II. Literature Review

An overview of the extent of employment-related websites lends credence to the idea that the Internet might fundamentally alter the dynamics of the labor market. As detailed by Nakamura et al. (2007), the Internet has facilitated a large number of employment innovations for both job seekers and those seeking employees.[18] Specific websites (e.g., Monster.com), employment portal websites for major corporations, streamlined online application systems, and many other innovations have greatly reduced the costs of looking for jobs, looking for employees, and exchanging resumes or filling out applications. Although it is slightly hazardous to generalize from such "first order" effects to characteristics of the resulting equilibria, it would be quite surprising if these cost-reducing innovations did not result in improved job matching and decreased search cost and duration. (Complicating this simple picture somewhat, however, is the overwhelming evidence suggesting the majority of Internet job seekers are currently employed.) Nakamura et al. (2007) provide a battery of statistics suggesting online job seeking is widespread and viewed as effective by the users: Monster.com received over 18 million distinct visitors in September, 2004; 92% of the largest North American corporations had employment sections on their corporate websites as early as 2000; 87.6% of surveyed men ages 25-34, and 93.8% of women of the same ages, reported using an Internet jobsite in 2007; 41.8% of surveyed men ages 25-34, and 39.3% of women from the same cohort, reported they successfully used the Internet in finding their current or most recent job in 2007.[19] Using a different sample, Stevenson (2008) reported that, "...workers

[18] A. Nakamura, K. Shaw, R. Freeman, E. Nakamura and A. Pyman, *Jobs Online*, in D. Autor ed. STUDIES OF LABOR MARKET INTERMEDIATION (2009) at 27-65.

[19] *Id*.

believe that the Internet is helping them find jobs. … [A]mong those that began a job in mid-2002, 22% credited the Internet as the primary means by which they found their job … [O]ver half of those surveyed felt that the Internet was an effective method of job search … ."[20]

In stark contrast to these general observations, specific studies using employment data suggest that the Internet is either of limited effectiveness, or else is worse than useless. The widely discussed findings of Kuhn and Skuterud (2004), which utilized longitudinal observations on Internet use and subsequent employment for a group of unemployed persons, found that once allowance is made for different values of relevant covariates, use of the Internet actually appears to reduce the prospects of job seekers slightly.[21] They remark, "[o]nce observable differences between Internet and other searchers are held constant, however, we find no differences in unemployment durations, and in some specifications even significantly longer durations among Internet users," and later, "[w]e conclude that either (a) Internet job search is ineffective in reducing unemployment durations or (b) Internet job searchers are adversely selected on unobservable characteristics: further research is needed to disentangle these two possibilities."[22] The analysis of Fountain (2005: 1253) offers only a very slightly more positive assessment: "[r]esults suggest the Internet's contribution to an unemployed searcher's information pool may afford a small advantage only to the extent that other job searchers are not using it."[23]

The perceptions of job searchers appear quite at variance with the (admittedly limited) evidence on the effectiveness of the Internet for obtaining employment. Assuming that the poor performance of the Internet in facilitating job search is confirmed by later research, one could say that this misalignment of perception and reality presents a pattern familiar in the literature of psychology, especially with regard to the notions of well-being and depression. Feelings of powerlessness and an inability to control events or one's environment are conventional features of psychological descriptions of depressive disorder.[24] If, as is often alleged, the Internet provides users with virtual communities that

---

20 Stevenson (2008), *supra* n. 4 at 3.

21 Kuhn and Skuterud (2004), *supra* n. 4.

22 Kuhn and Skuterud (2004), supra n. 4 at 219.

23 C. Fountain, *Finding a Job in the Internet Age*, 83 SOCIAL FORCES 1235-1262 (2005).

24 B. Frey and A. Stutzer, *What Can Economists Learn from Happiness Research?*, 40 JOURNAL OF ECONOMIC LITERATURE 402-435 (2002) at 51.

offer support, encouragement, and connection, then use of the Internet might lead to higher subjective evaluations of the job search process than a factual reading of the record would merit. This explanation would depend, of course, on the ability of the Internet to provide such affirmation.

Although many studies have established the danger of excessive or compulsive use of the Internet, especially among young people, the potential therapeutic value of online activities has also received attention.[25] This focus has arisen from the widely-accepted findings of Fernandez and Harris (1992) documenting the effects of social networks on perceived well-being, mental health, and life success, including employment status.[26] In general, social bonding and communication, which is facilitated by the use of the Internet for some people, increases the perception of control of the personal environment, and reduces the severity and duration of episodes of depression associated with either unemployment or unsatisfactory job performance. Studies by Hoybye, Johansen, and Thomsen (2009), Houston, Cooper, and Ford (2002), Shaw and Gant (2002) and many others strongly suggest that, when used correctly, the Internet can significantly improve mental health and outlook for many people facing traumatic events.[27] Most recently, Ford and Ford (2008), in POLICY PAPER NO. 38, employ a wide variety of empirical tests on a large sample of aged persons in the U.S. and find that Internet use by this group reduces depression to a sizeable degree.[28] Depression and joblessness are strongly linked because the transition from work to joblessness is highly stressful, and is therefore a trigger for depression and other emotional difficulties. Prause and Dooley (2001), for example, find that depression at time t is a valid predictor of unemployment at time $t+1$, and similarly that employment at time t is associated with less

---

25 *See, e.g.*, Sajjadian and Nadi, *Depression & Social Isolation in Adolescent and Young Adult Internet Users, Correlation with Time Duration of Internet Use*, 4 JOURNAL OF RESEARCH IN BEHAVIOURAL SCIENCES 33-38 (2006) and the extensive citations in Ford and Ford (2008), *supra* n. 1.

26 R. Fernandez and D. Harris, *Social Isolation and the Underclass*, in A. Harrell and G. Peterson eds. DRUGS, CRIME, AND SOCIAL ISOLATION: BARRIERS TO URBAN OPPORTUNITY (1992) at 257-293.

27 M. Hoybye, C. Johansenand T. Tjornhoj-Thomsen, *Online interaction: Effects of Storytelling in an Internet Breast Cancer Support Group*, 14 PSYCHO-ONCOLOGY 211-220 (2005); T. Houston, L. Cooper, D. Ford, *Internet Support Groups for Depression: A 1-year Prospective Cohort Study*, 159 AMERICAN JOURNAL OF PSYCHIATRY 2062-2068 (2002); L. Shaw and L. Gant, *In Defense of the Internet: The Relationship Between Internet Communication and Depression, Loneliness, Self-Esteem and Perceived Social Support*, 5 CYBERPSYCHOLOGY & BEHAVIOR 157-171 (2002).

28 *Supra* n. 1.

depression at time $t + 1$, more-or-less confirming the mutually causative role of mental state and employment status.[29] Thus, job loss could trigger depression, which reduces the prospects for re-employment, resulting in a vicious cycle and higher healthcare and social support costs.

III. Empirical Strategy

As observed by Autor (2001), Stevenson (2008) and others, it is possible that the Internet, while not necessarily producing an independently significant number of job matches, does tend to keep the jobless searching (perhaps using other, more effective technologies).[30] In light of our discussion above, the question immediately arises: does access and use of the Internet prevent job seekers from becoming discouraged, i.e. "non-searchers"? In response, our focus here is on the effects of Internet connectivity and use on the probability that the jobless become discouraged, i.e., cease looking for work for reasons arising from their perceptions of the nature of the labor market, rather than perhaps temporary and exogenous barriers such as illness in the family. Because the definition of a discouraged worker used by the BLS is quite stringent, dated and not necessarily reflective of the potential impacts of Internet use, we will additionally examine a somewhat modified notion of discouragement of our own construction.

A. *Data*

Data on Internet use, (un)employment, and other covariates of interest comes from the 2007 *Internet and Computer Use Supplement* to the *Current Population Survey*.[31] This data allows Internet use to be measured in three ways: Dialup use at home, Broadband use at home, and Public use (such as at a public library). This same data permits the classification of respondents as employed, unemployed, or marginally attached, including whether those persons indentified as marginally attached are discouraged or not. The survey also

---

[29] J. Prause and D. Dooley, *Effect of Favorable Employment Change on Psychological Depression: Two-Year Follow-Up Analysis of the National Longitudinal Survey of Youth*, 50 APPLIED PSYCHOLOGY: AN INTERNATIONAL REVIEW 282-304 (2001). Also see D. Dooley, J. Prause and K. Ham-Rowbottom, *Underemployment and Depression: Longitudinal Relationships*, 41 JOURNAL OF HEALTH AND SOCIAL BEHAVIOR 421-436 (2000).

[30] Autor (2001), *supra* n. 2; Stephenson (2008), *supra* n. 4.

[31] The data is accessed via DataFerrett (available at: http://dataferrett.census.gov). The supplement to the *Current Population Survey* has been collected in 1994, 1997, 1998, 2000, 2001, 2003, and 2007.

contains a large number of demographic and geographic variables on respondents.

1. *Internet Use*

For Internet use, we define an Internet user by a "Yes" response to the question "Internet use - any location."[32] Of all Internet users, we categorize home users by a "Yes" response to "Connect to Internet from home."[33] Finally, for home users, we distinguish between dialup and broadband use by responses to the question "Currently access - Internet using … (1) A regular dialup telephone; (2) DSL, cable modem, satellite, ….; and (3) something else."[34] We assume broadband users are all home users not using dialup. We thus have three treatments: (1) homes users with dialup (11% of the observations); (2) home users with broadband (48% of observations); and (3) Internet users using public connections (14% of observations), and denote these variables as DIALUP, BROADBAND, and PUBLIC, respectively. About 27% of the sample respondents do not use the Internet. There is no inherent ordering between these treatments (e.g., is home dialup more or less intense than public Internet use?) and we do not impose one.

2. *Labor Market Status*

In forming our sample, we exclude all employed respondents. Our interest focuses on the effect of Internet use on search efforts by the jobless (rather than use by those employed persons who access the Internet to look for more appealing jobs), and on whether Internet use keeps jobless persons in the labor force. Of the jobless persons in our sample, we classify them variously as unemployed, discouraged, or marginally attached in the same manner as the Bureau of Labor Statistics ("BLS") and National Bureau of Economic Research ("NBER").[35] Unemployed persons are in the labor force, while the marginally attached are not. In the data, the unemployed are identified by an "Unemployed-On Layoff" or "Unemployed-Looking" response to the "Labor Force-employment status" survey question.[36] The division of the marginally

---

32 Survey variable: HENET1.

33 Survey variable: HENET3.

34 Survey variable: HENET4.

35 Bureau of Labor Statistics, *supra* n. 10.

36 Survey variable: PEMLR.

attached into the discouraged and "just marginal" classes is summarized in Table 2. Following the BLS definitions, the discouraged are answering the question "reason not looking" with one of the offered responses: a) believes no work available; b) couldn't find any work; c) lacks necessary schooling or training; d) believes employers think he/she is too young or too old; and e) believes other types of discrimination preclude finding a job.[37] Those classified as "just marginal", on the other hand, face problems such as: a) can't arrange for child care; b) has family responsibilities precluding work; c) is in school or receiving training; d) has ill-health or a disability; e) has transportation problems; and f) all other responses.

Table 2. Jobless Classifications

Response	*n*	BLS	Information-related (Authors)
Believes No Work Available	71	Discouraged	Discouraged
Couldn't Find Any Work	113	Discouraged	Discouraged
Lacks Schooling/Training	23	Discouraged	Discouraged
Emp's Think Too Young/Old	31	Discouraged	Just Marginal
Other Discrimination	8	Discouraged	Just Marginal
Can't Arrange Child Care	35	Just Marginal	Discouraged
Family Responsibilities	296	Just Marginal	Just Marginal
In School/Training	213	Just Marginal	Just Marginal
Ill-Health, Disability	187	Just Marginal	Discouraged
Transportation Problems	42	Just Marginal	Discouraged
Other	422	Just Marginal	Just Marginal
Sum	1,441		
Sample Size	4,229		
Unemployed		2,788	2,788
Discouraged		246	471
Just Marginal		1,195	970

After excluding observations with missing data, the full sample consists of 4,229 responses. For Internet use, the summary statistics are: 27% do not use the Internet at all, 48% use Broadband at home, 14% access the Internet at public (not home) sites, and 11% use Dialup service. As for employment, 66% are unemployed and 44% are marginally attached. Of this latter group, 83% are "just marginal" (i.e., marginally attached but *not* discouraged) and 17% are

[37] The full documentation for the question is: PEDWRSN, Labor Force-(not in discouraged) reason not looking: (-1) Not in Universe; (1) Believes No Wrk Avl In Line Lk Or Area; (2) Couldn't Find Any Work; (3) Lacks Necessary Schooling/Training; (4) Employers Think Too Young Or Too Old; (5) Other Types Of Discrimination; (6) Can't Arrange Child Care; (7) Family Responsibilities; (8) In School Or Other Training; (9) Ill-Health, Physical Disability; (10) Transportation Problems; (11) Other – Specify.

discouraged. Of the jobless, 66% are unemployed, 28% are just marginal, and 6% are discouraged.

The BLS definition of a discouraged worker, although well-known, extensively studied, and thoroughly debated, is not self-evidently useful if one seeks to analyze the potential role of the Internet in job search activities. In particular, we believe that the Internet likely has a rather complex effect on the status of the unemployed worker, as discussed in section II. Thus, to examine better the potential effects of the Internet on the job search process, we offer an alternative definition of discouragement that highlights the potential informational effects of Internet use. In the final column of Table 1, we have reclassified the responses to construct a new measure of discouragement that is intended to better reflect these factors. For example, "believing no work available" or "couldn't find any work" are clearly employment obstacles that Internet use may help overcome. Similarly, information on schooling and training requirements are available online, as are educational programs. These three responses are included in the discouraged definition in the BLS classification and clearly fall into a class of circumstances that the Internet may aid in resolving. In our new, alternative definition, three responses are moved into the discouraged class, including the inability to find childcare, having poor health or a disability, and transportation problems. Finding childcare and transportation can be facilitated by Internet use, and the disabled can work from home using an Internet connection. We question whether a person's beliefs about age or other forms of discrimination can be modified by Internet use. Thus, we move age-related and other discrimination out of the discouraged class. In this modified definition of discouraged, 11% of the sample is discouraged and 23% is just marginal.[38] We will perform our statistical analysis using both the traditional BLS definition of discouragement, and our proposed alternative definition.

B. *Causal Effects*

Labor markets have been the focus of much of the treatment effects literature in economics. This paper is partially intended as an addition to that literature, so we view Internet use as the treatment (T) and jobless categorization as the outcome (Y). As with the bulk of this existing literature, we do not have experimental data; the data just described is observational, collected by the Census Bureau. As a consequence, the assignment of the treatment is not

---

[38] Increasing the sample size of "discouraged" also has benefits of a purely empirical nature.

random (as one would of course wish), but is based on the choices of the individuals in the sample. If the factors influencing treatment choice also influence the outcome, then there is a great risk of obtaining a biased measure of the treatment effect using simple statistical tests that ignore this characteristic of the sample. The primary technical challenge of this PAPER then is to develop credible estimates of the treatment effects of the various types of Internet use given the nature of the sample.

Selection bias is most easily illustrated in the potential outcomes framework based on the Rubin Causal Model.[39] Let there be a dichotomous treatment T_i, such as participation in a program (or use of the Internet). Suppose that this treatment will affect some outcome, say wages, Y_i. The observed outcome for individual i can be written as

$$\begin{aligned} Y_i &= \begin{cases} Y_{1i} & \text{if } T_i = 1 \\ Y_{0i} & \text{if } T_i = 0 \end{cases} \\ &= Y_{0i} + (Y_{1i} - Y_{0i})T_i \end{aligned} \tag{1}$$

where Y_{1i} - Y_{0i} is the causal effect of the treatment T_i. (Presumably, the outcomes follow some distribution in the population). The core difficulty with measuring the causal treatment effect is that (in most cases) only Y_{1i} or Y_{0i} is observed in a sample since an individual either received, or did not receive, the treatment. In other words, the treatment effect of interest is Y_1 versus Y_0 for individual i, but in practice individual i is either treated or untreated, so either Y_1 or Y_0 is not observable. Consequently, the direct computation of the treatment effect is precluded, and we are forced to compare the outcomes of a sample of individuals receiving the treatment to a sample of individuals not receiving the treatment.

After some algebraic manipulation, the average effect of the treatment in a sample is

$$\begin{aligned} E[Y_i \mid T_i = 1] - E[Y_i \mid T_i = 0] = E[Y_{1i} - Y_{0i} \mid T_i = 1] + \\ E[Y_{0i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 0] \end{aligned} \tag{2}$$

where $E[\,]$ is the expectation. The first term on the right hand side is the causal effect of interest, which is the difference in outcomes for individual i upon

[39] Rubin (1974), *supra* n. 13; Imbens and Wooldridge (2009), *supra* n. 13; Angrist and Pischke (2009), *supra* n. 14.

receiving the treatment. Added to this causal effect, however, is an additional term that equals the average difference in the untreated states of those who were treated and those who were not. The second term on the right hand side can be thought of as a *selection bias*. If the treated have a different value of Y_0 from the untreated, then the average difference in outcomes in a sample will be biased. For example, it is sensible to expect more intelligent persons will earn higher incomes, at least on average. It is also true that more intelligent persons are more likely to get a college education. The difference in average incomes between those with and without a college degree includes a selection bias component since, without a college degree, the incomes of the more intelligent will be higher. Generally, if the treated are likely to do better (worse) than the untreated in any case, then the selection bias is positive (negative) and the estimated treatment effect will be too large (small). Resolving the problem requires some procedures and/or assumptions that ensure the selection bias is zero. As is well known, random assignment of the treatment solves the problem since the assignment is independent of potential outcomes. This is the basis of sample design in laboratory science. In observational data, however, there is no a priori reason to expect independence between the assignment of the treatment and the outcomes, so something must be done ex post to account for the bias.

As detailed by Imbens and Wooldridge (2009), the most common way to proceed when estimating the causal treatment effect in observational studies is to appeal to the concepts of (1) unconfoundedness (or conditional independence) and (2) covariate overlap.[40] Unconfoundedness implies that, conditional on observed covariates X_i, the treatment assignment probabilities are independent of potential outcomes, or

$$\{Y_{0i}, Y_{1i}\} \perp T_i \mid X_i , \tag{3}$$

where the symbol "$\perp$" denotes independence, so the expression reads, "the outcomes are independent of the treatment given the conditioning covariates." If we have a sufficiently rich set of observable covariates, regression analysis including the variables X_i leads to valid estimates of causal effects. Since the X_i must be observed to be included in the model, this approach is often referred to as "selection-on-observables".

In this PAPER, we employ regression analysis to estimate the treatment effects. Selection bias in a regression framework is directly analogous to the

40 Imbens and Wooldridge (2009), *supra* n. 13, at 23-8.

potential outcomes approach. Consider the regression approach to measuring the causal effect,

$$Y_i = \alpha + \theta T_i + e_i \tag{4}$$

where e_i is the random disturbance term.[41] Evaluating the conditional expectations, we see that

$$E[Y_i \mid T_i = 1] - E[Y_i \mid T_i = 0] = \theta + \left(E[e_i \mid T_i = 1] - E[e_i \mid T_i = 0]\right) \tag{5}$$

where the right hand side equals the treatment effect (θ) plus the selection bias term in parenthesis. The similarities between Expressions (2) to (5) are apparent. In regression, the selection bias appears (or can be written as) as a correlation between the disturbance term, e_i, and the treatment variable, T_i. This selection arises because of the non-zero difference between the no-treatment potential outcomes between those receiving and those not receiving the treatment, or $E[Y_{0i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 0]$, which is the same as in the potential outcomes approach. Appealing to unconfoundedness, we can decompose the disturbance term e_i into a linear function of observable covariates X, so that

$$e_i = X'\lambda_i + v_i \ , \tag{6}$$

then by construction of the least squares estimates (v_i is uncorrelated with X_i) and by the conditional independence from Expression (3), we can estimate the regression model

$$Y_i = \alpha + \theta T_i + X_i'\lambda + v_i \ , \tag{7}$$

and interpret θ as the causal treatment effect, since the residual v_i is uncorrelated with T_i and X_i.

A second condition for the measurement of the causal effect is covariate overlap. Imbens and Wooldridge (2009) observe that, once one commits to the unconfoundedness assumption, the issue of covariate overlap is the "main

41 *See*, e.g., Angrist and Pischke (2009), *supra* n. 14 at 58-9; Imbens and Wooldridge (2009), *supra* n. 13, at 26.

problem facing the analyst."[42] Imbens and Wooldridge (2009) define the overlap condition as

$$0 < p(T_i = 1 \mid X_i = x) < 1, \text{ for all } x \tag{8}$$

where "p" indicates probability.[43] This condition implies that the support of the conditional distribution of X_i given $T_i = 0$ overlaps completely with the conditional distribution of X_i given $T_i = 1$. Put simply, covariate overlap implies that the covariate distributions for the treated and untreated groups are sufficiently alike, which is important, given the inherent extrapolations between the groups made in regression analysis. From above, recall that in practice we only observe Y_1 or Y_0 for any individual. Thus, we must use the observed outcomes from the untreated observations to project Y_0 onto the treated observations in order to compute an average treatment effect (Y_{1i} - Y_{0i}). If the untreated observations are very different demographically and geographically from the treated, the projection will be a poor one. A number of studies have shown that a lack of covariate overlap is a major concern in estimating treatment effects.[44]

Unlike the unconfoundedness assumption for which there is no direct empirical test, covariate overlap can be evaluated in a relatively straightforward manner. Imbens and Wooldridge (2009), for example, recommend evaluating the normalized differences for each covariate,

$$\Delta x = \frac{\left|\overline{X}_1 - \overline{X}_0\right|}{\sqrt{V_1 + V_0}} \ , \tag{9}$$

where the $\overline{X}_i$ and V_i are the sample means and variances for the treated and untreated groups.[45] If these normalized differences exceed 0.25, then the regression estimates tend to be sensitive to model specification.[46] This threshold

---

42 Imbens and Wooldridge (2009), *supra* n. 13, at 43.

43 Imbens and Wooldridge (2009), *supra* n. 13, at 26.

44 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); J. Heckman, H. Ichimura, and P. Todd, *Matching as an Econometric Evaluation Estimator*, 65 REVIEW OF ECONOMIC STUDIES 261–94 (1988).

45 Imbens and Wooldridge (2009), *supra* n. 13, at 24.

46 *Id.*

value is exceeded for a few covariates in our data. As detailed below, in order to remedy this problem and satisfy the overlap assumption we follow Dehejia and Wahba (1999) and Crump et al. (2009) and employ propensity score methods to trim the data prior to estimation.[47] This technique deletes observations that exhibit a high likelihood of exhibiting the presence of any given treatment, since such observations are unlike observations one would observe in a sample in which treatments were randomly assigned. Duplicating the character of such a sample is the goal.

C. *Specification Details: Treatment and Outcomes*

In the prototypical framework, a single treatment (e.g., a labor program) is evaluated on an outcome that is often continuous in nature (i.e., wage or income). Our task is a bit different in that we have a trichotomous outcome—unemployed, discouraged, and just marginal—and multiple treatments—Internet use at home using either dialup or broadband and public Internet use. The trichotomous outcome is handled by use of multinomial logit estimation (unordered), which is an established and widely used estimation technique.[48]

The multinomial logit model specifies that

$$p_{ij} = \frac{\exp\left(T_i'\theta_j + X_i'\beta_j\right)}{\sum_{l=1}^{m}\exp\left(T_i'\theta_j + X_i'\beta_j\right)}, \tag{10}$$

where p_{ij} is the probability observation i falls into outcome category j (j = 1, … , m), the X_i are the covariates for each observation i, and the β_j the estimated coefficients for each outcome category j.[49] There are three possible treatments in T_i, and thus three coefficients θ_j for each j. To identify the model, the β_j are set to zero for one of the outcomes. The unemployed outcome is set as the base case, so all coefficients measure the probability of falling into either the discouraged or just marginal category relative to the unemployed category. If Internet use reduces job search costs and keeps jobless people in the labor force, then the coefficients on the Internet use variables will be negative.

---

47 Dehejia and Wahba (1999), *supra* n. 44 ; Crump *et al.* (2009), *supra* n. 16.

48 Cameron and Trivedi (2005), *supra* n. 14 at Ch. 15.

49 *Id.*

The problem of multiple treatments is challenging. In the presence of multiple treatments, Lechner (2002) proposes pair-wise comparisons of treatment types and outcomes using propensity score methods.[50] The propensity score, which is the conditional probability of receiving the treatment (given the X's), can be estimated either "structurally" using multinomial or ordered logit (or probit) models, or using "reduced form" single equation logit models in pair-wise comparisons. Matching algorithms are then applied to the estimated propensity scores to compute the pair-wise treatment effects. Estimation of propensity scores for our three treatments is not problematic; multinomial logit can be used for that purpose. Given our trichotomous outcome, however, the second-stage matching approach is not ideal; our trichotomous outcome is perhaps better modeled with multinomial logit. Therefore, we choose to combine propensity score methods and regression, which is an increasingly common approach in empirical research.[51] Specifically, we trim the data (rendering an estimation sample $\mathbb{A}$) using the propensity score to improve covariate overlap. Our trimming is based on the rule of thumb proposed by Crump et al. (2009).[52] Specifically, we first estimate propensity scores, $\hat{s}(X_i)$, then trim the sample to improve covariate balance by keeping only observations where $0.10 \leq \hat{s}(X_i) \leq 0.90$. As discussed below, by trimming the data in this way, the normalized differences are below threshold for all covariates.

This estimation strategy is not without costs, since trimming the data changes what is being estimated.[53] Specifically, our approach does not estimate the average causal effect for the population, but rather estimates the *conditional average treatment effect* ("CATE") for the subpopulation $\mathbb{A}$ ("CATE-$\mathbb{A}$"). Nevertheless, Crump et al. (2009) recommend estimating CATE-$\mathbb{A}$ because these are easier to estimate and (potentially) more precise than population estimates.[54] In any case, care must be taken when extrapolating CATE-$\mathbb{A}$ to the general population.

---

50 M. Lechner, *supra* n. 16.

51 Imbens and Wooldridge (2009), *supra* n. 13 at 38-46.

52 Crump et al. (2009), *supra* n. 16. The same approach is discussed in Imbens and Wooldridge (2009), *supra* n. 13 at 45-47 and Angrist and Pischke (2009), *supra* n. 14 at 90-1.

53 Imbens and Wooldridge (2009), *supra* n. 13 at 46.

54 Crump *et al.* (2009), *supra* n. 16; *see also* Imbens and Wooldridge (2009), *supra* n. 13 at 45-6; Angrist and Pischke (2009), *supra* n. 14, at 90-91.

IV. Results

Our empirical strategy is implemented as follows. First, we discuss the unconditional treatment effects that are estimated by considering only means differences in outcomes across the three treatments and outcomes. By the discussion above, there is a risk that this measure of the average treatment effect is biased, since neither unconfoundedness nor overlap is addressed. Second, in an effort to correct for this bias, we add covariates to the model and compute the conditional average treatment effect ("CATE"). Third, we trim the sample to improve covariate overlap following Crump et al. (2009), thereby potentially addressing both unconfoundedness and overlap.[55] This approach requires a two-step estimation technique. In the first stage, we estimate a propensity score, which is simply the conditional expectation of a sample respondent receiving the relevant treatment (the predicted probability of a logit regression). In the second stage, we estimate the multinomial logit model, but trim the sample to exclude observations exhibiting extreme values of the propensity score (thus creating subsample $\mathbb{A}$) using the rule of thumb proposed by Crump et al. (2009).[56] The estimated treatment effect from this model is CATE-$\mathbb{A}$, the conditional average treatment effect for the subsample $\mathbb{A}$.

A. *Unconditional Treatment Effect*

To begin, we estimate the unconditional treatment by regressing the outcomes on the Internet use variables alone. This unconditional treatment is estimated by including only the treatment dummy variables in the multinomial logit model, which can be written as,

$$p_{ij} = \frac{\exp\left(T_i'\theta_j\right)}{\sum_{l=1}^{m}\exp\left(T_i'\theta_j\right)}. \tag{11}$$

The predicted outcomes from the unconditional estimates are summarized in Table 3. Detailed results are provided in the Appendix Table A-1 and A-2. These results should be interpreted with care, since by the argument above the estimated effects could be substantially biased.

55 *Id.*

56 *Id.*

For the BLS definition of discouragement, the coefficients on the Internet use variables (i.e., the treatment effects) are all negative and mostly statistically different from zero. Internet use reduces both discouraged and just marginal classifications, but the response for discouraged is larger. (Conversely, one may conclude that Internet use increases the probability that the individual will be actively searching for work.) The average treatment effects, measured as the percentage difference in the predicted outcomes, are very large and consistently statistically significant in the case of the discouragement outcome. Broadband and Public use are found to reduce discouragement by over 60%, and dialup has a treatment effect of about 40%. The effects of Internet use on just marginal are small and not all statistically different from zero.

Joint tests on the results using the BLS definitions are as follows. First, we can reject the null hypothesis that Internet use of any type has no effect (prob(χ^2) < 0.01) in the entire model. Second, we can reject the null hypothesis that Internet use has no effect in the discouraged equation (prob(χ^2) < 0.01) and the just marginal equation (prob(χ^2) < 0.01). Third, in the discouraged equation, we can reject the null hypothesis that the effect of Dialup equals that of Broadband or Public use (Dialup has a smaller effect) (prob(χ^2) < 0.01 in both cases), but we cannot reject the null hypothesis that the effect of Broadband equals that of Public Use (prob(χ^2) < 0.45). We can summarize as follows: Internet use of *all* types reduce discouragement, with Broadband and Public use having the same effect, which is larger than the effect of Dialup.

Table 3. Unconditional Treatment Effects
(n = 4,229)

	ML Results			Treatment Effect		
Bureau of Labor Statistics Definitions (n = 4,229; Pseudo-R^2 = 0.011)						
	Coef.	St. Err.	T-Stat	Untreated	Treated	Difference
Discouraged						
Dialup	-0.543*	0.213	-2.55	0.105	0.064	-39%
Broadband	-1.151*	0.153	-7.51	0.105	0.038	-64%
Public	-1.347*	0.255	-5.28	0.105	0.033	-69%
Just Marginal						
Dialup	-0.030	0.121	-0.25	0.298	0.306	2.4%
Broadband	-0.202*	0.083	-2.42	0.298	0.279	-6.5%
Public	-0.386*	0.118	-3.27	0.298	0.245	-18%
Information-related (Author) Definitions (n = 4,229; Pseudo-R^2 = 0.016)						
Discouraged						
Dialup	-0.513*	0.162	-3.16	0.194	0.123	-37%
Broadband	-1.065*	0.115	-9.28	0.194	0.077	-61%
Public	-1.295*	0.188	-6.87	0.194	0.064	-67%
Just Marginal						
Dialup	0.113	0.132	0.86	0.209	0.247	18%
Broadband	0.004	0.092	0.04	0.209	0.240	15%
Public	-0.168	0.128	-1.32	0.209	0.214	2.2%

* Significant 5% level or better.

The results are different for the Information-related definitions of jobless status. While Internet use reduces Information-related discouragement (as we defined it above), it has no effect on the just marginal type. These results are consistent with our earlier conjecture, since the new definition of discouragement was motivated by the expected effects of Internet use. Despite the change in definition, the treatment effects for discouragement are similarly sized. Broadband use at home and Public use by over 60%, and Dialup use reduces discouragement by nearly 40%.

Joint tests are as follows. First, we can reject the null hypothesis that Internet use of any type has no effect (prob(χ^2) < 0.01) in the entire model. Second, we can reject the null hypothesis that Internet use has no effect in the discouraged equation (prob(χ^2) < 0.01), but not the just marginal equation (prob(χ^2) = 0.29). Third, in the discouraged equation, we can reject the claim that the effect of Dialup equals that of either Broadband or Public use (Dialup has a smaller effect) (prob(χ^2) < 0.01 in both cases), but we cannot reject the null hypothesis that the effect of Broadband equals that of Public Use (prob(χ^2) < 0.23). We can again summarize as follows: Internet use of *all* types reduces discouragement (but not just marginal attachment), with Broadband and Public use having the same effect, which is larger than the effect of Dialup.

B. *Conditional Treatment Effect Ignoring Overlap*

Next, we estimate the multinomial logit model while including the covariates X_i to obtain the CATE, described earlier. The list of covariates includes: a dummy variable equal to 1 if there are children 18 or younger in the home; a dummy variable equal to 1 if the respondent is male; a dummy variable equal to 1 if the respondent has a college education; a dummy variable equal to 1 if the respondent does not have a high school degree; a dummy variable equal to 1 if the respondent is Caucasian; a dummy variable equal to 1 if the respondent is an immigrant; a dummy variable equal to 1 if the respondent lives in a metro area; a dummy variable equal to 1 if the respondent is a veteran; a dummy variable equal to 1 if the respondent is currently in school; a set of income dummy variables indicating incomes ≤ \$20,000, \$20,000 to \$40,000, \$40,000 to \$60,000, and \$60,000 to \$100,000 (with a "> \$100,000" dummy omitted); set of dummy age variables indicating persons 20 years or younger, between 20 and 40 years, and between 40 and 60 years (with a "> 60 years" dummy left out to avoid the dummy trap). In all, there are 16 covariates in X_i, three treatment dummy variables in T_i, and a constant term.

As in the unconditional model, there are 4,229 observations. The multinomial logit model results are summarized in Table 4 (details are provided in the Appendix at Table A-1 and A-2. Of the 16 covariates in X_i, 10 in the discouraged equation and 12 in the just marginal equation have coefficients statistically different from zero at the 10% level or better). The Pseudo-R^2 of the multinomial logit model for BLS definitions rises from 0.011 in the unconditional case to 0.068—a sizeable increase. For the Information-related definition, the Pseudo-R^2 rises from 0.016 to 0.077, which again is a sizeable increase in goodness of fit.

Table 4. Conditional Treatment Effects
(n = 4,229)

	ML Results			Treatment Effect		
Bureau of Labor Statistics Definitions (n = 4,229; Pseudo-R^2 = 0.068)						
	Coef.	St. Err.	T-Stat	Untreated	Treated	Difference
Discouraged						
Dialup	-0.511*	0.234	-2.18	0.091	0.061	-33%
Broadband	-0.947*	0.183	-5.16	0.091	0.042	-54%
Public	-1.115*	0.257	-4.33	0.091	0.036	-60%
Just Marginal						
Dialup	-0.191	0.134	-1.42	0.310	0.285	-8.1%
Broadband	-0.286*	0.099	-2.88	0.310	0.274	-12%
Public	-0.319*	0.125	-2.56	0.310	0.269	-13%
Information-related (Author) Definitions (n = 4,229; Pseudo-R^2 = 0.077)						
Discouraged						
Dialup	-0.424*	0.173	-2.45	0.163	0.117	-28%
Broadband	-0.787*	0.136	-5.80	0.163	0.086	-47%
Public	-1.049*	0.193	-5.44	0.163	0.068	-58%
Just Marginal						
Dialup	-0.105	0.148	-0.71	0.232	0.229	-1.6%
Broadband	-0.156	0.110	-1.42	0.232	0.229	-1.5%
Public	-0.127	0.136	-0.94	0.232	0.238	2.6%

* Significant 5% level or better.

While the conditional models perform better, and many of the covariates in X_i are statistically significant determinants of the outcomes, the treatment effects remain large and are similar, though typically smaller, to the unconditional treatment effects in Table 3. For the BLS definitions, the coefficients on the Internet use variables (treatment effects) are all negative and, again, mostly statistically different from zero. Internet use reduces both discouraged and just marginal classifications, but the response for discouraged is again larger. The average treatment effects on discouragement are, as before, very large—over 50%—for both Broadband and Public use. Dialup continues to have a large effect, reducing discouragement by about one-third.

The joint tests also tell a similar story. We can reject the null hypothesis that Internet use of any type has no effect (prob(χ^2) < 0.01) in the entire model, and that Internet use has no effect in either the discouraged equation (prob(χ^2) < 0.01) or the just marginal equation (prob(χ^2) = 0.02). Dialup has a smaller effect than either Broadband or Public use (the null of equality is rejected at prob(χ^2) < 0.053 in both cases). Again, we cannot reject the null hypothesis that the effect of Broadband equals that of Public Use (prob(χ^2) < 0.55).

For the Information-related definitions of joblessness, the regression results are generally similar; coefficients and treatment effects are smaller than in the unconditional case but remain large. The treatment effects for Broadband (-47%)

and Dialup (-28%) are about one-quarter smaller than in the conditional versus the unconditional case. For Public use (-58%), the treatment effect shrinks only by about 13%. Generally, such drops in estimated treatment effects are to be expected, and reflect the underlying selection mechanism, which motivates the estimation strategy. All the effects are statistically different from zero. The joint tests on the coefficients render similar results, so these are not discussed here (see Table A-2 for details).

C. *Conditional Treatment Effect including Overlap Condition*

We now turn to the issue of covariate overlap, using propensity score methods to address the issue. To begin, the normalized difference (Expression 9 above) is used to assess covariate overlap in the full sample. Given three treatments, there are six pair-wise comparisons for which overlap can be evaluated. With 20 covariates in Z_i and six pairs, we have 120 normalized differences. For the full sample, there are 13 normalized differences that exceed the threshold (0.25). The details are provided in the Appendix at Table A-3. Most of the differences are found in the education and income variables. Correcting for this circumstance is our immediate goal.

Following Lechner (2002) and Crump et al. (2008), we use propensity score methods to trim the data to improve covariate overlap. For a more thorough treatment of overlap, and to produce a better prediction of Internet use (which is measured by the propensity score), we add to the list of covariates in X_i a few more variables including: a dummy variable equal to 1 if the respondent is married; a variable measuring household size, and three regional dummy variables (with the fourth excluded to avoid the dummy trap). This larger set of covariates is labeled Z_i.

Given three treatments, there are six propensity scores to estimate, and we do so using multinomial logit with the model,

$$p_{it} = \frac{\exp(Z_i'\gamma_t)}{\sum_{l=1}^{m}\exp(Z_i'\gamma_t)}. \tag{12}$$

The predictions from this model, $\hat{p}_{it}$, provide a treatment class probability for each observation, so each observation has three associated treatment probabilities (one for Dialup, one for Broadband, and one for Public use). Estimates from the propensity score model are detailed in Table A-4 in the Appendix. Between treatment probabilities are computing using $p_{i1}/(p_{i1}+p_{i2})$, and so forth. There are six such probabilities. Following Crump et al. (2008), we select a subsample $\mathbb{A}$ from the full sample based on $0.10 \le p_{it} \le 0.90$ for all i and t.

The subsample $\mathbb{A}$ has only 2,562 observations, so approximately 40% of the sample is excluded. While the sample is much smaller, the trimming is successful in the sense that the normalized differences for all covariates, across all pair-wise comparisons, fall below the threshold level.

Summary statistics on outcomes and treatments for sample $\mathbb{A}$ are similar to those of the full sample. Using the BLS definition, about 6% of the sample is discouraged and 26% is just marginal. (For the full sample, these percentages were 6% and 28%.) Likewise, the Internet use categories are similar. In sample A, about 12% use Dialup, 43% use Broadband, and 16% use Public Internet service, where in the full sample these percentages were 11%, 48%, and 14%, respectively. Using our proposed, alternative definition, the (Information-related) discouraged account for 11% of the sample, a proportion identical to that in the full sample, while the just marginal are 21% of the sample, slightly down from 23% in the full sample. So, while subsample $\mathbb{A}$ is much smaller than the full sample, the treatment and outcome shares are not much changed.

With the overlap condition satisfied, we can now estimate the treatment effect CATE-$\mathbb{A}$ by estimating the ML model using sample $\mathbb{A}$. The results are summarized in Table 5 in the familiar format with detailed results in Tables A-1 and A-2. Trimming the sample has affected the estimates in some ways. For example, only six of the 15 covariates in the discouraged equation are statistically different from zero at the 10% level or better, which is about half as many as in the full sample. Ten covariates have statistically significant coefficients in the just marginal equation, which is similar to the full sample. The results from the joint tests are similar, though it is not possible to reject the hypothesis that the treatment effect of Dialup and Broadband are identical (prob(χ^2) = 0.105), perhaps due to the wider confidence interval on the Dialup coefficient (which is now significant only at the 10% level).

Using the BLS definitions, the Internet use treatments are all negative and statistically different from zero at the 10% level or better. The largest effect is for Public use (-68%), with Broadband also having a large effect (-52%). We cannot, however, reject the hypothesis that the effect of Public and Broadband are equal (prob(χ^2) = 0.18). These are sizeable treatment effects and indicate that Internet use is a potent force for keeping the jobless in the labor force. Both Broadband and Public use are statistically significant in the just marginal equation as well, but the treatment effects are small (about -2.5%).

Table 5. Conditional Treatment Effects for Subsample A

	ML Results			Treatment Effect		
	Coef.	St. Err.	T-Stat	Untreated	Treated	Difference
Bureau of Labor Statistics Definitions (n = 2,562; Pseudo-R^2 = 0.066)						
Discouraged						
Dialup	-0.473**	0.273	-1.73	0.092	0.065	-30%
Broadband	-0.917*	0.207	-4.44	0.092	0.044	-52%
Public	-1.343*	0.333	-4.03	0.092	0.029	-68%
Just Marginal						
Dialup	-0.264	0.164	-1.61	0.293	0.254	-13%
Broadband	-0.336*	0.120	-2.79	0.293	0.248	-15%
Public	-0.371*	0.151	-2.46	0.293	0.246	-16%
Information-related (Author) Definitions (n = 2,562; Pseudo-R^2 = 0.071)						
Discouraged						
Dialup	-0.475*	0.211	-2.25	0.159	0.111	-30%
Broadband	-0.665*	0.157	-4.24	0.159	0.096	-40%
Public	-1.033*	0.231	-4.47	0.159	0.068	-57%
Just Marginal						
Dialup	-0.191	0.181	-1.05	0.224	0.208	-7.2%
Broadband	-0.300*	0.133	-2.26	0.224	0.196	-13%
Public	-0.265	0.165	-1.61	0.224	0.208	-7.5%

* Significant 5% level or better.

For the Information-related definitions, the estimates continue to support a strong effect of Internet use on job search. All the treatment effects are statistically different from zero in the discouraged equation. The largest treatment effects are for Broadband (-40%) and Public Use (-57%); statistically, the two effects are identical (prob(χ^2) = 0.12). Dialup has a smaller effect relative to the other treatments (-0.30), but the effect is still large in absolute terms and statistically significant. Having Dialup Internet access reduces the probability a jobless person abandons the labor force by nearly one-third. Notably, we cannot reject the hypothesis that the effect of Dialup is equal to that of Broadband (prob(χ^2) = 0.37), but we can reject the hypothesis that Dialup has the same effect as Public use (prob(χ^2) = 0.047).

As before, the treatment effects are not very large in the just marginal equation and typically not statistically significant. The coefficient on Broadband at home is now statistically different from zero, but the joint test that all treatments are zero is not rejected (prob(χ^2) = 0.14). The treatment effects remain relatively small, with the largest effect for Broadband at -13%.

D. *Summary of Empirical Results*

Table 6 summarizes the average treatment effects across our three models. There are a few results of note. First, the effects of Internet use on discouragement are large across all models and all definitions of discouragement. While the conditional treatment effects are smaller than the unconditional, the difference is not very large in most cases. Generally, selection bias appears to account for about 20% of the unconditional treatment effect. Taken as a group, these findings strongly suggest that use of the Internet by any technology, but especially Broadband and Public use, motivates the jobless to continue active job searches and stay in the labor force.

Table 6. Comparison of Estimated Average Treatment Effects

	Discouraged			Just Marginal		
	Unc. ATE	CATE	CATE-A	Unc. ATE	CATE	CATE-A
BLS Definitions						
Dialup	-39%*	-33%*	-30%*	2.4%	-8.1%	-13%
Broadband	-64%*	-54%*	-52%*	-6.5%*	-12%*	-15%*
Public Use	-69%*	-60%*	-68%*	-18%*	-13%*	-16%*
Information-related (Author) Definitions						
Dialup	-37%*	-28%*	-30%*	18%	-1.6%	-7.2%
Broadband	-61%*	-47%*	-40%*	15%	-1.5%	-13%*
Public Use	-67%*	-58%*	-57%*	2.2%	2.6%	-7.5%

Significance level (*, 5%) (**, 10%).

A skeptic may suggest that Internet use reduces discouragement because Internet use makes it easy to review online job classifieds and other employment lists. However, the BLS definition of job search requires active and serious search efforts and does not include "merely reading about job openings that are posted in newspapers or on the Internet."[57] The effect of Internet use, therefore, is more than casual job search.

As described in detail earlier, the treatment effects for discouraged are all statistically different from zero and sizeable. In all cases, the effect of Broadband at home is statistically equal to the effect of Public use, which is an important finding for public policy. Public use of the Internet is often disparaged in the policy debate, but such use is shown here to be as effective as home Broadband use for some purposes. Further, we find that Dialup use reduces labor market discouragement, though it does so to a lesser extent than does broadband use of either home or public types. The effect is large, typically reducing discouragement by about one-third. We are unaware of any other research quantifying differences in outcomes based on Broadband versus Public or Dialup use, so this paper may be the first.

---

[57] *How the Government Measures Unemployment*, *supra* n. 7.

E. *Additional Analysis*

This paper contains a significant amount of statistical detail. Nevertheless, these results represent only a portion of our complete analysis. There are a few additional points worth summarizing briefly. First, there remain some questions about the proper calculation of standard errors when applying propensity score trimming.[58] Since the propensity score is estimated in the first stage, it is not clear that the traditional standard errors reported by the econometric software are legitimate. To address this concern, we bootstrapped the standard errors for the CATE-𝔸 estimates (400 replications). The bootstrapped test statistics were nearly identical to those reported, so we do not summarize them here. This result is mildly surprising but also encouraging.

Second, the classification of the jobless and unemployed contains an "Other" category, which represents a non-trivial part of our sample. This open response is, of course, not obviously assignable to either the Information-related discouragement or the just marginal category. Eliminating the "Other" response altogether does not, however, impact the results substantially, particularly for discouragement since the category is not assigned to that group. Moving "Other" to the discouragement category also has very little effect on the results. Thus, our conclusions are not materially affected by alternative treatments of this response.

Third, the propensity scores here are estimated using multinomial logit. As noted by Lechner (2002), it is also possible to estimate the propensity scores in a pair-wise manner using logit regression. We did so and the propensity scores were highly correlated with those from the multinomial logit estimation. Trimming on the pair-wise propensity scores produced normalized differences below threshold for all but two of the covariates, and even then the remaining offending differences were only around 0.275. The estimated treatment effects (that is, CATE-𝔸) were essentially identical to those reported here. The "structural" approach to estimating the propensity scores has some benefits, so we employ that approach.

Fourth, when preliminary results from this PAPER were presented to members of the national broadband team at the FCC, a question was raised about whether the treatment effects may be different for low-income

58 *See, e.g.,* Angrist and Pischke (2009), *supra* n. 14 at 86-91.

Americans.[59] To answer that question, we extended the multinomial logit model by including as additional regressors interactions terms in each equation where we multiply the lowest income category (income ≤ 20,000) by the treatment dummy variables. A joint test of significance on these additional regressors is a direct test of an equal effect for low-income persons, with a null hypothesis that low-income persons have the same treatment effect. We cannot reject this hypothesis at traditional significance levels in any version of the model.[60] The treatment effect for low-income Americans is no different than average.

Finally, we have assumed that all potential confounders are observed and included in the covariates X_i. We recognize that some may challenge this assumption. We note, however, that the estimated treatment effects, while different, are consistently large in both the unconditional and conditional estimations. Given the three treatments, an attempt to capture unobservables using Instrumental Variables seems hopeless. Finding suitable instruments to explain differences in the choice of Dialup and Broadband at home and Public Internet use would be exceedingly difficult. Of course, we encourage others to try, but our own efforts were unfruitful in this regard. All empirical analyses are subject to criticism and this PAPER is no exception. We encourage policy makers to consider our results as one piece of a portfolio of evidence on this topic, and we encourage researchers to pursue alternative estimation procedures in an effort to improve the estimated treatment effects.

V. Conclusion and Extensions

In this POLICY PAPER, we test the hypothesis that Internet use reduces job search costs, thereby keeping individuals from abandoning the labor force altogether due to Discouragement or other reasons. Discouragement is defined both according to the BLS definitions, and by using a new, modified definition constructed to represent better the plausible effects of Internet use.

Using the BLS definitions of discouragement and just marginally attached persons, the econometric results uniformly suggest a large effect of Internet use on job search efforts. Using broadband at home or in a public setting reduces the

---

[59] http://fjallfoss.fcc.gov/ecfs/document/view?id=7020351947.

[60] Using BLS definitions, the test statistic for the Discouraged is (χ^2 = 2.85, Prob = 0.42) and for the Just Marginal is (χ^2 = 2.73, Prob = 0.43). For the Information-related definitions, the test statistic for the Discouraged is (χ^2 = 3.11, Prob = 0.38) and for the Just Marginal is (χ^2 = 1.24, Prob = 0.74).

probability of abandoning the labor market due to discouragement by about 50% (or more). The effect of Dialup use is smaller (about 30%), but is still statistically different from zero. Public use is at least as effective as home broadband use for both types of discouragement, a finding with potentially large policy ramifications.

Stronger results are obtained when switching to a newly-constructed "Information-Related" discouragement definition. Using this alternative definition, Internet use reduces the probability of discouragement but generally has no effect on the probability of just marginal attachment. This latter result is encouraging and expected given that the definition of discouragement was intended to reflect the likely consequences of Internet use on labor market choices.

Our results provide several useful insights for policymakers: on the demand-side, our results show that programs to promote Internet use may equate to more employment; on the supply-side, our results demonstrate that if the cost to provide broadband to every American home proves too expensive in the end, then the promotion of shared connections in unserved and underserved areas may, in fact, be as effective as home Broadband use for some purposes. Dialup Internet service can be an effective tool, though perhaps not as effective as Broadband use either at home or in public locations.

Appendix

Table A-1. Multinomial Logit Results (BLS Definitions)

	Unconditional		Full Sample		Trimmed Sample	
	Disc.	Marg.	Disc.	Marg.	Disc.	Marg.
Treatments						
Dialup (θ_D)	-0.543	-0.030	-0.511	-0.191	-0.473	-0.264
	(0.21)*	(0.12)	(0.23)*	(0.13)	(0.27)**	(0.16)
Broadband (θ_B)	-1.151	-0.202	-0.947	-0.286	-0.917	-0.336
	(0.15)*	(0.08)*	(0.18)*	(0.10)*	(0.21)*	(0.12)*
Public Use (θ_P)	-1.347	-0.386	-1.115	-0.320	-1.343	-0.370
	(0.25)*	(0.12)*	(0.26)*	(0.13)*	(0.33)*	(0.15)*
Kids at Home	…	…	-0.438	0.012	-0.236	0.025
			(0.20)*	(0.09)	(0.24)	(0.11)
Male	…	…	0.029	0.512	-0.163	0.562
			(0.15)	(0.08)*	(0.18)	(0.10)*
College	…	…	-1.516	-1.657	0.467	0.050
			(0.26)*	(0.16)*	(0.43)	(0.21)
No High School	…	…	-1.774	-1.523	0.729	0.274
			(0.21)*	(0.13)*	(0.49)	(0.24)
White	…	…	-1.596	-1.321	-0.151	-0.050
			(0.20)*	(0.13)*	(0.21)	(0.11)
Metro Area	…	…	0.224	0.097	-0.184	-0.262
			(0.22)	(0.11)	(0.19)	(0.11)*
Immigrant	…	…	0.476	0.367	-0.256	0.360
			(0.26)**	(0.13)*	(0.30)	(0.14)*
Veteran	…	…	-0.115	0.036	-0.534	0.260
			(0.15)	(0.08)	(0.45)	(0.24)
In School	…	…	-0.256	-0.270	-0.359	0.949
			(0.16)	(0.09)*	(0.49)	(0.18)*
Income 20	…	…	0.225	0.214	-0.694	-0.415
			(0.20)	(0.11)*	(0.25)*	(0.13)*
Income 20-40	…	…	-0.447	0.416	-0.435	-0.423
			(0.28)	(0.15)*	(0.22)*	(0.13)*
Income 40-60	…	…	-0.666	-0.466	-0.630	-0.228
			(0.18)*	(0.11)*	(0.32)*	(0.16)
Income 60-100	…	…	-0.533	-0.475	-1.246	-0.850
			(0.19)*	(0.11)*	(0.76)	(0.36)*
Age 20	…	…	-0.822	-0.211	-2.002	-1.872
			(0.27)*	(0.12)**	(0.36)*	(0.23)*
Age 20-40	…	…	-0.538	-0.177	-1.942	-1.593
			(0.26)*	(0.12)	(0.29)*	(0.19)*
Age 40-60	…	…	-0.657	0.942	-1.635	-1.482
			(0.33)*	(0.12)*	(0.28)*	(0.19)*
Constant	-1.737	-0.693	0.193	0.499	0.148	0.687
	(0.10)*	(0.07)*	(0.35)	(0.20)*	(0.55)	(0.30)*
n	4,229		4,229		2,562	
Pseudo-R^2	0.0118		0.068		0.066	
$\theta_D = \theta_B = \theta_P = 0$	68.52*	13.42*	37.75*	10.22*	29.55*	9.58*
$\theta_D = \theta_B$	7.52*	2.35	3.75**	0.63	2.63	0.21
$\theta_D = \theta_P$	7.14*	6.47*	3.74**	0.85	5.02*	0.33
$\theta_B = \theta_P$	0.56	2.84**	0.35	0.08	1.47	0.06

Robust standard errors in parenthesis. Significance level (*, 5%) (**, 10%).

Table A-2. Multinomial Logit Results (Author Definitions)

	Unconditional		Full Sample		Trimmed Sample	
	Disc.	Marg.	Disc.	Marg.	Disc.	Marg.
Treatments						
Dialup (θ_D)	-0.513	0.113	-0.424	-0.105	-0.475	-0.191
	(0.16)*	(0.13)	(0.17)*	(0.15)	(0.21)*	(0.18)
Broadband (θ_B)	-1.065	0.004	-0.787	-0.156	-0.665	-0.300
	(0.11)*	(0.09)	(0.14)*	(0.11)	(0.16)*	(0.13)*
Public Use (θ_P)	-1.295	-0.168	-1.049	-0.127	-1.033	-0.265
	(0.19)*	(0.13)	(0.19)*	(0.14)	(0.23)*	(0.16)
Kids at Home	…	…	-0.355	0.100	-0.204	0.088
			(0.14)*	(0.10)	(0.16)	(0.13)
Male	…	…	0.122	0.585	0.091	0.619
			(0.11)	(0.08)*	(0.13)	(0.11)*
College	…	…	0.350	0.020	0.453	-0.060
			(0.17)*	(0.11)	(0.31)	(0.23)
No High School	…	…	0.599	0.293	0.695	0.168
			(0.20)*	(0.14)*	(0.36)**	(0.26)
White	…	…	-0.030	0.034	-0.069	-0.068
			(0.12)	(0.09)	(0.15)	(0.12)
Metro Area	…	…	-0.240	-0.278	-0.244	-0.247
			(0.12)**	(0.10)*	(0.14)**	(0.12)*
Immigrant	…	…	0.042	0.300	-0.156	0.451
			(0.16)	(0.12)*	(0.22)	(0.15)*
Veteran	…	…	0.120	0.354	0.316	-0.124
			(0.20)	(0.16)*	(0.28)	(0.31)
In School	…	…	-0.790	1.135	-0.698	1.244
			(0.27)*	(0.12)*	(0.38)**	(0.20)*
Income 20	…	…	-0.359	-0.589	-0.199	-0.619
			(0.14)*	(0.11)*	(0.18)	(0.15)*
Income 20-40	…	…	-0.337	-0.564	-0.281	-0.501
			(0.15)*	(0.12)*	(0.17)**	(0.14)*
Income 40-60	…	…	-0.441	-0.248	-0.248	-0.320
			(0.19)*	(0.13)**	(0.23)	(0.18)**
Income 60-100	…	…	-0.532	-0.142	-1.719	-0.632
			(0.20)*	(0.13)	(0.75)*	(0.37)**
Age 20	…	…	-1.465	-1.722	-1.698	-2.022
			(0.22)*	(0.17)*	(0.29)*	(0.24)*
Age 20-40	…	…	-1.461	-1.638	-1.572	-1.722
			(0.17)*	(0.14)*	(0.24)*	(0.21)*
Age 40-60	…	…	-1.087	-1.549	-1.258	-1.682
			(0.16)*	(0.14)*	(0.23)*	(0.21)*
Constant	-1.123	-1.048	0.074	0.313	0.015	0.632
	(0.08)*	(0.08)*	(0.28)	(0.21)	(0.42)	(0.32)**
n	4,229		2,562		2,562	
Pseudo-R^2	0.016		0.077		0.071	
$\theta_D = \theta_B = \theta_P = 0$	106.2*	3.76	48.86*	2.11	29.47*	5.48
$\theta_D = \theta_B$	11.14*	0.82	4.45*	0.15	0.81	0.42
$\theta_D = \theta_P$	12.28*	3.56**	7.28*	0.02	3.96*	0.13
$\theta_B = \theta_P$	1.44	2.24	1.66	0.05	2.37	0.05

Robust standard errors in parenthesis. Significance level (*, 5%) (**, 10%).

Table A-3. Normalized Differences

Sample: (Full, A)

	0-D	0-B	0-P	D-B	D-P	B-P
Kids at Home	0.02, 0.00	0.13, 0.14	0.02, 0.03	0.15, 0.13	0.04, 0.03	0.11, 0.11
Male	0.09, 0.02	0.08, 0.05	0.05, 0.01	0.01, 0.03	0.04, 0.03	0.04, 0.06
College	0.09, 0.02	0.01, 0.02	0.02, 0.04	0.08, 0.04	0.07, 0.02	0.01, 0.06
No High School	0.29#, 0.10	0.07, 0.01	0.35#, 0.13	0.21, 0.12	0.06, 0.03	0.28#, 0.15
White	0.23, 0.12	0.02, 0.05	0.22, 0.07	0.25#, 0.17	0.01, 0.06	0.24, 0.11
Metro Area	0.12, 0.09	0.08, 0.03	0.10, 0.10	0.05, 0.06	0.22, 0.19	0.17, 0.13
Immigrant	0.10, 0.06	0.14, 0.10	0.12, 0.08	0.04, 0.04	0.03, 0.02	0.01, 0.02
Veteran	0.06, 0.02	0.16, 0.10	0.06, 0.01	0.10, 0.12	0.00, 0.03	0.10, 0.09
In School	0.22, 0.16	0.19, 0.21	0.32#, 0.21	0.03, 0.06	0.11, 0.05	0.13, 0.01
Income 20	0.45#, 0.25	0.07, 0.02	0.49#, 0.23	0.37#, 0.23	0.04, 0.02	0.41#, 0.20
Income 20-40	0.05, 0.06	0.04, 0.04	0.08, 0.02	0.09, 0.10	0.13, 0.09	0.04, 0.01
Income 40-60	0.23, 0.15	0.09, 0.06	0.19, 0.16	0.14, 0.09	0.04, 0.02	0.10, 0.11
Income 60-100	0.30#, 0.10	0.11, 0.05	0.42#, 0.09	0.20, 0.05	0.13, 0.01	0.33#, 0.04
Age 20	0.12, 0.06	0.15, 0.17	0.14, 0.09	0.03, 0.12	0.03, 0.03	0.01, 0.09
Age 20-40	0.11, 0.04	0.13, 0.09	0.02, 0.02	0.24, 0.13	0.09, 0.07	0.15, 0.06
Age 40-60	0.03, 0.05	0.09, 0.16	0.00, 0.04	0.12, 0.12	0.02, 0.01	0.10, 0.13
Married	0.19, 0.08	0.04, 0.03	0.20, 0.08	0.23, 0.11	0.01, 0.00	0.24, 0.11
HH Size	0.16, 0.03	0.18, 0.15	0.26#, 0.14	0.03, 0.13	0.11, 0.11	0.08, 0.02
Northeast	0.04, 0.00	0.02, 0.07	0.09, 0.04	0.06, 0.07	0.05, 0.04	0.11, 0.11
Midwest	0.01, 0.04	0.03, 0.01	0.00, 0.04	0.03, 0.04	0.00, 0.00	0.03, 0.03
South	0.04, 0.05	0.07, 0.01	0.12, 0.03	0.03, 0.04	0.08, 0.09	0.05, 0.04

0 = No Internet; D = Dialup; B = Broadband; P = Public Use.
# Exceeds 0.25

Table A-4. Propensity Score Model and Descriptive Statistics

	ML for Propensity Score (Base = No Internet)			Sample Means	
	Dialup	Broadband	Public Use	Full	A
Kids at Home	-0.007	0.397	0.052	0.244	0.265
	(0.16)	(0.14)*	(0.12)		
Male	0.197	0.086	0.061	0.512	0.519
	(0.12)	(0.11)	(0.09)		
College	-1.156	-0.934	-1.647	0.619	0.722
	(0.20)*	(0.21)*	(0.16)*		
No High School	-2.092	-1.197	-2.804	0.221	0.215
	(0.23)*	(0.23)*	(0.19)*		
White	0.595	-0.030	0.590	0.731	0.724
	(0.14)*	(0.11)	(0.10)*		
Metro Area	-0.440	-0.240	0.263	0.790	0.732
	(0.13)*	(0.12)**	(0.11)*		
Immigrant	-0.191	-0.611	-0.515	0.140	0.137
	(0.17)	(0.16)*	(0.13)*		
Veteran	-0.421	-0.550	-0.304	0.072	0.039
	(0.23)**	(0.26)*	(0.17)**		
In School	1.330	0.821	1.787	0.136	0.098
	(0.23)*	(0.21)*	(0.18)*		
Income 20	-0.895	-0.079	-0.988	0.267	0.269
	(0.17)*	(0.14)	(0.12)*		
Income 20-40	0.071	-0.050	-0.265	0.221	0.299
	(0.16)	(0.16)	(0.12)*		
Income 40-60	0.790	0.416	0.562	0.122	0.125
	(0.20)*	(0.21)*	(0.16)*		
Income 60-100	1.315	0.620	1.516	0.136	0.028
	(0.25)*	(0.27)*	(0.21)*		
Age 20	0.281	1.345	0.832	0.168	0.178
	(0.26)	(0.29)*	(0.20)*		
Age 20-40	-0.102	1.142	0.582	0.409	0.434
	(0.20)	(0.25)*	(0.16)*		
Age 40-60	0.031	0.832	0.496	0.319	0.317
	(0.19)	(0.24)*	(0.15)*		
Married	0.328	-0.052	0.398	0.342	0.298
	(0.15)*	(0.14)	(0.11)*		
HH Size	0.291	0.349	0.682	0.864	0.857
	(0.19)	(0.17)*	(0.14)*		
Northeast	-0.030	-0.261	0.078	0.200	0.179
	(0.18)	(0.17)	(0.13)		
Midwest	-0.004	-0.151	-0.006	0.240	0.251
	(0.17)	(0.15)	(0.12)		
South	-0.017	-0.332	-0.200	0.311	0.312
	(0.16)	(0.14)*	(0.12)**		
Constant	-0.105	-0.685	0.497		
	(0.34)	(0.37)**	(0.27)**		
n		4,229			
Pseudo-R^2		0.135			

Robust standard errors in parenthesis. Significance level (*, 5%) (**, 10%).