Effect of Internet Access on Job Search Costs

Matthew Borelli

4/25/2020

Introduction

The 21^{st} century has seen an incredible expansion of internet access in the United States. As of 2018, approximately 110 million households have broadband internet access, a 1564% increase over the 7.07 million households in 2000 (Statista). The internet has made many tasks easier to handle for Americans, from cutting out the need for bank trips to accessing large repositories of the world's knowledge within seconds. For this project, we want to examine the causal effect of increased internet access on job search costs. Specifically, we want to see if increased internet access has caused a decrease in the amount of time spent on job search activities including filling out job applications, sending out resumes, and contacting employers. The motivating idea behind this potential causal link is that the internet expidites the process of job search activities through websites like LinkedIn and text processors like Microsoft word. Therefore, as more people gain internet access across the country, we should expect to see that the amount time spent on job search activities has fallen.

We consider the expansion of internet access to be a natural experiment, where many people across the U.S. gained access to broadband internet over a short period of time. Particularly, we look at the years 2003 to 2007, the period of most rapid growth in the 21^{st} century. In 2003 27.74 million households had broadband internet access subscriptions, and in 2007 that number had increased to 71.7 million households. This seemingly exogenous internet access shock, cause by the decline in prices and subsequent popularity of home personal computers, could effect job search times. This is important as even though the percentage of the population with internet access has continually increased over the years, many of the U.S.'s poorest citizens still cannot afford internet access. If internet access does cause decreases in job search costs, poorer citizens are the most affected group as they would want to find better jobs but fall behind other applicants who do have internet access. Therefore, if a causal link is established, there stands to be good reasons to promote policies that expand broadband access, such as municipally-owned ISP's or expansion of public library services.

Data

The first data set used for this project is from the American Time Use Survey (ATUS). A randomly selected subpopulation of people who partake in the Current Population survey are asked to log a diary for a 24-hour period that details all of their activities with time amounts. However, ATUS does not have the same people year-to-year, so this is panel data and not longitudinal data. Thus, the data is aggregated by state and year means in order to create more consistent year-to-year data. Of interest to this topic is the time-use category of "job search activities". According to the ATUS-X data extract builder, this time category includes but is not limited to the following actions:

- Asking about job openings
- Checking vacancies
- Filling out job application
- Making phone calls to prospective employers

- Meeting with headhunter/temporary agency
- Picking up job applications
- Researching an employer
- Sending out resumes
- Submitting applications
- Writing/updating resumes

Each person in the data extract for this project has a listed time in minutes for job search activities, ranging from 0 to 670 minutes of search time. In the raw extract, unemployed and employed people are together with other categories such as students and retirees that would skew the data. As a processing step, we are going to consider four different subsets of the data:

1. Employed
2. Unemployed
3. Employed with non-zero job search time
4. Unemployed with non-zero job search time

The first two will include people that did not partake in any job search activities, so estimations will not only account for changes in time spent on job search activities due to internet access changing the amount of time needed, but also of more people deciding to spend time on job search activities due to the newfound ease of access provided by broadband access. This implies that in trying to estimate the causal effect of internet access on job search costs, we might be estimating two opposite causal effects together. Internet access lowers the amount of time needed to do the same amount of job searching, but could actually increase the amount of time that people spend searching on average due to how easy it is. To try and isolate change caused by internet access lowering the amount of time needed, we create subsets for employed and unemployed workers who search for work, so that when aggregated to the state and year level means we are accounting for any potential people who chose to enter the job market.

The other data source utilized in this project comes from Cunningham and Kendall's paper "Prostitution 2.0: The changing face of sex work". They used a data set that contains information on the number of high-speed broadband connections for each state for the years 1999-2007, of which we will use the data for 2003-2007. To note, we have to drop Hawaii from the data, as they have no recorded information for the number of residential connections until 2006. We use the number of residential connections per household as our measure of internet accesss. Expressing this number as a ratio controls for any changes in populations that might occur over the time period. Data also exists for the total number of broadband connections which includes businesses, but this is not relevant to this project.

Model

We want to test for two causal effects: the effect of internet access on entering the job search market, and the effect of internet access on job search costs. To model the first effect, we regress the percentage of people who spent time on job search activities. For both the set of employed and unemployed people, we estimate this using the following equations:

$$W_{its} = a + \theta(Broadband_{ts}) + X_{its} + d_t + d_s + e_{its}$$

$$W_{its} = a + \theta(Broadband_{ts}) + X_{its} + d_t + d_s + d_s * trend + e_{its}$$

Where W_{its} is the percentage of people who spend any time on job search activities. The i subscript denotes each obsetvation, the t subcript denotes the year, and the s subscript denotes the state. $Broadband_{ts}$ is the percent of residential households with broadband access for a given state and year. X_{its} is the vector of additional control variables including race, income level, and gender f. There are two fixed effects parameters: d_t for year fixed effects, and d_s for state fixed effects. The residuals from the regression are represented as

Table 1: Covariates Over Time for Employed Respondents

Variable	**2003**	**2004**	**2005**	**2006**	**2007**
Age	41.57	41.93	41.55	41.64	41.92
Male	0.5	0.49	0.49	0.48	0.49
Black	0.11	0.1	0.1	0.12	0.12
Hispanic	0.11	0.11	0.12	0.13	0.13
$0 - $25000	0.15	0.14	0.15	0.15	0.13
$25000 - $60000	0.36	0.34	0.33	0.33	0.33
$60000 - $100000	0.38	0.25	0.24	0.24	0.24
$100000+	0	0.14	0.15	0.17	0.18
Married	0.57	0.58	0.55	0.54	0.54

e_{its}. The two equations are mostly similar, with the only difference being the addition of the *trend* (set to 1 for 2003 through 5 for 2007) interacted with state fixed effects in the second equation. The causal effect is estimated by θ, the coefficient for household broadband access percentage.

To estimate the causal effect of internet access on job search costs, we will utilize ordinary least squares regression with state and year fixed effect. For each of the four groups mentioned above, we will estimate the following equations:

$$Y_{its} = a + \delta(Broadband_{ts}) + X_{its} + d_t + d_s + e_{its}$$
$$Y_{its} = a + \delta(Broadband_{ts}) + \beta X_{its} + d_t + d_s + d_s * trend + e_{its}$$

For these equations, the causal effect of interest is δ, the coefficient for the effect of residential broadband connectionon on job search time. By using state and year fixed effects, this will be an unbiased estimator of the causal effectl

Covariate Balance

Before estimating the causal effect parameters, we want to check that the covariates we are controlling for are balanced over the data, both for the employed and unemployed subpopulations. The following tables display the averages for the covariates in X_{its}: Age, sex, race, income brackets, and marriage.

OLS Estimation

Conclusion

Citations

Sandra L. Hofferth, Sarah M. Flood, Matthew Sobek and Daniel Backman. American Time Use Survey Data Extract Builder: Version 2.8 [dataset]. College Park, MD: University of Maryland and Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D060.V2.8

thttps://www.statista.com/statistics/183614/us-households-with-broadband-internet-access-since-2009/

Table 2: Covariates Over Time for Unemployed Respondents

Variable	**2003**	**2004**	**2005**	**2006**	**2007**
Age	36.52	38.97	37.18	37.93	38.18
Male	0.47	0.47	0.43	0.42	0.46
Black	0.21	0.2	0.26	0.24	0.22
Hispanic	0.12	0.13	0.16	0.18	0.16
$0 - $25000	0.35	0.34	0.39	0.37	0.38
$25000 - $60000	0.33	0.33	0.3	0.27	0.31
$60000 - $100000	0.22	0.16	0.13	0.17	0.14
$100000+	0	0.04	0.06	0.06	0.09
Married	0.33	0.39	0.31	0.31	0.29

Table 3: OLS Estimation for Populations

Outcome	**Employed**		**Unemployed**	
	1	**2**	**1**	**2**
res_percent	0.000223	0.0425	0.02862	119.2719
	(0.03433)	(0.1713)	(0.01933)	(112.7149)
Controls	x	x	x	x
State/Year FE		x		x

Standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 4: OLS Estimations for People Who Looked For Work

Outcome	**Employed**		**Unemployed**	
	1	**2**	**1**	**2**
res_percent	-26.368	-448.117	96.1271	116.4327
	(66.269)	(244.812)	(64.807)	101.0514
Controls	x	x	x	x
State FE		x		x
Year FE				

Standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01