VIEWPOINTS

MILKEN INSTITUTE
CENTER FOR JOBS AND HUMAN CAPITAL

Internet Usage Effect on Educational Attainment: Evidence of Benefits

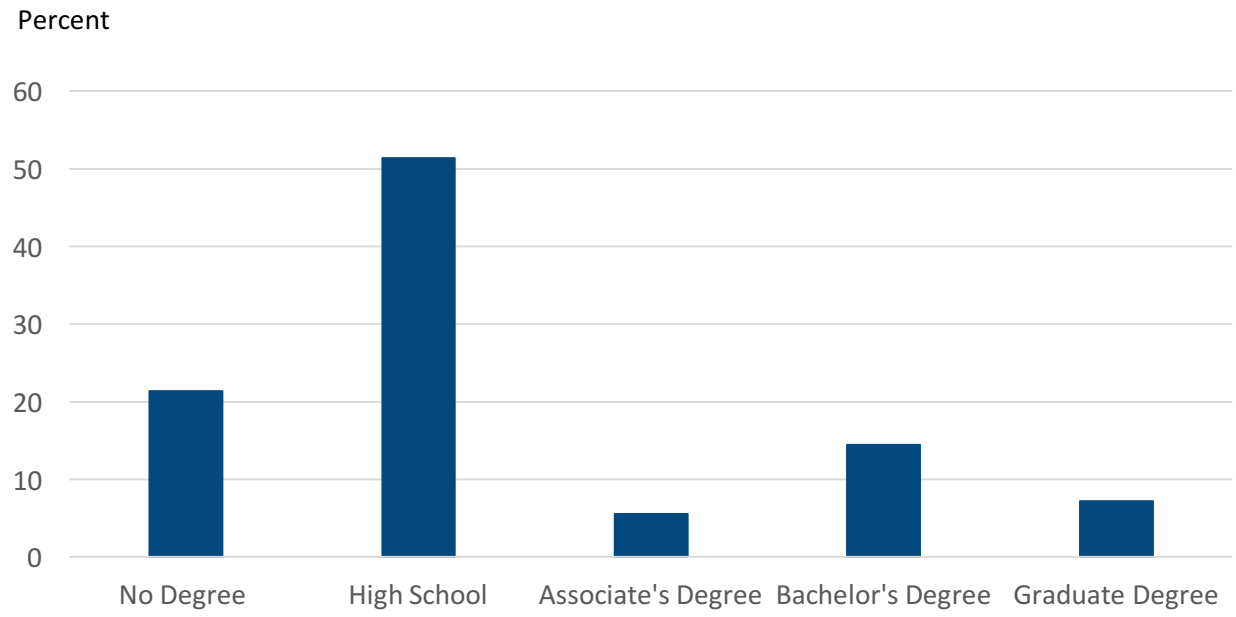
Joe Lee
July 2017

Introduction

Since the internet became a widely adopted platform for information exchange, there has been a variety of research into the effect the internet has on education. Some research finds increased success rates for minorities in K-12 when provided computers, while other research focuses on the role of mobile platforms and social media.[1,2,3] Education has seen large changes due to the internet, including modifications to curriculum and differences in the way education is delivered. This paper provides evidence of a positive relationship between the hours spent on the internet per week and educational attainment. This analysis shows that the relationship is robust to specification changes and displays stability in both magnitude and directionality of the statistical relationship. Controlling for outside influences in line with previous research, there is evidence that the internet can be a tool to improve educational outcomes. Expanding the scope of educational attainment research will give policy-makers a better understanding of how to increase economic opportunity for all.

[1] Fairlie, W, Robert. "Academic achievement, technology and race: Experimental evidence" Economics of Education Review 31, (2012): pg 663-679. https://people.ucsc.edu/~rfairlie/papers/published/eer%202012%20-%20minority%20computers.pdf
[2] Fairlie, W, Robert, London, A, Rebecca. "The Effects of Home Computers on Educational Outcomes: Evidence from a Field Experiment with Community College Students" The Economic Journal 122 (2012), http://escholarship.org/uc/item/4v87w469
[3] Bulman, George, Fairlie, W, Robert. "Technology and Education: Computers, Software, and the Internet" NBER Working Paper Series, (2016) no 22237, http://www.nber.org/papers/w22237.pdf

FIGURE 1. Highest Degree Attained



Source: U.S. General Social Survey, 2000-2016, highest degree attained, 18-89

Recently, past academic work has been used to qualify proposed policies seeking to increase educational attainment in pursuit of expanding economic prosperity. In 1974, Jacob Mincer's research produced a correlation that illustrated that education and experience have a positive effect on income.[4,5] Over time, changes in the economy created a need for a more skilled workforce. This shift is predicated on the expanding role of technology and the ever-increasing need for specialization. Acquiring skills through education can help people adapt to these changes. Figure 1 shows that between 2000 and 2016, 27.1 percent of the population had attained a bachelor's degree or higher. The internet has fundamentally transformed how people experience cultural and economic changes. As a new medium for individual opinions and information exchange, the internet has changed the way we communicate by increasing the speed and reach of ideas.

Internet usage has affected how economic activity concentrates in clusters, and industry interactions within these clusters.[6,7] As the shift from manufacturing occupations to knowledge-based occupations has changed the employment landscape in metro clusters, urban concentrations of high skilled service jobs have grown. The continued rise of science and technology has developed new fields in healthcare,

[4] Mincer, Jacob, and Solomon Polachek. "Family Investments in Human Capital: Earnings of Women." Journal of Political Economy 82, no. 2 (1974): 76-108. http://www.jstor.org/stable/1829993.
[5] Mincer, Jacob. "Schooling and Earnings" NBER (1974): http://www.nber.org/chapters/c1765.pdf
[6] Leamer, E, Leamer, Storper, Michael. "The Economic Geography of the Internet Age" Journal of International Business studies 32, no.4 (2001): 641-665, http://www.jstor.org/stable/pdf/3069470.pdf
[7] Forman, Chris, Goldfarb, AVI, Greenstein, Shane. "Geographic location and the diffusion of Internet technology" Electronic Consumer Research and Applications, 4 (2005) pg 1-13, http://s3.amazonaws.com/academia.edu.documents/43895172/ecra.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1490915087&Signature=iG8J%2F2aQ0LM3yngGQRMs4oPuBFI%3D&response-content-disposition=inline%3B%20filename%3DGeographic_location_and_the_diffusion_of.pdf

computer science, and engineering. These fields create new economic power centers and redefine the composition of skills needed to increase economic opportunity throughout the U.S. economy. Through education we can impart skills to an expanding workforce.

As a society, we have created new opportunities as technology advances, and the internet continues to redefine how the U.S. economy functions. The rise of social media, cord cutting, and video games are all observable changes. Though we have created many of new ways to spend free time, the internet has done more than just give us new ways to consume content.[8,9] The internet has reduced the cost of transmitting information by lowering the friction of access to that information.[10] New capabilities, such as an institution's capacity for posting lectures online to websites offering online certificates, is evidence that access to knowledge has increased. Beyond institutions and other educational outlets, people can learn musical instruments or learn to code through online tutorials. The internet has given a space to those people who wish to share information on any topic, both positive and negative.

The genesis of the internet swiftly and permanently ushered in a new era. However, there are sections of the U.S. that still have limited internet access. Previous research examining the effect of access to the internet in a residence shows a positive effect on students choosing to purse higher education.[11] However, there are notable differences in internet access and speed when comparing urban population centers and rural areas. Figure 2 shows 22.8 percent of U.S. households do not have some type of internet subscription. This includes dial-up, DSL, broadband, and mobile internet services. Cable internet subscriptions represent 41 percent of U.S. households, while the second largest subscription type is mobile broadband (38.1 percent). In major urban centers 12 percent of households don't have access to the internet compared with 20.5 percent of rural households. The differential of internet speed and access between urban and rural areas reduce the beneficial impact of internet usage on educational attainment of the population.

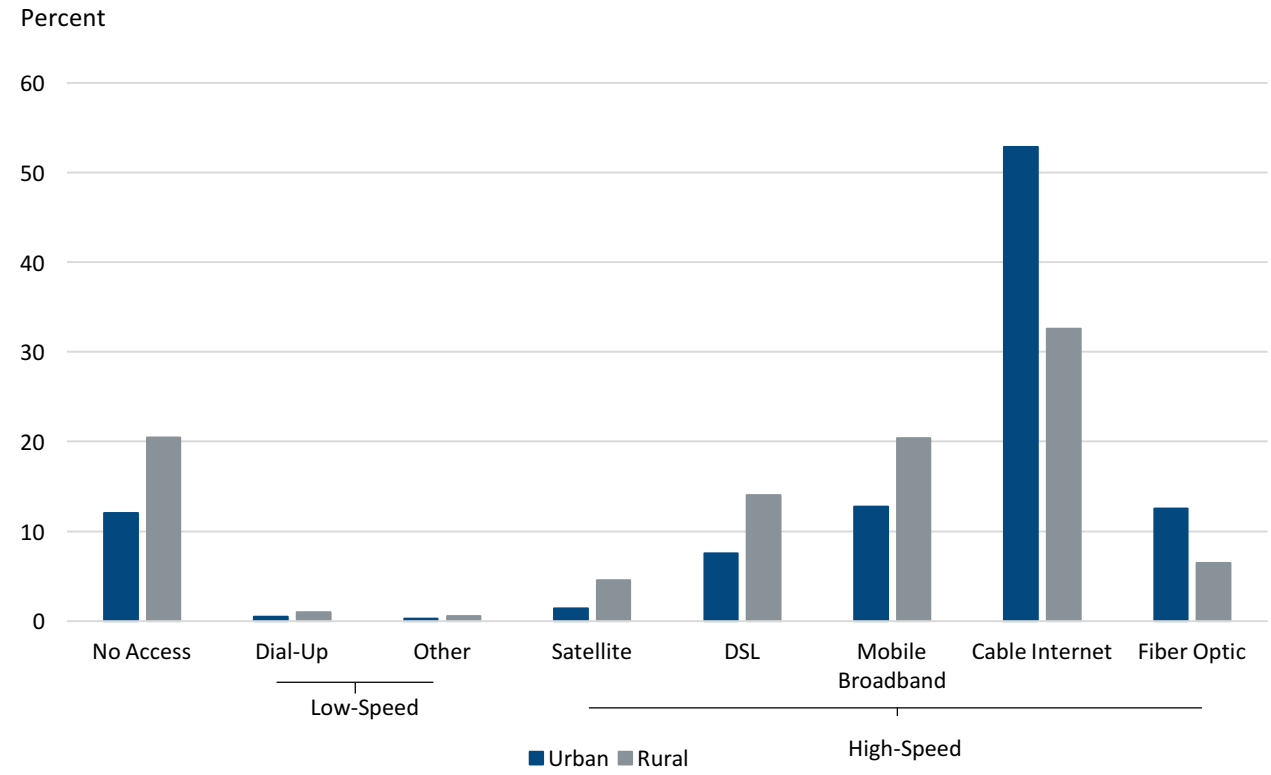
[8] Junco, Reynol. "The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement" Computers & Education 58, (2011): 162-171. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.716.205&rep=rep1&type=pdf

[9] Aguiar, et al., "Leisure Luxuries and the labor supply of Young Men", NBER working paper series, no 23552, (2017): http://www.nber.org/papers/w23552.pdf

[10] Torberg, Falch, Lujala, Päivi, Strøm, Bajarne. "Geographical constraints and educational attainment" Regional Science and Urban Economics 43, no. 1 (2012): 164-176. http://www.sv.ntnu.no/iso/torberg.falch/Articles/Geographical%20constraints.pdf

[11] Bowden, P, Mark, Doughney, James. "The importance of cultural and economic influences behind the decision to attend higher education" The Journal of Soci-Economics 41, (2012): pg 95-103. http://vuir.vu.edu.au/9370/1/Cultural%20and%20econ%20influences%20and%20decision%20to%20attend%20he.pdf

FIGURE 2. Urban vs. Rural Internet Subscriptions of U.S. Households



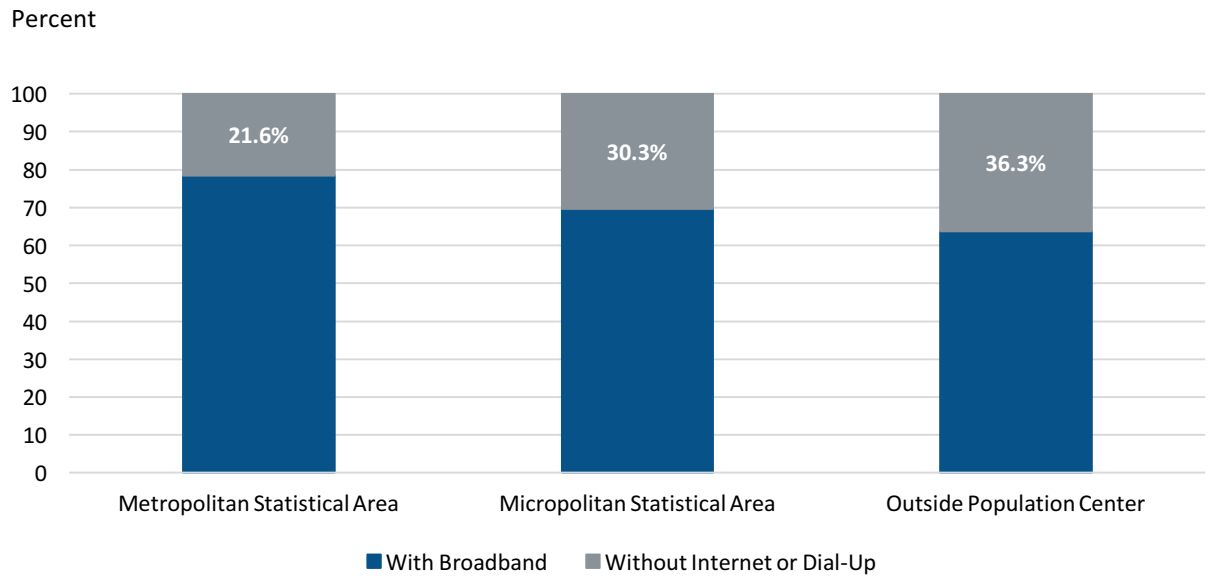
Source: IPUMS American Community Survey 2015

Urban vs. Rural Divide in American Internet Access

The U.S. has increasing concentrations of economic activity in urban areas, making larger metropolitan areas major economic forces. This concentration of economic activity generates innovation and can drive growth, which increasingly shifts resources toward these clusters. Slow internet reduces the efficiency of knowledge transfer through internet usage and is more common in rural areas of the U.S. than in urban areas.

Online educational resources can help reduce the cost of education through minimizing boarding, travel, materials, allocation of time, and facility expenditures due to online educational resources, however the effectiveness and quality of these programs are still subject to further inquiry. Some high school, college, and graduate programs are now conducted entirely online, opening access to education in a new way. Figure 3 shows that about 78 percent of households in rural America have access to high-speed internet capable of supporting online classrooms. Broadening access also allows people to overcome both geographic barriers and time constraints through increased flexibility of scheduling. The ability for someone in rural America to obtain a degree from a major institution on the other side of the nation provides an avenue for economic opportunity not previously available. However, the benefits of this new education access are lost to over 1 in 10 urban households and 1 in 5 rural households without an internet connection (see Figure 3).

FIGURE 3. Households with Broadband Internet Subscriptions by Geographic Category



Source: Census American Community Survey 2015

A student can quickly access program requirements and curriculum information on the internet. Access to a college's admission requirements allows potential applicants to assess prerequisite conditions to decide what courses they should complete or what SAT/ACT scores they need to maximize their likelihood of acceptance. This reduction of time and resources required to obtain information is provided by the ease of access to information. Even during the college experience, quick and user-friendly communication platforms enable peer interaction and teacher support.[12] Through increased peer interaction and access to teachers, the internet has increased the speed of communication and the spread of knowledge.

Data and Results

General Social Survey

The main results of this paper use the GSS National Data Program for the Social Sciences, which is a pooled cross-section of the U.S. population. A Bayesian Ordered Probit Model (a discrete choice model for the probability of predicting the choice where the order of outcomes is important) is used for this analysis examining the relationship between educational attainment and internet usage. The sample period from 2000 to 2016 encapsulates the rise of the internet into its current form. It covers major events from the collapse of the Dot Com Bubble through the Great Recession and its subsequent recovery years. Controlling for three major groups of variables, the model shows a positive relationship between the hours spent on the internet per week and the highest degree attained. The graduate

[12] Gikas, Joanne, grant, M, Michael. "Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media" Internet and Higher Education 19, (2013): 18-26 http://s3.amazonaws.com/academia.edu.documents/38919269/gikas_grant_mobile_devices.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1490891993&Signature=Zze0kmHjwaeqirWVH0Rek2NrayE%3D&response-content-disposition=inline%3B%20filename%3DMobile_Computing_Devices_in_Higher_Educa.pdf

degree category includes any higher education degree requiring a bachelor's degree for program entry. Family influences, demographics, time, and regional effects have been controlled for with the intent to improve this analysis by accounting for other confounding influences that might wash out the effect of internet usage. The main findings of this paper are reported in column three of Table 2.

Individual and Household Characteristics

The first set of control variables are related to household characteristics. These variables capture important external influences on an individual. The analysis controls for both parents' highest degree attained to explain how educational attainment is valued.[13,14] Both parent education variables are highly significant and show a positive relationship with educational attainment. These variables are highly correlated between generations and influence the expected outcome of education attainment in this model.[15]

Real individual income also shows a positive and highly significant impact on educational outcomes.[16] Higher personal income increases access to technology as well as someone's ability to invest in education. When looking at younger age ranges it is important to remember that people will generally have lower levels of income. The benefits from private education and financial support for higher education are increasingly influential. Resources devoted to higher education increases the likelihood of graduation by reducing financial constraints. This result is supported by previous research on the positive relationship between parent educational attainment, and financial resources available for education.[17]

In this vein, increases in household size have a highly significant negative effect on educational outcomes. As the number of people in a household increases, fewer resources are available to each person for education investment. The opportunity cost of an individual is increased as age increases because people will have higher incomes but also the need to maintain an income as household sizes increase financial constraints.

---

[13] Haveman, Robert,Wolfe, Barbara. "The Determinants Of Childern's Attainments: A Review of Methods and Findings" Journal of Economic Literature 33, no. 4 (1995): 1829-1878
https://www.researchgate.net/profile/Barbara_Wolfe/publication/44819892_The_Determinants_of_Children%27s_Attainments_A_Review_of_Methods_and_Findings/links/0c96053bc08dcac165000000/The-Determinants-of-Childrens-Attainments-A-Review-of-Methods-and-Findings.pdf

[14] Hung, Jin. "Intergenerational transmission of educational attainment: The role of household assets" Economics of Education Review 33, (2013): 112-123.
https://www.researchgate.net/profile/Jin_Huang23/publication/257106204_Intergenerational_transmission_of_educational_attainment_The_role_of_household_assets/links/550d8eb40cf2ac2905a7e724.pdf

[15] Taubman, Paul. "Role of Parental Income in Educational Attainment." The American Economic Review 79, no. 2 (1989): 57-61.
http://www.jstor.org/stable/1827730.

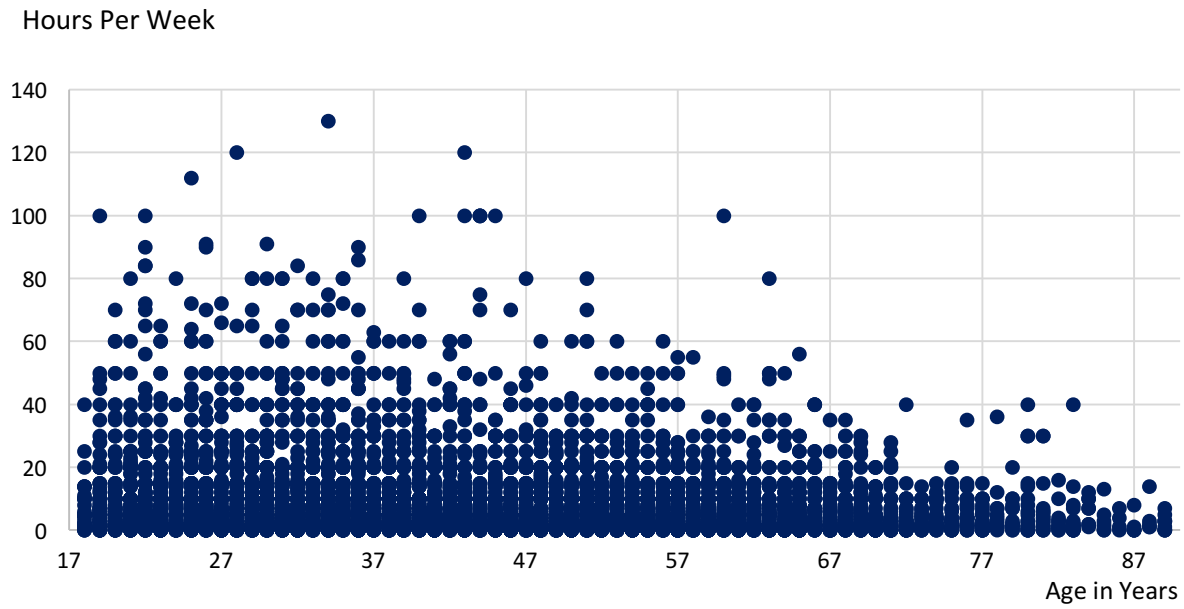
[16] Bowden, P, Mark, Doughney, James. "The importance of cultural and economic influences behind the decision to attend higher education" The Journal of Soci-Economics 41, (2012): pg 95-103.
http://vuir.vu.edu.au/9370/1/Cultural%20and%20econ%20influences%20and%20decision%20to%20attend%20he.pdf

[17] Hung, Jin. "Intergenerational transmission of educational attainment: The role of household assets" Economics of Education Review 33, (2013): 112-123.
https://www.researchgate.net/profile/Jin_Huang23/publication/257106204_Intergenerational_transmission_of_educational_attainment_The_role_of_household_assets/links/550d8eb40cf2ac2905a7e724.pdf

Demographic Characteristics

The set of variables representing demographics are intended to isolate cultural and generational influences that affect educational attainment. This analysis controls for age and age-squared. Age is highly significant, serving several purposes in this analysis. Age-squared is not significant which is unexpected (see appendix). The first use of age is to control for the increased likelihood of higher educational attainment due to increases in age. The second reason is to address concerns of directional causality of higher education leading to higher internet use. Age is correlated to both experience and seniority of occupations in combination with personal income, which is highly correlated to occupational choice. The use of these two variables in intended to mitigate the impact of multi-directional causality of internet usage and educational attainment. The third is to isolate the effect of people of different ages' tendency to use the internet.

FIGURE 4. Age vs. Hours on Internet Per Week



Source: General Social Survey 2000-2016

At a 99 percent confidence level, females have a greater likelihood of obtaining a higher degree. This is reflected by the larger female enrollment numbers in higher education institutions.[18,19] The race variable includes all non-white ethnicities and is highly significant with a positive relationship to higher educational attainment. While encompassing all minorities is not a great indicator and could be reflecting that the sample is 19 percent nonwhite and family income is positively skewed. For this data set, white individuals have a higher mean Ln(personal income) which is statistically significant at the 99 percent confidence level see Table 3.

[18] National Center for Education Statistics, Table 318.30 (2013 to 2016) https://nces.ed.gov/programs/digest/d13/tables/dt13_318.30.asp
[19] National Center for Education Statistics, Table 310 (2011) https://nces.ed.gov/programs/digest/d12/tables/dt12_310.asp

The reason for the aggregation is that all other minorities other than blacks are already aggregated in this dataset. In order to address influence of cultural differences in educational attainment, ethnicity, and religion variables are included.

Religion has been broken out into binary variables that are not aggregates of multiple religious doctrines. These variables are normalized against Protestants because they are the largest religious group in the sample. Judaism, Buddhism and Hinduism are all highly significant and have a positive relationship with educational attainment. These demographic variables are included to help control for cultural influences not accounted for by the nonwhite variable. [20]

Regional and Time Effects

The third group of variables used as controls are region and time effects. The regions are related to zip code regions which are normalized against the northeast. All regional dummy variables are highly significant and negatively signed, which indicates the northeast has an advantage for increased educational attainment. This outcome may be due to a number of factors such as population density, concentration of per capita GDP, employment opportunities that draw in people with higher levels of education, large numbers of higher educational institutions providing local human capital to labor markets due to in-state tuition lowering costs, or more opportunity for placement because of the amount of schools. The region variables are included to capture the variation of educational opportunity across the U.S.

There are macroeconomic effects to be controlled for as well, as the sample data covers two major recessions. Starting in 2000, I have included all years but 2008 due to a lack of internet usage data for that year. All time variables are significant and negatively signed. The time variables are included because periods of recession caused differences in the role of the internet in the economy. The structure of the data allows for new people to enter the sample which includes younger people who have earlier contact with the internet. The addition of year variables allows for these effects to be accounted for.

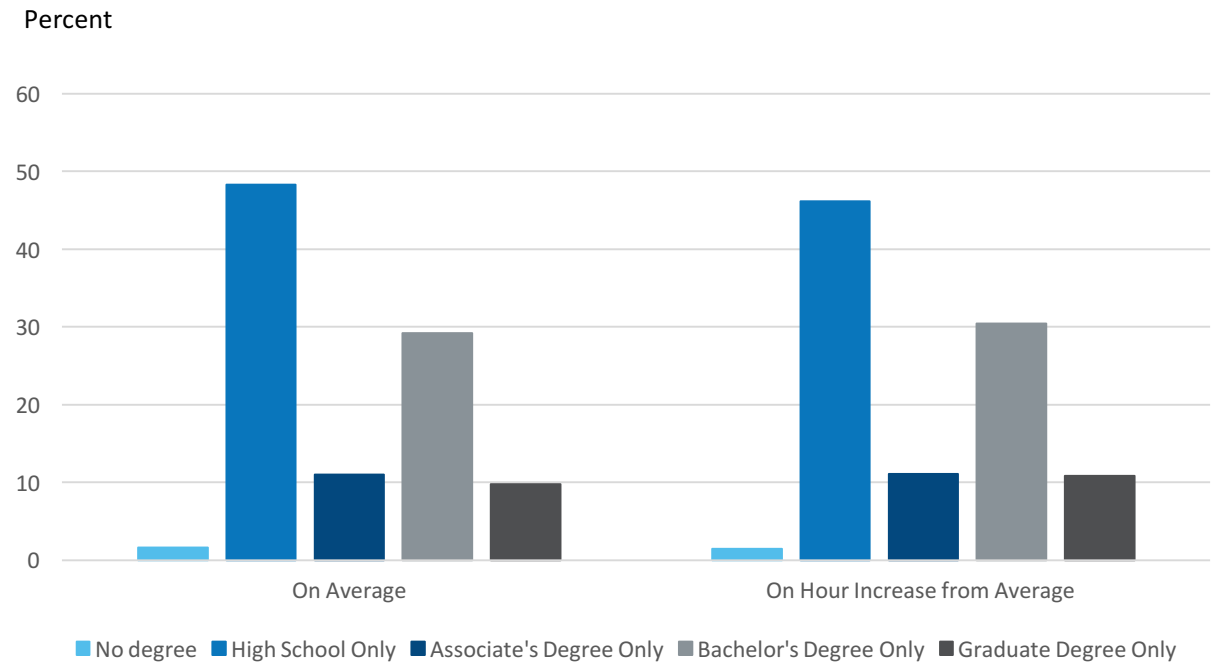
Internet Usage

The controls in this model are being used to isolate the effect that hours per week online has on educational attainment outcomes. The family and demographic characteristics are used to condition for external influences while time and regional effects are used to condition for more abstract factors. This analysis addresses concerns about possible reverse causality of internet usage and educational attainment by controlling for age and income. Income is highly correlated to occupation and helps control for differences in exposure to technology and the internet. This issue of directional causality is a common problem with all econometrics but has been addressed as much as possible with the data available for this analysis. The necessary robustness checks of these influences confirm that the effect of average weekly internet usage is significant and demonstrates a large positive effect on the probability

[20] Bowden, P, Mark, Doughney, James. "The importance of cultural and economic influences behind the decision to attend higher education" The Journal of Soci-Economics 41, (2012): pg 95-103.
http://vuir.vu.edu.au/9370/1/Cultural%20and%20econ%20influences%20and%20decision%20to%20attend%20he.pdf

of increased educational attainment, on average. Internet usage shows a remarkably stable effect with each additional specification of the full model. The increase of one hour weekly internet use implies over a one percent increase in the probability of attaining bachelor's degrees and graduate degrees, on average. Since the measure of internet usage is spread over the week, an increase of an hour on average is not a huge time investment. This is evidence that internet usage has a highly significant positive effect on educational attainment.

FIGURE 5. Effect of an Hour Increase of Average Weekly Internet Use on Educational Attainment Outcome



Source: Milken Institute

Conclusion

This analysis shows that the time spent on the internet has a positive effect on educational attainment conditional on family, demographic, regional and time effects. As households increase in size, the probability of a person achieving a lower level of education also increases. The model provides further evidence that parental educational attainment is highly correlated with an individual's educational attainment level. This analysis provides some evidence that more time spent on the internet can mitigate negative external effects for individuals in households with lower parental education levels, lower income, or large families.

In the future, being able to differentiate hours spent on the internet for entertainment rather than productivity would enhance this analysis. There is evidence that for people between the ages of 18-22, internet usage does not affect educational attainment. However, this age range is much less likely to have any degree beyond high school which truncates educational attainment outcomes which precludes

the usefulness of this subset of ages. The population shift to urban clusters generates increasingly disproportional economic activity which drives up skill requirements through competition in the labor market. This not only puts more emphasis on being able to perform at a higher level, but also requires individuals to signal a higher level of ability. Higher education can be a way to do both.

In the context of equitable access, spreading high speed internet access to both urban and rural residents can help increase the chances for improving economic opportunity. Given the importance of the internet in modern life, equal access to information and expanding access will be ever more important to encourage higher education.

Appendix, Tables, Data, and Methodology

TABLE 1. Descriptive Statistics Full GSS data

Variables	Observations	Mean	Standard Deviations
Age	62,245	45.99	17.51
Highest Degree	62,293	1.35	1.17
Father's degree	46,985	0.89	1.17
Mother's Degree	54,451	0.84	0.99
Household size	62,460	2.65	1.51
Female	62,466	0.56	0.50
Non-white	62,466	0.19	0.40
Ln(Family Income)	56,142	10.33	0.99
Ln(Respondent Income)	36,524	9.51	1.11
Ln(Mean Internet Hours Per Week)	11,119	1.52	1.19
Protestant	59,044	0.61	0.49
Catholic	59,044	0.29	0.45
Judaism	59,044	0.02	0.14
Atheist	59,044	0.05	0.23
Buddhist	59,044	0.002	0.04
Hindu	59,044	0.0017	0.04
Muslim	59,044	0.0024	0.05
Orthodox-Christian	59,044	0.0021	0.05
Christian	59,044	0.01	0.08
Other Religions	59,044	0.01	0.11
New England	62,466	0.05	0.21
Mid-Atlantic U.S.	62,466	0.14	0.35
East North Central U.S.	62,466	0.18	0.39
West North Central U.S.	62,466	0.07	0.26
South-Atlantic U.S.	62,466	0.19	0.39
East South Central U.S.	62,466	0.07	0.25
West North Central U.S	62,466	0.09	0.29
Mountain	62,466	0.06	0.24
Pacific	62,466	0.13	0.34
	Min	Max	
Year	1972	2016	

This data is the full data set without removal of any observations

TABLE 2. Regression Results of Bayesian Ordered Probit [21,22]

Variables	Base Degree	Full Degree	Full Degree	Full Degree
Ln(Internet Usage)	0.08***	0.09***	0.09***	0.10***
	(0.01)	(0.01)	(0.01)	(0.01)
Mother's Degree	0.18***	0.18***	0.18***	0.19***
	(0.02)	(0.02)	(0.02)	(0.02)
Father's Degree	0.21***	0.22***	0.21***	0.23***
	(0.01)	(0.01)	(0.01)	(0.01)
Ln(real personal income)$_{2000\ USD}$	0.29***	0.28***		
	(0.02)	(0.02)		
Ln(Family income per earner)			0.47***	
			(0.02)	
Ln(real personal income)$_{1997\ USD}$				0.32***
				(0.03)
Household Size	-0.04***	-0.04***	-0.03***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)
Female	0.22***	0.22	0.17	0.12***
	(0.03)	(0.01)	(0.03)	(0.03)
Non-white	-0.01	0.01	0.06	-0.01
	(0.04)	(0.04)	(0.04)	(0.04)
Age	0.03***	0.02***	0.04***	0.05***
	(0.005)	(0.01)	(0.01)	(0.01)
Age^2	-0.0001	-0.0001	-0.0003***	-0.0004***
	(0.00001)	(0.00001)	(0.00009)	(0.0001)
Catholic	0.06*	0.05	0.05	0.07**
	(0.03)	(0.03)	(0.03)	(0.03)
Judaism	0.64***	0.58***	0.58***	0.65***
	(0.10)	(0.10)	(0.10)	(0.10)
Atheist	-0.04	0.05	0.04	0.04
	(0.06)	(0.06)	(0.06)	(0.06)
Buddhist	0.71***	0.69***	0.70***	0.76***
	(0.17)	(0.18)	(0.18)	(0.18)
Hindu	1.11***	1.05***	0.83**	1.07***
	(0.21)	(0.21)	(0.21)	(0.21)
Muslim	0.51**	0.48**	0.51**	0.46**
	(0.17)	(0.21)	(0.227)	(0.21)
Other religion	0.04	0.03	0.09	0.06
	(0.15)	(0.16)	(0.16)	(0.16)
Orthodox- Christian	0.29	0.17	0.20	0.16
	(0.21)	(0.22)	(0.22)	(0.22)

[21] Imai, Kosuke, Gary King, and Olivia Lau. 2008. "Toward A Common Framework for Statistical Analysis and Development." Journal of Computational and Graphical Statistics, Vol. 17, No. 4 (December), pp. 892-913, http://j.mp/msE15c.
[22] Choirat, Christine; Christopher Gandrud, James Honaker; Kosuke Imai; Gary King; Olivia Lau. 2017. Zelig: Everyone's Statistical Software, Version 5.0-15, URL: http://ZeligProject.org.

Variables Cont.	Base Degree	Full Degree	Full Degree	Full Degree
Christian	-0.20	-0.13	-0.15	-0.11
	(0.13)	(0.13)	(0.13)	(0.13)
		(0.08)	(0.08)	(0.08)
Mid-Atlantic		-0.09	-0.09	-0.09
		(0.08)	(0.08)	(0.08)
East North Central		-0.24***	-0.23***	-0.23***
		(0.07)	(0.07)	(0.07)
West North Central		-0.25***	-0.23***	-0.27***
		(0.08)	(0.08)	(0.09)
South-Atlantic		-0.19**	-0.17**	-0.19**
		(0.08)	(0.08)	(0.08)
East South Central		-0.20**	-0.18*	-0.22***
		(0.09)	(0.09)	(0.09)
West South Central		-0.39***	-0.37***	-0.38***
		(0.08)	(0.08)	(0.08)
Mountain		-0.19**	-0.19**	-0.22***
		(0.08)	(0.08)	(0.08)
Pacific		-0.28***	-0.32***	-0.28***
		(0.08)	(0.08)	(0.08)
2002		-0.18***	-0.15***	-0.16***
		(0.06)	(0.06)	(0.06)
2004		-0.12***	-0.10	-0.11
		(0.06)	(0.06)	(0.06)
2006		-0.12***	-0.12**	-0.14**
		(0.06)	(0.06)	(0.06)
2010		-0.14***	-0.12	-0.19***
		(0.07)	(0.07)	(0.07)
2012		-0.26***	-0.22***	-0.30***
		(0.07)	(0.07)	(0.07)
2014		-0.31***	-0.27***	-0.37***
		(0.07)	(0.07)	(0.06)
2016		-0.28	-0.26***	-0.35***
		(0.06)	(0.07)	(0.06)
Constant	-2.12***	-1.70***	-4.03***	-0.37**
	(0.18)	(0.20)	(0.26)	(0.18)
Gamma 2	2.10***	2.13	2.15***	2.10***
	(0.04)	(0.03)	(0.04)	(0.04)
Gamma 3	2.39***	2.42***	2.45***	2.39***
	(0.04)	(0.03)	(0.04)	(0.04)
Gamma 4	3.42***	3.45***	2.45***	3.38***
	(0.05)	(0.03)	(0.04)	(0.04)
Observations	5,715	5,715	7,034	5,688

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 3. T-Tests General Social Survey 2000-2016

Two-Sample T-Test of Equal Variance by Non-White

Variables	T-score	P-Value
Ln(Real Respondent Income)	17.43	0.00

Data Cleaning

For any estimated parameters to have validity for statistical inference, there are some assumptions that need to be met. To maintain these assumptions and achieve a robust analysis, adjustments were made to the data set. The working subset of the GSS is created by first sub-setting the years where internet usage was prevalent, 2000 to 2016 except 2008 where there is no data. Outliers were then identified by using Dffits values to conform to the assumption of asymptotic normality. The base model specification with only the natural logs taken of internet usage and family income, not including age-squared, was used in this estimation. The regional or time binary variables in this model are kept in their raw collapsed form.

Two observations were removed due to reality constraints. Internet hours per week has two observations that are unreasonable, given that there are 168 hours in a week. Observations recording an average of 168 hours or greater per week are not be reasonable and are removed. Binary variables were created for all useable years and regions. This is done because there is no ordinal interpretation of either of these variables and binary variables allow the estimated parameters to have an interpretable outcome. The religion variables were broken down into binary variables for all recorded religions that are not aggregates. Years are normalized by 2000 and the religions are normalized by Protestants. Due to the lack of ordering and the subjectivity of religion and usefulness of separating out cultural influence, binary variables are more informative. Separating out years follows the same logic in that the aggregation of years into a single parameter may wash out macroeconomic effects and reduce the specification's ability to be interpreted. Regions are normalized against the northeast. Family income is in real 2000 dollars and is transformed with a natural logarithm for better functionality. The average hours spent per week on the internet is highly positively skewed and is corrected for by using a natural logarithm transformation. These changes in the model's functional form are intended to conform to the assumption of conditional mean zero of maximum likelihood estimation (MLE). All of the changes in the functional form are also intended to increase the power of the parameter estimation for interpretation.

Model Testing

An Ordered Probit Model was chosen because of the discrete nature of educational outcomes as well as the existence of a clear ordinal component of educational credentials. The full model passes frequentist specification tests of MLE assumptions. The model likelihood ratio chi-square (χ^2) statistic of 2153.63 and specification linktest with linear predicted value ($\hat{y}$) and linear predicted value squared ($\hat{y}^2$) displaying significance and non-significance, respectively. No signs of multicollinearity where observed

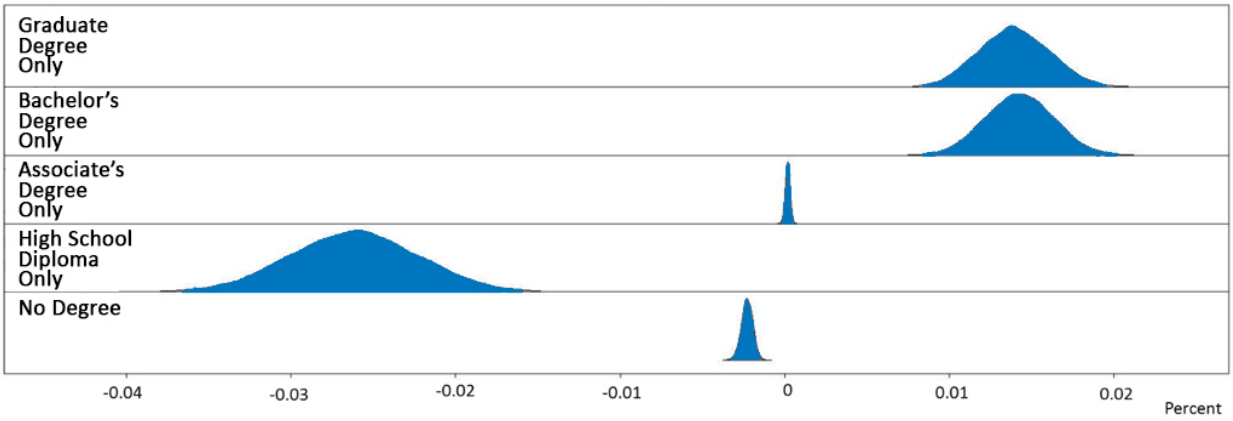
with all variance inflation factor (VIF) values under two when using no broken out binary variables. When breaking out region, year, and religion some variables have VIF increases but are all less than five. The one exception is age and age-squared, which is expected. Without the squared term there are no signs of multicollinearity. The provided set of survey weights corrects for any sampling issues that might occur. There were no major changes in any of the frequentist estimations using any of the provided sample weights. Large outliers where removed (see data cleaning section). The model does fit with in the standard MLE assumptions, but a Bayesian Ordered Probit was chosen to increase the precision of the estimation having established accuracy. Using Geweke and Raftery-Lewis diagnostic tests for convergence, I found that the variables are stationary and have weak auto correlation.[23] The acceptance rate of the Markov chain Monte Carlo (MCMC) iterations average is 0.5 percent and is supported by tests of MCMC convergence being close to one except classification cutoff points. A normal prior was used to reflect the underlying distribution of years of education, for which degree attainment is an alternative functional form.

Starting with the base model with only family dynamics and expanding to the full model shows the natural log of average hours per week spent on the internet never lost significance and maintains directionality and magnitude. This is confirmed by a relatively stable t-statistic with increases in the coefficient's magnitude. Overall, the sensitivity analysis demonstrates that the natural logarithm of average hours spent on the internet is robust in its positive and large effect on the probability of achieving higher educational attainment outcomes. More importantly, in the base model without year and regional controls, internet usage is highly significant and remains highly significant when adding in those controls. Internet usage remains stable in this specification throughout all robustness checks. Figure 6 shows the marginal effects of an hour increase in the average number of hours spent per week on the internet, all else equal, calculated at means. A one hour increase in the average weekly usage of the internet on average increases the probability of obtaining a bachelor's degree by 1.42 percent. This effect for a graduate degree is a 1.39 percent increase on average.

The unexpected result that age does not exhibit decreasing returns is an issue and through a set of robustness checks seems to be an anomaly with the main results see Table 2 columns 4 and 5. The robustness checks give more evidence that overall controlling for age, with the addition of the quadratic term, and the use of individual income does control for endogeneity of educational attainment and internet use as best as possible. With the different income variables used in the robustness checks, which increase sample size and representation of higher incomes, the quadratic term is significant at the 99 percent confidence level. The two different income variables are a Ln(nominal respondent's income) or a Ln(real income per earner) variable. There are reasons that any one of the income variables could fit into this model, but the reported results are still the best option for use as controls in this specification. The difference in the data of the income variables come from more variation at the higher end Ln(Family income) or the truncation of higher income earners and the categorical fictional form of the variable. The result of the robustness analysis gives the added power of additional observations also adds to the effectiveness of income and age controlling for endogeneity bias.

[23] ibid

FIGURE 6. The Effect of an Hour Increase in Internet Usage from Sample Averages



Source: Milken Institute

Acknowledgments

I would like to thank Ross DeVol, Minoli Ratnatunga, Jessica Jackson, and Ken Sagynbekov for their help in the development process of this paper.

About the Author

Joe Lee is a research analyst with the Milken Institute on the regional economics team. He specializes in labor economics with a focus on human capital and economic development. His recent work includes regional workforce issues and human capital interaction with urban clusters.

Before joining the Institute, Lee worked as a lab instructor in the department of economics at California State University, Long Beach (CSULB) and worked for Amazon in Seattle, Washington. Lee received his MA in economics from CSULB and graduated from The Evergreen State College.

About the Center for Jobs and Human Capital

The Milken Institute Center for Jobs and Human Capital promotes prosperity and sustainable economic growth around the world by increasing the understanding of the dynamics that drive job creation and promote industry expansion.

About the Milken Institute

The Milken Institute is a nonprofit, nonpartisan think tank determined to increase global prosperity by advancing collaborative solutions that widen access to capital, create jobs and improve health. We do this through independent, data-driven research, action-oriented meetings and meaningful policy initiatives.