Title

Have we finally bridged the digital divide? Smart phone and Internet use patterns by race and ethnicity

Author

Fairlie, Robert W

Publication Date

2017

DOI

10.5210/fm.v22i9.7919

Peer reviewed

Have We Finally Bridged the Digital Divide?
Smart Phone and Internet Use Patterns by Race and Ethnicity

Robert W. Fairlie
Professor of Economics, University of California, Santa Cruz
rfairlie@ucsc.edu

July 2017

Abstract

Two decades ago an influential article documented the alarming disparities that existed in access to computers and the Internet between African-Americans and whites (Hoffman and Novak 1998). Using the latest Census Bureau/Bureau of Labor Statistics data on computer and Internet access, I find that the "Digital Divide" has not been bridged and remains as large as it was two decades ago. African-Americans and Latino-Americans are less likely to use the Internet on smart phones, computer, tablets or other devices than are whites. A statistical decomposition analysis reveals that income and education inequalities are the leading causes of the disparities in access to technology. The findings have implications for policies that subsidize broadband to low-income families.

1.Introduction

Two decades ago an influential article documented the alarming disparities that existed in computer and Internet use between African-Americans and whites (Hoffman and Novak 1998). Four years later, the National Telecommunications and Information Administration (NTIA) documented the rapid growth in the use of the Internet over the previous few years and noted that Internet use among African-Americans and Latinos grew at a substantially faster rate over the previous year than for whites (NTIA 2002). Racial differences were not even discussed in the next NTIA report on Internet use in the United States (NTIA 2004).

The rapid rise in Internet use, especially through the use of smart phones, suggests that the Digital Divide might have been bridged. But, has it? Recent Census/Bureau of Labor Statistics (BLS) microdata are used here to examine whether the Digital Divide has finally been bridged and make comparisons to the racial disparities in computer and Internet use reported in Hoffman and Novak (1998). The analysis of the latest Census/BLS microdata is also expanded to include the rapid rise in the use of smart phones to access the Internet, and to include estimates of use among Latino-Americans, who now represent the largest minority group in the United States, and Asian-Americans, who also represent a rapidly growing segment of the population.

The focus here is on exploring the Digital Divide within the United States. Although disparities across ethnic and racial groups within the country are large, disparities across countries around the world are substantial, especially between developed and developing countries (International Telecommunications Union 2016).[1] More than half the world's population does not have Internet access.

2. Data

Patterns of Internet and Smart Phone use are estimated using microdata from the latest Computer and Internet Use Supplement to the Current Population Survey (CPS), which is from July 2015. The U.S. Census Bureau and U.S. Bureau of Labor Statistics conduct the survey. Interviewing approximately 50,000 households and 130,000 individuals, it is a good representation of the U.S. population, meaning that the data can reliably tell us much about individuals' computer, Internet and smart phone use. It is the primary source of information on technology use reported by the federal government.

3. Racial Differences in Smart Phone and Internet Use

Cell phone use is widespread, and racial disparities in cell phone use are relatively small. Table 1 reports estimates of cell phone and smart phone use by race and ethnicity. Seventy-seven percent of

[1] See Chinn and Fairlie (2007, 2010) and Mardikyan et al. (2015) for examples of studies exploring the determinants of the global Digital Divide.

whites use a cell phone compared with 74 percent of African-Americans, 71 percent of Latinos, and 79 percent of Asians. Of these cell phones used by Americans 53 to 57 percent are smart phones that are used to browse the web, send email, and access social networking sites, such as Facebook, Instagram and Twitter. Unconditionally, roughly 40 percent of blacks and Latinos use smart phones. Smart phone use among whites and Asians is higher at roughly 45 percent.

Has the rise in smart phone use eliminated the Digital Divide in overall Internet use? The answer is no – racial disparities in overall Internet use (i.e. from all devices) remain large. Presently, 31 percent of African-Americans and 34 percent of Latinos do *not* use the Internet on a smart phone, computer, tablet or any other device. Figure 1 displays racial/ethnic differences in smart phone use next to overall rates of Internet use (also see Table 1). Disparities are found for both measures with larger disparities found for overall Internet use. Rates of Internet use among African-Americans and Latinos are 9 and 13 percentage points lower than for whites, respectively. In 1996-97, 18 percent of whites used the Internet compared with 10 percent of blacks (Hoffman and Novak 1998) implying a roughly similar sized white/black gap. To eliminate the current racial and ethnic gaps in Internet use 4 million blacks and 6 million Latinos would have to gain access to the Internet.

African-Americans and Latinos are also more likely to rely only on smart phones than desktop, laptop or tablet computers to access the Internet at home than are whites. Among Internet users at home, 26 percent of blacks and 27 percent of Latinos exclusively use a smart phone compared with 18 percent of white and Asian Internet users. Although smart phones can be used for browsing the web, emailing, social networking, and entertainment (games, music, photos and video) they do not support the full range of activities possible with desktop or laptop computers.

Related to the differences in reliance on smart phones to access the Internet, do racial groups use the Internet for different activities? Table 2 reports common activities on the Internet by race and ethnicity. White Internet users are more likely than black and Latino users to report using the Internet for “higher-level” activities such as financial services, telecommuting and searching for medical information, but have a similar likelihood of using the Internet for more routine activities such as email, texting, social networking, web browsing, watching videos, and listening to music.

Compared to other locations, Internet use at home most likely represents the highest quality access in terms of availability and autonomy, potentially providing the most benefits to the user. Racial disparities in Internet use at home are even larger than for overall Internet use. Table 1 reports Internet and computer use and access at

home by racial and ethnic group. Only 61 percent of African-Americans and 59 percent of Latinos use the Internet at home compared with roughly three quarters of whites and Asians. Hoffman and Novak (1998) found that 15 percent of whites and 9 percent of blacks used the Internet at home in 1996-97.

Blacks and Latinos are also less likely to have a hi-speed connection to the Internet. Table 1 reports estimates. Seventy-one percent of blacks and 69 percent of Latinos have a hi-speed connection at home compared with roughly 80 percent of whites and Asians. The difference is driven by mobile plans and not dialup plans. Very few individuals of any race or ethnicity connect to the Internet through dialup plans.

Personal computers are valuable for many uses beyond the Internet, such as for word processing, financial spreadsheets, and specialized software. The racial divide in access to home computers remains large and relatively unchanged over the past two decades. In 1996-97, 44 percent of whites and 29 percent of African-Americans had access to a home computer (Hoffman and Novak 1998). Twenty years later, the percentage of whites with access to home computer increased to 78 percent and the percentage of blacks increased to 63 percent, but the gap remains the same in percentage points.[2]

[2] An earlier literature documents lower computer ownership among minorities. See Goldfarb and Prince (2008), Ono and Zavodny (2007), and Fairlie (2004) for example.

The constant racial gap of roughly 15 percentage points is alarming. It implies that to increase black levels of computer ownership to white levels the same relative number of blacks would need computers to eliminate the gap – nearly 7 million. The digital divide between Latinos and whites is equally large with only 63 percent of Latinos having access to home computers. Roughly 7 million Latinos would need home computers to eliminate the gap in ownership with whites.

4. Potential Causes of the Digital Divide

Why do African-Americans and Latinos have relatively low rates of Internet use? At a first pass, self-reported reasons for not subscribing to Internet service through a smart phone or computer are compared. Table 3 displays self-reported reasons for not having Internet access at home by race and ethnicity. The most common response among blacks and Latinos is that they "can’t afford it " with 31 percent of blacks and 32 percent of Latinos noting this as the most important reason for no service. In contrast, only 17 percent of whites report that they cannot afford Internet service. The most likely response among whites is that they “don’t need it.” Among all groups, only a very small percentage report not subscribing to Internet service

because they can use it somewhere else, it is not available where they live, or for privacy or safety concerns.[3]

Income disparities are likely to explain part of the Digital Divide. To explore how much income as well as other factors, such as family structure, education, occupation, and geographical location explain of the racial gaps in access to the Internet I perform a statistical decomposition analysis. The decomposition technique combines multivariate logit regression estimates and sample distributions to identify how much each factor independently contributes to the racial disparities in Internet access (Fairlie 2017). The Technical Appendix includes a description of the methodology of the non-linear decomposition technique.

Table 4 reports estimates from the decomposition technique. The first row of the table shows the magnitude of the gap between blacks/whites and Latinos/whites for Internet and computer access.[4] The largest factor found to explain racial disparities in Internet access through a smart phone or computer at home is income. Racial differences in income levels explain roughly one third of the black-white and Latino-white gaps in Internet access. African-Americans and Latinos have lower income levels than whites contributing to the gaps

[3] Interestingly, 22 percent of white households with Internet access note that they experienced an "online security breach, identity theft, or a similar crime." The rate was lower among black, Latino and Asian households (13 to 18 percent).

[4] The decomposition is not performed for the Asian-white comparison because the levels of access are similar (thus there is no disparity to decompose).

in Internet access, even after controlling for educational, occupational and other differences. The second most important factor is education (and largest for Latinos). Racial differences in education, which might capture exposure, interest and longer-term income potential, explain 9 percent of the gap for blacks and 36 percent of the gap for Latinos.[5] Latino levels of education are substantially lower than white levels.

Interestingly, the racial/ethnic gaps in Internet and computer access are not due to employment or occupational differences. The Digital Divide also does not appear to be due to regional differences. These have also been concerns about the Digital Divide. Finally, the negative contributions are interpreted as factors that are “favorable” for Internet access. For example, Latinos are younger on average than are whites and the gaps in Internet and computer access would be even larger if they had a similar age distribution as whites.

5. Conclusions

Even with the rapid expansion in smart phone use in the United States, the Digital Divide in the United States has not been bridged. The latest Census data indicate that 31 percent of African-Americans and 34 percent of Latinos, representing more than 30 million people, do not use the Internet on a smart phone, computer, tablet or other device. Across several different measures, the Digital Divide remains

[5] Income and educational differences explain similar shares of the gaps in computer ownership.

as large as it was two decades ago (Hoffman and Novak 1998). The constant gap, which is in many cases as large as 15 percentage points, translates into several million blacks and Latinos having to gain access to boost levels of access to white levels.

Although the continuing rise in smart phone use may bridge part of the gap, the Digital Divide is unlikely to disappear altogether soon. Smart phones do not represent a complete substitute for laptop or desktop computers for Internet use. Only a small percentage of smart phone users rely exclusively on their smart phones to access the Internet (90 percent of smart phone users also report accessing the Internet at home through a laptop, desktop or tablet computer).

A large percentage of blacks and Latinos report that they do not subscribe to Internet service because it is "too expensive." Supporting this suggestive evidence, statistical decomposition techniques reveal that income inequality is the leading cause of disparities in access to the Internet even after controlling for other factors. To address this concern, other countries have experimented with ambitious programs that provide subsidies for computer purchases among low-income families, and a growing number of state, school district and individual school programs are experimenting with one-to-one laptop or tablet programs that allow students to take computers home (Hall and Duch 2017).[6] Several recent programs also focus on increasing access to and

[6] A few examples of programs providing computers to low-income families include the Home Access Programme in England, the Euro 200 Program in Romania, and the Yo Elijo Mi PC Program in Chile. Additionally, extensive

reducing the costs of broadband service to low income families (e.g. the federal E-Rate and Lifeline Programs, and Comcast's Internet Essentials Program), but the future support of federal programs is unclear. Recently, the FCC has imposed some restrictions to low-income broadband subsidies.[7]

Cost is not the only factor, however, limiting access to the Internet. Income disparities do not explain the entire gap in technology access. Although also expensive, racial differences in cell phone and smart phone use are much smaller than differences in computer use. More research is needed on what other factors limit computer use, and how they are related to smart phone use. Additionally, policies addressing more than income constraints, such as community technology centers and comprehensive technology programs (e.g. the Smart Communities Program in Chicago), might be helpful (Servon 2002; Mossberger et al. 2012).

The lingering "Digital Divide" may have serious consequences for disadvantaged minority groups as Internet use becomes increasingly important in education, the labor market, commerce, health, communications, consumption, and political engagement.[8] The impacts

---

efforts to provide laptops to schoolchildren also exist in many developing countries (see http://one.laptop.org/about/countries).

[7] “At F.C.C., Obama-Era Rules on Chopping Block,” NY Times, April 5, 2017.

[8] Previous researchers have also identified disparities in computer skills as an understudied and important aspect of the digital divide and have referred to them as the "second-level" of the digital divide (Hargittai 2002; Servon 2002). Unfortunately, the CPS does not include information on skills. However, access to home computers appears to be strongly associated with computer skills (Fairlie 2012).

on job search, education and health may be especially acute. The primary method of job search now is through the Internet. Among all Internet users, for example, 37 percent of blacks and 32 percent of Latinos report using the Internet to search for jobs. Schools and colleges are also increasingly using technology in the classroom, remotely (e.g. MOOCs), and for instructional and administrative communications. Finally, more than half of Internet users search for health information online and more than one-quarter communicate with their physicians online.

References

Chinn, Menzie D. and Robert W. Fairlie. 2007 "The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration," *Oxford Economic Papers*, 59:16–44.

Chinn, Menzie D. and Robert W. Fairlie, 2010. "ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration," *Review of International Economics*, 18(1): 153-167.

Fairlie, Robert W. 2004. "Race and the Digital Divide," *Contributions to Economic Analysis & Policy, The Berkeley Electronic Journals*, 3(1), Article 15: 1-38.

Fairlie, Robert W. 2012. "The Effects of Home Access to Technology on Computer Skills: Evidence from a Field Experiment," *Information Economics and Policy* 24(3–4): 243–253.

Fairlie, Robert W. 2017. " Addressing Path Dependence and Incorporating Sample Weights in the Nonlinear Blinder-Oaxaca Decomposition Technique for Logit, Probit and Other Nonlinear Models," Stanford Institute for Economic Policy Research Working Paper No. 17-013.

Goldfarb, Avi, and Jeffrey Prince. 2008. "Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide." *Information Economics and Policy*, 20(1), 2-15, March.

Hargittai, Eszter. 2002. "Second-Level Digital Divide: Differences in People's Online Skills." *First Monday*. 7(4).

Hoffman, Donna L. and Thomas P. Novak. 1998. "Bridging the Racial Divide on the Internet." *Science* 17 April: 390-391.

Hull, Marie, and Katherine Duch. 2017. “One-to-one technology and student outcomes,” University of North Carolina Working Paper.

International Telecommunications Union. 2016. *ICT Facts and Figures 2016*.

Mardikyan, Sona, et al. 2015. "Examining the global digital divide: a cross-country analysis." *Communications of the IBIMA* 2015: 1.

Mossberger, Karen, Caroline J. Tolbert, and William Franko. 2012. *Digital*

Cities: The Internet and the Geography of Opportunity. New York: Oxford University
Press.

National Telecommunications and Information Administration. 2002. *A Nation Online: How Americans are Expanding Their Use of the Internet*. Washington, D.C.: U.S. Department of Commerce.

National Telecommunications and Information Administration. 2004. *A Nation Online: Entering the Broadband Age*. Washington, D.C.: U.S. Department of Commerce.

Ono, Hiroshi, and Madeline Zavodny, 2007. “Digital Inequality: A Five Country Comparison Using Microdata,” *Social Science Research,* 36 (September 2007): 1135-1155.

Servon, Lisa. 2002. *Bridging the Digital Divide: Community, Technology and Policy* (Blackwell).

Table 1: Cell Phone, Smart Phone, and Internet Use Rates by Race/Ethnicity

	White, non-Latino		African-American		Latino		Asian	
	Rate	Std. Err.	Rate	Std. Err.	Rate	Std. Err.	Rate	Std. Err.
Cell and smart phone use	77.2%	0.2%	73.8%	0.4%	70.6%	0.5%	78.5%	0.6%
Smart phone use	44.2%	0.2%	39.4%	0.5%	39.6%	0.5%	45.1%	0.7%
Internet use (any device, any location)	78.8%	0.2%	69.4%	0.4%	65.5%	0.5%	79.5%	0.6%
Internet use (any device at home)	73.9%	0.2%	60.5%	0.5%	58.5%	0.5%	75.1%	0.6%
Internet access at home	79.5%	0.2%	66.8%	0.4%	68.4%	0.5%	83.3%	0.5%
Computer use at home	69.4%	0.2%	53.7%	0.5%	47.9%	0.5%	71.8%	0.6%
Computer access at home	78.0%	0.2%	63.1%	0.5%	63.1%	0.5%	83.6%	0.5%
With access to Internet at home								
Access with mobile plan	60.8%	0.2%	65.1%	0.6%	66.7%	0.6%	60.8%	0.7%
Access with hi-speed service	78.5%	0.2%	71.3%	0.5%	68.5%	0.6%	79.6%	0.6%
Access with satellite service	3.7%	0.1%	2.6%	0.2%	3.6%	0.2%	2.4%	0.2%
Access with dialup service	0.6%	0.0%	0.6%	0.1%	0.8%	0.1%	0.3%	0.1%
Access with other service	0.7%	0.0%	0.6%	0.1%	0.9%	0.1%	0.3%	0.1%
Mobile plan for Intenet and no other plan	17.9%	0.2%	26.1%	0.5%	27.6%	0.5%	18.2%	0.6%
Sample size	71937		11245		9916		5282	

Table 2: Internet Uses by Race/Ethnicity

	White, non-Latino		African-American		Latino		Asian	
	Rate	Std. Err.	Rate	Std. Err.	Rate	Std. Err.	Rate	Std. Err.
Use the Internet for:								
Email	92.5%	0.1%	89.1%	0.5%	86.1%	0.6%	94.0%	0.5%
Texting	84.5%	0.2%	89.7%	0.5%	90.0%	0.5%	88.4%	0.7%
Social networking	69.6%	0.3%	70.9%	0.7%	72.3%	0.8%	72.0%	1.0%
Video conferencing	36.5%	0.3%	34.1%	0.7%	35.7%	0.9%	49.4%	1.1%
Web browsing	91.0%	0.2%	88.6%	0.5%	87.3%	0.6%	91.2%	0.6%
Watch videos	65.0%	0.3%	67.3%	0.7%	73.1%	0.8%	75.3%	1.0%
Listen to music	51.7%	0.3%	57.0%	0.7%	59.5%	0.9%	57.6%	1.1%
Maps	72.8%	0.3%	69.9%	0.7%	73.0%	0.8%	76.8%	1.0%
Telecommute	34.2%	0.3%	27.0%	0.8%	28.2%	0.9%	46.2%	1.3%
Job search	25.2%	0.2%	37.3%	0.7%	32.2%	0.8%	32.3%	1.1%
Online classes and training	20.8%	0.2%	21.4%	0.6%	20.6%	0.7%	26.4%	1.0%
Financial services	68.4%	0.3%	59.7%	0.7%	61.5%	0.9%	72.7%	1.0%
Online shopping	72.6%	0.3%	62.3%	0.7%	60.0%	0.9%	75.4%	1.0%
Household equipment	6.9%	0.1%	7.2%	0.4%	7.1%	0.5%	8.7%	0.6%
Health information	52.3%	0.3%	42.2%	0.7%	39.8%	0.9%	48.7%	1.1%
Communicate with physician	28.8%	0.3%	21.5%	0.6%	20.3%	0.7%	31.5%	1.1%
Health monitoring device	6.1%	0.1%	4.9%	0.3%	5.5%	0.4%	7.0%	0.6%
Sample size	71937		11245		9916		5282	

Table 3: Main Reason Do Not Have Internet at Home through Smart Phone, Computer or Other Device

	White, non-Latino		African-American		Latino-American		Asian-American	
	Rate	Std. Err.	Rate	Std. Err.	Rate	Std. Err.	Rate	Std. Err.
Don't need it	35.1%	0.4%	29.8%	0.7%	29.9%	0.8%	34.2%	1.6%
Not interested	23.7%	0.3%	17.7%	0.6%	16.4%	0.7%	19.7%	1.3%
Can't afford it	17.3%	0.3%	30.5%	0.7%	32.4%	0.8%	19.3%	1.3%
Not worth it	1.9%	0.1%	1.9%	0.2%	2.0%	0.2%	2.4%	0.5%
Can use it elsewhere	2.1%	0.1%	2.3%	0.2%	1.6%	0.2%	2.6%	0.5%
Not available in area	3.0%	0.1%	1.7%	0.2%	1.8%	0.2%	2.2%	0.5%
No computer or computer inadequate	6.6%	0.2%	8.7%	0.4%	9.0%	0.5%	9.0%	1.0%
Privacy concerns	1.3%	0.1%	0.6%	0.1%	0.2%	0.1%	0.7%	0.3%
Safety concerns	0.5%	0.1%	0.6%	0.1%	0.2%	0.1%	0.0%	0.0%
Moved	0.6%	0.1%	0.3%	0.1%	0.4%	0.1%	0.9%	0.3%
Other reason	7.8%	0.2%	5.9%	0.4%	6.0%	0.4%	9.0%	1.0%
Sample size	15139		3972		3165		900	

Table 4: Explanatory Factors for Racial Gap in Internet and Computer Access

	Internet Access		Computer Access	
	African-American	Latino	African-American	Latino
White/minority gap in rate	0.141	0.130	0.130	0.096
Contributions from racial differences in:				
Income	0.046	0.038	0.043	0.035
	(0.001)	(0.001)	(0.001)	(0.001)
	33.0%	29.4%	33.1%	36.2%
Education	0.012	0.046	0.012	0.040
	(0.000)	(0.001)	(0.000)	(0.001)
	8.5%	35.5%	9.1%	42.2%
Employment / Occupation	0.001	0.004	-0.001	0.001
	(0.000)	(0.001)	(0.001)	(0.001)
	0.4%	3.1%	-0.4%	1.3%
Demographic characteristics and family structure	0.001	-0.022	-0.004	-0.036
	(0.001)	(0.001)	(0.001)	(0.001)
	0.8%	-17.2%	-3.0%	-37.1%
Region	0.005	-0.001	0.003	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
	3.2%	-0.9%	2.1%	0.9%
Central city status	-0.006	-0.009	-0.007	-0.008
	(0.001)	(0.001)	(0.001)	(0.001)
	-4.3%	-6.6%	-5.2%	-8.8%
All included variables	0.058	0.056	0.046	0.033
	45.0%	43.2%	35.7%	25.6%
Sample size	98768	98768	98768	98768

Note: Standard errors are reported in parentheses below contribution estimates.

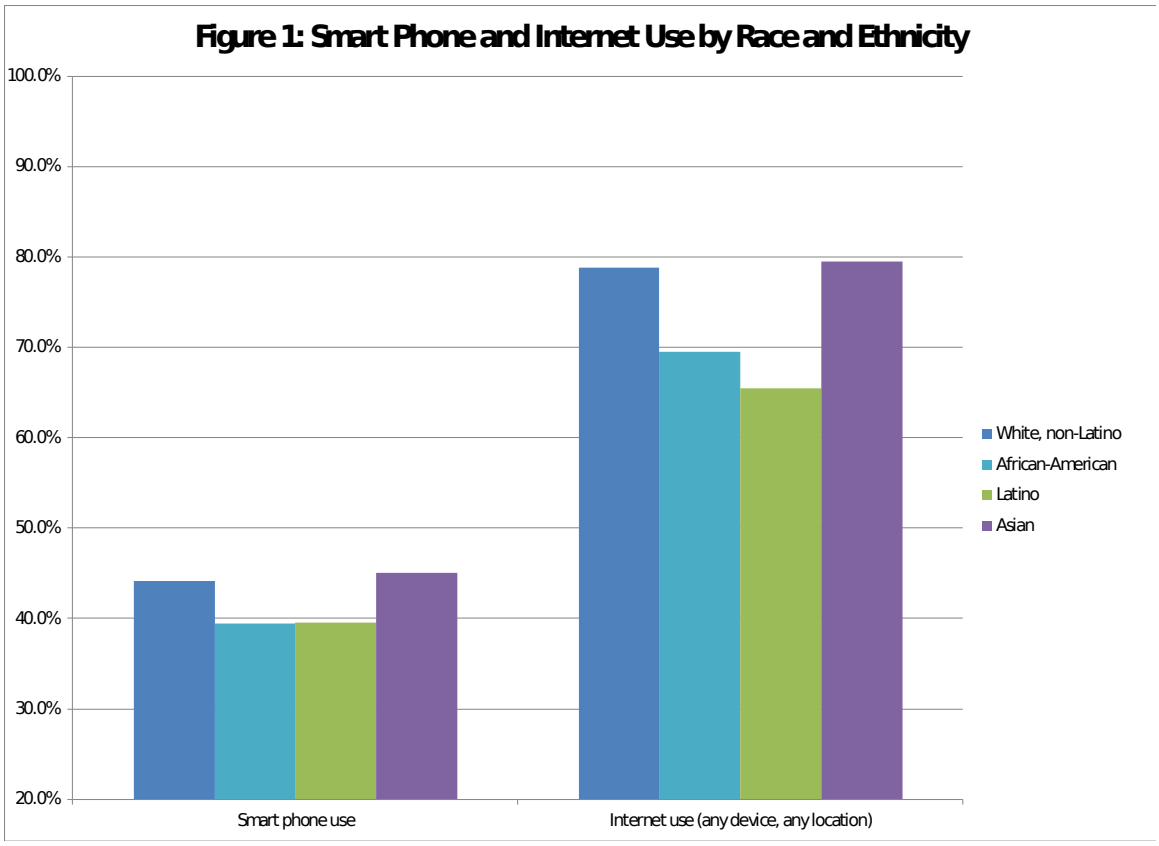
Figure 1: Smart Phone and Internet Use by Race and Ethnicity
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
Smart phone use
Internet use (any device, any location)
White, non-Latino
African-American
Latino
Asian

Technical Appendix: Description of Nonlinear Decomposition Method

The decomposition technique used to identify the causes of white-minority differences in Internet and computer access is presented. The technique decompose inter-group differences in mean levels of an outcome into those due to different observable characteristics or "endowments" across groups and those due to different effects of characteristics or "coefficients" of groups. For a linear regression, the standard Blinder-Oaxaca decomposition of the white/black gap in the average value of the dependent variable, Y, can be expressed as:

(1) $$\overline{Y}^W - \overline{Y}^B = \left[\left(\overline{X}^W - \overline{X}^B\right)\hat{\beta}^W\right] + \left[\overline{X}^B\left(\hat{\beta}^W - \hat{\beta}^B\right)\right],$$

where $\overline{X}^j$ is a row vector of average values of the independent variables and $\hat{\beta}^j$ is a vector of coefficient estimates for race *j*. The Blinder-Oaxaca technique is used for dependent variables that are estimated with linear regressions, however, an alternative technique is required for a non-linear regression, such as a logit regression, which is used to estimate the probability of computer and Internet access.

Following Fairlie (1999), the decomposition for a nonlinear equation, such as

$Y = F(X\ \hat{\beta}\ )$, can be written as:

(2) $\overline{Y}^{W} - \overline{Y}^{B} = ¿ ¿$

where N^j is the sample size for race j. This alternative expression for the decomposition is used because $\bar{Y}$ does not necessarily equal $F(\bar{X}\ \hat{\beta}\ )$.[9] In both (1) and (2), the first term in brackets represents the part of the racial gap that is due to group differences in distributions of X, and the second term represents the part due to differences in the group processes determining levels of Y.[10] The second term also captures the portion of the racial gap due to group differences in unmeasurable or unobserved endowments. Similar to most previous studies applying the decomposition technique, I do not focus on this "unexplained" portion of the gap because of the difficulty in interpreting results.

To calculate the decomposition, define $\bar{Y}^j$ as the average probability of the binary outcome of interest for race j and F as the

---

[9] Note that the Blinder-Oaxaca decomposition is a special case of (2).

[10] Contributions to the gap may be negative. As found here, the negative contributions from demographic characteristics and family structure are mainly due to the black and Latino populations being younger than the white population (which is a favorable characteristic in predicting who has access to the Internet and a computer).

cumulative distribution function from the logistic distribution.[11] Alternatively, for a probit model F would be defined as the cumulative distribution function from the standard normal distribution. Results are generally very similar when using a probit model.

An equally valid method of calculating the decomposition is to use the minority coefficient estimates, $\hat{\beta}^M$, as weights for the first term and the white distributions of the independent variables, $\bar{X}^W$, as weights for the second term. This alternative method of calculating the decomposition often provides different estimates, which is the familiar index problem with the Blinder-Oaxaca decomposition technique. A third alternative is to weight the first term of the decomposition expression using coefficient estimates from a pooled sample of the two groups (see Oaxaca and Ransom 1994 for example). This approach is followed to calculate the decompositions. In particular, I use coefficient estimates from a logit regression that includes a sample of all racial groups.

Using the pooled coefficients from a sample of all racial groups has the advantage over using the white coefficients because it captures the determinants for all groups and are more precisely

[11] A useful property of the logit regression that includes a constant term is that the average of the predicted probabilities must equal the proportion of ones in the sample. In contrast, the predicted probability evaluated at the means of the independent variables is not necessarily equal to the proportion of ones, and in the sample used below it is larger because the logit function is concave for values greater than 0.5.

estimated (because of the larger sample and more heterogeneity of firms).

The first term in (4.2) provides an estimate of the contribution of racial differences in the entire set of independent variables to the racial gap in the dependent variable. Estimation of the total contribution is relatively simple as one only needs to calculate two sets of predicted probabilities and take the difference between the average values of the two. Identifying the contribution of group differences in specific variables to the racial gap, however, is not as straightforward. To simplify, first assume that $N_B=N_W$ and that there exists a natural one-to-one matching of black and white observations. Using

coefficient estimates from a logit regression for a pooled sample, $\hat{\beta}^{¿}$, the independent contribution of X_1 to the racial gap can then be expressed as:

(3) $$\frac{1}{N^B}\sum_{i=1}^{N^B} ¿F(\hat{\alpha}^{¿}+X_{1i}^{W}\hat{\beta}_1^{¿}+X_{2i}^{W}\hat{\beta}_2^{¿})-F(\hat{\alpha}^{¿}+X_{1i}^{B}\hat{\beta}_1^{¿}+X_{2i}^{W}\hat{\beta}_2^{¿})¿$$

Similarly, the contribution of X_2 can be expressed as:

(4) $$\frac{1}{N^B}\sum_{i=1}^{N^B} ¿F(\hat{\alpha}^{¿}+X_{1i}^{B}\hat{\beta}_1^{¿}+X_{2i}^{W}\hat{\beta}_2^{¿})-F(\hat{\alpha}^{¿}+X_{1i}^{B}\hat{\beta}_1^{¿}+X_{2i}^{B}\hat{\beta}_2^{¿})¿$$

The contribution of each variable to the gap is thus equal to the change in the average predicted probability resulting from sequentially

switching the white characteristics to black characteristics one variable or set of variables at a time.[12] A useful property of this technique is that the sum of the contributions from individual variables will be equal to the total contribution from all of the variables evaluated with the full sample.

In practice, the sample sizes of the two groups are rarely the same and a one-to-one matching of observations from the two samples is needed to calculate (3) and (4). In this example, the black sample size is substantially smaller than the white sample size. To address this problem, first use the pooled coefficient estimates to calculate predicted probabilities, $\hat{Y}_i$, for each black and white observation in the sample. Next, draw a random subsample of whites with a sample size equal to N_B and randomly match it to the full black sample. The decomposition estimates obtained from this procedure depend on the randomly chosen subsample of whites. Ideally, the results from the decomposition should approximate those from matching the entire white sample to the black sample. A simple method of approximating this hypothetical decomposition is to draw a large number of random subsamples of whites, match each of these random subsamples of whites to the black sample, and calculate separate decomposition

[12] Unlike in the linear case, the independent contributions of X_1 and X_2 depend on the value of the other variable. This implies that the choice of a variable as X_1 or X_2 (or the order of switching the distributions) is potentially important in calculating its contribution to the racial gap. The estimates reported here are not sensitive to the reordering of variables.

estimates. The mean value of estimates from the separate decompositions is calculated and used to approximate the results for the entire white sample. All of the decompositions reported in this chapter use 100 random subsamples of whites to calculate these means.

Technical Appendix References

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Variables." *Journal of Human Resources*, 8: 436-455.

Fairlie, Robert W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment," *Journal of Labor Economics*, 17(1): 80-108.

Fairlie, Robert W. 2017. "Addressing Path Dependence and Incorporating Sample Weights in the Nonlinear Blinder-Oaxaca Decomposition Technique for Logit, Probit and Other Nonlinear Models," Stanford Institute for Economic Policy Research Working Paper No. 17-013.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14 (October), 693-709.

Oaxaca, Ronald, and Michael Ransom. 1994. "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics*, 61, 5-21.