PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES
www.phoenix-center.org

PHOENIX CENTER POLICY PAPER SERIES

Phoenix Center Policy Paper Number 58:

Digital Discrimination: Fiber Availability and Speeds by Race and Income

T. Randolph Beard, PhD
George S. Ford, PhD

(September 2022)

Phoenix Center Policy Paper No. 58
Digital Discrimination: Fiber Availability and Speeds by Race and Income

T. Randolph Beard, PhD*
George S. Ford, PhD†

Abstract: The lack of broadband in many rural and Tribal communities is widely recognized, but there are also claims of a lack of broadband availability in predominantly Minority and urban communities, sometimes labeled *digital redlining* or *digital discrimination*. Motivated by such claims, the bi-partisan Infrastructure Investment and Jobs Act of 2021 includes a specific provision to address digital discrimination and the Federal Communications Commission is currently contemplating formal rules. In this POLICY PAPER, we provide a definition of digital discrimination and describe the sort of empirical conditions and methods needed to quantify it. An empirical analysis of digital discrimination in fiber deployment and broadband speeds is performed. The results are encouraging—no systematic evidence of digital discrimination by race or income level is found.

TABLE OF CONTENTS:

---

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

† Chief Economist, Phoenix Center for Advanced Legal & Economic Public Policy Studies. The views expressed in this paper are the authors' alone and do not represent the views of the Phoenix Center or its staff.

I. Introduction

High-speed Internet service is increasingly viewed as an essential service, yet broadband, whether fixed or mobile, is neither ubiquitously available nor universally subscribed. Significant policy attention is devoted to availability gaps in rural America,[1] but some advocates worry policymakers are ignoring availability gaps in poor, urban, or Minority communities, going so far as to say that this "gap in broadband coverage in a poorer neighborhood is effectively a digital form of redlining."[2]

---

[1] According to the latest data from the Federal Communications Commission ("FCC"), at least one home in 95.6% of census blocks can obtain a broadband service, but there is a material difference between rural blocks (82.7%) and urban blocks (98.8%). *In the Matter of Inquiry Concerning Deployment of Advanced Telecommunications Capability to All Americans in a Reasonable and Timely Fashion,* FCC 21-18, FOURTEENTH BROADBAND DEPLOYMENT REPORT, 36 FCC Rcd 836 (rel. January 19, 2021) (available at: https://docs.fcc.gov/public/attachments/FCC-21-18A1.pdf).

[2] *See, e.g.,* E. Falcon, *The FCC and States Must Ban Digital Redlining*, Electronic Frontier Foundation (January 11, 2021) (available at: https://www.eff.org/deeplinks/2021/01/fcc-and-states-must-ban-digital-redlining); S. Tibken, *The Broadband Gap's Dirty Secret: Redlining Still Exists in Digital Form,* CNET (June 28, 2021) (available at: https://www.cnet.com/features/the-broadband-gaps-dirty-secret-redlining-still-exists-in-digital-form); G. Strain, E. Moore, and S. Gambhir, *AT&T's Digital Divide in California*, HASS INSTITUTE FOR A FAIR AND INCLUSIVE SOCIETY (2017) (available at: https://belonging.berkeley.edu/sites/default/files/haas_broadband_042417-singles.pdf);
Testimony of Christopher Lewis, President and CEO - Public Knowledge, Before the U.S. House of Representatives Committee on Energy & Commerce, Subcommittee on Communications and Technology, hearing on "Broadband Equity: Addressing Disparities in Access and Affordability" (May 6, 2021) (available at: https://docs.house.gov/meetings/IF/IF16/20210506/112553/HHRG-117-IF16-Wstate-LewisC-20210506-U1.pdf); Committee on Energy and Commerce, Opening

(Footnote Continued. . . .)

To address such concerns, the bi-partisan Infrastructure Investment and Jobs Act of 2021 ("Infrastructure Act") includes Section 60506, labeled "Digital Discrimination."[3] Section 60506(a) states that it shall be the policy of the United States, insofar as "technically and economically feasible," that subscribers "*within the service area of a provider*" should benefit from the "*equal opportunity to subscribe* to an offered service that provides comparable speeds, capacities, latency, and other quality of service metrics" at "comparable terms and conditions."[4] Section 60506(c) also contains a second statement of federal policy, "prohibiting deployment discrimination." Section 60506 makes no mention of broadband adoption, so adoption appears beyond the scope of this provision—a sensible reading of Section 60506 since broadband providers have little control over which consumers do or do not buy offered service.[5] Pursuant to this statement of federal policy, Section 60506(b) directs the Federal Communications Commission ("FCC") to adopt final rules no later than two years after enactment of the Infrastructure Act "to facilitate equal access to broadband internet access service, taking into account the issues of technical and economic feasibility" in order to prevent "digital discrimination of access based on the protected classes limited to income level, race, ethnicity, color, religion, or national origin."[6]

In March 2022, the Commission issued a *Notice of Inquiry* ("*NOI*") to begin the implementation of Section 60506 as directed by Congress.[7] The *NOI* makes clear that the Commission grasps the difficulty of crafting a program both to measure and to address digital discrimination (to the extent it exists) and appears to

---

Statement as Prepared for Delivery of Subcommittee on Communications and Technology Chairman Mike Doyle, Hearing on "Broadband Equity: Addressing Disparities in Access and Affordability" (May 6, 2021) (available at: https://energycommerce.house.gov/sites/democrats.energycommerce.house.gov/files/documents/Opening%20Statement_Doyle_CAT_2021.5.6.pdf).

3 Infrastructure Investment and Jobs Act § 60506, codified at 47 USC § 1754.

4 Section 60506(a) (emphasis supplied).

5 While some advocates want digital discrimination to adhere to adoption, doing so is plainly inconsistent with the statute. *See*, *e.g*., J. Mimura, *What is "Digital Discrimination,"* National Digital Inclusion Alliance (May 21, 2022) (available at: https://www.digitalinclusion.org/blog/2022/05/17/what-is-digital-discrimination).

6 Section 60506(b).

7 *In the Matter of Implementing the Infrastructure Investment and Jobs Act: Prevention and Elimination of Digital Discrimination,* FCC 22-21, NOTICE OF INQUIRY, __ FCC Rcd __ (rel. March 17, 2022) (available at: https://www.fcc.gov/document/fcc-initiates-inquiry-preventing-digital-discrimination).

indicate the Commission will take an economic approach to defining digital discrimination (as is seemingly required by the Infrastructure Act's deliberate mention of economic and technical feasibility considerations). In particular, the Commission sought advice about how to address the fact that some of the protected classes have a lower demand for broadband, and this lower demand reduces the "economic feasibility" of deploying networks.[8] Also, the Commission asks how it should handle the "underlying cost or geographic hurdles" that may make deployment "unprofitable," another factor falling into the economic and technical feasibility criteria.[9] Survey evidence and empirical research on broadband adoption show that income has a demonstrable effect on demand. Also, empirical research and survey evidence show that some racial minorities—in particular, Hispanic, Black, and Native Americans—are less likely to adopt fixed-service broadband services in the home.[10] Moreover, income is correlated with many factors that affect demand including, among other things, employment, education levels, and housing stability, which makes the determination of "income discrimination" extremely difficult since it is the discriminatory treatment of low-income households, and not these other correlated factors, that is mentioned in the statute. Minority population shares and income levels are also correlated with population density, which affects the cost of network deployment and upgrades. Quantifying "digital discrimination" is, therefore, an *extremely* challenging endeavor.

In this POLICY PAPER, we use a typical model of discrimination based on Heckman (1998) to aid the Commission in its investigation and to lay the

---

8 *Id*. at ¶ 24 ("If underlying cost or geographic hurdles exist in conjunction with demand in an area that makes it unprofitable, how should the Commission address such a situation?").

9 *Id*.

10 *See*, *e.g*., S. Atske and A. Perrin, *Home Broadband Adoption, Computer Ownership Vary By Race, Ethnicity in the U.S.*, Pew Research Center (July 16, 2021) (available at: https://www.pewresearch.org/fact-tank/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s) ("Black and Hispanic adults in the United States remain less likely than White adults to say they own a traditional computer or have high-speed internet at home …"); *Exploring the Digital Nation: Computer and Internet Use at Home*, National Telecommunications and Information Administration (2013) at p. vi (available at: https://www.ntia.doc.gov/files/ntia/publications/exploring_the_digital_nation_computer_and_internet_use_at_home_11092011.pdf) ("Lower income families, people with less education, those with disabilities, Blacks, Hispanics, and rural residents generally lagged the national average in both broadband adoption and computer use."); G.S. Ford, *Race and Broadband Adoption: A Decomposition Analysis*, PHOENIX CENTER POLICY BULLETIN No. 52 (May 2021) (available at: https://www.phoenix-center.org/PolicyBulletin/PCPB52Final.pdf).

groundwork for useful empirical analysis.[11] Digital discrimination, as we define it, is present when differences in some relevant outcome exists across communities *when the profitability of serving the communities is equal*.[12] (Both economic and technical feasibility create profit differences.) The outcomes we study here are fiber availability and download speeds in metropolitan areas (to exclude unmeasurable cost factors in rural areas), but the general approach is transferable to other outcomes. Due to the economic and technical feasibility requirements of the statute, and the Commission's concern about demand and costs, the analysis of digital discrimination requires addressing demand and cost differences across similar areas. Empirical work, therefore, will not be straightforward, though we provide some examples of the sorts of procedures that may be used. Moreover, it may be argued that the digital discrimination provision contained in Section 60506 is a peculiar addition to the Infrastructure Act given the billions of subsidy dollars provided by the Act to ensure broadband access across *all* areas. To the extent equal access is not present, one might conclude unequal access reflects a failure in subsidy programs, such as Universal Service or the Infrastructure Act, rather than a provider's digital discrimination.

It is important, as we see it, that the title of Section 60506 is denoted "*Digital Discrimination*", rather than *digital equality*, *digital equity*, or some other term indicating unconditional equivalence in service provision.[13] Discrimination implies more than mere differences among protected classes; discrimination requires differences in outcomes caused solely by membership in a protected class, other things constant. Economic, technical, and demographic characteristics, aside from the protected class being studied, must be held constant (which may be possible using multiple regression and other methods) across communities within an area to tease out the specific influence of a protected class, else differences may reflect economic and technical feasibility (i.e., profitability) unrelated to differences due to membership in a protected class. We recognize that advocates

---

[11] J.J. Heckman, *Detecting Discrimination*, 12 JOURNAL OF ECONOMIC PERSPECTIVES 101–16 (1998).

[12] *See*, *e.g.*, MEASURING RACIAL DISCRIMINATION, National Research Council, National Academy of Sciences (2004) ("racial discrimination occurs when a member of one racial group is treated less favorably than a similarly situated member of another racial group and suffers adverse or negative consequences (at p. 40)."

[13] The "equality" rather than "discrimination" view is embraced by some political constituencies. *See, e.g., AT&T's Digital Redlining*, National Digital Inclusion Alliance (2017) (available at: https://www.digitalinclusion.org/blog/2017/03/10/atts-digital-redlining-of-cleveland). This analysis compares outcomes across incomes but makes no effect to address differences in demand and costs that may influence profitability, which a proper analysis of redlining or discrimination requires.

sometimes view actions that have an unintended discriminatory effect as discriminatory (i.e., *indirect discrimination*), but such concerns, while relevant, are not an issue here as differences in profitability are a direct and not indirect influence on outcomes and "fulfill a business need."[14] Section 60506, in our view, is not a prohibition against profit maximizing behavior (else, the Infrastructure Act's subsidies would be unnecessary).[15]

II. Defining Digital Discrimination

The Infrastructure Act's "digital discrimination" concept is related to the racist historical practice of *redlining*, so it is worth reviewing how researchers study redlining to lay some groundwork. The term "redlining" originates in the Federal Housing Administration's ("FHA") practice, starting in the 1930s, of refusing government-backed mortgage insurance to homebuyers in predominantly Black neighborhoods—neighborhoods labeled by the federal government as "hazardous" and outlined on maps using the color red.[16] Private mortgage writers

---

[14] Practices that may lead to indirect discrimination must "fulfill a genuine business need," and profits certainly qualify. *See*, *e.g.*, *Griggs vs. Duke Power*, 401 U.S. 424, 432 (1971) (available at: https://www.naacpldf.org/case-issue/griggs-v-duke-power-co). Indirect discrimination happens when there is a policy that applies in the same way for everybody but disadvantages a group of people who share a protected characteristic. *See*, *e.g.*, *What is Direct and Indirect Discrimination?*, Equality and Human Rights Commission (last visited: June 20, 2022) (available at: https://www.equalityhumanrights.com/en/advice-and-guidance/what-direct-and-indirect-discrimination).

[15] Some advocates appear to believe the statute does prohibit profit maximization. *See*, *e.g.*, Comments of the Electronic Frontier Foundation, *et al*., FCC Docket No. 22-69 at p. 15 ("A driving factor of digital discrimination is the three-to-five year return-on-investment (ROI) formulas that major ISPs follow when determining where to invest fiber. *** This short time frame [] is discriminatory towards lower income households…") (available at: https://files.fcc.gov/ecfs/download/4b496fda-d216-46f3-ac5f-feb61739adb4?orig=true&pk=cb77b2ec-1a58-dbc6-139b-ad192cfd5d9b).

[16] *See*, *e.g*., R. Rothstein, THE COLOR OF LAW: A FORGOTTEN HISTORY OF HOW OUR GOVERNMENT SEGREGATED AMERICA (2018); T. Gross, *A "Forgotten History" Of How The U.S. Government Segregated America,* National Public Radio (May 3, 2017) (available at: https://www.npr.org/2017/05/03/526655831/a-forgotten-history-of-how-the-u-s-government-segregated-america); A. C. Madrigal, *The Racist Housing Policy That Made Your Neighborhood*, THE ATLANTIC (May 22, 2014) (available at: https://www.theatlantic.com/business/archive/2014/05/the-racist-housing-policy-that-made-your-neighborhood/371439); B. Mitchell and J. Franco, *HOLC "Redlining" Maps: The Persistent Structure of Segregation and Economic Inequality*, NCRC RESEARCH (2018) (available at: https://ncrc.org/wp-content/uploads/dlm_uploads/2018/02/NCRC-Research-HOLC-10.pdf); B. Little, *How A New Deal Housing Program Enforced Segregation*, HISTORY.COM (October 20, 2020)

(Footnote Continued. . . .)

engaged in similar discriminatory actions. Such practices were forbidden by the Fair Housing Act of 1968 and the Fair Housing Act of 1977. Section 805 of the Fair Housing Act of 1977, for instances, makes it

> … unlawful for any lender to discriminate in its housing-related lending activities against any person because of race, color, religion, national origin, sex, handicap, or familial status.[17]

The law treats these non-economic factors differently than economic concerns. As observed by the Federal Reserve:

> The prohibition against redlining does not mean that a lending institution is expected to approve all housing loan applications or to make all loans on identical terms. Denying loans or granting loans on more-stringent terms and conditions, however, must be justified on the basis of economic factors [ ].[18]

These economic factors include, but may not be limited to: (1) the applicant's income or credit history; (2) the condition, use, or design of the proposed security property (or of those nearby properties that clearly affect the value of the proposed security property), provided that such determinants are strictly economic or physical in nature; (3) the availability of neighborhood amenities or city services; and (4) the need of the lender to hold a balanced real estate loan portfolio, with a reasonable distribution of loans among various neighborhoods, types of property, and loan amounts.[19]

Redlining is present when people with *equal* economic characteristics (who thus may be expected to produce the same "profit" to the firm) experience *unequal*

(available at: https://www.history.com/news/housing-segregation-new-deal-program); A.M. Perry and D. Harshbarger, *America's Formerly Redlined Neighborhoods Have Changed, and So Must Solutions to Rectify Them*, BROOKINGS (October 14, 2019) (available at: https://www.brookings.edu/research/americas-formerly-redlines-areas-changed-so-must-solutions).

17 *Consumer Compliance Handbook*, U.S. Federal Reserve (viewed July 16, 2021) (available at: https://www.federalreserve.gov/publications/supervision_cch.htm), with specific reference to *Federal Fair Lending Regulations and Statutes, Fair Housing Act* (available at: https://www.federalreserve.gov/boarddocs/supmanual/cch/fair_lend_fhact.pdf); *See also* 42 U.S.C. § 3605 (available at: https://www.justice.gov/crt/fair-housing-act-2).

18 Consumer Compliance Handbook, *id.*

19 *Id*.

treatment based on non-economic factors such as race. Suppose the potential profits from serving groups A and B are identical, but group B is served less because of its racial composition.[20] Such an outcome would be consistent with redlining. Alternately, if the potential profits offered by groups A and B are different, then differential treatment may not constitute redlining if all treatment differences are attributable to the differences in expected profits.

Evidence on redlining requires dividing the differential treatment of racial minorities (or other relevant groups) into two components: (1) permissible economic factors (income and financial risk); and (2) impermissible non-economic factors (race). Under Section 60506, however, differential treatment based on income is not permissible, thus the Infrastructure Act treats income—a decidedly *economic* factor—as if it were a non-economic factor.[21] Empirical analysis can decompose these considerations to quantify redlining (the non-economic part of the difference). Numerous academic studies have done this, including Holmes and Horvitz (1994), Tootell (1996), Ong and Stoll (2007), and Cohen-Cole (2011).[22] These studies offer evidence on redlining by dividing the differential treatment of racial minorities (or other relevant groups) into two components: (1) permissible

---

[20] In this scenario, if a firm discriminates based on race, then it (or its customers) must pay to do so. G.S. Becker, THE ECONOMICS OF DISCRIMINATION (1957); K.M. Murphy, *How Gary Becker Saw the Scourge of Discrimination*, CHICAGO BOOTH REVIEW (June 15, 2015) (available at: https://review.chicagobooth.edu/magazine/winter-2014/how-gary-becker-saw-the-scourge-of-discrimination).

[21] In the banking industry, there is an issue of "source of income discrimination" (*e.g.,* wages, alimony, dividends), where lenders cannot refuse to loan money to households based on the source of income (*e.g.,* Housing Choice Vouchers). But discrimination related to the source of income is very different than the level of income. *See*, *e.g.*, A.K. Fasanelli and P. Tegeler, *Your Money's No Good Here: Combatting Source of Income Discrimination in Housing*, AMERICAN BAR ASSOCIATION (November 30, 2019) (available at: https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/economic-justice/your-money-s-no-good-here--combatting-source-of-income-discrimin/#:~:text=%22'Source%20of%20income%20discrimination',person's%20lawful%20form%20of%20income.%22).

[22] G.M.B. Tootell, *Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods?* 111 QUARTERLY JOURNAL OF ECONOMICS 1049-1079 (1996); A. Holmes and P. Horvitz, *Mortgage Redlining: Race, Risk and Demand*, 49 JOURNAL OF FINANCE 81-99 (1994); and P.M. Ong and M.A. Stoll, *Redlining or Risk? A Spatial Analysis of Auto Insurance Rates in Los Angeles*, 26 JOURNAL OF POLICY ANALYSIS AND MANAGEMENT 811-829 (2007) ("disentangling the two competing explanations for the higher premiums paid by those in poor and Minority communities: the higher rates are due to higher legitimate costs of insuring residents in poor and Minority communities as a result of the greater risks there, and the higher rates are the product of redlining."); E. Cohen-Cole, *Credit Card Redlining*, 93 REVIEW OF ECONOMICS & STATISTICS 700-713 (2011).

economic factors (income and financial risk); and (2) impermissible non-economic factors (race). Before turning to the empirical analysis, we first offer a simple conceptual framework for digital discrimination.

A. *A Conceptual Model*

In the broadband context, the model posited by various interest groups is that network deployment or service offerings arise from decisions made by providers who are motivated by profits and, perhaps, by racial animus (or animus toward any protected class).[23] Since empirical evidence suggests that race and income are potent determinants of fixed broadband adoption and, thus, economic feasibility, any differential treatment of these protected classes arises from two distinct channels: (1) the effect of lower demand on economic feasibility (i.e., profitability); and (2) discrimination against a protected class. Our method of analysis is constructed to accommodate these basic complications. Our conceptual framework is based on broadband deployment and speeds, but the model is equally useful for studying service quality, prices, and other outcomes of interest.

Following the general approach described by Heckman (1998), we represent the demand for service in area *i* as:[24]

$$D(X_i, r_i) \tag{1}$$

where X_i are demand drivers such as income, education, and so on, and $r_i \in [0,1]$ indicates membership in a protected class for the neighborhood in question, where for example $r_i = 1$ represents an area with an entirely Minority population. Costs to deploy service in the neighborhood are C_i so provider profit from offering service in the neighborhood is written as:

23 *See, e.g.*, D. Turner, *Digital Denied: The Impact of Systemic Racial Discrimination on Home-Internet Adoption*, Free Press (December 2016) (available at: https://www.freepress.net/news/press-releases/digital-denied-free-press-report-exposes-impact-systemic-racism-internet); B. Callahan, *AT&T's Digital Redlining Of Cleveland*, NATIONAL DIGITAL INCLUSION ALLIANCE (March 2017) (available at: https://www.digitalinclusion.org/blog/2017/03/10/atts-digital-redlining-of-cleveland); M. Eichensehr, *Baltimore City Council Members, 92 Other Officials Call on FCC to Address "Digital Redlining"*, BALTIMORE BUSINESS JOURNAL (March 2021) (available at: https://www.bizjournals.com/baltimore/news/2021/03/16/city-council-asks-fcc-to-tackle-digital-redlining.html); J. Eggerton, *Broadband Redlining Complaint Filed Against AT&T at FCC,* NEXT TV (August 2017), (available at: https://www.nexttv.com/news/broadband-redlining-complaint-filed-against-att-fcc-168100).

24 Heckman, *supra* n. 11.

$$\pi\big(D(X_i, r_i), C_i\big). \tag{2}$$

The treatment of the neighborhood of interest is whether appropriate broadband service is available there. Denote the probability that area i has broadband as Y, where:

$$Y = Y\Big(\pi\big(D(X_i, r_i), C_i\big), r_i^*\Big)\ , \tag{3}$$

where $r_i^* \in \{0, 1\}$ indicates the predominance of a protected class, *i.e.*, $r_i^* = 1$ if and only if $r_i > r_0$, where r_0 is the relevant cutoff for, say, Minority presence. Given this formulation, one can say that discrimination is occurring when:

$$Y\Big(\pi\big(D(X_i, r_i), C_i\big), r_i^* = 1\Big) < Y\Big(\pi\big(D(X_i, r_i), C_i\big), r_i^* = 0\Big). \tag{4}$$

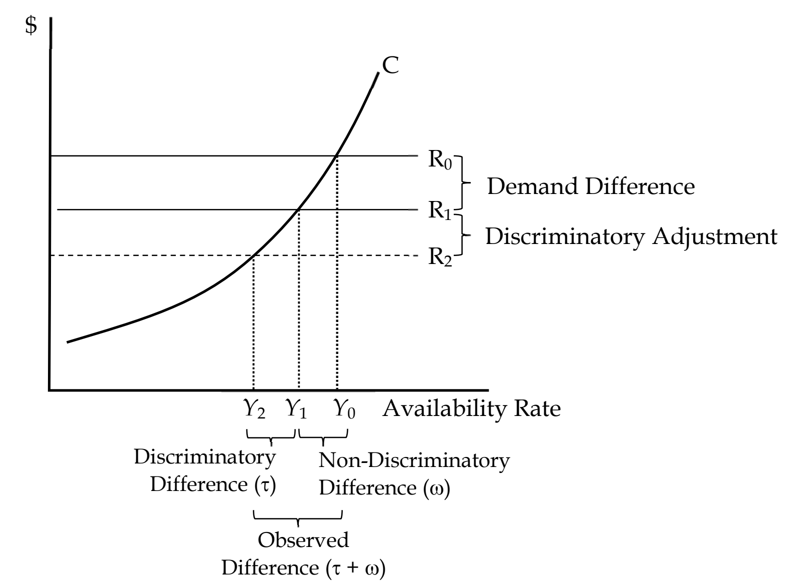
Given the quality of data available and other (unobservable) factors that may influence outcomes, one might require some minimum threshold difference to invoke a policy response (say, a few percentages points). Costly policy intervention presumably should occur only in response to "meaningful" differences in deployment.

An important element of this test is that r_i enters Expression (4) in two ways. First, the implied test for discrimination is based on r_i via r_i^*. Second, r_i enters the demand function directly. A failure to account for r_i in the demand function leads to a biased measure of discrimination. For instance, since minorities and low-income households have, on average, a lower adoption rate for fixed broadband in the home, ignoring the effect of race or income on demand will bias a statistical model toward a finding of digital discrimination. Our empirical approach aims to account for demand and cost differences, thereby addressing this misspecification and eliminating (or attenuating) bias in the measure of digital discrimination.

B. *A Graphical Explanation*

This analysis can be summarized graphically. In Figure 1, households (or areas) in two markets of the same size and identical cost conditions are ordered by the incremental cost of deployment (labeled C) along the horizontal axis. The vertical axis in measured in dollars. The markets are two types: (1) a predominantly non-Minority community ($r_i^* = 0$); and (2) a predominantly Minority community ($r_i^* = 1$) with lower demand for broadband than the non-Minority community. We assume the costs of serving the two are identical but the demand for broadband differs.[25] The horizontal line labeled R_0 is the net expected revenue per household from serving the predominantly non-Minority community. The profit maximizing availability level is Y_0. With a lower demand for broadband in the Minority community, the expected revenues per account are lower, as indicated by the horizontal line labeled R_1. (Predominantly non-Minority communities may also have demand R_1, a fact we exploit in the empirical analysis.) The broadband availability rate in the predominantly Minority community is Y_1. While availability is different in the Minority community ($Y_1 < Y_0$), *no digital discrimination is implied because the difference is due to demand and not race*—the two areas are not of equal economic feasibility (i.e., profitability).

Figure 1. Detecting Discrimination



Now, say that broadband providers discriminate against Minority communities, arbitrarily treating the "true" expected revenues as something less than R_1. Let the horizontal line R_2 reflect this discriminatory adjustment (an adjustment equal to $R_1 - R_2$). Availability is reduced to Y_2. The empirical problem is that what we observe in the data is the difference $Y_0 - Y_2$, a difference that does not measure discrimination but includes a non-discriminatory difference based on differences in demand (labeled ω) and the discriminatory difference (labeled τ). Rather, discrimination is measured as the difference $\tau = Y_1 - Y_2$, which cannot be

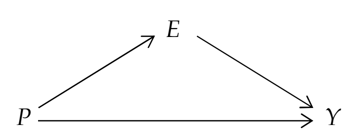
---

25 Ford, *supra* n. 10, shows that much of the difference in demand between Minority and White communities is unexplained by traditional demand factors.

directly observed. Detecting discrimination requires exposing $Y_1 - Y_2$ from the available data, a challenging task, especially for areas of widely-disparate incomes.

The difference ω measures the effect of demand and cost differences related to race (a sort of *selection bias* if ignored) and the difference τ is the *discrimination effect* based solely on the differential treatment of a protected class (e.g., a *treatment effect* on the treated).[26] Quantifying discrimination requires an empirical method that compares communities of both types where demand for broadband is equal (say, R_1), and to then test whether the availability rates in both communities are equal. If predominantly Minority areas have lower availability rates (or some other outcome of interest), then this difference may signal a discriminatory outcome.

Figure 2. DAG: Relationships of Interest



An alternative graphical approach is a simple Directed Analytic Graph ("DAG"), represented in Figure 2, that describes all causal relationships of interest.[27] The variable P represents membership in a protected class, Y is the outcome of interest, and E is a portmantua of economic conditions (demand and cost). There are two paths by which P influences Y. First, there is a direct causal effect ($P \rightarrow Y$). Second, there is a path from P to Y through E (a mediator) whereby P affects Y through its effect on E ($P \rightarrow E \rightarrow Y$). In the present context, the P variable might be race, Y might be the availability of fiber broadband, and E is demand. Race may have a direct effect on fiber deployment, which is the discriminatory effect (τ), and race influences fiber deployment through its effect on profitability (ω). The total effect of race on deployment is the sum of the two effects (τ + ω). In quantifying digital discrimination, we are interested in the direct effect of race on fiber deployment (τ) and not the indirect effect through demand

---

[26] *See*, *e.g.*, J.D. Angrist and J.S. Pischke, MOSTLY HARMLESS ECONOMETRICS: AN EMPIRICIST'S COMPANION (2008); G. Imbens and J. Wooldridge, *Recent Developments in the Econometrics of Program Evaluation*, 47 JOURNAL OF ECONOMIC LITERATURE 5-86 (2009); S. Cunningham, CAUSAL INFERENCE: THE MIXTAPE (2021).

[27] Cunningham, *id.*, at pp. 96-118.

(ω). As such, we must condition on economic conditions to limit the measured effect to the direct effect τ.

C. *A Proposed Definition*

Based on Expression (4) and the language of Section 60506, a sensible definition of define digital discrimination might be as follows:

> Digital discrimination occurs when differences in the deployment of and/or the quality, terms, and conditions of access to broadband services are not explained by differences in the profitability of serving the different areas, but instead reflect non-economic decisions to underserve protected classes in a manner that causes adverse or negative consequences.[28]

The protected classes are enumerated in the statute and include "income level, race, ethnicity, color, religion, or national origin." Economic and technical feasibility are related and, in effect, essentially determine profitability, though perhaps in different ways. An example of technical feasibility affecting service levels is the highly varied speeds capable over DSL networks due to loop length (e.g., the longer the loop, the slower the connection on average).[29] Equivalent speeds may be accomplished by shortening loops, or by building an entirely new network, both of which are costly options (often prohibitively so). In any geographic area, there may also be important variations in terrain, altitude, authority to provide service, and so on. Questions of economic and technical feasibility likely will be situationally specific, but all these qualifiers can create profitability differentials.

III. Data

Our empirical analysis requires data on broadband availability, demand, and costs. Data on fiber broadband deployment and maximum advertised download speeds at the census block level for the contiguous U.S. states (excluding Alaska

---

28 *See* Section 60506(c)(2) ("the predominant race or ethnicity composition of an area").

29 *See, e.g., Local-Loop and DSL: Reference Guide*, EXFO (2014) (available at: http://www.equicom.hu/wp-content/uploads/EXFO_Reference-Guide-Local-Loop-DSL-v1_en.pdf); K. Stordahl, *Long-term Penetration and Traffic Forecasts for the Western European Fixed Broadband Market*, ECONSTOR (2011) (available at: https://ideas.repec.org/p/zbw/itse11/52210.html).

and Hawaii) is obtained from the FCC's Form 477 data for December 2020.[30] We take a broad perspective here and analyze fiber deployment by all broadband providers.[31] We study speeds in the same manner. While the Infrastructure Act uses the term *area* in several (and sometimes inconsistent) ways, including the "service area of a provider" or just "area," we illustrate the empirical methodology with this broader definition. We believe there are some advantages to this approach. As the area of analysis gets smaller there is a greater risk of idiosyncratic influences on deployment, such as permission to enter buildings, government influences on deployment, peculiarities affecting costs, and so forth.

The five-year American Community Survey ("ACS") for years 2016-2020 provides demographic data at the block group level.[32] This data includes block-group level estimates of broadband adoption and computer ownership, which can be used to measure demand. With sufficiently rich demographic data only available at the block group level, the geographic area analyzed is the census block group.[33] For data available at the block level, aggregation from census blocks to block groups use household-weighted averages based on 2010 households, since the 2020 Form 477 blocks are based on 2010 cartography. Because of aggregation, any dichotomous indicator at the census block level becomes a share (on the unit interval).

The ACS data provides measures of racial composition, median household income, education levels, the fixed broadband penetration rate (based on subscriptions to DSL, cable, and fiber), the share of households with mobile-only broadband subscriptions, the share of homes with a computer, among other

---

30 As is well established, this indicator overstates availability since only one home in a Census block must be served to qualify as "served." *See, e.g.*, J. Busby and J. Tanberk, *FCC Reports Broadband Unavailable to 21.3 Million Americans, BroadbandNow Study Indicates 42 Million Do Not Have Access*, BROADBANDNOW (May 27, 2021) (available at: https://broadbandnow.com/research/fcc-broadband-overreporting-by-state); *see also* G.S. Ford, *Quantifying the Overstatement in Broadband Availability from the Form 477 Data: An Econometric Approach* , PHOENIX CENTER POLICY PERSPECTIVE No. 19-03 (July 11, 2019) (available at: https://www.phoenix-center.org/perspectives/Perspective19-03Final.pdf).

31 Some providers offer both fiber and traditional services to a census block, but we include only the fiber-provisioned blocks in such cases.

32 Data available at: https://www.nhgis.org.

33 Mixing data of disparate aggregations can lead to biased coefficients, so it is often best to aggregate to the lowest level of aggregation of the data. G.S. Ford, *Challenges in Using the National Broadband Map's Data,* PHOENIX CENTER POLICY BULLETIN No. 27 (March 2011) (available at: https://www.phoenix-center.org/PolicyBulletin/PCPB27Final.pdf).

variables. To account for costs, we use several variables. CostQuest's estimate of cost at the census block level provides direct estimates of costs.[34] The U.S. Department of Agriculture's ("USDA") Rural-Urban Continuum Code ("RUCC")—a classification scheme that distinguishes geographic areas by the population size, the degree of urbanization, and adjacency to a metro area—is used to limit the sample to metropolitan areas.[35] Population density, the share of homes identified by the Census Bureau as being in a "rural" census block (though in a metropolitan area), and an indicator variable for block groups that lie within a census place (i.e., city), are also used as cost variables.

Census block groups with missing data are excluded.[36] To reduce cost variations that may be difficult to measure, and to avoid areas that may have received subsidies for broadband deployment, we limit the analysis to block groups within census tracts that are classified as metropolitan by the RUCC.[37] Given the unique character of most Tribal areas, we also exclude block groups within Tribal census tracts, which are few (only 312 block groups are eliminated). There are 127,983 census block groups in the final sample.

IV. Empirical Specification for Digital Discrimination

While there are several protected classes, our analysis looks at race and income. The statute states that discrimination should be based on the "service areas of a provider" or "an area." For Section 60506(a)(1) the term "service area of a provider" is used, while Section 60506(a)(2) uses the term "given area." It remains undecided how the Commission will define either term. Here, we take a broad approach and look simply at differences among census block groups. For deployment, we limit our attention to fiber deployment, which seems necessary

---

[34] The CostQuest cost data contains six-categories of costs, and one category is assigned to each census block. The categories include costs less than $30, between $30 and $40, between $40 and $50, between $50 and $52.50, between $52.50 and 200, and greater than $200. The small group ($50-52.50) is based on the Commission's threshold for cost support. We combine this group with the $40-$50 category to reduce the count from six to five categories. When aggregated to the block group level, the variables become population-weighted shares of these cost indicators.

[35] Data available at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx The data are at the census tract level.

[36] Missing data affects about 27% of block groups, with missing data on fixed broadband adoption leading to the bulk of the effect.

[37] Since most tracts are so classified, the effect on sample size is relatively small (about 82% of the full sample is retained). Including subsidized areas may reduce the ability to detect discrimination by including in the sample uneconomic deployments.

because more than 90% of households already have access to "broadband services" as defined by the FCC (at least 25 Mbps download, 3 Mbps upload), and deployment at the 100/20 Mbps level is nearly as high.[38] Also, fiber deployment was the focus of debates regarding digital redlining or digital discrimination. Sample mean fiber deployment rates average about 46%, allowing some variation in the outcome. For speeds, we look at maximum advertised download speeds, as reported to the Commission in the Form 477 data.

A. *Identifying Digital Discrimination*

Quantification of discrimination requires comparing communities within a provider's service area that are alike in profitability (or economic and technical feasibility) but differ in racial composition or income levels. The empirical task, therefore, is to address selection bias to expose the discriminatory effect τ from the total (and observed) difference $\tau + \omega$. To do so, we propose to identify τ using a matching estimator, comparing collections of block groups of predominantly protected and unprotected classes with equal D_i and C_i, thus eliminating (or, at least attenuating) selection bias.[39]

Coarsened Exact Matching ("CEM") is used to create matched samples.[40] Matching is straightforward for binary and many categorical variables, but exact matching for continuous variables is difficult due to the curse of dimensionality (and both D_i and C_i are continuous). CEM mimics exact matching by coarsening continuous variables that can then be more easily matched. The degree of coarsening can be adjusted to improve sample balance. Applying CEM to this data ensures that the D_i and C_i between protected and unprotected classes have near equal means, similar variances, and one hopes comparable distributions. A

---

[38] G.S. Ford, *A Quality Check on Form 477 Data: Errors, Subsidies, and Econometrics*, PHOENIX CENTER POLICY PERSPECTIVE No. 21-05 (October 27, 2021) (available at: https://www.phoenix-center.org/perspectives/Perspective21-05Final.pdf).

[39] That is, we use matching to satisfy the Conditional Independence Assumption ("CIA"). Angrist and Pischke, *supra* n. 26. *Also see*, *e.g.*, MEASURING RACIAL DISCRIMINATION, *supra* n. 12 at p. 146 ("Matching methods provide an alternative to multivariate linear regression as a way to control for variables that are likely to matter for an outcome in observational studies. Matching consists of comparing outcomes of two paired individuals (or groups) who are comparable on relevant observed attributes except for race. Matching attempts to mimic the experimental setting in the same way as paired testing. To the extent that (1) the observed factors capture the relevant variables affecting the outcome and (2) the comparability is close, racial differences in the outcome variable in a matching study can be attributed to discrimination.").

[40] S.M. Iacus, G. King and G. Porro, *Causal Inference without Balance Checking: Coarsened Exact Matching*, 20 POLITICAL ANALYSIS 1-24 (2012).

standard rule of thumb for ensuring adequate overlap of the covariate distributions is that the Standardized Difference is less than 0.25, and a ratio of variances near unity is likewise desirable.[41]

Table 1. Means by Share of Minority Population

Minority Population	Fiber	Density '000	Income '000	Fixed BB Adoption Rate	Mobile BB Only
0 to 10%	0.406	1.68	93.10	0.756	0.089
10 to 20%	0.474	2.80	88.01	0.773	0.088
20 to 30%	0.473	2.95	78.98	0.754	0.098
30 to 40%	0.472	3.02	71.59	0.731	0.107
40 to 50%	0.476	3.08	65.99	0.707	0.118
50 to 60%	0.490	3.36	61.17	0.685	0.127
60 to 70%	0.482	3.80	57.69	0.659	0.135
70 to 80%	0.519	4.34	54.01	0.632	0.143
80 to 90%	0.518	4.86	49.83	0.594	0.157
90 to 100%	0.517	6.04	44.75	0.546	0.170

We can get a sense of the complexity of the problem and the presence of selection bias with some simple descriptive statistics. Table 1 summarizes some interesting statistics about race, fiber deployment, household density, income, broadband adoption, and persons only using mobile broadband (and not fixed). We divide Minority population shares into ten groups by shares and compute the means of the variables for each group. The table shows that fiber availability *rises* as Minority share rises. This positive relationship between fiber deployment and minority share does not imply the absence of digital discrimination. Observe, for instance, that population density (lower costs, on average) rises in Minority share, a relationship that might explain (in part) the higher deployment rates. On the other hand, the adoption of fixed broadband and income fall in minority share, and the share of mobile broadband only households rises with minority share, which reflect lower demand and thus lower profitability. As minority share rises costs fall and demand falls, on average, so the simple positive correlation between deployment and minority share does not, taken alone, say much about digital discrimination. To measure the effect of race on deployment, systematic differences in profitability between minorities and non-minorities must be accounted for to avoid confounding discrimination with demand- and cost-based factors.

41 The Standardized Difference is computed as $(\overline{x}_0 - \overline{x}_1)/(s_0^2 + s_1^2)^{0.5}$. Imbens and Wooldridge, *supra* n. 26 at pp. 43-4.

B. *The Statistical Test*

Our empirical approach mimics the analysis illustrated in Figure 1. We employ a somewhat simple identification strategy but recognize there are other methods that may be useful for this problem. Also, we suspect the analysis of digital discrimination may focus on a provider's service area, but we set that issue aside to illustrate our method in larger samples. Obtaining statistically sufficient block group sample sizes in many service areas may be challenging.

What we observe in the data are the average fiber availability rates or average download speeds. A test of the difference between these means may be conducted using the bivariate regression model,

$$Y_i = (\tau + \omega) r_i^* + \varepsilon_i , \tag{5}$$

where Y_i is the outcome of interest, r_i^* is a "treatment" indicator set equal to 1 for block groups that are predominantly populated by a protected class (0 otherwise), and ε_i is the econometric disturbance term. (We retain the τ and ω notation for the coefficients as these are estimates of the quantities of interest.) Equation (5) allows a hypothesis test on unconditioned means differences (i.e., the difference in the observed means). The estimated coefficient on r_i^* is $\tau + \omega$, which includes both a non-discriminatory (selection bias, ω) and a discriminatory component (τ).

Restricting the analysis to a sample of block groups that have equal D_i and C_i, where such differences account for the selection bias, then we may estimate,

$$\tilde{Y}_i = \tau \tilde{r}_i^* + \nu_i , \tag{6}$$

where $\tilde{Y}_i$ and $\tilde{r}_i^*$ indicate values in a sample of equal demand and costs and ν_i is the econometric disturbance term. The task is to construct a sample across protected and unprotected classes with equal demand and costs, which may be challenging given the acknowledged relationships between race and income and demand and costs.

C. *Measuring Demand and Cost*

Measuring demand and costs presents challenges. Several variables are available measuring both, and these measures are often highly correlated. To fit the conceptual framework above, to address multicollinearity, and to reduce the number of variables for matching, we use factor analysis to generate the variables

D_i and C_i.[42] For demand, we include four variables that measure or influence demand including the fixed broadband adoption rate, the mobile broadband adoption rate, the share of persons with a tertiary education, and the share of homes with a computer.[43] Only one factor has an eigenvalue exceeding 1.0 and the Kaiser-Meyer-Olkin measure of sampling adequacy ("KMO") is 0.82, which is "meritorious" by the standard definition.[44] Factor loadings are summarized in Table 2.

Table 2. Factor Analysis

Demand, D_i	**Loading**	**Cost, C_i**	**Loading**
Fixed Adoption	0.845	Cost Group 1	-0.784
Mobile Adoption	0.864	Cost Group 3	0.747
Tertiary Education	0.745	Cost Group 5	0.211
Computer in Home	0.707	ln(Density)	-0.791
		Rural Blocks	0.856
KMO Statistic	0.820	KMO Statistic	0.804

Cost variables include population density, the share of rural blocks within the block group, an indicator for whether the block group lies within a census place, and two cost variables from the CostQuest data (the first group, the third group, and the fifth group).[45] Again, only one factor has an eigenvalue exceeding 1.0 and the KMO measure of sampling adequacy is 0.80, again quite large. The cost factor C_i is negatively correlated with household density, so costs are lower when density is high.

D. *Coarsened Exact Matching*

Selection bias is addressed using a matching estimator. Since the D_i and C_i are continuous, CEM is applied. We match on both D_i and C_i and to ensure matches are drawn from areas with similar fiber deployment rates we also include the average fiber deployment in the county in which the block group is located (F_i) as a matching variable. The matched samples create similar distributions of demand

---

42 *See*, *e.g*., B.B. Tabachnick and L.S. Fidell, USING MULTIVARIATE STATISTICS (2011).

43 All variables are measured at the block group level.

44 Standard rules of thumb for the KMO are: < 0.50 unacceptable; 0.50 to 0.59 poor; 0.60 to 0.69 mediocre; 0.70 to 0.79 middling; 0.80 to 0.89 meritorious; and 0.90 to 1.00 excellent.

45 All variables except for the Urban-Rural Codes are measured at the block group level. The CostQuest cost variables sum to 1.0, so including two provides a reasonable approximation of the cost distribution. *See supra* n. 34 for definitions of the cost categories.

and costs, which permits (in theory) the identification of τ. We note that the Standardized Differences are typically large in the unmatched samples, but well below the 0.25 threshold in the matched samples (see the results tables below).

E. *Defining "Predominantly"*

As this empirical analysis estimates the differences in outcomes between areas that are predominantly composed of a protected class and those that are not, we must define which census block groups are *predominantly* of a protected class and which are not. Our analysis covers race and income, so definitions are required for both.

Minority communities are defined based on the summed population shares of Hispanics and (non-Hispanic) Black Americans, since it is these minorities on which attention has been focused. Also, both racial groups have below average adoption rates for fixed broadband, and both have relatively large population shares.[46] The U.S. share of Black and Hispanic Americans is about 32%.[47] In studies of redlining and other forms of racial discrimination, predominance thresholds for Minority groups often fall in the 50% to 90% interval (a wide range).[48] Rather than choose a specific threshold and in an effort to evaluate heterogeneity in the estimates of τ and ω, we divide the minority shares into ten groups: 0% to 9.9%; 10% to 19.9%, …, 90% to 100%. We can then compare outcomes across the various categories, for example comparing communities with less than 10% Minority share (predominantly non-Minority) and those with 90% or more Minority shares (predominantly Minority). Statistical inference is possible both on the individual groupings and jointly across the groupings.

This approach has several advantages. First, the predominance shares cannot be cherry-picked. Second, comparing multiple groupings allows for an assessment of the heterogeneity in the differences. Third, the challenges in finding communities of equal demand and costs across disparate predominance shares

46 Asian and Multi-racial Americans have adoption rates like those of White Americans. *See* Ford, *supra* n. 10.

47 Data available at: https://www.census.gov/quickfacts/fact/table/US/PST045221.

48 L. Rawlings, L. Harris, and M.A. Turner, S. Padilla, *Race and Residence: Prospects for Stable Neighborhood Integration*, NEIGHBORHOOD CHANGE IN URBAN AMERICA SERIES, Urban Institute (March 2004) (available at: https://www.issuelab.org/resources/7608/7608.pdf); J.S. Rugh, *Why Latinos Were Hit Hardest by the US Foreclosure Crisis*, 93 SOCIAL FORCES 1139-1184 (2015); *NonWhite School Districts Get $23 Billion Less Than White Districts Despite Serving the Same Number of Students*, EDBUILD (2019) (available at: https://edbuild.org/content/23-billion).

can be directly observed, and we believe this observation may be the most important result we provide. For expositional purposes, we limit the groupings for predominantly non-Minority to 30% or less (since the mean Minority share is about 30%).

For income, measured as median household annual income, we take a similar approach, categorizing block groups into six groupings: (1) less than $25,000; (2) $25,000 to $50,000; (3) $50,000 to $75,000; (4) $75,000 to $100,000; (5) $100,000 to $150,000; and (6) $150,000 to $250,000 (the top-coded income). This approach permits large spreads in income and, in turn, a large spread in poverty levels. Average median income for the lowest income group is about $19,000 with a mean poverty rate of 46%. The average median income for the highest income group is about $188,000 with a mean poverty rate of 3.0%.

V. Empirical Results

Our empirical analysis aims to quantify the means differences in the availability of fiber-optic broadband networks and broadband speeds between predominantly Minority and majority areas, which are measured by the τ coefficient. Equations (5) and (6) are used for this purpose. Since fiber deployment is a rate on the unit interval (i.e., it is bounded by zero and one), we estimate the model using a Generalized Linear Model ("GLM") of the binomial family with a logit link. The GLM coefficients are not directly interpretable, so we provide the estimated fiber availability rates.[49] Clustered standard errors at the county level are used to address heteroskedasticity and allowing for correlated disturbances at the county level.[50]

A. *Quantifying Race Discrimination in Fiber Deployment*

We begin with an analysis of digital discrimination based on race. As detailed above, if τ is negative (positive), then fiber networks are less (more) present in areas of a predominantly protected class. The null hypothesis is "no digital discrimination" ($\tau = 0$). Equation (5) provides the means difference including

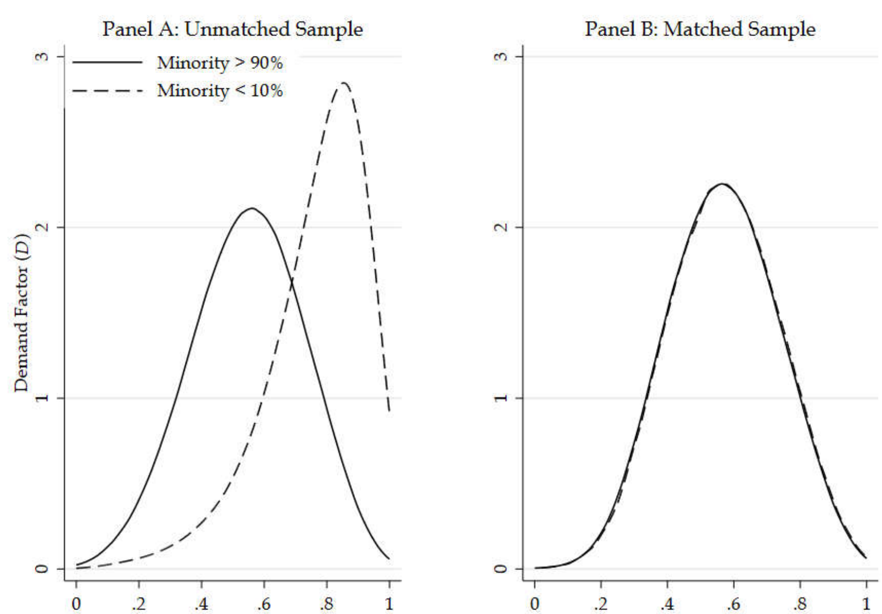
49 Stata (ver. 17) is used for all statistical analysis.

50 Due to clustering, we limit the sample to counties with 10 or more observations since there must be sufficient observations in each cluster to produce meaningful results.

selection bias ($\tau + \omega$) while Equation (6) provides the means difference without selection bias (τ).

Before turning to the estimates, we first offer a visual demonstration of the matching approach. In Figure 3, the distributions of the demand factor D_i are illustrated for the unmatched sample (Panel A) and the matched sample (Panel B). The two groups are minority shares of 90% or more and minority shares of less than 10%. In the unmatched sample, we see wide disparities in the distribution of D_i between the two groups. Given such wide differences, comparing the means of fiber deployment (or other outcomes) across the groups will surely reflect selection bias. In the matched sample, however, the distribution of D_i between the groups is essentially the same. A comparison of mean fiber deployment rates using the matched sample provides a more accurate assessment of discrimination.

Figure 3. Demand (D_i), Unmatched and Matched, Race



Results based on Equations (5) and (6) are summarized in Table 3. The table indicates the base share of minority population, which are lower levels of minority population shares. The fiber availability rates for each group are also provided along with the differences. Statistical significance of the differences is indicated. For each minority share comparison, the results from the unmatched sample ($Y_1 - Y_0 = \tau+\omega$) and the matched samples ($Y_1 - Y_0 = \tau$) are provided in sequence. Observation counts and the share of the sample retained in matched samples are provided. In the final column, the Standardized Differences for the matching variables (D_i, C_i, and F_i) are provided.

Table 3. Fiber Deployment Results for Race
(Unmatched and Matched Samples)

Minority Share (r = 0)	Minority Share (r = 1)	Y_1, Y_0	$\tau + \omega$ / τ	Obs	Matched Share	Stan. Diff. (D_i, C_i, F_i)
0-10%	50%-60%	0.495, 0.414	0.081***	46,971		0.49,0.78,0.22
		0.512, 0.504	0.007	34,443	0.733	0.01, 0.03, 0.01
	60%-70%	0.486, 0.414	0.072***	46,319		0.65,0.86,0.24
		0.500, 0.505	-0.005	30,803	0.665	0.01, 0.04, 0.01
	70%-80%	0.522, 0.414	0.108***	46,361		0.82,0.91,0.31
		0.542, 0.523	0.019	29,000	0.626	0.01, 0.05, 0.02
	80%-90%	0.519, 0.414	0.105***	46,755		1.04,0.97,0.36
		0.535, 0.536	-0.000	27,359	0.585	0.00, 0.07, 0.02
	90%-100%	0.518, 0.414	0.104**	48,950		1.32,1.06,0.35
		0.528, 0.541	-0.014	26,358	0.538	0.02, 0.13, 0.02
10-20%	50%-60%	0.495, 0.479	0.016	26,724		0.58,0.34,0.05
		0.512, 0.515	-0.003	22,284	0.834	0.03, 0.02, 0.00
	60%-70%	0.486, 0.479	0.007	26,072		0.74,0.43,0.07
		0.501, 0.515	-0.014	20,526	0.787	0.03, 0.03, 0.00
	70%-80%	0.522, 0.479	0.043**	26,114		0.92,0.50,0.14
		0.537, 0.531	0.006	19,839	0.76	0.03, 0.04, 0.00
	80%-90%	0.519, 0.479	0.040	26,508		1.14,0.58,0.19
		0.533, 0.542	-0.009	19,290	0.728	0.03, 0.06, 0.00
	90%-100%	0.518, 0.479	0.039	28,703		1.44,0.70,0.19
		0.519, 0.535	-0.016	19,504	0.68	0.02, 0.09, 0.00
20-30%	50%-60%	0.495, 0.476	0.018	19,606		0.44,0.21,0.04
		0.514, 0.51	0.004	16,740	0.854	0.03, 0.00, 0.00
	60%-70%	0.486, 0.476	0.010	18,954		0.60,0.30,0.06
		0.503, 0.513	-0.010	15,582	0.822	0.03, 0.01, 0.00
	70%-80%	0.522, 0.476	0.046**	18,996		0.78,0.37,0.12
		0.540, 0.523	0.016	15,099	0.795	0.04, 0.02, 0.00
	80%-90%	0.519, 0.476	0.043*	19,390		1.01,0.45,0.18
		0.534, 0.531	0.003	14,957	0.771	0.03, 0.04, 0.00
	90%-100%	0.518, 0.476	0.042	21,585		1.30,0.58,0.17
		0.522, 0.522	0.000	16,201	0.751	0.01, 0.07, 0.00

Stat Sig. *** 1% , ** 5%, * 10%

The table is interpreted as follows. The first two rows of the table compare block groups with less than 10% minority population to block groups with at least 50% but less than 60% minority population. For the unmatched sample in the first row, the difference in fiber availability ($\tau + \omega$) is positive, large, and statistically different from zero; the predominantly Minority mean is 0.495 and the predominantly non-Minority mean is 0.414 for a difference of 0.081 percentage points. Predominantly minority areas have more fiber deployment than predominantly non-minority areas, suggesting a sort of *reverse* digital discrimination, but this difference includes selection bias (ω). These results comport with the statistics in Table 1. The second row summarizes the results from the matched sample, where 73% of the full sample is retained. The difference in mean fiber deployment is now much smaller ($\tau = 0.007$) and statistically indistinguishable from zero. There is no economically- or statistically-significant digital discrimination.

For the comparisons on a base of 0-10% minority share for the *unmatched* samples, all the differences are positive and statistically different from zero, with differences near 10-percentage points. In the *matched* samples, where selection bias (ω) is eliminated (or, attenuated), none of the differences are statistically different from zero, the differences are small, and the signs of the differences differ across the comparisons. When using base shares of 10-20% or 20-30%, very few means differences are statistically significant and those that are significant are always positive and always include selection bias (i.e., unmatched samples).

It is also possible to conduct a joint test of significance for all comparisons in each base share.[51] The test is based on a sample-size weighted average effect. For the base share of 0-10% Minority, the average means difference is 0.094 with a χ^2 statistic of 12.42, which is statistically different from zero at the 1% level. For the matched samples, the average means difference is 0.002 with a χ^2 of 0.01 and a probability of 0.94. Selection bias (ω) accounts for all the observed difference. For the base share of 10-20% Minority, the average means difference is 0.011 with a χ^2 statistic of 1.81 (prob = 0.18) for the unmatched and -0.003 with a χ^2 of 0.09 (prob = 0.77). Finally, for the base share of 20-30%, the mean difference is 0.007 with a χ^2 of 2.28 (prob = 0.13) for the unmatched sample and 0.0006 for the matched sample with a χ^2 of 0.01 (prob = 0.92). Accounting for selection bias, there is no evidence of digital discrimination.

There is no evidence of meaningful digital discrimination in any of the comparisons; the influence of selection bias is apparent. These results suggest that studies looking only at simple means comparisons in unmatched samples are prone to biased estimates of digital discrimination. Quantifying digital discrimination requires some method to address selection bias, else the conclusions drawn from an analysis are invalid.

A few other results are worth mentioning. First, as the minority shares become more disparate, the share of observations retained in the matched samples decline.

---

51 Imbens and Wooldridge, *supra* n. 26 at p. 33-4.

At more extreme shares, it becomes more difficult to find suitable matches. Also, the Standardized Differences on the matching variables are often very large in the unmatched samples, but all such differences are small in matched samples.

B. *Quantifying Income Discrimination in Fiber Deployment*

Table 4 summarizes the results for digital discrimination based on income. The table is interpreted as before. With income, we see very large Standardized Differences in D_i and C_i for both base income categories, with huge disparities in D_i. Importantly, the share of observations retained in matching are low especially across widely disparate income groupings. For instance, only 12.2% of the sample is retained when analyzing incomes below $25,000 to those above $150,000 and $150,000 (only 1,666 block groups are retained in the matched sample. Thus, quantifying income discrimination, a somewhat puzzling addition to the traditional set of protected classes, may be challenging as it will be difficult to craft a useful counterfactual.

Table 4. Fiber Deployment Results for Income (Unmatched and Matched Samples)

Income Level (r = 1)	Income Level (r = 0)	Y_1, Y_0	$\tau + \omega$ τ	Obs	Matched Share	Stan. Diff. (D_i, C_i, F_i)
$0-25k	$50-75k	0.396, 0.437	-0.041	42,005		1.65,0.56,0.00
		0.401, 0.438	-0.036	28,727	0.684	0.02, 0.02, 0.01
	$75-100k	0.396, 0.470	-0.074**	29,568		2.20,0.66,0.12
		0.468, 0.475	-0.006	15,820	0.535	0.03, 0.03, 0.00
	$100-150k	0.396, 0.533	-0.138***	25,846		2.81,0.69,0.31
		0.546, 0.549	-0.003	8,721	0.337	0.21, 0.04, 0.00
	$150-250k	0.396, 0.620	-0.224***	13,662		3.37,0.80,0.57
		0.664, 0.637	0.026	1,666	0.122	0.17, 0.15, 0.01
$25-50k	$50-75k	0.418, 0.437	-0.019*	66,742		0.77,0.26,0.04
		0.421, 0.432	-0.011	63,121	0.946	0.03, 0.01, 0.00
	$75-100k	0.418, 0.470	-0.053***	54,305		1.31,0.35,0.16
		0.441, 0.465	-0.024	45,478	0.837	0.09, 0.02, 0.00
	$100-150k	0.418, 0.533	-0.116***	50,583		1.90,0.33,0.36
		0.477, 0.492	-0.015	32,525	0.643	0.22, 0.03, 0.01
	$150-250k	0.418, 0.620	-0.203***	38,399		2.44,0.38,0.63
		0.583, 0.577	0.006	10,852	0.283	0.28, 0.08, 0.00

Stat Sig. *** 1% , ** 5%, * 10%

As with race, there is no evidence of digital discrimination on income levels. While large, negative differences are sometimes found in the unmatched samples (low-income areas have less fiber), these differences appear to reflect only selection bias. In the matched samples, which exposes τ, the means differences are small and often positive, and none are statistically different from zero. As for the joint test, the null hypothesis of the difference for $0-25,000 base income (-0.095) is rejected at the 1% level for the unmatched sample (χ^2 = 10.67), but the null is not rejected for the matched sample difference of 0.022 (χ^2 = 0.53). Likewise, for the $25-50,000 base income group, the mean difference of -0.055 is statistically significant at the 1% level (χ^2 = 21.76) and for the matched sample the small difference of -0.011 is not statistically different from zero (χ^2 = 0.57). The null hypothesis of "no digital discrimination" cannot be rejected. Again, simplistic comparisons of fiber deployment offer a biased indicator of discrimination. While fiber deployment rates may differ in a simple means comparison, these differences appear only to reflect differences in profitability.

C. *Quantifying Discrimination in Broadband Speeds*

Thus far we look only at fiber deployment. Fiber is not the only technology capable of very-high speed broadband. While pushing fiber deeper into their networks, the cable industry, the largest provider of broadband services in the nation typically have deployed hybrid fiber-coax networks which can deliver fiber-type speeds to consumers. Therefore, comparing average download speeds is another way to study differential treatment.

Using the same empirical format, we replace fiber technology deployment with (the natural log of) maximum advertised download speeds in the census block group. A few download speeds exceed 1 Gbps, and for these we set the value at 1 Gbps to avoid distorting the results.[52] Results are summarized in Table 6. Since speeds are continuous, the model is estimated by OLS with clustered standard errors at the county level.

[52] This modification of the data has very little effect on the results.

Table 5. Download Speed Results for Race (Unmatched and Matched Samples)

Minority Share (r = 0)	Minority Share (r = 1)	Y_1, Y_0	$\tau + \omega$ / τ	Obs	Matched Share	Stan. Diff. (D_i, C_i, F_i)
0-10%	50%-60%	1086, 1022	64.1***	46,971		0.49,0.78,0.22
		1044, 1060	-15.8	34,443	0.733	0.01, 0.03, 0.01
	60%-70%	1084, 1022	62.7***	46,319		0.65,0.86,0.24
		1029, 1054	-25.0	30,803	0.665	0.01, 0.04, 0.01
	70%-80%	1091, 1020	71.0***	46,361		0.82,0.91,0.31
		1029, 1052	-23.4	29,000	0.626	0.01, 0.05, 0.02
	80%-90%	1090, 1018	71.9***	46,755		1.04,0.97,0.36
		1024, 1043	-19.7	27,359	0.585	0.00, 0.07, 0.02
	90%-100%	1094, 1013	81.4***	48,950		1.32,1.06,0.35
		1020, 1026	-6.30	26,358	0.538	0.02, 0.13, 0.02
10-20%	50%-60%	1030, 1038	-7.40	26,724		0.58,0.34,0.05
		1023, 1049	-26.2**	22,284	0.834	0.03, 0.02, 0.00
	60%-70%	1027, 1035	-8.70	26,072		0.74,0.43,0.07
		1014, 1045	-31.4**	20,526	0.787	0.03, 0.03, 0.00
	70%-80%	1032, 1032	-0.80	26,114		0.92,0.50,0.14
		1016, 1045	-29.4**	19,839	0.760	0.03, 0.04, 0.00
	80%-90%	1031, 1030	0.20	26,508		1.14,0.58,0.19
		1003, 1021	-18.2**	19,290	0.728	0.03, 0.06, 0.00
	90%-100%	1035, 1025	9.50	28,703		1.44,0.70,0.19
		999, 1011	-12.4	19,504	0.680	0.02, 0.09, 0.00
20-30%	50%-60%	1028, 1026	2.50	19,606		0.44,0.21,0.04
		1021, 1031	-10.3	16,740	0.854	0.03, 0.00, 0.00
	60%-70%	1024, 1023	1.20	18,954		0.60,0.30,0.06
		1007, 1028	-21.1**	15,582	0.822	0.03, 0.01, 0.00
	70%-80%	1028, 1018	9.10	18,996		0.78,0.37,0.12
		1001, 1020	-18.8**	15,099	0.795	0.04, 0.02, 0.00
	80%-90%	1026, 1016	10.0	19,390		1.01,0.45,0.18
		997, 1015	-18.6**	14,957	0.771	0.03, 0.04, 0.00
	90%-100%	1029, 1010	19.2	21,585		1.30,0.58,0.17
		996, 1009	-13.6	16,201	0.751	0.01, 0.07, 0.00

Stat Sig. *** 1% , ** 5%, * 10%

Table 5 shows that the average maximum download speeds are approximately 1 Gbps for all groups across all comparisons. In the matched samples, most of the differences are both statistically and economically insignificant. While a few estimates are statistically significant in matched samples, the differences are trivially small and likely related to the differences in the way speeds are reported (e.g., some cable providers report 1 Gbps speeds while others report 940 Mbps). Even when the differences are statistically significant, the difference in speeds is no more than about 3%. In any case, a few Mbps difference on a near 1 Gbps connections have no meaningful effect on consumers.

Table 6. Download Speed Results for Income (Unmatched and Matched Samples)

Income Level (r = 1)	Income Level (r = 0)	Y_1, Y_0	$\tau + \omega$ τ	Obs	Matched Share	Stan. Diff. (D_i, C_i, F_i)
$0-25k	$50-75k	1020, 1063	-42.9***	42,005		1.65,0.56,0.00
		1001, 1001	0.10	28,727	0.684	0.02, 0.02, 0.01
	$75-100k	1022, 1045	-22.7*	29,568		2.20,0.66,0.12
		999, 1005	-6.30	15,820	0.535	0.03, 0.03, 0.00
	$100-150k	1024, 1010	14.2	25,846		2.81,0.69,0.31
		1034, 1012	22.2	8,721	0.337	0.21, 0.04, 0.00
	$150-250k	1039, 997	42.0	13,662		3.37,0.80,0.57
		1054, 1024	30.2	1,666	0.122	0.17, 0.15, 0.01
$25-50k	$50-75k	1020, 1034	-13.4**	66,742		0.77,0.26,0.04
		1018, 1018	-0.30	63,121	0.946	0.03, 0.01, 0.00
	$75-100k	1030, 1024	6.40	54,305		1.31,0.35,0.16
		1022, 1026	-4.00	45,478	0.837	0.09, 0.02, 0.00
	$100-150k	1050, 1007	43.3***	50,583		1.90,0.33,0.36
		1017, 1022	-4.70	32,525	0.643	0.22, 0.03, 0.01
	$150-250k	1083, 1011	72.6***	38,399		2.44,0.38,0.63
		1113, 1057	55.9	10,852	0.283	0.28, 0.08, 0.00

Stat Sig. *** 1%, ** 5%, * 10%

In Table 6, the results for download speeds based on income are summarized. As with race, all the speeds are at the 1 Gbps level and the differences are mostly statistically insignificant, trivially small, and likely undetectable by consumers. Such small differences may be, in part, attributed to the reporting vagaries of providers. The joint tests by base group indicate trivial and mostly statistically insignificant differences. For the $0-25,000 group, the average difference is 0.013 Mbps for the unmatched sample (prob = 0.22) and -0.0017 for the matched sample (prob = 0.86). For the $25-50,000 base group, the average difference is -0.013 Mbps for the unmatched sample (prob = 0.011) and -0.001 for the matched sample (prob = 0.88). There is no meaningful difference in speeds either by race or income, so there is no evidence of digital discrimination.

VI. Caveats

The analysis above aims to provide a working definition of digital discrimination and provides some empirical tests for such discrimination. We find no evidence of systematic or meaningful digital discrimination with respect to fiber deployment or download speeds. We rely on demand and cost data to address selection bias, which at times presents challenges especially for income discrimination. Also, fiber networks are still being deployed and have not reached maturity, so what is observed in any particular release of the Form 477 data may not reflect reality six-months later (and the data is already nearly two-year out of date). Likewise, there may be patterns in deployment unrelated to race or income, such as regulatory barriers or an area's provider using a different technology, that may be correlated with race and income. Any analysis of fiber deployment that ignores continuing network expansion will be parochial at best.

We also suspect that the Commission will, in some cases, address digital discrimination on a provider-by-provider, market-by-market basis. Here, we look for systematic discrimination across many markets and all providers. The results are useful, but there may be instances of disparate treatment in particular market areas that require explanation. That said, a single instance of disparate treatment does not imply discrimination, absent good evidence of equal profitability. In some respects, we believe our approach may be more informative, since the analysis of small areas will be prone to idiosyncratic influences that may be difficult to ascertain and quantify.

At the center of our analysis is the idea that discrimination is costly to a broadband provider. As profit maximizing firms, discrimination seems improbable as profits are foregone when equally profitable areas are treated differently based on something like racial animus. It seems likely that cost and demand differences will explain differential treatment, and such differences do not reflect discrimination. Costs and demand may be difficult to measure in small areas, complicating the Commission's task. (Normally, one might substitute income for demand, but that is precluded by adding income to the traditional set of protected classes.) Yet, differences in demand and costs must be accounted for, since the Infrastructure Act uses the term *digital discrimination* and not *digital equity* or *digital equality*, and it includes references to economic and technical feasibility, both of which speak to profitability. Discrimination, at least in economics, is a term of art, and implies differential treatment of equally-profitable groups based on membership in some sort of protected class, and this definition seems to comport with the statute and the Commission's concerns expressed in its *NOI*.

VII. Conclusion

In this POLICY PAPER, we provide a conceptual framework for analyzing the Infrastructure Act's digital discrimination provisions. We compare fiber broadband deployments and download speeds between predominantly Minority and non-Minority communities and between low-income and high-income communities. Our analysis aims to separate economic factors from race and income, since discrimination requires differential treatment for equally profitable consumers. Our findings are encouraging—we find no meaningful evidence of digital discrimination in either race or income for fiber deployments or for download speeds. Discrimination is costly to the firm (i.e., forgone profits), so these results are consistent with profit-maximizing behavior by providers.

Digital discrimination may take many forms; we look only at fiber deployment and download speeds. There are other outcomes of potential interest and other ways to interpret the statute. Discrimination, however, implies differential treatment that reflects factors beyond the economic and technical. In areas where demand is low or costs are high (or both), broadband providers cannot profitably provide service, but there is no discrimination implied by the failure to do so. A gap in coverage is a necessary but not a sufficient condition for digital discrimination. Banning rational economic behavior, as some advocates have proposed, would be an extreme and dangerous interpretation of the digital discrimination provision. Certainly, if Congress intended coverage shortfalls to constitute discrimination, then it would not have provided $42.5 billion in subsidy dollars to expand the availability of broadband service. Likewise, if affordability issues constituted discrimination, then Congress would not have provided in the Infrastructure Act nearly $20 billion in low-income subsidies for the Affordable Connectivity Program ("ACP") and Tribal Broadband Connectivity Program, among others.[53]

The solution to broadband availability shortfalls is to subsidize broadband deployment in marginal areas. The Infrastructure Act provides such subsidies in large measure. Once these subsidies are fully allocated and networks are built, which will take years, then any shortfall in coverage might be better blamed on insufficient funding or improper allocation of available funds, both of which are in the lap of government.[54] By some estimates, the $42.5 billion in subsidies will be adequate to serve almost all unserved households, and the new low-earth orbit satellite networks may help reduce the costs of deployment to very remote areas.[55]

---

[53] D. Goovaerts, *Broadband Gets $65 Billion in U.S. Infrastructure Bill – Here's What Happens Next*, FIERCETELECOM (November 8, 2021) (available at: https://www.fiercetelecom.com/telecom/broadband-gets-65-billion-u-s-infrastructure-bill-here-s-what-happens-next).

[54] The federal government does not have a good track record for subsidy allocation. *See*, *e.g.*, T.R. Beard, G.S. Ford, M. Stern, *Bridging the Digital Divide: An Empirical Analysis of Public Programs to Increase Broadband Adoption*, 67 TELEMATICS AND INFORMATICS 101754 (February 2022) (available at: https://www.sciencedirect.com/science/article/abs/pii/S0736585321001933) (and citations therein).

[55] *See*, *e.g.*, G.S. Ford, *Assessing Broadband Policy Options: Empirical Evidence on Two Relationships of Primary Interest*, PHOENIX CENTER POLICY PERSPECTIVE No. 21-04 (July 28, 2021)

(Footnote Continued. . . .)

Plainly, subsidizing network construction in already served areas should be avoided, as it does nothing to increase availability to unserved homes and businesses.

(available at: https://www.phoenix-center.org/perspectives/Perspective21-04Final.pdf) (and citations therein).