https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 71, No. 7, July 2025, pp. 5419–5440
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Can Improvements to Mobile Internet Service Help Reduce Digital Inequality? An Empirical Analysis of Education and Overall Data Consumption

Karthik Babu Nattamai Kannan,[a,*] **Eric Overby,**[b] **Sridhar Narasimhan**[b]

[a] Cox School of Business, Southern Methodist University, Dallas, Texas 75275; [b] Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30308

*Corresponding author

Contact: kbnk@smu.edu, https://orcid.org/0000-0001-5125-9897 (KBNK); eric.overby@scheller.gatech.edu, https://orcid.org/0000-0001-5285-4436 (EO); sri.narasimhan@scheller.gatech.edu, https://orcid.org/0000-0003-2163-3088 (SN)

Received: December 5, 2022
Revised: September 15, 2023; February 26, 2024
Accepted: April 25, 2024
Published Online in Articles in Advance: October 9, 2024

https://doi.org/10.1287/mnsc.2022.03770

Abstract. Quality internet access is critical to participating in contemporary society. Unfortunately, many households—particularly those of low socioeconomic status and/or those in rural areas—do not have quality internet access. Some have no access, whereas others are reliant on their mobile data plans for internet access (i.e., they are "smartphone dependent"). This generates inequality in internet access and use. Given the smartphone dependence of many disadvantaged households, we explore whether improvements to mobile internet service can help reduce digital inequality. We focus on a specific improvement: access to unlimited mobile data. For access to unlimited data to help reduce digital inequality, it must generate larger data consumption increases for disadvantaged households than for advantaged ones, including for data likely to enhance welfare, such as online education content. It is not obvious that this will be the case. Accordingly, we use detailed subscriber-level data from a major telecommunications company to examine changes in the consumption of education and other content after subscribers switch to unlimited mobile data plans. We find that although all subscribers increase their consumption, the increases are significantly larger for disadvantaged subscribers, both for overall content and for education content. This is an important finding given that identifying programs that generate disproportionate data consumption increases for disadvantaged households—including for education and other "enhancing" content—is a necessary step for reducing digital inequality.

History: Accepted by Hemant Bhargava, information systems.
Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mnsc.2022.03770.

Keywords: digital inequality • smartphone dependence • mobile internet • difference in differences • data caps

> Internet access is no longer nice-to-have, but need-to-have for everyone, everywhere. … When we need access to the internet, we aren't thinking about how much data it takes to complete a task, we just know it needs to get done. It's time the FCC take a fresh look at how data caps impact consumers and competition. (Jessica Rosenworcel, Chairwoman, Federal Communications Commission (Federal Communications Commission 2023a))

Introduction

Access to the internet has become a de facto requirement for participating in contemporary society. People without quality internet access may have difficulty completing school assignments, they may be unable to apply for jobs, and they may miss out on broad swaths of contemporary culture (Agarwal et al. 2009, Ransbotham et al. 2016). The importance of internet access became vividly apparent at the beginning of the coronavirus disease 2019 (COVID-19) pandemic when school and work routines shifted from physical buildings to the internet (Sen and Tucker 2020). The U.S. Federal Communications Commission (FCC) has made it clear that providing quality internet access to every citizen is a top priority (Federal Communications Commission 2023a). Unfortunately, there are many people in the United States and around the world who lack quality internet access, particularly those of low socioeconomic status (SES) or those who live in rural areas (Greenstein and Prince 2009, Federal Communications Commission 2022). Some have no access, whereas others are limited by poor connectivity or usage restrictions, such as monthly data caps. Thus, those who might benefit the most from quality internet access—and its ability to provide access to educational content, jobs, and other information—may be the least likely to have it. Inequality in internet access and the resulting "homework gap" in education—which refers to school-age children lacking the internet access they need to complete schoolwork at home (Auxier and Anderson 2020, Federal Communications Commission 2022)—are well documented (Anderson et al. 2022).

One reason for inequality in internet access is that although over 90% of U.S. households have internet access (Perrin and Atske 2021), many households' only access method is via their smartphone and their mobile internet service plan (Perrin 2021, Pew Research Center 2021). This "smartphone dependence" is particularly common for households of low socioeconomic status. This puts these households at a disadvantage compared with households that also have home internet service. This begs the question. Could improvements to mobile internet service help reduce digital inequality? Although there are multiple ways to improve mobile internet service, such as providing faster speeds and more stable connections, we focus on providing access to unlimited mobile data. This focus is consistent with the FCC's June 2023 investigation of whether data caps limit access to important information, particularly for low-income and other disadvantaged households (see the introductory quote (Federal Communications Commission 2023a)). If access to unlimited mobile data is to help reduce inequality, then two conditions must hold. First, access to unlimited mobile data must produce larger increases in data consumption for disadvantaged households than for advantaged households. Second, these larger increases should be for content likely to enhance welfare, such as education content. This is because improvements in internet access can put disadvantaged households at a further disadvantage if the access is used for nonenhancing purposes, such as consuming digital entertainment or gaming (Van Dijk 2017). It is unclear a priori whether these conditions will hold. On one hand, because the mobile internet is often the only method of connectivity for disadvantaged households, eliminating usage restrictions on mobile internet service might be particularly helpful—and yield disproportionate consumption increases—for them. On the other hand, most of the research on digital inequality concludes that improvements to internet service often yield disproportionate increases for *advantaged* households, partly because they are more likely to use the improvements to consume educational and other enhancing content (Bonfadelli 2002, Hargittai and Dobransky 2017). Considering these competing theoretical possibilities, we pose the following research question. After gaining access to unlimited mobile data, do disadvantaged households (i.e., those of low socioeconomic status and/or in rural areas) increase their consumption of educational and other content more than advantaged households do?

We address this question by studying the adoption of unlimited mobile data plans by subscribers of a large telecommunications company. In January 2016, the company announced a program that allowed existing subscribers to switch to unlimited data plans. Our data contain subscriber-month observations from January 2015 to December 2016, including subscribers' data consumption and whether they switched to the unlimited plan. Our data also contain subscribers' income range and zip code, which we use to measure socioeconomic status and urban versus rural. Using a difference-in-differences approach, we find that switching to unlimited data leads to an approximately 12-gigabyte (GB)/month increase in data consumption for rural households and those of low socioeconomic status, which is significantly larger than the corresponding increase for urban households and those of high socioeconomic status, which is ~8.5 GB/month. We find a similar pattern for education data consumption. We find an approximately 24-megabyte (MB)/month increase for rural households and those of low socioeconomic status, which is significantly larger than the corresponding increase for urban households and those of high socioeconomic status, which is ~15 MB/month. The 9-MB/month additional increase for disadvantaged households corresponds to roughly one to two additional electronic textbooks per month or a larger number of educational materials with smaller file sizes. The disproportionate increase in education data consumption may help disadvantaged households narrow gaps in educational outcomes (Fairlie et al. 2010, Rothwell 2022). We extend our results by showing that smartphone dependence is a likely mechanism driving the larger increase in data consumption for disadvantaged households. We also find that increases are larger for households with children who are likely to benefit from increased consumption of education content.

Our key finding is that access to unlimited mobile data generates disproportionate data consumption increases for disadvantaged households, including—and critically—for education data. This finding is novel and nonobvious given that many studies have shown that improvements in internet access generate disproportionate increases for *advantaged* households. Our results are important because identifying and implementing programs that yield disproportionate consumption increases for disadvantaged households are necessary to reduce digital inequality. Our results inform the FCC's June 2023 inquiry into the effect of data caps on disadvantaged households, and they have implications for future initiatives to improve mobile internet service. This is because as telecommunication companies and policymakers consider future improvements to mobile internet service, they will be interested in whether these improvements can reduce inequality in internet use. We show that the answer is likely to be yes.

Literature Review

Two literature streams are particularly relevant for our study: (1) research on digital inequality and (2) research on data caps. We review each stream and discuss how our study extends existing research.

Research on Digital Inequality

Digital inequality refers to differences across groups in the adoption and use of information and communication technology (DiMaggio et al. 2004, Racherla and Mandviwalla 2013). We focus on inequality in internet use, which research has shown to contribute to inequality in economic, educational, and social opportunities (see Lythreatis et al. 2022 for a recent review). Early studies of digital inequality focused on differences in physical internet *access* (i.e., some groups had access to the internet, whereas others did not). Reasons for this "first-level digital divide" included a lack of a computer and/or internet connection as well as a lack of interest in digital technology (Helsper and Van Deursen 2015). Although the first-level divide persists, it has narrowed significantly over the years (e.g., Harris et al. 2017). Accordingly, research focus shifted to differences in how groups with physical access *use* the internet (e.g., some groups use the internet more and/or more productively than other groups) (Bonfadelli 2002, Greenwood and Agarwal 2016). One reason for this "second-level digital divide" is differences in digital literacy, which refers to the skills needed to use the internet. These skills include medium-related skills—which describe the ability to use computers, mobile phones, etc. to access the internet—and content-related skills—which describe the ability to find relevant and accurate information on the internet (Van Dijk 2017). The "third-level digital divide" relates to gaps in how effectively internet users translate their use into beneficial outcomes (Helsper and Van Deursen 2015).

The Role of Geography and Socioeconomic Status. Two factors that contribute to digital inequality—in terms of access, use, and outcomes—are geography and socioeconomic status. Regarding physical internet access, households in rural areas and/or households with low socioeconomic status have historically had lower access than urban and high socioeconomic status households. One reason that access differs geographically is that telecommunications companies have greater incentives to build the physical infrastructure for internet service in urban areas than in rural areas given population density and economies of scale (Prieger 2013, Greenstein 2015). Differences in access across socioeconomic status exist because low-income households may not have service at their residences and/or may be unable to pay (Hsieh et al. 2008, Hargittai and Dobransky 2017).

Regarding internet use and outcomes, several studies have shown that households of high socioeconomic status are more likely to use the internet for "enhancing" activities than households of low socioeconomic status, thereby perpetuating digital inequality. This includes using the internet for work-related, education-related, and healthcare-related information (see Robinson et al. 2020 for a review). For example, a key promise of massively online open courses (MOOCs) is to make high-quality educational materials available to anyone via the internet. However, MOOCs mainly serve well-educated, affluent individuals rather than the disadvantaged (Hansen and Reich 2015). Related, households of low socioeconomic status are more likely than those of high status to use the internet for hedonic purposes (Hsieh et al. 2008). Hargittai and Dobransky (2017, p. 195) describe this "usage gap" quite pointedly: "Digital inequality research has established that people from less privileged societal positions are less likely to be internet users, and when they do go online, they tend to partake in fewer capital-enhancing activities than their more privileged counterparts." These findings are consistent with the knowledge gap hypothesis, which states that "as the infusion of mass media information into a social system increases, segments of the population with higher socioeconomic status tend to acquire this information at a faster rate than the lower status segments, so that the gap in knowledge between these segments tends to increase" (Tichenor et al. 1970, pp. 159–160). They are also consistent with the Matthew effect (e.g., Rigney 2010), which is often summarized as "the rich get richer and the poor get poorer." Studies have shown that this pattern holds—and may be exacerbated—when users access the internet via mobile phone instead of a personal computer (Pearce and Rice 2013, Tsetsi and Rains 2017). There is some evidence of a similar usage gap between urban and rural users (e.g., Yu 2010, Robinson et al. 2020), although it is less acute than the gap between high and low socioeconomic groups. A key reason that we study consumption of education data—in addition to consumption of data overall—is because education data are the type of "enhancing" content that prior research has shown that disadvantaged households fall behind in.

The Homework Gap. The "homework gap" refers to inequality in internet access as it relates to access to education content. As such, it is a specific type of digital inequality. The homework gap exists because some school-age children—typically from disadvantaged households—lack the necessary internet access to complete schoolwork at home (Auxier and Anderson 2020, Federal Communications Commission 2022). The homework gap is distinct from the "achievement gap," which occurs when a group of students of a given race/ethnicity, socioeconomic class, etc. significantly outperforms another group in terms of educational achievement (National Center for Education Statistics 2022). In other words, the homework gap does not reflect gaps in educational outcomes—only gaps in access to online education content. However, the homework gap affects the achievement gap. Research has shown that students with poor access to digital education resources—who are often from disadvantaged households—are less likely to receive high grades (Robinson et al. 2018).

Programs for Addressing Digital Inequality. Two ways to address digital inequality are (1) to provide internet access to disadvantaged households that do not have it and (2) to improve the quality and usability of the internet service that disadvantaged households already have. Programs to provide access to households that do not have it include free or subsidized internet service, such as Comcast's Internet Essentials program (Hsieh et al. 2008, 2010; Venkatesh and Sykes 2012; Rosston and Wallsten 2020; Zuo 2021; Federal Communications Commission 2022; Ogbo 2022), as well as government incentives for telecommunications companies to add network infrastructure to connect rural and low socioeconomic status areas (Hauge and Prieger 2015, National Telecommunications and Information Administration 2023). Other programs help disadvantaged households adopt home broadband service by connecting them to internet service subsidies and training programs (Manlove and Whitacre 2019, Beard et al. 2022). Some programs are targeted to the homework gap, such as providing students with home broadband service via Educational Broadband Spectrum, Wi-Fi mesh networks, and Wi-Fi hot spots (Reisdorf et al. 2019, Federal Communications Commission 2023b). Although it is clearly important to provide internet access to unconnected households, only 7% of U.S. households are not online based on a 2021 survey (Perrin and Atske 2021). This reflects the narrowing of the first-level digital divide noted above. Thus, an arguably more fruitful way to reduce digital inequality is to improve the quality and usability of the internet service that households already have (Gonzales et al. 2021). Prior research has examined some efforts, such as shifting from a dial-up connection to a broadband connection (Hitt and Tambe 2007), providing higher-quality content (Viard and Economides 2014), and training households on how to make better use of their existing internet connections (Xiong and Zuo 2019). We study a heretofore unstudied (to our knowledge) approach for reducing digital inequality by improving pre-existing internet service: access to unlimited mobile data via the removal of the data cap.

How Our Study Extends Prior Research. Many existing studies that evaluate programs designed to reduce digital inequality focus on access to internet service rather than on changes in how households use the internet (e.g., Manlove and Whitacre 2019, Beard et al. 2022). Studies that investigate how households use the internet typically rely on cross-sectional questionnaire data or case studies (Maceviciute and Wilson 2018), making it difficult to study changes over time and across groups. We address these gaps by using a nationwide, multimonth, household-level data panel to study whether a program that provides access to unlimited mobile data leads to differential changes in internet use for disadvantaged households versus advantaged households. This responds to recent calls for panel studies of digital inequality (e.g., Hargittai 2021, p. 5). We study not only changes in overall data consumption but also, changes in education data consumption given that the type of internet use is as important or more important than the amount of internet use when studying digital inequality (Van Dijk 2017).

Research on Data Caps

Although unlimited mobile data plans are increasingly popular, capped plans persist (Holslin and Parrish 2023). Data caps are sometimes imposed early in a technology's life cycle, as was the case with 3G wireless (Segall 2011) and is currently the case with many satellite internet services (Christiansen 2023). In June 2023, the FCC identified data caps as being particularly constraining to disadvantaged households such that removing them might have disproportionate benefits for these households. The FCC initiated an inquiry into data caps, which underscores the importance of studying their effects (Federal Communications Commission 2023a). Existing research on data caps has focused mainly on firms' pricing strategies (Sen et al. 2019, Chillemi et al. 2020) and how customers use their data budgets (Xu et al. 2019). A study particularly relevant to our study examines how low-income internet users in India use their data budgets when they have a daily versus monthly data cap (Ramdas and Sungu 2024). These users often exhaust most of their monthly data budget early in the month, preventing them from accessing "life-improving" information later in the month. Results show that the daily cap acts as a commitment device that helps users budget their data consumption more effectively, which leads to an increase in consumption of life-improving information.

How Our Study Extends Prior Research. We extend Ramdas and Sungu (2024)—and other studies that examine data caps—in several ways. First, whereas Ramdas and Sungu (2024) study the effect of switching from a monthly data cap to a daily data cap, we study the effect of removing the data cap entirely. This distinction is important because the Ramdas and Sungu (2024) results do not offer a clear prediction of what might occur when data caps are removed. On one hand, Ramdas and Sungu (2024) show that when internet users are faced with data shortages because of burning through data early in the month, they reduce their consumption of education and health content later in the month. If data shortages are eliminated (e.g., via removal of the data cap as in our setting), then we might expect an increase in consumption of education data, with this increase being disproportionately large for disadvantaged households, such as those that Ramdas and Sungu (2024) study. On the other hand, Ramdas and Sungu (2024) show that when internet users do not feel data constrained (in their case, early in the month), they binge on digital entertainment and compulsively check social media. If data caps are removed (as in our setting), then we might expect no increase in consumption of education data given the revealed preference for other types of data. In other words, once the commitment device that curbs undesirable behavior (the daily data cap) is removed, the corresponding desirable behavior may disappear. This theoretical ambiguity illustrates the need for our study. Second, our results apply to a different set of users. By construction, the Ramdas and Sungu (2024) intervention applies to users who remain subject to a data cap; the intervention that we study applies to users who do *not* remain subject to a data cap. The latter group is potentially much larger and warrants dedicated analysis. Third, we observe effects for both advantaged and disadvantaged users, which allows us to test empirically whether the intervention that we study can help close the gap between the two. By contrast, because Ramdas and Sungu (2024 observe effects for disadvantaged users only, they cannot comment empirically on whether the intervention they study can close gaps. Fourth, in addition to studying both advantaged and disadvantaged households, we study a larger sample of users ($n = 250{,}904$ versus $n = 929$) over a longer time period (104 versus 12 weeks) than Ramdas and Sungu (2024) (and related studies). This improves the generalizability of our results.

Theoretical Motivation and Tension

Improving mobile internet service for those who already have it by providing unlimited mobile data is an attractive approach because it does not require households to invest in new equipment given that most households already have a smartphone with internet access (Perrin and Atske 2021). Furthermore, several papers have shown that expansion of internet service improves overall economic and social outcomes (Ganju et al. 2016). However, if access to unlimited mobile data is to reduce digital inequality, then it must generate disproportionate increases in data consumption for disadvantaged households, including for "enhancing" content, such as education data. This is because if all households experience proportionate increases, then the gap between advantaged and disadvantaged households will persist. Whether access to unlimited mobile data can help reduce the gap, does not affect the gap, or even exacerbates the gap is unclear a priori. Given the nature of our study, we focus on the gap in internet use, including the amount and type of use (e.g., accessing education content). Later, we discuss the link between use and outcomes.

On one hand, access to unlimited mobile data may reduce the usage gap by generating disproportionate increases in data consumption for disadvantaged households. The logic for this is as follows. First, many disadvantaged households are "smartphone dependent" and lack other options for internet access (Tsetsi and Rains 2017). According to recent surveys (Perrin 2021, Pew Research Center 2021), approximately 27% of adults with annual incomes below $30,000 are "smartphone-only" internet users compared with 6% of adults with incomes above $75,000. Similar differences exist regarding use of the internet for education purposes. Twenty-four percent of U.S. teens from households with incomes less than $30,000 cannot complete their homework because of the lack of a computer or internet connection at home (Anderson et al. 2022). This is three times more than households with incomes exceeding $75,000. Many teens from low-income households who *are* able to complete online homework must do so using their mobile phones. Thirty-five percent of low-income teens have to complete homework on their mobile phone, which is almost two times more than high-income teens. Given the smartphone dependence of many disadvantaged households, mobile data caps are likely to be binding for them in a way that they are not for advantaged households. Whereas advantaged households can use their home broadband service if/when their mobile data are exhausted, disadvantaged households—if smartphone dependent—cannot. Thus, disadvantaged households might take fuller advantage of—and experience larger consumption increases from—access to unlimited data, thereby reducing the gap in both overall and education data consumption. Indeed, policymakers in several countries have presented improvements to mobile internet service as a remedy for digital inequality given the high levels of smartphone dependence among disadvantaged households (Correa et al. 2021). Second, disadvantaged households spend a relatively large portion of their income on phone service, forcing them to pay a higher relative price for service compared with advantaged households. Behavioral economics research on mental accounting (e.g., Just and Wansink 2011) suggests that this will cause disadvantaged households to value the service more highly, potentially leading to greater data consumption increases (thus, reducing the gap) once the data cap is removed.

On the other hand, access to unlimited mobile data may widen the usage gap between advantaged and disadvantaged households. Much of the existing literature—reviewed above—suggests that this will be the case. The logic for this is as follows. First, as discussed above, research on digital inequality suggests that members of disadvantaged households often lack the digital literacy skills to take advantage of improvements in data/internet access, which widens the gap between them and advantaged households. In our setting, we can assume that all households have adequate *medium-related* digital literacy skills given that we observe all of them using their mobile phones to consume internet data. However, disadvantaged households may lack the *content-related* digital literacy skills to expand their content consumption, including of educational materials (Correa et al. 2021). This is consistent with research that shows that smartphone-only users use the internet for a relatively narrow range of activities, such as accessing social networking sites (Pearce and Rice 2013). Second, when given better access to education (and other) data, members of disadvantaged households may not have the same social expectation and/or support systems as members of advantaged households to take advantage of it (e.g., Helsper 2021). This follows from research that shows that socioeconomic status is positively correlated with educational expectations (e.g., Zhang et al. 2023) and that influence from personal connections affects overall internet use more strongly for advantaged users than disadvantaged users (Hsieh et al. 2008). These differing social expectations and support systems could widen the gap both for overall data consumption and for education data consumption.

It is also possible that we will see mixed effects. In particular, the mechanisms for reducing the gap (e.g., the lack of a nonmobile option for smartphone-dependent households and mental accounting factors) may dominate those for exacerbating the gap (e.g., differing content-related digital skills and social expectations) for overall data consumption but not for education data consumption. On balance, this mixed result seems the most likely theoretically given that research consistently shows that advantaged households use the internet to consume educational and other "enhancing" content more so than disadvantaged households. If this is the case, then we will see (1) disproportionate increases for disadvantaged households in overall data consumption but (2) disproportionate increases for advantaged households in education data consumption (or proportionate increases for both type of households in education data consumption). However, because there are plausible theoretical explanations for any pattern of results, it is important to conduct empirical testing.

Data and Estimation Strategy

Empirical Context and Data Overview

We conduct an observational study with archival data to examine whether access to unlimited mobile data yields disproportionate increases in data consumption for disadvantaged households. We use subscriber-level data from January 2015 to December 2016 provided by a large U.S. telecommunications company. In January 2016, the company announced a program that allowed existing subscribers to switch to an unlimited mobile data plan (i.e., with no mobile data cap). Prior to introducing this plan, the company offered only mobile data plans with a data cap. The company introduced the new plan in response to (1) competition from other companies' introduction of unlimited mobile data plans, which were becoming more popular at the time (Snider and Baig 2017); (2) customer demand for more data; and (3) improved network capacity (Ericsson 2017). The plan provides unlimited mobile data to all individuals subscribed via the same mobile account (e.g., a family's account might cover multiple family members). The plan charged a base price for the first line and an additional charge for each extra line. Pricing of the unlimited plan was the same for all subscribers. On average, subscribers who switched to the unlimited plan paid \$10 more per month than they had before they switched, partly because many previously had plans with relatively large data caps.[1] To continue with the unlimited plan after the first three months, households had to subscribe to at least one additional service beyond wireless; otherwise, they would be placed into a plan with a data cap. Households could also voluntarily switch back to plans with data caps. Ninety-two percent of households that switched to the unlimited plan kept it throughout our analysis period, and results are consistent if we drop households that did not, as we discuss below. The plan did not permit tethering (i.e., using the phone as a hot spot for another device). Thus, all data consumption was related to activity on the subscribers' phone(s). Some subscribers switched to the unlimited mobile data plan immediately, some switched to it later, and some did not switch during the analysis time period. We leverage this variation in a difference-in-differences setup to measure the increase in data consumption after switching to unlimited data, with our focus on whether the increase varies across advantaged versus disadvantaged households.

We measure monthly data use at the household level *i*—which is the sum of data use for all lines in that household—to create a household/month panel. We excluded households that changed addresses during the study period; this allows us to cleanly categorize each household as urban, rural, etc. We also only include households that were subscribers prior to the start of and after the conclusion of our study period.

This yields a stable set of households that we observe throughout the study period.

Household-Level Variables. We measure household-level variables from the subscriber data, including data consumption, income, geographic location, plan information, household characteristics, and whether the household switched to the unlimited data plan. $Data_{it}$ is each household i's overall data use in month t. The telecommunications company categorized approximately one third of each household i's data use based on the websites visited or apps used. We use the company's categorization system and focus on $Data\_Edu_{it}$, which is data use in the education category. This consists of data associated with visits to websites for educational lessons and courses, such as khanacademy.org, udemy.com, scholastic.com, and purplemath.com, as well as educational reference materials, such as reference.com, wikipedia.org, and dictionary.com. Because of the telecommunications company's data storage procedures, the education category data were only available from November 2015 to September 2016. $Income_i$ is recorded in the subscriber data as a range. We categorize household i by socioeconomic status into *Low SES*, *Mid SES*, or *High SES* if its $Income_i$ was below \$20,000, between \$20,000 and \$125,000, and above \$125,000, respectively. We use income to measure socioeconomic status because prior studies have shown it to be the strongest indicator (DiMaggio et al. 2004). % $Rural_i$ is based on household i's zip code as recorded in the subscriber data. We matched zip codes to zip code tabulation areas (ZCTAs) and used the percentage of rural households for the ZCTA from the 2010 Census. We categorize each household's location as "*Rural*" (99%–100% rural), "*Mostly Rural*" (50%–99% rural), "*Mostly Urban*" (1%–50% rural), and "*Urban*" (0%–1% rural). The three socioeconomic groups and the four geographic groups yield a total of 12 geosocial groups: *Urban-Low SES*, *Urban-Mid SES*, *Urban-High SES*, etc. *Dec15 Plan Charges*$_i$ is the amount that household i paid for mobile service in December 2015, which is the month before the introduction of the unlimited mobile data plan. *Dec15 No. of Lines*$_i$ is the number of lines for household i in December 2015, which proxies for household size. *Dec15 Above Cap*$_i$ is a dummy variable indicating whether household i exceeded its data cap in December 2015. *Estimated No. of Children*$_i$ is the number of children in household i, which is estimated for the telecommunications company by a market research firm. This variable is recorded as zero for households with no children and when the number of children in the household is undetermined, which we discuss further below. $Unlimited_{it}$ is a dummy variable indicating whether household i had the unlimited data plan in month t.

Zip Code-Level Variables. We use zip code-level measures from the American Community Survey in our robustness checks and mechanism analysis. % $Broadband_k$ is the percentage of households in zip code k with broadband internet service of any type. % *Mobile Broadband Only*$_k$ is the percentage of households in zip code k with *only* mobile broadband service. % *Mobile and Other Broadband*$_k$ is the percentage of households in zip code k with both mobile and other broadband service, typically home service. % *Smartphone Only*$_k$ is the percentage of households in zip code k that have only a smartphone. % *Smartphone and Other Device*$_k$ is the percentage of households in zip code k that have both a smartphone and another computing device(s). % *Bachelors Degree or Higher*$_k$ is a zip code-level measure of educational attainment.

Estimation Strategy

A concern with measuring the increase in data consumption after a household switches to unlimited data is a potential selection bias; specifically, households choose to switch as opposed to being randomly assigned. As a result, households that switch to unlimited data may be systematically different from households that do not switch to unlimited data in unobserved ways that might bias our estimation. Although we cannot completely eliminate this potential selection bias, we take the following steps to mitigate it.

First, following Hosanagar et al. (2014) and Jung et al. (2019), we include in our analysis only those households that switched to unlimited data at some point between January 2016 and December 2016. We use households that switched in June 2016 or later ("late-16" households) as counterfactuals for households that switched in January 2016, February 2016, or March 2016 ("early-16" households). This accounts for unobserved factors that determine *whether* households switch to unlimited data (because all households in our analysis switched, just in different months), although it does not account for unobserved factors that determine *when* households switch (which we examine in the Robustness Checks and Alternative Analyses section below). We use the 17-month period from January 2015 to May 2016 for our analysis window. During this window, none of our counterfactual "late-16" households had yet switched to unlimited data. This allows us to cleanly estimate the increase in data consumption for the early-16 households after they switch (at least for the first few months).

Second, to strengthen the validity of the counterfactuals, we match each early-16 household to a counterfactual late-16 household. This increases the likelihood that a change in data use after switching to the unlimited plan is because of the switch and not because of systematic differences between households. Our matching procedure is as follows. First, consider early-16 households that switched in January 2016. We used propensity score matching to match these to late-16 households based on their data use ($Data_{it}$) in June through December 2015 (i.e., in the seven months before they switched to the unlimited plan), $Dec15\ Plan\ Charges_i$, $Dec15\ No.\ of\ Lines_i$, and the geosocial groups defined above. Our approach of matching on pretreatment values of $Data_{it}$ follows from Chabé-Ferret (2017), who recommends this approach when several observations of the pretreatment outcome variable are available.[2] The matching procedure ensures that the early-16 households are matched to late-16 households with similar pretreatment data use patterns, similar willingness to pay for mobile service, similar household sizes—including the likely presence of children—and similar socioeconomic and geographic backgrounds. We also matched based on the zip code-level variables: % $Broadband_k$, % $Mobile\ Broadband\ Only_k$, % $Mobile\ and\ Other\ Broadband_k$, % $Smartphone\ Only_k$, and % $Smartphone\ and\ Other\ Device_k$. Although only available at the zip code level, matching on these variables increases the likelihood that early-16 households are matched to late-16 households with similar levels of smartphone dependency. We set aside the matches for this January "cohort," so they were not used in the next step. Next, consider households that switched in February 2016. We used the same procedure after incrementing (by one) each month used for matching on pretreatment values of $Data_{it}$. We proceeded analogously for the March switchers. We dropped early-16 households for which we could not find a match. In cases in which we found more than one match for an early-16 household, we randomly selected one of the matches. As shown in Figure A1 and Table A1 in the Online Appendix, our approach yields good balance. We use two link functions for the propensity score: (1) a logit link function, on which our focal results are based, and (2) a random forest link function, the results of which are similar and appear in Table A2 in the Online Appendix.

Third, we include household fixed effects in our regressions. This controls for unobserved household characteristics that are time invariant, such as demographics and geographic location, which do not vary for households in our sample, as noted above.

Fourth, we conduct a series of robustness checks to assess selection bias. These are discussed below.

Despite our efforts, it is possible that our estimate of the increase in data consumption after switching to unlimited mobile data is biased. However, our primary interest is in the differential changes in data consumption between disadvantaged and advantaged households. As long as any bias not accounted for by our estimation strategy affects our estimates in a similar way for different household groups, then our conclusions about differences between groups will be valid.

The matching process yields 250,904 households. Table 1 shows the distribution of the matched sample across socioeconomic and geographic groups. Even the smallest cell—mostly rural and low SES—has more than 800 households. As such, we believe that we have enough data to generate valid conclusions for each socioeconomic and geographic group. Table 2 and Table A3 in the Online Appendix show summary statistics and correlations for the matched sample. To limit the possibility of outliers skewing the results, we winsorize $Data_{it}$ and $Data\_Edu_{it}$ at the 1st and 99th percentiles, although results are similar if we do not winsorize. Figure 1 illustrates our study design.

Table 1. Number of Households in the Matched Sample by Socioeconomic and Geographic Group

SES \ Geography	Urban	Mostly urban	Mostly rural	Rural	Total
Low SES	1,824	3,180	860	1,434	7,298
Mid SES	51,986	78,710	13,988	15,524	160,208
High SES	33,698	44,832	3,124	1,744	83,398
Total	87,508	126,722	17,972	18,702	250,904

Main Results

Model-Free Analysis

Figure 2 plots the monthly data use of the households that switched to unlimited data in February 2016 and their matched counterfactuals that switched in June or later. We use February households for this analysis because that allows us to examine the education category data for three months prior to switching and four months postswitching, including the switching month. Figure 2 shows that in the months before February 2016, monthly data use was essentially the same for both groups. Starting in February 2016, there is an increase in data use for the households that switched relative to their counterfactuals; note that February typically has lower than average data use because it has 28 days. This suggests that switching to unlimited data increases data use, both in aggregate and for the education category.

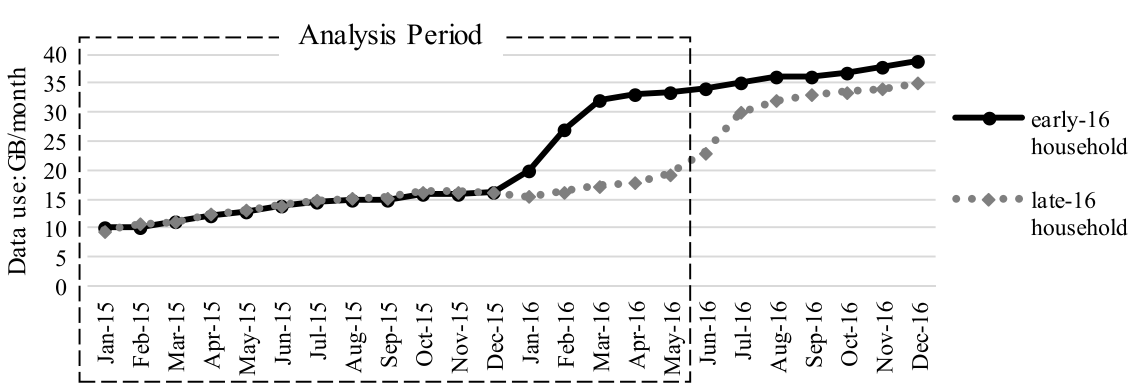
Our primary interest is in whether this increase varies between disadvantaged and advantaged households. Figure 3 decomposes the data shown in Figure 2 by high versus low socioeconomic status and urban versus rural households. The results indicate that switching is associated with a larger increase in data use both in aggregate and for the education category for rural households and those of low socioeconomic status. The "Low SES" panel for the education category shows a slightly higher level of data use for early-16 households prior to switching. Despite this, there is a pronounced increase for these households after switching. The more formal analysis below indicates that the larger increase in education data use for low SES households compared with high SES households is not an artifact of a pre-existing difference in early-16 versus late-16 low SES households in our sample.

Table 2. Summary Statistics for the Matched Sample

Variable	Description	Mean	Standard deviation	Min	Max
$Unlimited_{it}$	Indicator for whether household i has the unlimited plan in month t	0.12	0.32	0.00	1.00
$Income_i$	Household i's income category. 1: <\$20,000; 2: \$20,000–\$39,000; 3: \$40,000–\$74,000; 4: \$75,000–\$125,000; 5: >\$125,000. In our analysis, we coarsen these into 3 groups: low SES, mid SES, and high SES.	3.79	1.10	1.00	5.00
% $Rural_i$	Percentage of rural households in household i's zip code	19.67	30.04	0.00	100.00
Dec15 Plan Charges$_i$	Monthly amount household i paid for mobile service in December 2015 by category. 1: <\$50; 2: \$50–\$100; 3: >\$100; see table notes.	2.00	0.73	1.00	3.00
Dec15 No. of Lines$_i$	Number of lines for household i in December 2015	3.62	1.43	1.00	10.00
Dec15 Above Cap$_i$	Indicator for whether household i exceeded its data cap in December 2015	0.08	0.26	0.00	1.00
% $Broadband_k$	Percentage of households in zip code k with broadband internet service of any type	79.60	11.30	0.00	100.00
% *Mobile Broadband Only*$_k$	Percentage of households in zip code k with *only* mobile broadband service	8.05	4.51	0.00	100.00
% *Mobile and Other Broadband*$_k$	Percentage of households in zip code k with both mobile and other broadband service, typically home service	43.61	12.75	0.00	100.00
% *Smartphone Only*$_k$	Percentage of households in zip code k that have only a smartphone	4.00	2.88	0.00	100.00
% *Smartphone and Other Device*$_k$	Percentage of households in zip code k that have both a smartphone and another computing device(s)	69.87	11.59	0.00	100.00
% *Bachelors Degree or higher*$_k$	Percentage of residents in zip code k with a bachelors' degree or higher	31.93	17.14	0.00	100.00
Estimated No. of Children$_i$	Estimated number of children in household i	1.30	1.89	0.00	9.00
$Data_{it}$ (GB)	Household i's overall data use in month t	14.09	11.00	0.74	61.76
$Data\_Edu_{it}$ (MB)	Household i's education data use in month t	15.21	55.43	0.00	444.89

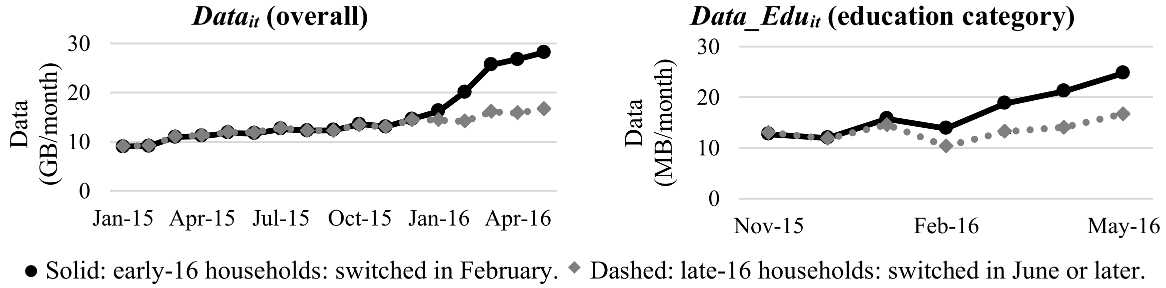
Notes. The matching procedure yields a total of 250,904 households. As illustrated in Figure 1, we observe $Data_{it}$ over 17 months in our analysis, yielding a total of 4,265,368 household/month observations. We observe $Data\_Edu_{it}$ for seven months, yielding a total of 1,756,328 household/month observations. We winsorize $Data_{it}$ and $Data\_Edu_{it}$ at the 1st and 99th percentiles. Winsorizing does not drop observations; instead, it reduces the impact of outliers by replacing extreme values with the values for the 1st or 99th percentile, whichever is applicable. Although we used the continuous value for *Dec15 Plan Charges*$_i$ in the matching procedure, we report the descriptive statistics for an ordinal version of the variable here to remain in compliance with our nondisclosure agreement.

Figure 1. Illustration of the Study Design and Estimation Strategy



Notes. This figure is for illustrative purposes only. The late-16 household that switches to unlimited data in June 2016 is used as the counterfactual for the early-16 household that switches in January 2016. Both have similar data use before the early-16 household switches. The difference between the two between January 2016 and May 2016 represents the increase associated with switching to unlimited data. Only the 17-month period between January 2015 and May 2016 is used in the main analysis because the counterfactuals switch after May 2016.

Figure 2. Average Data Use of Households That Switched to Unlimited Data in February 2016 and Their Matched Counterfactuals That Switched in June 2016 or Later



Notes. Overall data use is available from January 2015. Education data use is available from November 2015. We included error bars, reflecting the standard error of the mean. However, most of them are too small to be visible. Black circles and black solid lines indicate early-16 households that switched in February. Gray diamonds and gray dashed lines indicate late-16 households that switched in June or later.

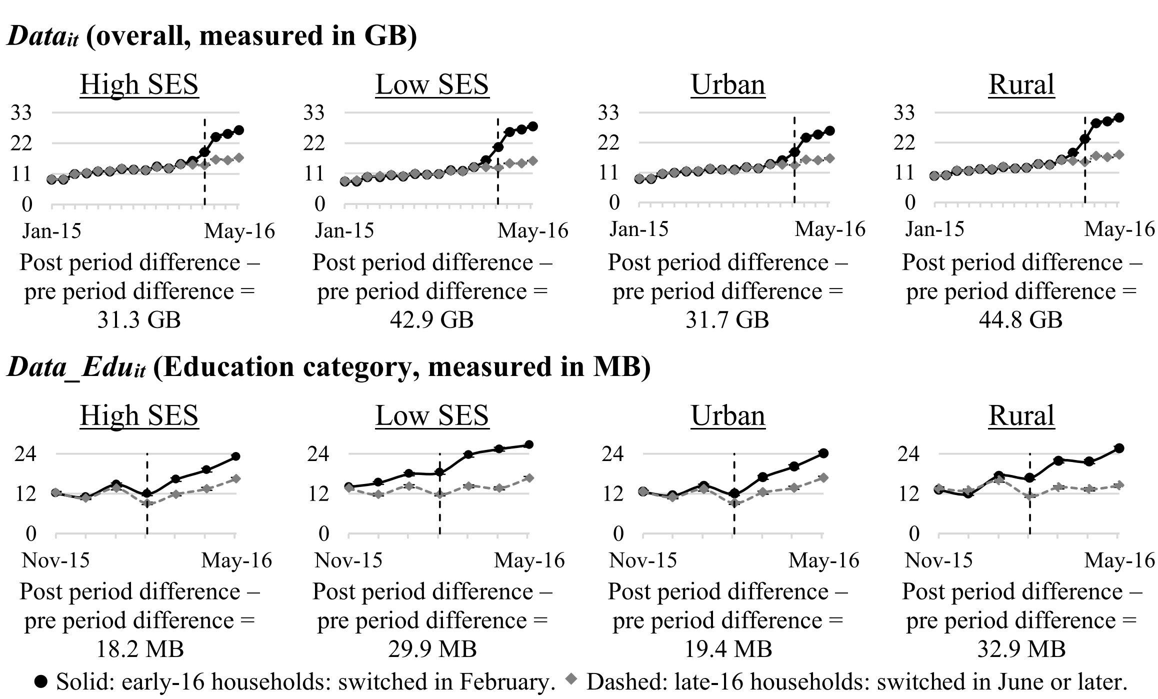
Main Regression Specification and Results

Our main difference-in-differences regression specification is shown below:

$$Y_{it} = \alpha + \beta \; Unlimited_{it} + h_i + \sum_{p=1}^{3} (T_t \times Cohort(p)_i) + \sum_{j=1}^{12} (T_t \times Geosocial(j)_i)) + \varepsilon_{it}. \quad (1)$$

Y_{it} is the data used by household i in month t, either overall ($Data_{it}$) or for the education category ($Data\_Edu_{it}$). $Unlimited_{it}$ is a dummy variable set to one for household i in the months t on or after it switched to unlimited data and zero otherwise. α is the constant term, h_i are household fixed effects, and T_t are month fixed effects. $Cohort(p)_i$ are dummy variables that reflect household i's cohort: January, February, or March. $Geosocial(j)_i$ are dummy variables that reflect household i's geosocial group. We interact the month fixed effects with $Cohort(p)_i$ and $Geosocial(j)_i$ to capture time trends for each cohort and geosocial group.[3] ε_{it} is the error term, clustered by household, as advised by Abadie et al. (2023)

Figure 3. Average Data Use of Households That Switched to Unlimited Data in February 2016 and Their Matched Counterfactuals That Switched in June 2016 or Later by Socioeconomic Status and Geography



Notes. Vertical dashed lines indicate when the early-16 households switched to unlimited mobile data. Overall data use is available from January 2015. Education data use is available from November 2015. We included error bars, reflecting the standard error of the mean. However, most of them are too small to be visible. The postperiod difference minus the preperiod difference is (1) the difference in average data use between households that switched in February and their counterfactuals in the post period minus (2) the difference in average data use between households that switched in February and their counterfactuals in the pre period.

given that "treatment" is assigned at the household level. We also clustered the errors in alternative ways as discussed below. We extend Model (1) by interacting $Unlimited_{it}$ with dummy variables for the geographic and socioeconomic groups to examine whether the change in data use is larger (or smaller) for disadvantaged households.

Table 3 shows the results of Model (1). Columns (1)–(3) in Table 3 show results for $Data_{it}$, and columns (4)–(6) in Table 3 show results for $Data\_Edu_{it}$. Column (1) in Table 3 shows that the average increase in overall data consumption after switching to unlimited data is 9.82 GB/month. Column (4) in Table 3 shows that the average increase for education data is 6.14 MB/month. Both are large increases over the preperiod trend. Columns (2), (3), (5), and (6) in Table 3 show the differential increases based on geographic location and socioeconomic status. The baseline group in columns (2) and (5) in Table 3 is *Urban* households, for which the increase is 8.59 GB/month for overall data and 5.13 MB/month for education data. The baseline group in columns (3) and (6) in Table 3 is *High SES* households, for which the increase is 8.42 GB/month for overall data and 5.11 MB/month for education data. The coefficients for the interaction terms represent the additional increase in data consumption for households from other groups—over and above the increase for *Urban* and *High SES* households. The additional increases for *Mostly Urban*, *Mostly Rural*, and *Rural* households are 1.50, 2.90, and 3.49 GB/month (or 17%, 34%, and 41% larger than the baseline increase) for overall data, respectively, and 1.08, 3.15, and 3.36 MB/month (or 21%, 61%, and 65% larger) for education data, respectively. The additional increases for *Mid SES* and *Low SES* households are 2.04 and 3.10 GB/month (or 24% and 37% larger than the baseline increase) for overall data, respectively, and 1.47 and 2.89 MB/month (or 29% and 57% larger) for education data, respectively. Each of the interaction coefficients is statistically different from the others ($p < 0.01$) except for two cases: the *Mostly Rural* and *Rural* interaction coefficients in column (5) in Table 3 ($p = 0.81$) and the *Mid SES* and *Low SES* interaction coefficients in column (6) in Table 3 ($p = 0.19$). Overall, the increases—both for overall data and for education data—are disproportionately large for rural households and those of low socioeconomic status.

The increases in education data consumption are underestimated because the telecommunications company only categorized approximately one third of the overall data, as noted above. Thus, assuming that the one third sample was representative, the average increases in education data consumption for *High SES* and *Low SES* households as shown in Table 3 are approximately 15 MB/month (i.e., $5.11 \times 3 \approx 15$) and

Table 3. Regression Results for Specification (1)

	Dependent variable: $Data_{it}$ (GB)			Dependent variable: $Data\_Edu_{it}$ (MB)		
Explanatory variable	(1)	(2)	(3)	(4)	(5)	(6)
$Unlimited_{it}$	9.82***	8.59***	8.42***	6.14***	5.13***	5.11***
	(0.04)	(0.06)	(0.06)	(0.17)	(0.26)	(0.24)
$Unlimited_{it} \times Mostly\ Urban_i$		1.50***			1.08***	
		(0.08)			(0.33)	
$Unlimited_{it} \times Mostly\ Rural_i$		2.90***			3.15***	
		(0.15)			(0.66)	
$Unlimited_{it} \times Rural_i$		3.49***			3.36***	
		(0.15)			(0.66)	
$Unlimited_{it} \times Mid\ SES_i$			2.04***			1.47***
			(0.07)			(0.31)
$Unlimited_{it} \times Low\ SES_i$			3.10***			2.89**
			(0.23)			(1.09)
Constant	14.1***	14.3***	14.2***	16.9***	17.1***	17.0***
	(0.09)	(0.09)	(0.09)	(0.49)	(0.49)	(0.50)
n (household-months)	4,265,368	4,265,368	4,265,368	1,756,328	1,756,328	1,756,328
n (households)	250,904	250,904	250,904	250,904	250,904	250,904
Number of months	17	17	17	7	7	7
R^2 (with household fixed effects)	0.72	0.72	0.72	0.53	0.53	0.53

Notes. Because education data use is only available starting in November 2015, columns (4)–(6) are based on observations from November 2015 to May 2016. Columns (1)–(3) are based on observations from January 2015 to May 2016. Fixed effects for households and months, interacted with cohort and geosocial group, are included. $Mostly\ Urban_i$, $Mid\ SES_i$, etc. do not appear as stand-alone terms because they are fully absorbed by the household fixed effects. For Models (2) and (5), the baseline reference group is *Urban*. For Models (3) and (6), the baseline reference group is *High SES*. Clustered standard errors by household are in parentheses.

$^{**}p < 0.05$; $^{***}p < 0.01$.

24 MB/month ($(5.11 + 2.89) \times 3 \approx 24$), respectively. Assuming an electronic textbook size of ~5 MB,[4] that equates to an additional three textbooks per month for high SES households and an additional five textbooks per month for low SES households. The analogous increases of electronic handouts, if we assume that they are ~100 KB each, are 154 and 246 per month for high and low SES households, respectively. The increases may be further underestimated because data included in other categories, such as tutorial videos on YouTube, may be educational. We discuss this potential measurement issue below.

Table 4 shows the results of Model (1) after logging the dependent variables. Results are similar to the main results; they show that the increases in consumption of education data and other data are significantly larger for disadvantaged households than for advantaged households.

Leads/Lags Model. A key test of the validity of our model is whether the pretreatment trends in data use are parallel for the early-16 and late-16 households before the early-16 households switched to unlimited data. We test for parallel pretreatment trends by using a leads/lags model in which we adjust Model (1) by replacing $Unlimited_{it}$ with a series of lead and lag dummy variables, denoted $Unlimited_{i(t\pm\tau)}$. For early-16 "treated" households, each $Unlimited_{i(t\pm\tau)}$ lead (lag) dummy equals one for observations in month t if month t is τ months before (after) switching to unlimited data and zero otherwise. For late-16 "control" households, the $Unlimited_{i(t\pm\tau)}$ dummies are always zero. The lead dummies allow us to test for a difference in data use between early-16 and late-16 households in the months before the early-16 households switched to unlimited data, thereby allowing us to check for parallel pretreatment trends. The lag dummies allow us to see how the change in data use evolves after switching.

Figure 4 plots the leads/lags coefficients and their 95% confidence intervals for both the $Data_{it}$ and $Data\_Edu_{it}$ models. The lead coefficients are small and essentially zero. Each of the confidence intervals for the lead coefficients for the $Data\_Edu_{it}$ model spans zero, as do some of those for the $Data_{it}$ model, although the coefficients in this model are precisely estimated with narrow confidence intervals. The lag coefficients are large and increasing over time. This provides evidence that the difference in data use only exists after the early-16 households switch to unlimited data (i.e., there is no meaningful pre-existing difference in the matched sample). We also estimated the leads/lags model for the *Low SES*, *High SES*, *Urban*, and *Rural* household groups separately. Figure A2 in the Online Appendix shows that the increases after switching to unlimited data for these groups are not artifacts of pre-existing differences.

Empirical Extensions and Exploration of Underlying Mechanisms

Deeper Exploration of the Increase in Education Data Consumption

Data Consumption as a Function of the Number of Children. A key reason that we explore changes in education data consumption as well as overall data

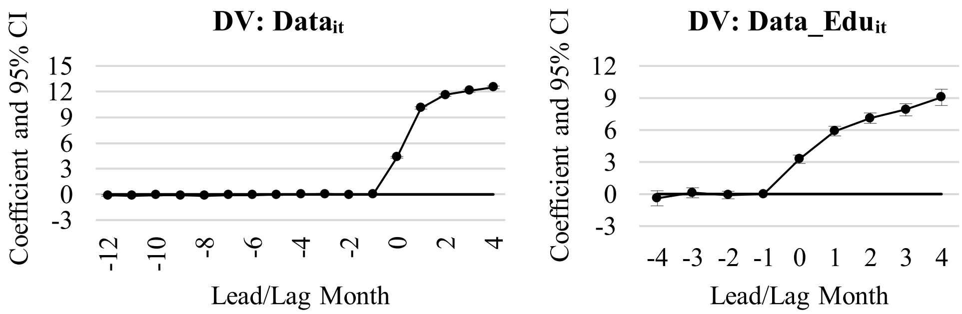
Table 4. Regression Results for Logged Dependent Variables

	Dependent variable: $\log(1 + Data_{it})$			Dependent variable: $\log(1 + Data\_Edu_{it})$		
Explanatory variable	(1)	(2)	(3)	(4)	(5)	(6)
$Unlimited_{it}$	0.43*** (0.00)	0.38*** (0.00)	0.37*** (0.00)	0.25*** (0.00)	0.23*** (0.01)	0.23*** (0.01)
$Unlimited_{it} \times Mostly\ Urban_i$		0.06*** (0.00)			0.03*** (0.01)	
$Unlimited_{it} \times Mostly\ Rural$		0.12*** (0.01)			0.07*** (0.02)	
$Unlimited_{it} \times Rural_i$		0.14*** (0.01)			0.09*** (0.02)	
$Unlimited_{it} \times Mid\ SES_i$			0.09*** (0.00)			0.03*** (0.01)
$Unlimited_{it} \times Low\ SES_i$			0.16*** (0.01)			0.05** (0.02)
Constant	2.47*** (0.00)	2.48*** (0.00)	2.48*** (0.00)	1.52*** (0.01)	1.53*** (0.01)	1.52*** (0.01)
n (household-months)	4,265,368	4,265,368	4,265,368	1,756,328	1,756,328	1,756,328
R^2 (with household fixed effects)	0.82	0.82	0.82	0.48	0.48	0.48

Notes. Each of the interaction coefficients is statistically different from the others ($p < 0.10$), except for the *Mostly Rural* and *Rural* interaction coefficients in column (5) ($p = 0.44$) and the *Mid SES* and *Low SES* interaction coefficients in column (6) ($p = 0.38$). Sample and regression details are as discussed in Table 3.

$^{**}p < 0.05$; $^{***}p < 0.01$.

Figure 4. Coefficients and 95% Confidence Intervals (CIs) for the Lead and Lag Terms



Notes. We withhold the −1 lead dummy to avoid the dummy variable trap, thereby using −1 as the "baseline." The confidence intervals for the $Data_{it}$ analysis are narrow, so much so that the error bars that represent them are barely visible in the figure. Given the January 2015 to May 2016 analysis period for the $Data_{it}$ analysis, the earliest possible lead dummy is −14, and the latest lag dummy is +4. This is because we can go back 14 months for households that switched to unlimited data in March 2016 and forward 4 months for households that switched in January 2016. We aggregate all leads 12 or more months before switching into the −12 lead. For the $Data\_Edu_{it}$ analysis, the analysis period is November 2015 to May 2016, yielding leads/lags from −4 to +4. DV, dependent variable.

consumption is that the *type* of data consumption is important to consider when investigating digital inequality. Because households with school-age children are more likely to consume education data than those without children, we explore if the effects of switching to unlimited data are greater for households with children. The raw measures of increased data consumption for households with and without children are not directly comparable because the former is likely to have more lines. To account for this, we divide $Data_{it}$ and $Data\_Edu_{it}$ by *Dec15 No. of Lines*$_i$ to construct $DataPerLine_{it}$ and $DataPerLine\_Edu_{it}$. We interact *Estimated No. of Children*$_i$ with $Unlimited_{it}$ and rerun our regression using $DataPerLine_{it}$ and $DataPerLine\_Edu_{it}$ as the dependent variables. As noted above, *Estimated No. of Children*$_i$ is zero when the household has no children or if the number of children is undetermined. For robustness, we rerun our analysis after dropping these households; results are shown in Table A5 in the Online Appendix, and they are consistent. As shown in Table 5, the increase in data consumption per line after switching to unlimited data grows by 0.09 MB for education data and 0.11 GB for overall data for each additional child in a household. This indicates that the increase in education data consumption is largest for households that are likely to have the most demand for it (viz., those with children).

Education Data Consumption as a Percentage of Overall Data Consumption. We explore whether access to unlimited mobile data might have shifted the *percentage* of mobile data used for education content (e.g., whether households allocate more or less of their overall data consumption to education content after switching to unlimited data). To do this, we reran specification (1) using % *Education Data*$_{it}$, which is $(Data\_Edu_{it})/Data_{it}) \times 3 \times 100$, as the dependent variable. We multiplied the ratio by three because the telecommunications company categorized data use for only approximately one third of the overall data, as noted above. We get qualitatively similar results if we do not apply the multiplication factor. We multiplied the ratio by 100 so that it represents a percentage rather than a proportion.

Table 6 shows that switching to unlimited data results in minimal to no change in the percentage of mobile data used for education content. As per columns (2) and (3) in Table 6, the coefficient for the baseline group (urban or high SES households) is insignificant, and the coefficients for the other geosocial groups do not significantly differ from the baseline group. The pooled effect

Table 5. Regression Results for Data per Line, Including Interactions for the Number of Children

	Dependent variable: $DataPerLine_{it}$ (GB)		Dependent variable: $DataPerLine\_Edu_{it}$ (MB)	
Explanatory variable	(1)	(2)	(3)	(4)
$Unlimited_{it}$	2.89 (0.01)***	2.75 (0.02)***	1.82 (0.06)***	1.70 (0.08)***
$Unlimited_{it}$ × *Estimated No. of Children*$_i$		0.11 (0.01)***		0.09 (0.03)***
Constant	4.14 (0.04)***	4.13 (0.04)***	5.38 (0.19)***	5.37 (0.19)***
n (household-months)	4,265,368	4,265,368	1,756,328	1,756,328
R^2 (with household fixed effects)	0.68	0.68	0.57	0.57

Note. Sample and regression details are as discussed in Table 3.
***$p < 0.01$.

Table 6. Regression Results for Education Data Consumption as a Percentage of Overall Data Consumption

	Dependent variable: % *Education Data*$_{it}$		
Explanatory variable	(1)	(2)	(3)
$Unlimited_{it}$	−0.01 (0.01)**	−0.01 (0.01)	−0.01 (0.01)
$Unlimited_{it} \times Mostly\ Urban_i$		0.00 (0.01)	
$Unlimited_{it} \times Mostly\ Rural_i$		0.01 (0.02)	
$Unlimited_{it} \times Rural_i$		0.02 (0.02)	
$Unlimited_{it} \times Mid\ SES_i$			−0.00 (0.01)
$Unlimited_{it} \times Low\ SES_i$			−0.01 (0.04)
Constant	0.38 (0.01)***	0.39 (0.01)***	0.38 (0.01)***
n (household-months)	1,756,328	1,756,328	1,756,328
R^2 (with household fixed effects)	0.50	0.50	0.50

Note. Sample and regression details are as discussed in Table 3.
**$p < 0.05$; ***$p < 0.01$.

is significant but small: 0.01 percentage points (or 0.0001 when represented as a proportion). This finding can be understood in conjunction with our main finding by considering two hypothetical households constructed for illustrative purposes, both of which switched to unlimited mobile data: (1) a high SES household that increased its overall data consumption from 10 to 16 GB/month and its education data consumption from 20 to 30 MB/month and (2) a low SES household that increased its overall data consumption from 10 to 20 GB/month and its education data consumption from 20 to 38 MB/month. Both households experienced significant and meaningful increases in education data consumption—with a larger increase for the low SES household—even though both households' *percentages* of education data decreased slightly by approximately 0.01 percentage points. The minimal effect of unlimited data on the *percentage* of education data consumed indicates that disadvantaged households that increased their consumption of education data by x% also increased their consumption of other data by approximately x%. This indicates that access to unlimited mobile data yields disproportionate increases for disadvantaged households in the consumption of education data *and* other types of data, which we discuss further in the Discussion section.

Exploration of Underlying Mechanisms

We next explore the potential mechanisms driving the larger increases in education and other data consumption by disadvantaged households. As theorized above, this may be because disadvantaged households are often "smartphone dependent" and lack other internet access options, such as fixed broadband service at their homes. We do not directly observe whether a household is smartphone dependent, but we can proxy for this at the zip code level using % *Mobile Broadband Only*$_k$ and % *Smartphone Only*$_k$. Columns (1) and (3) in Table 7 show the results after we include a % *Mobile Broadband Only*$_k$ interaction term. Households in zip codes with higher levels of smartphone dependence for internet access experience larger increases in both education data consumption and overall data consumption after switching to unlimited mobile data. These increases can be as large as 20 GB/month and 20 MB/month of overall data and education data, respectively, in zip codes with 100% smartphone dependence. By the same token, households in zip codes with higher levels of mobile *and* fixed line internet access as measured by % *Mobile and Other Broadband*$_k$ experience smaller increases in data consumption; see columns (2) and (4) in Table 7. Table 8 shows that results using interaction terms including % *Smartphone Only*$_k$ and % *Smartphone and Other Device*$_k$

Table 7. Regression Results: Mechanism Analysis Based on Internet Access Options

	Dependent variable: $Data_{it}$ (GB)		Dependent variable: $Data\_Edu_{it}$ (MB)	
Explanatory variable	(1)	(2)	(3)	(4)
$Unlimited_{it}$	8.22 (0.07)***	13.9 (0.13)***	4.49 (0.29)***	9.13 (0.53)***
$Unlimited_{it} \times$ % $Mobile\ Broadband\ Only_k$	0.20 (0.01)***		0.20 (0.03)***	
$Unlimited_{it} \times$ % $Mobile\ and\ Other\ Broadband_k$		−0.10 (0.00)***		−0.07 (0.01)***
Constant	14.2 (0.09)***	14.3 (0.09)***	16.9 (0.49)***	17.0 (0.49)***
n (household-months)	4,265,368	4,265,368	1,756,328	1,756,328
R^2 (with household fixed effects)	0.72	0.72	0.53	0.53

Note. Sample and regression details are as discussed in Table 3.
***$p < 0.01$.

Table 8. Regression Results: Mechanism Analysis Based on Smartphone Use

Explanatory variable	Dependent variable: $Data_{it}$ (GB) (1)	Dependent variable: $Data_{it}$ (GB) (2)	Dependent variable: $Data\_Edu_{it}$ (MB) (3)	Dependent variable: $Data\_Edu_{it}$ (MB) (4)
$Unlimited_{it}$	8.45 (0.06)***	16.3 (0.22)***	5.06 (0.25)***	10.5 (0.91)***
$Unlimited_{it}$ × % *Smartphone Only*$_k$	0.34 (0.01)***		0.26 (0.05)***	
$Unlimited_{it}$ × % *Smartphone and Other Devices*$_k$		−0.09 (0.00)***		−0.06 (0.01)***
Constant	14.2 (0.09)***	14.3 (0.09)***	16.9 (0.49)***	17.0 (0.49)***
n (household-months)	4,265,368	4,265,368	1,756,328	1,756,328
R^2 (with household fixed effects)	0.72	0.72	0.57	0.57

Note. Sample and regression details are as discussed in Table 3.
***$p < 0.01$.

are similar. Overall, these results indicate that smartphone dependence is a key mechanism driving the larger increase in data consumption for disadvantaged households.

We conducted a similar analysis in which we interacted $Unlimited_{it}$ with % *Bachelors Degree or Higher*$_k$. Table 9 shows that the increases in data consumption after switching to unlimited mobile data are higher in zip codes with lower levels of education. This is consistent with our other results that disadvantaged households experience disproportionate increases in data consumption after gaining access to unlimited mobile data. Although we do not have data on whether disadvantaged households improve educational performance after switching to unlimited mobile data, this finding suggests that unlimited mobile data can help provide disadvantaged households with the tools needed to improve educational performance. We explore this further in the Discussion section.

Robustness Checks and Alternative Analyses

Potential Measurement Error

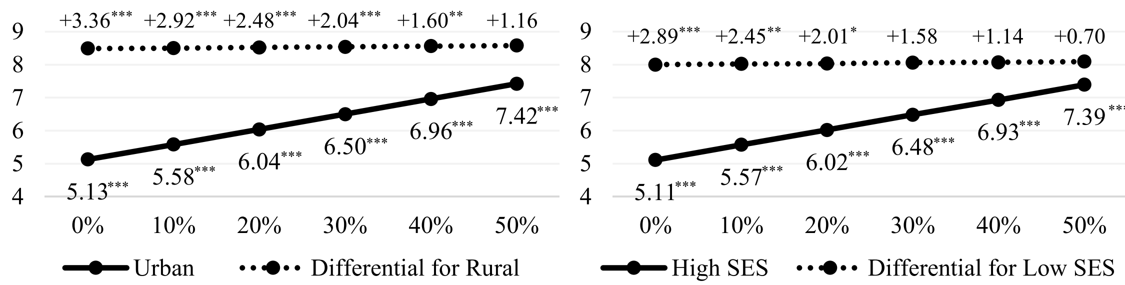
Our $Data\_Edu_{it}$ measure directly reflects the education category used by the telecommunications company to categorize data use. It captures education data consumption, although it may not capture *all* education data consumption. For example, $Data\_Edu_{it}$ does not capture consumption of educational videos on YouTube. Thus, $Data\_Edu_{it}$ is likely a lower bound on education data consumption. If any underestimation in $Data\_Edu_{it}$ is randomly distributed across households, then the larger increases in education data consumption for disadvantaged households will be even larger than we document. For example, if education data consumption for all households is really two times what we measure, then instead of education data consumption increasing by 15 and 24 MB/month for high SES and low SES households, respectively, it would increase by 30 and 48 MB/month, respectively. However, if our measure systematically underestimates education data consumption for some households but not others, then that could explain the larger increases for disadvantaged households, thereby threatening our conclusion. The most consequential version of this for our results is if education data consumption was underestimated for advantaged households but not for disadvantaged households. Although this pattern of mismeasurement is unlikely, we test how sensitive our results might be to it. We rerun our main regression after artificially inflating $Data\_Edu_{it}$ for the advantaged households only (urban and high SES households) in 10-percentage-point increments. As shown in Figure 5, for the additional increases for rural and low SES households—which are over and above the increases for urban and high SES households—to become insignificant, we would have to systematically underestimate $Data\_Edu_{it}$ for urban and high SES households by 40%–50% and 20%–30%, respectively. Because this pattern and magnitude of mismeasurement are unlikely, our results appear robust to this potential measurement issue.

Table 9. Regression Results: Mechanism Analysis Based on Educational Attainment

Explanatory variable	Dependent variable: $Data_{it}$ (GB) (1)	Dependent variable: $Data\_Edu_{it}$ (MB) (2)
$Unlimited_{it}$	12.5 (0.08)***	8.75 (0.32)***
$Unlimited_{it}$ × % *Bachelors Degree or higher*$_i$	−0.09 (0.00)***	−0.08 (0.01)***
Constant	14.2 (0.09)***	17.0 (0.49)***
n (household-months)	4,265,368	1,756,328
R^2 (with household fixed effects)	0.72	0.53

Note. Sample and regression details are as discussed in Table 3.
***$p < 0.01$.

Figure 5. Sensitivity Analysis for Potential Systematic Mismeasurement of Education Data



Notes. In this sensitivity analysis, we artificially inflated $Data\_Edu_{it}$ for urban and high SES households only. The solid lines represent the increase in education data consumption for urban and high SES households after switching to unlimited data at different levels of artificial inflation, which are shown on the *x* axis. As $Data\_Edu_{it}$ is inflated for urban and high SES households, the differential increase in education data consumption for rural and low SES households—represented by the dotted lines—gets smaller and eventually becomes insignificant. The values for 0% artificial inflation represent the main results and mirror those shown in Table 3. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Other Robustness Checks and Alternative Analyses

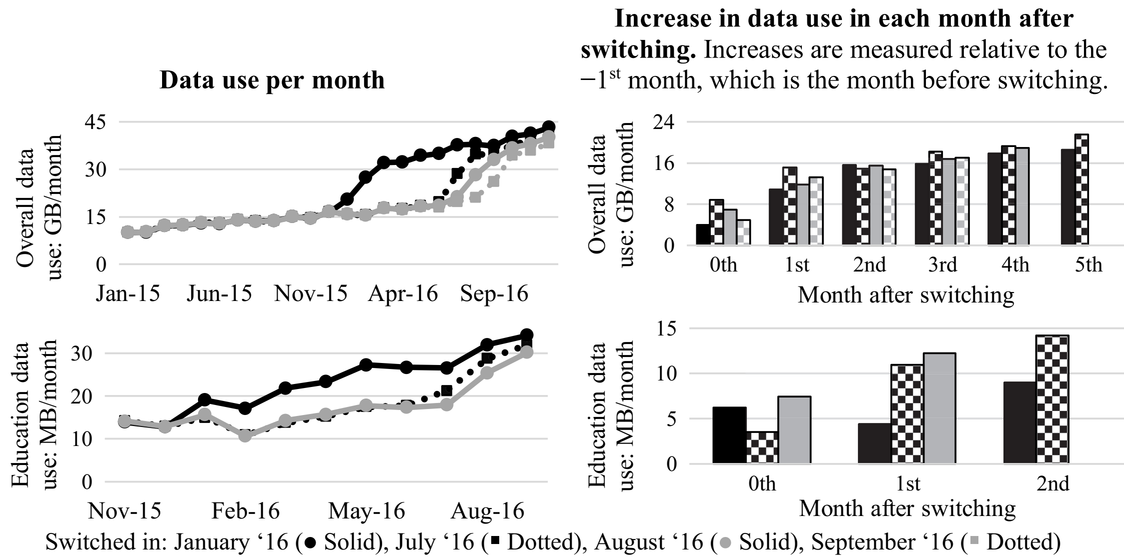
We conducted several additional robustness checks and alternative analyses. First, we further explored the possibility of selection bias by conducting analysis based on (1) whether households exceeded their data cap prior to switching to the unlimited plan and (2) growth in data consumption prior to switching to the unlimited plan. These analyses are presented in Tables A6 and A7 in the Online Appendix and support our main findings. Second, we conducted a random implementation test to assess whether our standard errors might be inconsistent because of serial correlation in the dependent variable. These results are consistent and appear in Table A8 in the Online Appendix. Third, we used households that did not switch to unlimited data during our sample period as counterfactuals for those that did using the same matching procedure as in our main analysis. This allowed us to include all months in our sample in the analysis, thereby yielding a longer time period over which to examine the effect. As shown in Figures A3 and A4 and Table A9 in the Online Appendix, the results are similar to our main results, with the average effect continuing to grow over the additional months. Fourth, we confirmed that our results are not sensitive to the break points that we used to define the geographic and socioeconomic groups. For the geographic group analysis, we interacted the raw value of % $Rural_i$ with $Unlimited_{it}$. For the socioeconomic group analysis, because our data only report income in ranges, we interacted $Unlimited_{it}$ with all five reported ranges rather than three as in our focal analysis. Results remain consistent; see Table A10 in the Online Appendix. Fifth, we considered the possibility of a bias because of the timing of when households switched. For example, it is possible that early-16 households switched early because of their need for more data. If this was the case, then we should see a larger increase in data consumption for early switchers than for late switchers. Figure 6 shows that this is not the case; households that switched in July, August, and September had similar increases in data consumption after switching as did households that switched in January. To further test whether our results are impacted by differences in treatment timing, we ran our focal analysis for the January cohort separately (Goodman-Bacon 2021) and decomposed our results for the households that switched in January, February, and March (Callaway and Sant'Anna 2021). Results that are shown in Figures A5 and A6 and Table A11 in the Online Appendix are consistent with our main results. Sixth, we removed households that switched to the unlimited plan but then reverted to a capped plan during our analysis period, along with their matched controls, from the analysis. Results remain consistent and appear in Table A12 in the Online Appendix. Seventh, we believe that household is the appropriate level for clustering the standard errors. However, we also clustered the standard errors at other levels—including at the five-digit, three-digit, and two-digit zip code levels as well as at the five-digit zip code + household level via two-way clustering. We find consistent results, as illustrated in Table A13 in the Online Appendix. Eighth, we performed inverse hyperbolic sine transformation on the dependent variables as an alternative to log transformation. Results are shown in Table A14 in the Online Appendix, and they are consistent. Ninth, we explored heterogeneity based on whether households that switched to unlimited data were above or below the median in data consumption in December 2015. As shown in Tables A15 and A16 in the Online Appendix, the differential increases in data consumption are consistent for both subsamples.

Discussion, Implications, Limitations, and Opportunities for Future Research

Discussion

The problem of inequality in internet access—and the resulting homework gap in education—is well known

Figure 6. Data Use for Households That Switched to the Unlimited Mobile Data Plan in January, July, August, and September



Notes. To understand the figures in the right panels, consider the households that switched to unlimited data in August 2016. The zeroth month after switching shows their average increase in data consumption from July to August, the first month shows the increase from July to September, etc. The figures in the right panels illustrate that increases are similar regardless of when households switched to unlimited data. Education data for the households that switched in September 2016 are not shown because September is the last month in which we observe education data. Black circles and black solid lines indicate that the switch took place in January of 2016. Black squares and black dotted lines indicate that the switch took place in July of 2016. Gray circles and gray solid lines indicate that the switch took place in August of 2016. Gray squares and gray dotted lines indicate that the switch took place in September of 2016.

(Auxier and Anderson 2020). There are multiple approaches to address the problem. Given the high costs and long lead time of providing fixed broadband service relative to mobile broadband service (e.g., installing fiber-optic lines to multiple households is likely more costly than installing a mobile telecommunications tower to serve those households), many of these approaches focus on improving mobile broadband service (e.g., Federal Communications Commission 2020). However, if improvements to mobile internet service are to reduce digital inequality, then they must yield larger increases in content consumption for disadvantaged households than for advantaged households, including for welfare-enhancing content, such as education content. Otherwise, the gap will persist or potentially widen.

To our knowledge, no prior research has examined whether improvements to mobile internet service yield larger increases in data consumption for disadvantaged households than for advantaged households. This is important to study empirically because either outcome is possible theoretically. On one hand, advantaged households may benefit relatively little from mobile internet service improvements because they are likely to have home broadband service that already meets their needs. Because many disadvantaged households do not have home broadband service and are smartphone dependent, improvements to mobile internet service may disproportionately help them "catch up" in terms of data consumption. However and on the other hand, digital inequality research has often shown that advantaged households experience larger consumption increases, particularly for "enhancing" content, like education data (Bonfadelli 2002, Hargittai and Dobransky 2017).

We examine these competing theoretical possibilities empirically and find significantly larger increases in education and other data consumption for disadvantaged households compared with advantaged households after gaining access to unlimited mobile data. The disproportionate increases in education data consumption may narrow gaps in educational outcomes, assuming that certain conditions hold, some of which we discuss here. First, one condition is that the increase in education data be used for the completion of school assignments. Completing schoolwork leads to better school performance, increased likelihood of graduation, and higher test scores (Fairlie et al. 2010, Hampton et al. 2021). Second, another condition is that the education data be used for high-quality education content. This follows from survey research indicating that online educational materials "will probably not generate gains for students nor equitable opportunities across groups of students unless the tools themselves are of high quality" (Rothwell 2022). One way for online educational materials to be of high "quality" is for them to provide learning experiences that are not available in traditional classrooms, such as programs that adapt to students' learning style and pace (Lei and Zhao 2007). Third, another condition is that any benefits attributable to the increase in consumption of education data are not countervailed by harms attributable to the increase in consumption of other types of data. This is of concern because use of smartphones in general often has negative effects on academic performance, in part because their recreational use (e.g., for games) distracts students from schoolwork (Gerosa et al. 2022). Fourth, an enabling factor (not necessarily a condition) for translating increased access to education data into better academic performance is the availability of nonteacher "mediators." A study in India showed that disadvantaged children provided with access to online educational materials via an internet kiosk achieved similar test scores as advantaged children but only when a nonteacher "mediator" encouraged the students to further explore what they were learning (Mitra and Dangwal 2010). If these (and other) conditions and enablers exist—which we suspect they do for at least some of the households in our data—then access to unlimited mobile data will help close the gap not only in education data consumption but also, in educational outcomes. Indeed, fewer connectivity disruptions—such as those caused by data shortages and computer breakdowns—are correlated with higher grades (Gonzales et al. 2021).

Implications

Our results have implications for (a) telecommunications policy on data caps and improvements to mobile internet service and (b) the zero-rating and network neutrality debate.

Telecommunications Policy on Data Caps and Improvements to Mobile Internet Service. Our results have direct implications for the FCC inquiry initiated in June 2023 into whether data caps restrict disadvantaged households' access to online education content. Although we do not test whether data caps restrict access to education content per se, we effectively show the inverse: that eliminating data caps opens up access to education content that disadvantaged households take advantage of. This finding will be useful for the FCC if it decides to regulate data caps as a way to combat digital inequality. This is because the FCC must discuss "how the agency chose its proposed solution to the problem" as part of its rulemaking process (see https://www.fcc.gov/about-fcc/rulemaking-process). The FCC can cite our paper as objective, peer-reviewed evidence that regulating data caps can help reduce digital inequality by generating disproportionate increases in consumption of education (and other) data by disadvantaged households. More broadly, telecommunications companies continually implement programs designed to improve mobile internet service, including programs that affect data availability, coverage area, connection stability, and download speeds. Although our results are specific to data availability, they suggest that other improvement programs may also generate disproportionate consumption increases for disadvantaged households, which can help reduce digital inequality.

Zero Rating and Network Neutrality Policy. Because data caps restrict access to content, it is important to consider whether consumption of education and other enhancing content should be excluded from data caps. This practice is known as zero rating. Zero rating is central to the network neutrality debate about whether telecommunication companies must treat all network traffic equally or whether they can prioritize certain types of content, be that via zero rating, by transmitting it faster, etc. (Easley et al. 2018). Zero rating is controversial because it privileges certain content. It is illegal in some countries because of concerns that it allows telecommunication companies to pick which websites and services will be successful and/or to control the flow of information by making some content cheaper or more accessible (Robertson 2018). However, zero rating education and other socially enhancing content, such as health information, is considered by many to have benefits that outweigh the costs. For example, Colombia's zero rating of education data during the COVID-19 pandemic was considered a success and a model for other countries (Organisation for Economic Co-operation and Development 2021, Shaji 2022). As with other policies, a key question when considering zero rating education content is whether it will help reduce digital inequality by generating disproportionate data consumption increases by disadvantaged households. Our results suggest that it will.

Zero rating education data seems feasible for telecommunications companies given that they already zero rate certain content types (van Schewick 2022). Furthermore, although we find significant and meaningful increases in consumption of education content—particularly for disadvantaged households—these increases are a fraction of the increase in overall data consumption. Thus, zero rating education data may not strain telecommunications networks. However, there are several practical challenges to zero rating education data, including (1) political feasibility and (2) difficulty determining what content qualifies. First, establishing stable zero rating and other network neutrality regulations has been difficult in the United States given that Democrats tend to favor regulations, whereas Republicans do not. However, allowing telecommunications companies to zero rate education data—if they so choose—may be an opportunity for compromise. Right-leaning regulators may support it because it affords telecommunication companies discretion over how to manage their networks. Left-leaning regulators may support it because it should increase access to education data for households, particularly disadvantaged households at risk for falling further behind. Second, a certification system(s) would be needed to classify what content counts as educational. These systems can build upon existing systems used by telecommunications companies (such as the one we leverage for our study) and websites (such as YouTube's Learning channel) as well as educational content accreditation systems at the federal, state, and local levels. These systems should be granular enough to distinguish educational and noneducational content from the same website/service/app given that many provide both. This is a challenging task given that these systems will create costs, yield false positives and false negatives, disadvantage educational content that is not certified, and be subject to political influence (Hernandez 2023). However, these issues exist with other educational content accreditation systems, which manage to function nonetheless. How to optimally design systems for classifying online education data is an opportunity for future research. Despite the challenges, our findings indicate that zero rating educational content is likely to disproportionately help disadvantaged households, assuming that even a fraction of the disproportionate increase in education data consumption translates into improved educational outcomes. Implementing and publicizing programs that help these households may generate reputational benefits for telecommunications companies that outweigh the costs.

Limitations and Opportunities for Future Research

Our research has limitations, some of which we list here. First, we do not observe outcomes that result from increased data consumption. For example, we do not observe whether children in households that increase their consumption of educational content via their mobile data plans improve their test scores. This is a common limitation of research exploring digital inequality (e.g., Santillana et al. 2020). Gathering archival (i.e., nonsurvey) data on data consumption and educational outcomes (grades, test scores, etc.) at the student level is difficult given privacy concerns with student information and the need to match the consumption data to the educational outcome data. However, future researchers might supplement archival data consumption data (such as what we use) with self-reported measures of educational outcomes or vice versa. This would permit a direct test of the link between education data consumption and outcomes. Second, we cannot comment on the long-run effects of access to unlimited mobile data. Our focal approach only permits analysis of the effects for the first 5 months after switching, although we extend this to 12 months in the analysis in which we compare households that switched with those that did not during our study period. Third, there is a possibility that selection issues bias our estimates of the increase in education and other data consumption after switching to unlimited data, despite our efforts to mitigate this. However, as long as this bias is similar across geosocial groups, then our conclusion that access to unlimited mobile data yields disproportionate data consumption increases for disadvantaged households will be valid. Future research might use randomized field experiments to study the effects of improvements to mobile internet service. It is also important to note that although we show that providing access to unlimited mobile data yields disproportionate increases in data consumption for disadvantaged households, other factors occurring simultaneously might lead to disproportionate increases for advantaged households, such as better home broadband service that only high-income households can afford. Despite this, identifying programs that generate disproportionate increases for disadvantaged households is a necessary—if not sufficient—step for reducing inequality.

Conclusion

Although there are many ways to improve mobile internet service, such as providing faster and more stable connections, we focus on access to unlimited mobile data. Using panel data from a major telecommunications company, we show that disadvantaged households experience larger increases in overall and education data consumption after gaining access to unlimited mobile data compared with advantaged households. On average, disadvantaged households increase their consumption of overall (education) data by approximately 12 GB/month (24 MB/month), whereas advantaged households increase their consumption by 8.5 GB/month (15 MB/month). The disproportionate increase for disadvantaged households indicates that access to unlimited mobile data helps to reduce inequality in internet use. The larger increases in education data for disadvantaged households translate into roughly two additional digital textbooks per month and a larger number of smaller educational materials. We cannot be sure that additional resources such as these will close the gap in educational outcomes or address other harms caused by digital inequality. However, identifying and implementing programs that yield disproportionate data consumption increases for disadvantaged households—such as access to unlimited mobile data—are important steps toward that end.

Endnotes

[1] As will be shown (see Table 2), subscribers in our sample used more than 10 GB/month of data on average prior to switching, which would necessitate a "large" plan circa 2016 (Asterino 2016).

[2] This approach is further motivated by Abadie et al. (2015), and it is often used for causal inference using the synthetic control method (see Abadie et al. 2010, Cunningham and Shah 2018, and Abadie 2021, among others).

[3] We also estimated the model with the following modifications: (1) month fixed effects without interactions with the cohort or geosocial group and (2) three-way interactions between the month fixed effects, $Cohort(p)_i$, and $Geosocial(k)$. Results are shown in Table A4 in the Online Appendix, and they are qualitatively unchanged.

[4] The file size of educational materials varies. See https://open.umn.edu/opentextbooks/textbooks/basic-algebra-with-applications for a 1.5-MB textbook and https://www.amazon.com/High-School-Math-Made-Simple-ebook/dp/B003NX7N4C/for a 6-MB textbook. See https://www.k5learning.com/free-math-worksheets/second-grade-2/addition/add-2-2-digit-numbers-no-regrouping for a 12-kilobyte (KB) worksheet.

References

Abadie A (2021) Using synthetic controls: Feasibility, data requirements, and methodological aspects. *J. Econom. Literature* 59(2): 391–425.

Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *J. Amer. Statist. Assoc.* 105(490):493–505.

Abadie A, Diamond A, Hainmueller J (2015) Comparative politics and the synthetic control method. *Amer. J. Political Sci.* 59(2):495–510.

Abadie A, Athey S, Imbens GW, Wooldridge JM (2023) When should you adjust standard errors for clustering? *Quart. J. Econom.* 138(1):1–35.

Agarwal R, Animesh A, Prasad K (2009) Research note—Social interactions and the "Digital Divide": Explaining variations in Internet use. *Inform. Systems Res.* 20(2):277–294.

Anderson M, Faverio M, Mcclain C (2022) How teens navigate school during COVID-19. *Pew Research Center* (June 2), https://www.pewresearch.org/internet/2022/06/02/how-teens-navigate-school-during-covid-19/.

Asterino M (2016) Verizon wireless deals. 2016's best plans and phones. *Payette Forward* (October 4), https://www.payetteforward.com/verizon-wireless-deals-2016-best-plans-phones/.

Auxier B, Anderson M (2020) As schools close due to the coronavirus, some U.S. students face a digital "homework gap." *Pew Research Center* (March 16), https://www.pewresearch.org/fact-tank/2020/03/16/as-schools-close-due-to-the-coronavirus-some-u-s-students-face-a-digital-homework-gap/.

Beard RT, Ford GS, Stern M (2022) Bridging the Digital Divide: An empirical analysis of public programs to increase broadband adoption. *Telematics Informatics* 67:101754.

Bonfadelli H (2002) The internet and knowledge gaps: A theoretical and empirical investigation. *Eur. J. Comm.* 17(1):65–84.

Callaway B, Sant'Anna PHC (2021) Difference-in-differences with multiple time periods. *J. Econometrics* 225(2):200–230.

Chabé-Ferret S (2017) Should we combine difference-in-differences with conditioning on pre-treatment outcomes? TSE Working Paper No. 17-824, Toulouse School of Economics, Toulouse, France.

Chillemi O, Galavotti S, Gui B (2020) The impact of data caps on mobile broadband Internet access: A welfare analysis. *Inform. Econom. Policy* 50:100843.

Christiansen P (2023) Your guide to internet data caps: How to stay within your limits. *SatelliteInternet.com* (April 10), https://www.satelliteinternet.com/resources/internet-data-caps/.

Correa T, Pavez I, Contreras J (2021) Digital inequality and mobiles: Opportunities and challenges of relying on smartphones for digital inclusion in disadvantaged contexts. Hargittai E, ed. *Handbook of Digital Inequality* (Edward Elgar, Cheltenham, UK), 62–74.

Cunningham S, Shah M (2018) Decriminalizing indoor prostitution: Implications for sexual violence and public health. *Rev. Econom. Stud.* 85(3):1683–1715.

DiMaggio P, Hargittai E, Celeste C, Shafer S (2004) *From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality* (Russell Sage Foundation, New York).

Easley RF, Guo H, Krämer J (2018) Research commentary—From net neutrality to data neutrality: A techno-economic framework and research agenda. *Inform. Systems Res.* 29(2):253–272.

Ericsson (2017) Ericsson mobility report November 2017. (November), https://www.ericsson.com/4ac62d/assets/local/reports-papers/mobility-report/documents/2017/emr-november-2017-shifting-mobile-data-plans.pdf.

Fairlie RW, Beltran DO, Das KK (2010) Home computers and educational outcomes: Evidence from the NYLS97 and CPS. *Econom. Inquiry* 48(3):771–792.

Federal Communications Commission (2020) 5G fund. Retrieved November 2, 2022, https://www.fcc.gov/5g-fund.

Federal Communications Commission (2022) Homework gap and connectivity divide. Retrieved October 29, 2022, https://www.fcc.gov/about-fcc/fcc-initiatives/homework-gap-and-connectivity-divide.

Federal Communications Commission (2023a) Chairwoman Rosenworcel proposes to investigate how data caps affect consumers and competition. *FCC News* (June 15), https://docs.fcc.gov/public/attachments/DOC-394416A1.pdf.

Federal Communications Commission (2023b) Emergency connectivity fund. Retrieved September 5, 2023, https://www.fcc.gov/emergency-connectivity-fund.

Ganju KK, Pavlou PA, Banker RD (2016) Does information and communication technology lead to the well-being of nations? A country-level empirical investigation. *MIS Quart.* 40(2):417–430.

Gerosa T, Gui M, Büchi M (2022) Smartphone use and academic performance: A pervasiveness approach beyond addiction. *Soc. Sci. Comput. Rev.* 40(6):1542–1561.

Gonzales AL, Yan H, Read GL, Brown A (2021) What's missing? How technology maintenance is overlooked in representative surveys of digital inequalities. Hargittai E, ed. *Handbook of Digital Inequality* (Edward Elgar, Cheltenham, UK), 9–30.

Goodman-Bacon A (2021) Difference-in-differences with variation in treatment timing. *J. Econometrics* 225(2):254–277.

Greenstein S (2015) *How the Internet Became Commercial: Innovation, Privatization, and the Birth of a New Network* (Princeton University Press, Princeton, NJ).

Greenstein S, Prince J (2009) Internet diffusion and the geography of the Digital Divide in the United States. Avgerou C, Mansell R, Quah D, Silverstone R, eds. *The Oxford Handbook of Information and Communication Technologies* (Oxford University Press, Oxford, UK), 168–195.

Greenwood BN, Agarwal R (2016) Matching platforms and HIV incidence: An empirical investigation of race, gender, and socioeconomic status. *Management Sci.* 62(8):2281–2303.

Hampton KN, Robertson CT, Fernandez L, Shin I, Bauer JM (2021) How variation in Internet access, digital skills, and media use are related to rural student outcomes: GPA, SAT, and educational aspirations. *Telematics Informatics* 63:101666.

Hansen JD, Reich J (2015) Democratizing education? Examining access and usage patterns in massive open online courses. *Science* 350(6265):1245–1248.

Hargittai E, ed. (2021) *Handbook of Digital Inequality* (Edward Elgar, Cheltenham, UK).

Hargittai E, Dobransky K (2017) Old dogs, new clicks: Digital inequality in skills and uses among older adults. *Canadian J. Comm.* 42(2):195–212.

Harris C, Straker L, Pollock C (2017) A socioeconomic related "digital divide" exists in how, not if, young people use computers. *PLoS One* 12(3):e0175011.

Hauge JA, Prieger JE (2015) Evaluating the impact of the American Recovery and Reinvestment Act's BTOP on broadband adoption. *Appl. Econom.* 47(60):6553–6579.

Helsper E (2021) Network and neighborhood effects in digital skills. Hargittai E, ed. *Handbook of Digital Inequality* (Edward Elgar, Cheltenham, UK), 75–96.

Helsper EJ, Van Deursen AJAM (2015) The third-level Digital Divide: Who benefits most from being online? Robinson L, Cotten SR, Schulz J, Hale TM, Williams A, eds. *Communication and Information Technologies Annual, Studies in Media and Communications*, vol. 10 (Emerald Group Publishing Limited, Leeds, UK), 29–52.

Hernandez J (2023) Florida rejects some social studies textbooks and pushes publishers to change others. *NPR* (May 10), https://www.npr.org/2023/05/10/1175232763/desantis-florida-textbooks-social-studies-schools.

Hitt L, Tambe P (2007) Broadband adoption and content consumption. *Inform. Econom. Policy* 19(3):362–378.

Holslin P, Parrish K (2023) Which internet service providers have data caps? *HighSpeedInternet.com* (June 26), https://www.highspeedinternet.com/resources/which-internet-service-providers-have-data-caps.

Hosanagar K, Fleder D, Lee D, Buja A (2014) Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. *Management Sci.* 60(4):805–823.

Hsieh JJPA, Rai A, Keil M (2008) Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. *MIS Quart.* 32(1):97–126.

Hsieh JJPA, Rai A, Keil M (2010) Addressing digital inequality for the socioeconomically disadvantaged through government initiatives: Forms of capital that affect ICT utilization. *Inform. Systems Res.* 22(2):233–253.

Jung J, Bapna R, Ramaprasad J, Umyarov A (2019) Love unshackled: Identifying the effect of mobile app adoption in online dating. *MIS Quart.* 43(1):47–72.

Just DR, Wansink B (2011) The flat-rate pricing paradox: Conflicting effects of "all-you-can-eat" buffet pricing. *Rev. Econom. Statist.* 93(1):193–200.

Lei J, Zhao Y (2007) Technology uses and student achievement: A longitudinal study. *Comput. Ed.* 49(2):284–296.

Lythreatis S, Singh SK, El-Kassar AN (2022) The digital divide: A review and future research agenda. *Tech. Forecasting Soc. Change* 175:121359.

Maceviciute E, Wilson TD (2018) Digital means for reducing digital inequality: Literature review. *Informing Sci. Internat. J. Emerging Transdiscipline* 21:269–287.

Manlove J, Whitacre B (2019) An evaluation of the Connected Nation broadband adoption program. *Telecomm. Policy* 43(7):101809.

Mitra S, Dangwal R (2010) Limits to self-organising systems of learning—The Kalikuppam experiment. *British J. Ed. Tech.* 41(5): 672–688.

National Center for Education Statistics (2022) Methodology studies—Achievement gaps. Retrieved August 10, 2023, https://nces.ed.gov/nationsreportcard/studies/gaps/.

National Telecommunications and Information Administration (2023) High-speed internet for all. Accessed August 24, 2023, https://www.internetforall.gov/.

Ogbo E (2022) Broadband voucher programs: Evaluating the Alabama Broadband Connectivity Program. Preprint, submitted August 2, https://papers.ssrn.com/abstract=4178771.

Organisation for Economic Co-operation and Development (2021) Ten principles for effective and equitable educational recovery from COVID. Accessed September 5, 2023, https://www.ei-ie.org/en/item/25231:effective-and-equitable-educational-recovery-10-principles.

Pearce KE, Rice RE (2013) Digital divides from access to activities: Comparing mobile and personal computer internet users. *J. Comm.* 63(4):721–744.

Perrin A (2021) Mobile technology and home broadband 2021. *Pew Research Center: Internet, Science & Tech* (June 3), https://www.pewresearch.org/internet/2021/06/03/mobile-technology-and-home-broadband-2021/.

Perrin A, Atske S (2021) 7% of Americans don't use the internet. Who are they? *Pew Research Center* (April 2), https://www.pewresearch.org/short-reads/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/.

Pew Research Center (2021) Internet/broadband fact sheet. *Pew Research Center* (January 31), https://www.pewresearch.org/internet/fact-sheet/internet-broadband/.

Prieger JE (2013) The broadband digital divide and the economic benefits of mobile broadband for rural areas. *Telecomm. Policy* 37(6):483–502.

Racherla P, Mandviwalla M (2013) Moving from access to use of the information infrastructure: A multilevel sociotechnical framework. *Inform. Systems Res.* 24(3):709–730.

Ramdas K, Sungu A (2024) The digital lives of the poor: Entertainment traps and information isolation. *Management Sci.*, ePub ahead of print September 27, https://doi.org/10.1287/mnsc.2024.4989.

Ransbotham S, Fichman RG, Gopal R, Gupta A (2016) Special section introduction—Ubiquitous IT and digital vulnerabilities. *Inform. Systems Res.* 27(4):834–847.

Reisdorf BC, Yankelevich A, Shapiro M, Dutton WH (2019) Wirelessly bridging the homework gap: Technical options and social challenges in getting broadband to disconnected students. *Ed. Inform. Tech.* 24(6):3803–3821.

Rigney D (2010) *The Matthew Effect: How Advantage Begets Further Advantage* (Columbia University Press, New York).

Robertson A (2018) India just approved net neutrality rules that ban "any form" of data discrimination. *The Verge* (July 11), https://www.theverge.com/2018/7/11/17562108/india-department-of-telecommunications-trai-net-neutrality-proposal-approval.

Robinson L, Wiborg Ø, Schulz J (2018) Interlocking inequalities: Digital stratification meets academic stratification. *Amer. Behav. Sci.* 62(9):1251–1272.

Robinson L, Schulz J, Blank G, Ragnedda M, Ono H, Hogan B, Mesch GS, et al. (2020) Digital inequalities 2.0: Legacy inequalities in the information age. *First Monday* 25(7).

Rosston GL, Wallsten SJ (2020) Increasing low-income broadband adoption through private incentives. *Telecomm. Policy* 44(9): 102020.

Rothwell J (2022) Better technology produced better learning outcomes during pandemic. *Gallup.com* (February 3), https://news.gallup.com/opinion/gallup/388502/better-technology-produced-better-learning-outcomes-during-pandemic.aspx.

Santillana M, Sraubhaar J, Schrubbe A, Choi J, Strover S (2020) Digital inequalities: Homework gap and techno-capital in Austin, Texas. *First Monday* 25(7).

Segall L (2011) Verizon ends unlimited data plan. *CNNMoney* (July 6), https://money.cnn.com/2011/07/05/technology/verizon_data_plan/index.htm.

Sen A, Tucker CE (2020) Social distancing and school closures: Documenting disparity in internet access among school children. Preprint, submitted April 26, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3572922.

Sen S, Joe-Wong C, Ha S, Chiang M (2019) Time-dependent pricing for multimedia data traffic: Analysis, systems, and trials. *IEEE J. Selected Areas Comm.* 37(7):1504–1517.

Shaji A (2022) Is zero-rating a threat to human rights? *Human Rights Pulse* (January 11), https://www.humanrightspulse.com/mastercontentblog/is-zero-rating-a-threat-to-human-rights.

Snider M, Baig EC (2017) Verizon unlimited data plan signals price war to come. *USA Today* (February 13), https://www.usatoday.com/story/tech/news/2017/02/13/verizon-unlimited-data-plan-signals-price-war-come/97850832/.

Tichenor PJ, Donohue GA, Olien CN (1970) Mass media flow and differential growth in knowledge. *Public Opinion Quart.* 34(2): 159–170.

Tsetsi E, Rains SA (2017) Smartphone internet access and use: Extending the digital divide and usage gap. *Mobile Media Comm.* 5(3):239–255.

Van Dijk JAGM (2017) Digital Divide: Impact of access. Rössler P, Hoffner CA, Zoonen L, eds. *The International Encyclopedia of Media Effects* (John Wiley & Sons, Ltd., Hoboken, NJ), 1–11.

van Schewick B (2022) Facebook, Google & big telecoms want to keep violating net neutrality in Europe. Regulators should stop them. Retrieved November 2, 2022, https://cyberlaw.stanford.edu/blog/2022/05/facebook-google-big-telecoms-want-keep-violating-net-neutrality-europe-regulators.

Venkatesh V, Sykes TA (2012) Digital Divide initiative success in developing countries: A longitudinal field study in a village in India. *Inform. Systems Res.* 24(2):239–260.

Viard VB, Economides N (2014) The effect of content on global internet adoption and the global "Digital Divide." *Management Sci.* 61(3):665–687.

Xiong J, Zuo M (2019) Older adults' learning motivations in massive open online courses. *Ed. Gerontology* 45(2):82–93.

Xu L, Duan JA, Hu YJ, Cheng Y, Zhu Y (2019) Forward-looking behavior in mobile data consumption and targeted promotion design: A dynamic structural model. *Inform. Systems Res.* 30(2):616–635.

Yu L (2010) How poor informationally are the information poor? Evidence from an empirical study of daily and regular information practices of individuals. *J. Documentation* 66(6):906–933.

Zhang M, Yun H, Hu Y (2023) The influences of socioeconomic status on parental educational expectations: Mediating and moderating effects. *Sustainability* 15(16):12308.

Zuo GW (2021) Wired and hired: Employment effects of subsidized broadband internet for low-income Americans. *Amer. Econom. J. Econom. Policy* 13(3):447–482.