# DIGITAL EQUITY IN K-12 EDUCATION: CONCEPTUALIZATION AND ANALYSIS OF STUDENTS' DIGITAL OPPORTUNITY

Cary Ka Wai Jim, B.S., M.A.T.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2022

APPROVED:

Hsia-Ching (Carrie) Chang, Major Professor
Robin K. Henson, Minor Professor
Joseph Oppong, Committee Member
Sarah Evans, Committee Member
Jiangping Chen, Chair of the Department of Information Science
Kinshuk, Dean of the College of Information
Victor Prybutok, Dean of the Toulouse Graduate School

Jim, Cary Ka Wai. *Digital Equity in K-12 Education: Conceptualization and Analysis of Students' Digital Opportunity*. Doctor of Philosophy (Information Science), May 2022, 124 pp., 11 tables, 22 figures, references, 167 titles.

Although digital equity is a recognized challenge in our K-12 school system, there is little research in using a holistic framework to investigate pre-conditions necessary for K-12 students to participate in digital learning and online processes. A conceptual framework of students' digital opportunity (SDO) is developed to represent the essential components of digital connectivity. The four key components are broadband internet availability, broadband usage, digital device ownership, and speed quality. A composite measure of SDO was created to quantitatively represent and measure the differences across 3,138 counties in the United States. Furthermore, spatial autocorrelation was applied to evaluate if the distribution of the SDO score is associated with geographical characteristics at the county level. The result showed the presence of significant county-level clusters with concentrations of high or low SDO scores. While the spatial analysis provided evidence of where the gaps in digital opportunities are located, there are underlying factors at the micro level that would need further investigation. This study suggests a collective approach between private and public entities to address the K-12 digital equity issue. The necessary conditions presented in the SDO model must be addressed first in order to bring change to K-12 students and schools in terms of obtaining high quality and reliable broadband internet and digital devices for learning with technology. Two research outputs are available from this research to allow others to further evaluate digital equity among K-12 schools and students.

TABLE OF CONTENTS

# LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER 1

# INTRODUCTION

## 1.1 Overview

Inadequate access to high-speed (broadband) internet and computer technology limits a person’s opportunity to participate in today’s information society. American school children are expected to become 21st century citizens who possess the skills and knowledge to navigate and succeed in a society that requires technological proficiency. However, digital equity continues to be a challenge because of persistent gaps in obtaining reliable internet and digital devices for learning. Currently, we lack a measure to evaluate digital equity among K-12 students in public education. Previous research pointed to the differences in personal beliefs and systematic barriers that prohibited digital participation or adoption of technology. The goal of this research is to evaluate the current digital disparity of K-12 public school students in the United States while using spatial methods to determine how policy can target the specific regions across our country.

## 1.2 Research Objectives

The term “digital divide” first appeared during the mid-1990s when the U.S. Presidential Office began several initiatives to prepare the country for a shift to an “information society” where knowledge and ability in telecommunications would determine how people live and conduct their daily lives. In academic literature, the term “digital divide” is commonly used to describe the barriers faced by people in accessing the internet or information communications technology (ICT) within or across countries. In policy work, the dynamics of the digital divide are often examined with connections to the political and economic systems within a country (Council of Economic Advisers, 2015, 2016; Mossberger et al., 2003; Servon, 2002). Currently, there is a continual disparity of access and use of the internet or ICT across different countries

(OECD, 2019; Pick & Sarkar, 2015; Schweik et al., 2018;  Yu, 2006). In sum, the nature of the digital divide has been viewed and experienced differently by regions, cultures, or philosophical groundings. Some research may lean toward a specific aspect of the digital divide while other research focuses on different population groups which makes comparison difficult. This research will examine the digital divide in the context of broadband internet, equipment ownership, and speed quality at the county and state levels.

The idea of a digital world, or virtual world, which breaks down all physical barriers to the access of knowledge and information may sound very promising. The reality is, we as human beings, still need a medium (physical access) and an understanding to utilize and participate in this virtual environment. For school-aged children, their access to the internet and ICT will mostly depend on their immediate environment, such as schools, home, or other community resources. Even when students have the aptitude to acquire digital skills, they may not have the physical access or support (e.g., role models) to show them how to utilize technologies and what is possible in this digital world. Students may not have the same degree of opportunity to participate due to limitations in their immediate environment. The main research objective of this study is to holistically evaluate the various factors that contribute to the unequal distribution of digital technologies among K-12 students, especially considering the interplay between geography, education, and digital technologies.

## 1.3 Statement of the Problem

The recent coronavirus pandemic (Covid-19 in 2020) drastically changed how traditional instruction is delivered in our K-12 public education system. Shelter-at-home orders and school closures have forced students and other people, including school staff and parents, to re-evaluate their current home access to the internet and ICT for essential activities. School teachers can also experience the digital divide if their homes are not equipped with ICT and reliable broadband internet to deliver seamless instruction online. For the U.S. general population, subscription cost and service availability continue to affect household adoption of broadband internet (Cohron, 2015; Horrigan, 2010; Liu et al., 2018). Demographic characteristics and socioeconomic differences are often used as components to explain usage variation (Campos-Castillo, 2015; Cohron, 2015; Horrigan, 2010). These layers of barriers for the adult population portray the indirect effects of the digital divide for school-aged children. Warf (2013) pointed out the lack of reliable access to a PC (personal computer) and internet at home, as well as social and technical support, is stratified by ethnicity and family income for students who experience the digital divide. Therefore, even when some students in public schools have the opportunities to take part online and are exposed to technology-supported learning, their access at home is not consistent for ongoing development of their digital skill. The student's family situation plays a vital role. The expectation of students using technology at home and participating in online learning will only increase in the future. Therefore, we are not only determining the internet and technology needs for today's learning, but we should also consider how students will transition to a growing digital world once they leave the K-12 system.

One challenge in evaluating the digital divide among K-12 students is the lack of a suitable measure that allows comparison over time and space. The current view of the "digital divide" has moved beyond the classic binary division (Servon 2002; Warschauer, 2002) of access where internet or ICT is divided between those who have or have-not. A continuum or multidimensional perspective to understand the digital divide has been proposed by various scholars (Hilbert, 2014; Sicherl, 2019; Vehovar et al., 2006). For example, the Pew Internet & American Life Project conducted a National Random Digit Dial Survey in 2002 (Lenhart &

Horrigan, 2003) through telephone interviews and focus groups with the adult population. Their results indicated a need to view internet access as a continuum rather than a dichotomous division. However, for practicality and research purposes, which variables or factors to model quantitatively or measure digital disparity are not widely discussed. One of the reasons for this is that academic researchers come from different disciplines or backgrounds and have shaped their research of the digital divide from different perspectives. There are also constraints on data availability at both macro and micro levels when examining the digital divide.

The last challenge is the semantic differences in terminology to describe what is the digital divide. The philosophy, ideology, and assumptions used were often not explicitly presented in previous quantitative or qualitative studies. Other terms that were used interchangeably with the "digital divide" are digital inequality, digital inclusion, digital exclusion, information inequality, or information poverty. However, the underlying factors and perspectives have some key differences and therefore have clouded the viewpoint of this issue. For example, Mori (2011) discussed how "digital inclusion" as a concept actually combines digital divide and social inclusion, but the lack of a framework for "digital inclusion" has limited scientific study or identification of analytical elements for practical application in public policy analysis. The inconsistent definitions created barriers for theory or concept development and further limited quantitative work in developing a construct for measurement and evaluation purposes. This is problematic in advancing the conceptualization and methodology for examining the digital disparity among K-12 students in our country and around the world.

## 1.4 Significance of the Study

To move forward from our current understanding of the digital disparity experienced by K-12 students, I proposed a construct named "students' digital opportunity" to unify the different components required before K-12 students can meaningfully participate in any form of digital learning. This students' digital opportunity (SDO) construct is composed of the key factors predominately discussed in the digital divide literature: broadband access, equipment ownership, and speed quality. The conception of this construct is positioned from the "digital equity" perspective (Gorski, 2007) rather than the "divide" viewpoint. This is a significant contribution to the existing literature because unlike previous constructs, the SDO is a multi-component measure. The Students' Digital Opportunity (SDO) measure quantifies the digital disparity among K-12 students across counties in the U.S. The SDO measure is dynamic such that it can be used in future development. As technology advances, the nature of each component may change, and thus other forms of internet connection or device types can still be evaluated in this model. The SDO measure can also be integrated or modified with other indicators (at the same geographical level or timeframe) to accommodate the specific population or environment.

Second, this work is a methodological innovation in social science where various sources of open data are linked to create a multi-component measure. The SDO construct can be used as an independent index or be combined with other feature variables, such as spatial elements, as shown in this work. It is especially unique because this study is a data-driven research supported by conceptions in existing literature while using measurement techniques to develop a robust measure. The methods demonstrated in this study serve as a methodological example and allow replication by other researchers. Third, by designing this research with a focus on practical implication, the outputs of this study support educational policy work and decision-making across counties and states within the United States.

## 1.5 Research Questions

This research aims to explore and evaluate digital opportunities among K-12 public school students. There are two research questions addressed in this study. The first question focused on the construct development and validation by applying the SDO measure to evaluate how it is distributed: How does students' digital opportunity (SDO) distribute across the United States at the county and state level? The second question focused on the relationship of geographical characteristics with the SDO distribution to discover structural differences of digital equity: Are there any geographical associations with the distribution of the Students' Digital Opportunity (SDO) measure across the United States?

## 1.6 Design Rationales

The rationale for using a quantitative research methodology and data science techniques to investigate digital equity is because we need a low-cost, quantifiable way to measure and evaluate digital disparity across the United States in K-12 education. While digital equity in K-12 education is a known condition (Horrigan, 2015; Kuttan & Peters, 2003), there is not a consistent and quantifiable way to capture the differences of digital connectivity among the K-12 student population. Thus, it is difficult to establish a baseline for understanding the current state of digital equity in K-12 education. Without a proper evaluation of the current state of digital opportunities across the country, it becomes difficult to identify which areas are truly in need and how resources should be allocated equitably.

## 1.7 Research Methodology

Several methods and techniques are used in this quantitative study of digital equity in K-12 education. First, there is a data science component of this research where secondary datasets were acquired, stored, and transformed with the programming language R before data analysis. Second, conceptual models and frameworks in relation to digital infrastructure and connectivity are reviewed to inform model building in this research. Third, a series of data-driven analyses:

exploratory data analysis and descriptive statistics are applied to support the evaluation of the data structure and characteristics to inform the research design and development. Fourth, factor analysis with reliability & validity tests are used to assess if the proposed construct reflects what it is designed to measure followed by spatial analysis of significant clusters of SDO values. During the analysis, data visualization methods are used to generate plots, graphs, and maps to allow visual examinations and decision-making during the data-driven research process. The outputs on this work are openly accessible for others to replicate (see GitHub link in Chapter 3).

## 1.8 Limitations of the Study

While fixed broadband internet, PC ownership, and speed quality are used to represent digital connectivity in this study, there are other modes of broadband connection or digital equipment used by the general population. For example, mobile internet services and alternative broadband connections (e.g., white space radio frequency) are being developed to expand their coverage. Although these non-fixed broadband connections are susceptible to interferences (e.g., weather or building structures) and may result in varying connection stability or speed quality, they are less costly to build and can address short-term needs. The type of digital device used to access the internet, e.g., tablet, smartphone, or laptop, would affect the processing of data during an internet session. Therefore, the users' perception of their internet connection and how users interact with the digital devices are not addressed in this study. These are related to the end user experience depending on the types of connection and equipment used.

Another limitation of this research lies in evaluating the K-12 students' digital skills in terms of their utilization of the internet and digital content for learning. Digital equity from a perspective of having inclusive and accessible content online within education is not well-known. According to Subramony (2011), there is an imbalance of information production versus information utilization due to the varying levels of access and skills children may have based on their surrounding factors. This is the human aspect of digital equity that we do not have a clear conception for quantitative research. This type of data sources at the national level can be added to enhance the students' digital opportunity (SDO) construct.

## 1.9 Summary

This chapter provided an overview of the dissertation and described the logic of this literature informed, data-driven research. In the following chapters, a review of the literature, methodology, and analytical techniques relevant to the conditions of digital opportunity will be explored in Chapters 2 and 3. The definitions of concepts and conditions related to digital infrastructure and connectivity can be found in Chapter 2. Chapter 3 provides in-depth information of the data sources, key variables, and quantitative methods used in this study. Chapter 4 shares the results of the various analyses, including the decisions involved in the process and the limitations of the design and results. Chapter 5 presents the discussion of this work as a whole and its implications for theory and methodology development. The applicability of the SDO model and suggestions for future research will be addressed in Chapter 5.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Introduction

While education is a necessity for all children, the resources available to each student have positioned them differently in their educational trajectories. Researchers, think tanks, and government agencies continue to find gaps among U.S. students in grade K-12 especially concerning internet connectivity and utilization of technology for their learning and educational needs (Calabrese & Nasr, 2020; Dolan, 2016; Edgerton & Cookson, 2020; Federal Communications Commission, 2021a; Gao & Hayes, 2021; García & Weiss, 2020; KewalRamani et al., 2018; Reisdorf et al., 2019). From an equity perspective, the fundamental question is this: How can we provide children the digital necessities and opportunities to ensure they can participate in technology-supported learning and have reliable internet access, functioning equipment, and ongoing support at school and at home?

While internet or ICT access may be seen as an accomplished task (Hibert, 2014) at the national level across countries, usage remains a concern in the literature of the digital divide. There are many complementary factors between access and usage: skills and capacity, cultural attitude, choice, social influence, or local environment (Hilbert, 2014, p. 822). For example, traditional studies of the digital divide may use the level of saturation in terms of mobile phone subscription or ICT spending per capita as indicators of technology progress in a country (Hüsing & Selhofer, 2002; International Telecommunication Union, 2003). However, these statistics suggest a certain level of ICT saturation in society, but not how these technologies are being used or if they are high quality. The following sections will present the context of related factors, definitions of related terminology, previous metrics, and methodological considerations that inform this work.

## 2.2 Context of the Research and Related Factors

### 2.2.1 Schools and Libraries

Academic scholars have indicated that public schools and libraries are not spared from the digital divide (Bertot, 2009; Cohron, 2015; Dolan, 2016; Goolsbee & Guryan, 2006; Mandel et al., 2010; Whitacre & Rhinesmith, 2015). Schools and libraries are considered a part of the social infrastructure for children and families. Public education in the United States is funded and governed at the state and local level. Each year, approximately 8% of the total funds come from federal sources and the rest are from state, local, and other private sources (U.S. Department of Education, 2021) to provide free education for K-12 students. Currently, there is a program to support school libraries to foster literacy at early childhood by the U.S. Department of Education federal grant. The public libraries receive support from the Institute of Museum and Library Services (IMLS) through the Library Services and Technology Act to provide STEM activities and coding for students (American Library Association, 2019). While the funding mechanism for public libraries is different to K-12 public education, they have been considered an important point of access to information and technology (Kinney, 2010). Within the public library system, there are known differences between rural area libraries and those in urban or suburban areas in terms of their resources and digital capacity (Barack, 2005; Bertot, 2009; Cohron, 2015; Mandel et al., 2010). Robinson et al. (2020) emphasized that the lack of infrastructure and cost are the biggest barriers in our country and government leadership must be in place to address the persisting digital inequalities because often community-based or voluntary efforts have yielded poor results. Even for those who can pay for high-speed internet, broadband options are limited by locations, such as the rural areas. For example, in quasi-rural Illinois, choices of internet services are limited to mobile network connection or satellite, which are insufficient to complete day-to-day online tasks, and residents use their local library as a supplemental option (Schmidt & Power, 2020). This unique strategy of using a library as a public access point is not uncommon even in urban areas (Gonzales, 2016). However, Warf (2013) emphasized that libraries also face budget and space constraints even though they are often the only place to offer access to PCs (Personal Computer) and internet in impoverished communities. There is a widespread belief that technological advancement can address the current digital inequality in education through innovative programs or intervention to improve students' learning. The reality is that digital technology or ICT were not originally built with equity in mind. They should be viewed as a tool like a piece of scientific equipment, (e.g., a microscope) which needs intentional integration for a meaningful learning experience. But first, the student should learn how to use this "equipment" appropriately for the task at hand. Typical learning activities that involve the internet and digital devices are keyboarding, gathering information from online search engines, and using software application to obtain learning materials and conduct schoolwork. These learning activities all require stable internet access, suitable devices, and a basic understanding of how to use these tools to complete the tasks successfully.

Selwyn et al. (2001) expressed from an educational policy perspective that "what IT cannot do by itself is to change the dispositional constraints or alter the social determinants of participation to necessarily provide a genuinely educational experience" (p. 264). Therefore, even when access to the internet and ICT discrepancies have been addressed, there is still much work to do to tackle the unequal conditions of digital equity in the long term. Simply placing more expectations onto children to use more technology, for example assigning digital homework as practices, will not solve the issue of unequal digital skills among K-12 students. A recent study analyzed Google Trends search data (Bacher-Hicks et al. 2021) to determine how socioeconomic status (SES) relates to search patterns for online learning resources before and during Covid-19. Their results indicated a substantially higher search intensity among high SES areas and suggested a widening academic gap for students in less advantaged areas since online learning is a key component of future schooling. The types of ICT tools available to the students is another factor that affects the students' learning experience. There is still a fundamental difference between students who can have a keyboard enabled device compared to those who solely rely on a tablet or smartphone. These differences may seem like not much of a problem for students in the lower grades, but once students have progressed into middle or high school, more sophisticated technology is needed to conduct proper research, composition, or even coding for their core subjects. From these subtle differences between students' access to the device types, we can see how important it is to introduce and model the use of technology in the classroom and ensure this practice can be replicated at home for students to improve their digital skills to truly take advantage of what digital technology can offer for their learning.

Therefore, schools and libraries play an intricate role at the community level as anchor institutions because they are unlikely to move and are mission-driven entities to improve the human capital and social welfare of their communities.

### 2.2.2 Policy and Infrastructure

One of the gaps in research in determining the digital divide within our country is because of the differences in policy and regulation at the national, state, and local levels. To a certain extent, human activities are closely associated with their immediate societal structure and governing bodies and may knowingly or unknowingly be constrained to certain types of resources in their surroundings. Regulation of broadband technology is one of the factors that cannot be ignored when discussing the digital divide. As residents at the local level, we are under certain rules and regulations towards what options we have and how we can receive or use resources and services in our communities. There are implicit results from local or regional policy that could stimulate or prohibit access and utilization of digital infrastructure among their residents. For example, a popular idea in addressing the digital divide is to make broadband a public utility similar to electricity and water supply in a broad sense. However, this is not a feasible recommendation for certain parts of the country because some states prohibit local municipalities from offering telecommunication services such as broadband internet to their residents. According to Connected Nation - Texas (2021), Texas did not begin to allow electric cooperatives to provide municipal broadband to residential customers until 2019. While this change in utilities regulation may address the supply side of broadband internet in some parts of Texas, we are not certain if the adoption of municipal broadband has improved usage, or if this policy will benefit school-aged children and their families. Gonzales (2016) described that "physical access is more than the ability or will to purchase an internet subscription, but rather hinges on a complicated web of resources, such as transportation, education, employment, or interpersonal relationships" (p. 236). Thus, for children in households with limited resources, policy change may not benefit them directly if other surrounding factors are not addressed simultaneously. Other scholars (van Dijk & Hacker, 2003) proposed that a policy approach should be centered around *social inclusion* and make it an objective of *equal distribution* of resources or life chances.

### 2.2.3 Social, Cultural, and Human Factors

Local cultures and social practices influence people's perceptions and attitudes toward digital technology and how they utilize the internet resources as part of their daily lives. The digital divide research agenda over the past three decades may have shifted from access to usage. However, the discussion of social or cultural differences continues to be used by researchers to explain the variation in internet use and online participation for K-12 students (Kuttan & Peters, 2003; Monroe, 2004; Rafalow, 2021a; Warschauer, 2003). Dutton and Reisdorf (2019) conducted a survey research with a group of adults (n = 995, aged 18-92) in the state of Michigan and found that cultures, to a degree, influence demand and interest of the internet and other digital choices.

A recent book titled *Digital Divisions: How Schools Create Inequality in the Tech Era* (Rafalow , 2021b) presents an ethnography study of three comparable middle schools (2 public and 1 private) in California. The author discovered how students learned differently with digital technology due to instructional differences where teachers socialized students into particular kinds of digital strategies which may have led to unequal gains. Subramony (2011) warned educational technologists that insensitive ICT solutions implemented with pre-existing structural inequities, language barriers, and the lack of culturally suitable role models will further alienate the disadvantaged students. "Formal education runs behind because means [to learn digital skills] are lacking, and teachers are not sufficiently trained or motivated" (van Dijk & Hacker, 2003, p. 326). A 1-to-1 (one laptop per child) program was implemented in a low-income, minority school system in Birmingham, Alabama with a goal to eliminate digital inequality (Cotten et al. 2011). An important finding from their work is that students who perceived more laptop usage by their teachers reported greater frequency in their own laptop use and displayed positive attitudes toward technology. From an education perspective, purposeful digital strategies to address human factors can be influential and beneficial to both teachers and students. While the Cotten et al. (2011) study did not focus on digital equity for teachers and how their classroom practices of technology use may affect students, there are several studies that examine these topics (see Graves & Bowers, 2018; Peck et al., 2015; Prescott, 2020).

### 2.2.4 Information and Its Value

In the field of library and information science, the value of having access to the internet and technology utilization are often associated with the value of information. “At its core, the digital divide is more than just an issue about providing citizens with access to computers and the internet – it is about leveling the playing field in regard to information diffusion (Cohron, 2015, p. 84). Hilbert (2014) argued that “entry to the digital realm consists of obtaining access and being able to use the information resources effectively” (p. 821). Information as a primary good is needed by everybody to function in an information society. However, it is also a positional good where it becomes increasingly important in terms of economic, social, and cultural competition (van Dijk & Hacker, 2003, p. 324). Yu (2006) presented his thoughts on information inequality, where the flow of information from the resource pool is influenced by layers of factors: first by political and economic factors, then cultural and social factors, and last by personal factors (p. 235). The result of this information diffusion process is that the disadvantaged groups or individuals were limited in all three aspects which is more challenging to overcome due to the structural differences in society. Even though our society seems to be flooded with mobile devices (smartphones) to access the internet, it does not equate to having informational capacity among all people (Hilbert, 2014). As more public information and services are being digitized, stored, and delivered through the web, user information behavior is not simply how users navigate the web, but more fundamentally, if they have the tools and capacity to access the information they need? (Note: I prefer to use the term capacity rather than ability, because the term “ability” often implies a person’s cognitive and mental functions. Human capital and social capital can be viewed from a capacity perspective which is external to a person’s innate ability).

In sum, there are several common themes derived from the literature in this section. First, schools as anchor institutions are related to K-12 students’ digital connectivity and technology use. Public libraries as social entities provide supplemental resources and internet access for K-12 students who do not have home-based internet access. The public libraries also serve other population groups within their immediate community. Second, merely providing ICT equipment and broadband internet at public schools is not sufficient to promote meaningful use of digital technology for true benefits. Third, the complexity of the digital divide is often layered with social, economic, political factors at the macro-, meso-, and micro- level. For instance, personal beliefs, perceptions, attitudes, and culture at the micro-level influence the dynamics of the digital divide within school settings and at home. Last, the inconsistent terminology and conceptualization of the digital divide affect policy work and create challenges in effectively evaluating and tracking the evolution of the digital divide.

## 2.3 Definition

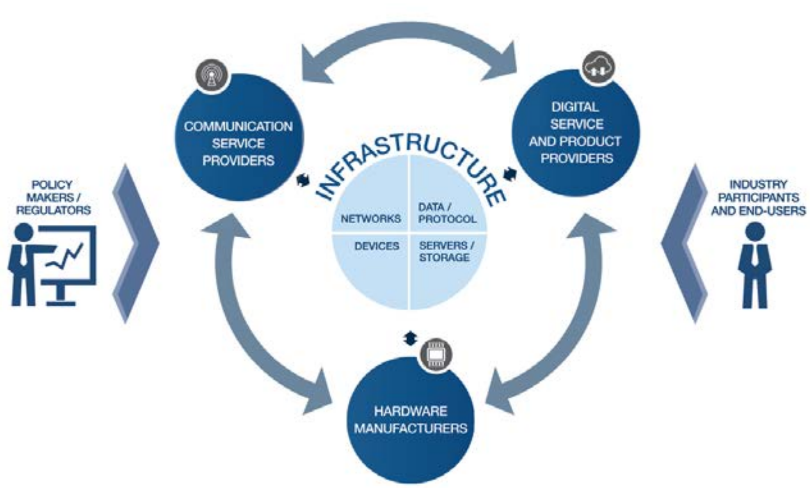
Due to the need to clarify and explain terminology in association with the conception of students’ digital opportunity model, the following sections provide descriptions of these terms as well as an overview of the interconnection between digital technology, infrastructure, broadband, and geography.

### 2.3.1 Digital Infrastructure

"The physical infrastructure of connectivity - consists of undersea, underground, and above-ground cables; tower sites, data centers, and satellites; the invisible spectrum used for wireless communication; and the variety of equipment that interconnects the world through the Internet" (Strusani & Houngbonon, 2020). The digital infrastructure facilitates the products and services between internet service providers, hardware, manufacturers, and digital services to bring the internet to end users (World Economic Forum, 2014). There are several categories of providers within the digital infrastructure such as communication service providers (e.g., fixed line, wireless telecommunications, cable companies and bandwidth providers), digital services (content, media, and IT service companies), digital products (hardware, software, and devices), and hardware manufacturers (infrastructure equipment, device, software, and component manufacturers). Figure 1 displays the interconnected roles of the various key components within the digital infrastructure to enable internet-based connectivity and services for users and other stakeholders.

Figure 1

*Stakeholders in Digital Infrastructure*




Source: World Economic Forum, 2014.

There are social components that are parallel to the physical assets within the digital infrastructure that are not illustrated in Figure 1. Schools, libraries, and community centers are currently viewed as social infrastructure that extends the digital infrastructure to neighborhoods. Public libraries and community-based facilities are viewed as 'hubs' for students and other population groups to access technology (e.g., computers and software) and the internet outside of their homes. However, having an access point is not a complete solution to ensure digital connectivity across different regions. At the consumer level, there are different types of internet connection, which can affect users' experience of the internet and how they utilize it. Grubesic and Murray (2002) analyzed xDSL in the early 2000's and pointed out high quality digital transmissions depended on the medium of this transfer; fiber-optic cable offers both high capacity and high-quality transmission (p. 201) in comparison to a copper transmission system. Thus, having the physical access point is only an entry point. The types of connection, equipment's specifications, and other technical and non-technical factors (e.g., local weather interference) have to be aligned to ensure a seamless online experience. I will further elaborate the differences of internet connection types in the next section.

### 2.3.2 Broadband Internet – High Speed Internet

The term "broadband" also known as wideband is a technical term that describes the carrying of multiple communication channels in a single wire or cable, which in a broader sense, refers to high-speed data transmission over the internet with a variety of technologies (Henderson, 2017). However, the meaning of "broadband" is also evolving because broadband, as a technology, has changed from the analog telephone-based systems in early days to the digital forms that are available today. The most common ways of referring to the connection types are fiber broadband, cable broadband, or mobile broadband. Kruger (2018) emphasized "how broadband is defined and characterized in statute and in regulation can have a significant impact on federal broadband policies and how federal resources are allocated to promote broadband deployment in unserved and underserved areas" (p. 1). The current definition for internet speed through a wireline broadband connection is 25/3 Mbps (megabits per second) download and upload speed as of 2015 (Ford, 2018). An older definition of broadband is at a speed of 4 Mbps download and 1 Mbps upload in 2010 (Federal Communications Commission, 2015). Since then, the benchmark for broadband internet by speed has not changed. "Bandwidth is the amount of data that can be transmitted over a connection in a given amount of time, latency is the time it takes for a data packet to make the round trip between the user's computer to a server located somewhere else" (Liu et al., 2018, p. 1). Speed is not only about bandwidth, but also latency as well. (See Noam 2011 for further discussion on the differences between speed, bandwidth, and data throughput). Content in digital media and communication today tends to include more speed-intensive features, such as high-definition video, dynamic web pages, and video conference applications where the 25 download/3 upload Mbps benchmark may no longer be enough to meet the needs of everyday use of the internet and related technologies. This is especially a challenge when multiple devices and users are sharing the same internet connection at home during the pandemic.

To gain internet connection at home, there are several components on the consumer end. Broadband internet is a paid service similar to the traditional telephone connection for residential users. It requires hardware such as a modem or router to connect users to the network at their location. This modem or router is often another cost the customer will pay in order to access the internet. Another requirement is that users will have to use their own computer equipment, such as a personal computer (PC) or laptop that has a port for wired connection, or an internal Wi-Fi card, to connect to the signal broadcast from the modem or router to receive wireless connection. The paid service offered by the communication service providers (i.e., Internet Service Providers, ISPs) is a subscription agreement between the consumer and the ISP. There are different types of internet connection available for household use, such as fixed broadband internet, mobile, or satellite connection and their availability depends on the home's geographical location. Fixed broadband includes non-dial-up connection such as DSL (copper telephone line), cable modem, fiber-optics, and satellite connection (KewalRamani et al., 2018; Kruger, 2018). Dial-up connections were used mostly in the 1990s and early 2000s and referred to as the analog connection. Nowadays, non-dial-up connection is what most Americans use for their internet connectivity. Mobile broadband represents internet connection through a mobile network. This form of connection is further divided into 5G, 4G, 3G, or LTE, which are signal types for smartphones, portable hotspots (Wi-Fi hotspots), or USB Wi-Fi adapter/modem sticks to allow access to the internet. This type of connection allows users to connect to their nearest mobile towers above ground. Reliability of mobile connection is affected by the strength of the mobile signals, which varies by the device's distance to the mobile tower and surrounding physical structures (e.g., buildings and floor levels such as the basement). As such, American households could be using different types of connections even within the same neighborhood. There are speed and price differences between the fixed broadband vs. mobile broadband, and subscription rates vary by the ISPs and their offered internet plans. Although the existing technologies are being improved, there are parts of the United States that simply do not have any digital infrastructure to utilize broadband internet. More technologies are being explored currently to find alternatives, such as using wireless spectrum (e.g., TV whitespace frequency or radio waves), to connect families in areas that do not have fixed (from the ground) broadband or mobile connection.

The deployment of broadband services in our country has no mandate for providing equitable distribution and therefore commercial internet service providers, such as ISPs, are free to choose their operation areas and coverages based on demand (Grubesic & Murray, 2002, p. 203). Since broadband is not randomly distributed across geography but rather deployed in areas where it is favorable to the ratio of demand to costs, it complicates the understanding of how broadband relates to economic outcomes (Ford, 2018, p. 775). One of the proposed solutions to address the digital divide is to use mobile phone networks instead of landline connections at home. This is a common approach in developing countries due to their infrastructural limitations across their lands. However, for American homes and children, internet access problems are more common in rural areas that do not have high-speed broadband landline or mobile networks due to the low population density in these regions. Even when access is provided, adoption in rural areas can still be a concern. Whitacre and Mills (2007) suggested demand for high-speed access should be stimulated in rural areas because building infrastructure capacity in rural areas alone is unlikely to bridge the gaps in high-speed internet access and use. The next section will further discuss the differences in broadband deployment and use due to geographical characteristics.

### 2.3.3 Geographical Differences, Urbanicity and Rurality

Grubesic and Murray (2002) pointed out a gap in the literature where local level geography is often overlooked, but local geography can have a major impact on infrastructure accessibility, connection speed, and connection quality (p. 199). This is also important when we examine the digital divide among K-12 students because their schools are part of the social infrastructure. The schools' institutional characteristics can have an influence on the students' experience as well as their access and use of technology.

Some scholars are not optimistic about the digital divide between urban and rural areas, even though public investment focused on infrastructure expansion and technology access has been in place since the late 1990s. Chakraborty and Bosman (2005) argued that the lack of measure to analyze the spatial divide between regions of the information rich or poor have limited our conceptual understanding of the digital divide (p. 397). A systematic literature review (Salemink et al., 2017) on digitization in rural parts of advanced countries (mainly European Union and the United States) identified two main themes: *connectivity issues* in relation to material inequalities and economic outcomes, and *inclusion issues* from a human perspective including knowledge, attitudes, skills, and aspirations. A key finding from their literature analysis is that a majority of social science research of digital inequality (inclusion issues) assumes ubiquitous connectivity and non-spatial factors which limit the applicability of their work to rural communities (Salemink et al., 2017, p. 366).

Mills & Whitacre (2003) analyzed the Census 2001 Current Population Survey and found the rate of internet use at school in rural areas compared to urban schools is different. They emphasized that ensuring access to digital technologies in school is essential to avoid the intergenerational divide (p. 239). The location of where students use digital learning tools is also a crucial factor in terms of exposure to technology and if they have the support needed to navigate the digital environment to truly benefit from these learning opportunities. Reisdorf et al. (2019) emphasized that "students will not realize the value of broadband without institutional (school) broadband access or institutional efforts to facilitate its use in the home" (p. 3826). Thus, broadband access at home and at school needs to be considered together in solving the digital inequality among school-aged children. For low-income groups, Gonzales (2016) identified the frequent disrupted experience as a barrier for these users. The fluctuating state of access of the internet between home and other public places created a challenge for low-income groups to "maintain" access.

In the UK, the divide in rural areas presented a similar challenge as in the American rural areas. Philip and Williams (2019) used an ethnographic study of three households in southwest Shropshire where a satellite broadband project was deployed. Their findings indicated that one of the important aspects of broadband initiatives should be "fit for purpose" to improve the livelihood of families and businesses in rural areas. The digital divide experienced by the households in their case study was not due to lack of internet use, but the quality of the broadband connection needed to support their online activities was insufficient. The geo-spatial aspect of the digital disparity among K-12 students warrants investigation in this study. Several techniques in evaluating spatial related phenomena will be described in Chapter 3.

### 2.3.4 Digital Divide

The description and definition of the digital divide continues to evolve since the mid-1990s. "Inequalities in Internet availability and service are broadly categorized under the term "digital divide" (Grubesic & Murray, 2002, p. 199). A majority of American literature credited the creation of the term "digital divide" due to the shift of the telecommunication policies in the mid-1990s (Cohron, 2015; Mori, 2011; Rogers, 2016). During the Clinton administration, several reports indicated telephone subscription and penetration is no longer enough for individuals in the information age (Jayakar, 2011). The National Telecommunications and Information Administration (1995) official report, *Falling Through the Net: A Survey of the Have Nots in Rural and Urban America*, focused on evaluating universal service in America, where computers, modems, and telephones are considered an integral part of how the people will access, accumulate, and assimilate information. Personal computers and modems are rapidly becoming the keys to the information vault (NITA, 1995, para. 3). The “information disadvantaged” was coined to describe this disparity of access (NITA, 1995, para. 5). The digital divide was clearly a technical challenge in the early 1990s when the internet was accessed through dial-up services and limited computer technologies (Howard et al., 2010). While governmental groups continue to assess and evaluate this matter from a policy and economic perspective, scholars in academia have further refined the definition of the digital divide. The divide is traditionally viewed as a binary divide between those who have or do not have access to telecommunication technology. However, there is a large variation in which types of technology scholars are referring to in these studies (Howard et al., 2010; Looker & Thiessen, 2003; Middleton & Chambers, 2010; Rice & Katz, 2002). As we discussed previously, there are many components within a digital infrastructure and types of broadband internet connections vary. Therefore, each part of the system could vary and influence the observed digital divide. Simply, even in the modern experience of using a personal computer, there are different brands, different designs, different hardware, and different software applications. Therefore, compressing the differences into a binary term such as the “divide” is certainly limiting our perceptions of this matter.

Servon (2002) suggested a three-dimensional view of the digital divide: access, training, and content. “Access is a necessary precondition but then engenders a need for training in order to use the tools. Once people have facility with the tools, they demand content that serves their interests and meets their needs” (p. 8). Kuttan and Peters (2003) defined the digital divide as “gaps in technology, access to technology (specifically the internet), education, and technology training between and within specific populations" (p. 3). Mossberger et al. (2003) argued that little is known on skills, needs, attitudes, or experience among the disadvantaged groups and proposed that the digital divide consists of "multiple divides: an access divide, a skills divide, an economic opportunity divide, and a democratic divide" (p. 2).

Internationally, scholars from different continents also examined the digital divide within their regions (for example, in the UK: Philip et al., 2017; Riddlesden & Singleton, 2014; Spain: Pérez-Amaral et al., 2021; Japan: Nishida et al., 2014; China: Shenglin et al., 2017). To develop a scientific conceptual model, van Dijk and Hacker (2003) discussed the digital divide as a form of information inequality with four kinds of barriers to access: mental access, materials access, skills access, and usage access. They challenged the common belief of providing a person with a personal computer and internet connection in order to solve the information inequality problem. However, there are neglected views in the four barriers they described. The following list further explains each barrier and how this contributes to the framing of the digital divide (van Dijk & Hacker, 2003):

- The mental access barrier is formed by the elementary digital experience due to the lack of interest, computer anxiety, and unattractiveness of the new technology (p. 315) which has been associated with the elderly, illiterates, or the unemployed.
- The material barrier represents the lack of possession of computers and network connections (p. 315) which has been the main focus in public opinion and policy.
- The skills barrier is caused by the lack of digital skills or the lack of user-friendliness tools in which inadequate education or social support further prohibit skill development. Digital skills are not limited to operating a computer or network connection, it also involves the ability to search, select, process, and apply information (p. 316).
- The lack of usage opportunity creates a usage barrier. Due to the free choice perspective by postmodern society, this barrier has not been viewed as important to social and educational policies (p. 316).

The usage gap suggested by van Dijk and Hacker (2003) on how certain groups within a population systematically benefit from advanced digital technology and applications for work and education is not new. Attewell (2001) suspected the “already-disadvantaged children may be dominated by games at home and unsupervised drill-and-practice or games at school, while affluent children enjoy educationally richer fare with more adult involvement” (p. 257). Within the U.S., the policy agenda has been mainly focused on overcoming the access barrier (physical access) by setting up digital infrastructure and encouraging technology ownership. Howard et al. (2010) pointed out that between rural and urban areas, the divide was persistent across income, education, or race/ethnic group. But there are little programs in place to teach skills to utilize and engage with the content on the internet (Howard et al., 2010, p. 116). The evolving nature of the internet and computer technology has complicated the digital divide. Thus, “ICTs never stand still but instead thrive on the dynamic disequilibria between user competencies and income… new divides are constantly being produced by the new and rapid product cycles” (Charkraborty & Bosman, 2005, p. 397). For users, it becomes an endless race to “catch” the next technology while the old ones become obsolete.

Another perspective is to break down the digital divide into three stages: access to technology, effective use of technology, and social integration & tangible results from technology (Hilbert, 2013, p. 822). Campos-Castillo (2015) re-evaluated the digital divide in the U.S. between 2007 and 2012 with the Health Information National Trends Survey (HINTS) of the adult population conducted by the National Cancer Institute. This work focused on internet access and how it varied by race, gender, and the intersection between these two factors. Although the work did not examine the second or third level of the digital divide, Campos-Castillo (2015) argued that determining the first level of digital divide should be continued because there is still a lot of ambiguity about “who has internet access” even in the United States.

Due to the limitation of the survey items, it was not possible to understand if the divides are changing across time. While her work provided a more recent understanding of racial and gender differences in terms of internet access, it is unclear if the K-12 student population in the U.S. experienced similar kinds of disparities due to the survey only including adults. A challenge in addressing this gap is the limitation of public data or enterprise information to evaluate internet access and use at home or at school.

### 2.3.5 Digital Inequality

Another common term that has been used interchangeably with the digital divide is "digital inequality." Katz and Gonzalez (2016b) described how digital inequality and social inequality are intertwined; and contextual factors in a person's local environment are far more important factors on one's technology use. Labels such as the "digital immigrants" or "digital natives" presented a diverging view of the generational gap between adults and children in our current society. These labels are problematic and say very little about how digital connectivity must be utilized to be meaningful to children and their families. Robinson et al. (2020) discussed how Silicon Valley in California is one of the perfect examples of digital inequalities. Silicon Valley, as a technology epicenter in our country, has many families who do not have the means or skills to access high-speed internet even though it is widely available in that region (see NTIA 2014 for this discussion). Dutton and Reisdorf (2019) suggested intervention efforts [for digital inequality] should focus on shaping attitudes and beliefs, which are more subject to change than fixed demographic factors (p. 19). As discussed in the previous section, social and cultural aspects are ingrained structurally and personally in which there are no simple solutions. For this work, socioeconomic characteristics and demographic information were considered as part of the model design.

## 2.4 Previous Index or Measure

As we have discussed in the previous sections, a large range of work from academia, government agencies, and think tanks have examined the digital divide. But there is not an effective measure to use within the U.S., especially among K-12 students. Currently, there is little work on developing benchmarks for measuring the digital divide within countries (Howard et al., 2010). Chakraborty and Bosman (2005) also pointed out the limitation of previous empirical studies of the digital divide that heavily relied on conventional statistical methods (simple comparison) to measure differences. Counting the number of internet users per capita as a measure of technology diffusion will not reveal the differences in how technology diffuses by socioeconomic status (Howard et al., 2010, p. 111). Standard measures of distributional inequality have rarely been used to analyze racial and economic disparities in the digital divide research literature (Chakraborth & Bosman, 2005, p. 396). Another limitation of utilizing previous indices or measures is due to their specific policy driven design, which provides little insights on technology adoption or adaptation over time (Howard et al., 2010). The global digital divide has been studied by researchers around the world where several indicators have been established.

### 2.4.1 Statistical Indicators Benchmarking the Information Society (SIBIS)

One of the early efforts in creating statistical measures to benchmark progress of the information society has been led by European countries (Corrocher & Ordanini, 2002). The Statistical Indicators Benchmarking the Information Society (SIBIS) project originated from the European Union (EU) members, EU accession countries, the United States and Switzerland (European Commission, 2001). The SIBIS project (Empirica, 2003) provided a list of statistics and indicators from two EU surveys, General Population Survey (GPS) 2002 & Decision Maker Survey (DMS) 2002 in the following topics:

- Basic access and usage
- Information security
- eCommerce
- eWork
- eGovernment
- eHealth
- Digital literacy
- Learning and training
- Digital divides

While these indicators are primarily created by the two EU surveys, the “digital divide” as a topic in this analysis used a different instrument to derive a composite index. This composite index DIDIX is discussed below.

### 2.4.2 Digital Divide Index (DIDIX)

The digital divide index, proposed by Hüsing and Selhofer (2002), focused on four indicators derived from the Eurobarometer public opinion survey (conducted by European Commission) as a proxy to measure ICT adoption within societies and subgroups. Each indicator and their arbitrary weights to calculate the compound DIDIX index are as follows:

- Percentage of computer users at a given location (30%)
- Percentage of people who use a computer at home (20%)
- Percentage of internet users at a given location (30%)
- Percentage of people who use internet at home (20%)

The constituent risk groups of the DIDIX to represent the disadvantaged groups are:

- Gender – women
- Age – people who are 50 years of age or older
- Education – low education group, people who finished formal school at age 15 or below
- Income – low-income group, which is the lowest quartile from the survey

The application of this index is then used to determine the varying degree of the digital divide among their identified groups by differences in gender, age, education, and income level. Due to the different time period in the survey data, their index was used to compare between 1997 survey results and 2000 survey results. The authors noted in their initial findings that little has changed for the digital divide among people with low levels of education between the two time periods. The gap for senior age groups has widened, and the digital divide for gender(female) or low-income groups were slightly reduced. This index (DIDIX) was later revisited to represent the ratio between the compound ICT adoption indicator in a certain risk group and the value of the corresponding indicator in the total population (Hüsing & Selhofer, 2004). The formulation of the revised DIDIX is shown below:

- Computer use - 50%
- Internet use (at all) - 30%
- Internet use at home, and access at home - 20%

The disadvantaged groups categorization from the previous version has not changed, and a 25% weight should be applied to each group.

### 2.4.3 Digital Access Index by ITU

The ITU Digital Access Index is often presented as the world's first global ICT ranking (International Telecommunication Union, 2003) for international comparison. This index measures the overall ability of an individual in a country to access and use ICT which is composed of five categories:

- Infrastructure
- Affordability
- Knowledge
- Quality
- Usage

Similar to the other indices discussed in this paper, the ICT variables (Infrastructure, quality, and usage) are derived from survey items to represent the number of subscribers in telephone, cellular, and internet services, along with the rate of internet users over a fixed number of individuals within a region. The knowledge variable is based on adult literacy rate and school enrollment information from primary to tertiary institutions. The affordability variable is based on the internet access price as a percentage of the country's gross national income. The ITU organization computed the Digital Access Index and ranked 178 countries. They pointed to education and affordability as two key factors to boost new technology adoption among the developed and developing countries.

### 2.4.4 Technological Capacity Measure

While there are well-established theories such as the diffusion of innovation theory and how to use the S-shaped diffusion process to explain the adoption of ICT devices or technologies, little is known about the capacity aspects of the ICT devices. Hilbert (2014) emphasized the importance of measuring the digital divide in terms of "have much" and "have little" in technological capacity (p. 821). ICT is defined as all tools and technology that mediate information and communication: store information through time, transmit information through space, or transform/compute information, in both analog and digital form (Hilbert, 2014, p. 823). The proposed direct measure of communication capacity includes ICT access, usage, and impact (Hilbert, 2014). The total technology capacity is the sum of the products of the number of installed devices by their respective performances as yearly averages. The unit of measure is in kilobits per second (kbps).

$$\begin{aligned} Technological\ capacity_{\ of\ group\ g} \\ &= Number\ of\ technological\ devices\ t_{kyu} \\ &* performance\ per\ technological\ device\ t_{kyu} \end{aligned}$$

where $t_{kyu}$ represents the subtypes of technologies with different performance for a given year. There are three groups (g) of technology: telecommunication, storage, and computation.

Their results, based on the comparison across countries, found that global level disparity between 2006-2010 has become more equal as well as within countries. However, this does not imply that gaps have been closed. Hilbert (2014) argued that information inequality has only begun to diminish, which means we have not reached the information revolution; ICT would need to be matured and become a general-purpose technology, similar to electricity or automobiles in the 21st century. While Hilbert's measure is innovative in assessing the ICT capacity in relation to digital inequality, this method is very difficult to replicate. For example, the referenced variables in the formula were sourced from 1100 more data sources to make a comprehensive dataset. While the work is impressive and allows international comparison, it is not replicable by others without access to data through the international level (inter-governmental) organizations.

### 2.4.5 Measure of Distribution of Internet Users

Howard et al. (2010) used the Gini coefficient to determine the distribution of internet users in Canadian provinces and American states and as an equality benchmark for each country between 1997-2009. The Gini coefficient was estimated two ways in this work:

1) an estimate of internet users across all provinces or states while treating each geographical layer as equivalent, which allows a comparable unit to measure the distribution

2) an estimate of internet users weighted by the population size at each province or state to measure distribution within each province or state.

The second approach uses the level of concentration in each state/province to determine where the concentrated internet users are located. This Gini coefficient is a common method used by economists to represent the distribution of income within a country to create an index to evaluate income inequality (Howard et al., 2010). Due to the limitation of data availability and data quality from the different surveys, the authors acknowledged the challenge in truly using this measure as a benchmark and could only examine it as a trend line of the Gini coefficient values for descriptive purposes (p. 122).

## 2.5 Methodology and Data Considerations

### 2.5.1 Previous Model or Theory

Even though the combination of surveys with official statistics has been used in the last 20 years, little progress has been made to further refine quantitative indicators/measures to evaluate the digital divide across the country in the U.S. Most of the current dialogues about the digital divide continue to focus on specific social groups and their readiness to ICT. As I have presented the complexities of the digital divide with regard to political, social, and economic differences, the same factors will also affect K-12 students as they progress in our current society. Therefore, a unified model/conceptual framework to study and analyze the digital divide among K-12 students should be a top priority for our country. This section will present several commonly discussed models in relation to the digital divide and how it informs the research design and method in this study.

First, the binary view of the digital divide that originated from popular conception and the political agenda in the 1990s (Hargittai, 2002; Selwyn, 2004) will not be applicable in this study. This study investigates digital opportunity from a multidimensional perspective and uses a quantitative approach to collect and analyze the data. However, it is challenging to decide which factors or variables to include in this study due to inconsistencies within the digital divide literature. As Selwyn (2004) pointed out, while substantial policies are being put in place to combat the digital divide, much of the debate remains conceptually oversimplified and theoretically under-developed (p. 343).

One of the most common theories that is applied to the digital divide research is the diffusion of innovation theory (Vehovar et al., 2006). However, using technology adoption to explain the digital divide is limited to a deterministic view and does not explain how old and new technology contribute to the divide process (Hüsing & Selhofer, 2004). Hilbert (2014) explained the digital divide as an evolving process, where each innovation will reopen the digital divide due to the nature of how adoption rates vary over time following an S-curve pattern from the perspective of the diffusion of innovation theory. This may lead readers to think that the digital divide will never be solved. However, my thinking is that the design (including the variables) is an important step in this quantitative analysis, and it should have some flexibility for substitution to meet the evolving objectives and to support an up-to-date understanding.

Van Dijk and Hacker (2003) suggested future research with large scale surveys using official government statistics where longitudinal and time-series data will be best to test hypotheses about trends on computer use and internet penetration (p. 316). At this time, there is not an ongoing collection of broadband usage data except for the released datasets for the years 2019 and 2020; the design of the SDO variable is meant to allow use across time.

Previously, there were different attempts to categorize the types of digital divide globally or within a country (see Section 2.3). However, none of the previous conceptualizations have logically and structurally presented the different layers of the digital divide. This is an important aspect to consider before developing quantitative analysis, especially using advanced statistical techniques in which the sequence of events and the nested layer of the variables/factors due to social structure will affect the outcome of the analysis. Therefore, this study has considered the multi-level and geographical characteristics of the variables during the analysis and interpretation.

### 2.5.2 Three Levels of Digital Divide

The three levels of the digital divide will be used as the conceptual foundation for the development of the quantitative models in this study. The following is a brief explanation including examples of previous work that pointed to the three levels of the digital divide.

Attewell (2001) described two digital divides in a commentary for the *Sociology of Education* journal in which the first is access and the second is computer use among school-aged children. He pointed out that poor neighborhood schools had less equipment with slow connection and cautioned that not all uses of computers have equivalent education benefits because of social differences in how computers are used at school and at home (Attewell, 2001, p. 253). This work implicitly indicated a social stratification as well as institutional differences (public schools and the families in their neighborhood) among school-aged children in terms of differences in their internet access and computer use. [Factor 1: varying degree of internet access by school location, equipment ownership by school location, and use of computer for different tasks between school and home]. The earliest discussion of the three levels digital divide perspective is from Dolnicar et al. (2004) who discussed the terminology and methodological challenges in measuring the digital divide among European Union (EU) countries. Their representation of the first digital divide referred to the population segments that use the internet and those who do not. At this level, there are further differences, such as dual digital divide (2$^{nd}$ level), due to personal internet or obstacles to use the internet (Dolnicar et al., 2004, p. 422). The second digital divide is an experience gap that could exist even after the first digital divide (access) has been addressed because of socio-demographic variations among users. The third digital divide is based on the idea of fast and slow access to the internet among users, in which users with fast access have immediate advantages and slow users will be de-privileged (Dolnicar et al., 2004, p. 423). Since that time, there has not been much discussion of the digital divide conceptualization concerning how each division or level may have changed over time. It is not until recent years that scholars have revisited the conceptualization of the levels of divide.

Scheerder et al. (2017) conducted a systematic literature review of the 2$^{nd}$ and 3$^{rd}$ level of the digital divide between 2011 and 2016. The inconsistency of different terms for the same thing continues to be an issue in the findings presented by Scheerder et al. (2017). Similar to the other three levels of digital divide, they concluded the unequal distribution of internet access among individuals as the first level, demographic factors associated with skills and use at the second level, and beneficial outcomes are at the third level. Ragnedda (2019) described his work in Africa, where the development of the digital divide is not simply about internet access (first level) but rather focuses on motivation, skills, and purpose of use (second level), as well as social and cultural benefits (third level). An interesting view from the opinions discussed by Ragnedda (2019) is that in "advanced countries, where Internet penetration is really high, inequalities at the base of the societal structures, such as education and gender, do not influence the access to ICTs" (p. 31). I would argue that there is no empirical evidence to support his statement as we have observed the continuing differences between social groups in the United States, especially among the minority groups (See Campos-Castillo, 2015; Katz & Gonzalez, 2016).

More recently, van Deursen and van Dijk (2019) suggested a shift in conception of the first level of digital divide from internet connection to materials access (p. 355) because not all equipment (e.g., computer devices and accessories) and software applications provide the same online opportunities. [Factor 2: varying types of equipment or software and the user experience offered by them] While the development of the digital divide conceptualization is informative, none of the work directly evaluates the digital divide among K-12 students in terms of their digital opportunity. One recent study for school-aged children is by Aydin (2021), who analyzed the 2018 International Computer and Information Literacy Survey (ICILS) by identifying factors from the first and second level of the digital divide among Korean and Chilean students. Their findings mainly evaluate background variables (i.e., gender, parents' level of education, computer experience, and internet connection at home) to students' achievement score in the ICILS survey. Other models such as relative deprivation theory (RDT) suggested by Helsper (2017) are used to understand a person's social and temporal contexts while Brandtzæg et al. (2011) offers a typology of internet user types. For the general population, having internet access and digital skills to navigate the digital world could possibly produce individual-level or community-level benefits, for example, achieved economic status or influence. For school-aged children, their economic benefits may not be observed immediately and therefore, we should determine what types of outcomes will be most meaningful to assess gains [Outcome: social and economic benefits for children]

### 2.5.3 Challenges of Using Self-Reported Data

As seen in the previous index and measure section, most of the previous empirical research utilized national or local surveys to gather internet or ICT access and use across countries. These methods help determine the digital divide in some basic requirements in terms of internet access or ownership of devices. However, these approaches were limited in providing a broad understanding of the issue. For example, the context of digital technology use may not be discussed as part of the digital divide. Van Dijk and Hacker (2003) pointed out that most survey data on computer use and internet penetration are too unreliable to assess the existence or development of the digital divide due to the sampling approach (mostly from marketing research methods) used to collect these data (p. 316). For example, the State of State in Michigan survey conducted in 2016 (Dutton & Reisdorf, 2019) gathered users' and non-users' attitudes and beliefs to understand digital inequalities in Michigan. However, the nature of the study is exploratory and even though the results are interesting, it is difficult to use these results in a meaningful application that can inform policy or support decision-making to address the divide. Hilbert's (2014) discussion of using proxies for internet access, such as measuring the percentage of access to equipment or technology in an area, provided a limited view on access divide because technology and equipment are changing. For example, an area may have a high percentage of reported ICT ownership, but the capacity of these ICT tools could vary from dated equipment to the latest consumer product. The Reisdorf et al. (2017) discussion reflected the same concern. For example, survey items focused on the latest technological innovations (e.g., mobile phones, social media platforms), but ignored some of the older technologies such as personal computers (p. 115) and therefore limited our understanding of how technology changes over time with factors such as the person's access or usage variation.

The self-reported data from internet service providers to the FCC also have similar issues where broadband availability does not reflect the true access for the population in the United States. It is the research objective of this study to demonstrate how to approach survey-based data (e.g., U.S. Census data) with other available public data to construct a more robust measure and model to evaluate the digital divide.

### 2.5.4 Other Micro-Level Studies within the U.S.

At the micro-level, researchers focus on specific groups such as those who are in disadvantaged positions where a social capital or social network perspective is often discussed (Salemink et al., 2017). Within the United States, there is a focus on marginalized communities as well as different demographic groups. The following section will focus on recent work in our country; older research can be found in Grubesic and Murray (2002), Hsieh et al., (2008), or Whitacre and Mills (2007).

Katz and Gonzalez (2016) evaluated what it means to have meaningful connectivity among parents and children of Mexican heritage communities in Arizona, California, and Colorado. Their findings indicate that while families in this study may have similar demographic features, their choices of technology adoption or connectivity are different based on their local environment. Thus, digital inequality is not the same even for people in the same ethic groups or cultural heritage. Another study implemented a digital literacy program in West Virginia (which is a mountainous and rural state with a high number of households in poverty) that used games to influence students' technology engagement in schools (Reynolds & Chiu, 2016). While this study concluded with a positive result of using games to address the digital divide among middle and high school students, the sample size (n = 242) is concerning due to the use of multi-level and multivariate analysis to test their hypothesis. The authors did not share the actual distribution of the race and ethnicity of the sample, even though race and ethnicity were used as factors to estimate the effects of this intervention. For the low-income adult population in the Midwest, Gonzales (2016) identified the reasons for low adoption were mostly due to participants' prior negative experiences, for example, disconnected services, broken hardware, or other barriers to utilizing public internet access.

Our current understanding of digital opportunity for students is very limited. Identifying the basic requirements for reliable internet access and suitable devices helps form a baseline for further evaluation. If not, we will continue to find differences or gaps and fail to notice that we are not comparing the same thing. International organizations, such as the United Nations 2030 Agenda for Sustainable Development Goals 4 (ensure inclusive and equitable quality education and promote lifelong learning opportunities for all) is concerned with the learning lost during COVID-19 across developing and developed countries (United Nations Statistics Division, 2021). Before the pandemic, most American students appeared to have access to the internet and digital devices for their learning, but people quickly learned that this was not the reality. COVID-19 pandemic has only magnified the digital inequality that has been ongoing within K-12 education for some time. Currently, the lack of explicit targets in technology for education is similar to the problem of lacking a national educational curriculum across technology-related subjects. Since there is so much variation in policy decisions in each state, this study cannot include that information meaningfully in the research design. However, future research that narrows the research focus to within-county or within-state variation will be able to incorporate policy work in addressing digital equity in K-12 schools.

## 2.6 Summary

While much of the previous research had extensively described the issue of the digital divide or digital equity, none of this work had quantitatively analyzed and evaluated broadband internet access and use along with device ownership and speed quality as proposed in this study. The similar but different use of terminologies surrounding digital equity issues added another layer of complexity. A discussion of the different factors and terms along with their definitions has been provided. Having consistent definitions of relating factors and concepts are essential in measurement development. The final set of variables used in the study will be further explained in Chapter 3. A discussion of previous models and conceptual frameworks displayed the inconsistency in evaluating the digital divide. Thus, the primary purpose of this research is to construct a holistic framework of digital opportunities based on the key factors in literature. This conceptual framework is further operationalized to create a tangible measure for comparison at the county level. This study is the first to provide empirical evidence as to how K-12 schools and students in different U.S. counties are different in terms of their digital opportunities. Chapter 3 provides the rationale for using a quantitative approach to answer the research questions and describes the techniques used in data analyses.

# CHAPTER 3

# RESEARCH METHODOLOGY

## 3.1 Introduction

This study used a quantitative approach to investigate digital opportunities of K-12 public school students across the United States. A set of public governmental data and broadband open data were used to establish a composite measure (Students' Digital Opportunity, SDO) for quantitative analysis. A preliminary study for an open data challenge (Jim et al., 2021a, 2021b) was conducted to evaluate broadband access and use at the county level with at least one operating K-12 school during the 2017-2018 school year. Several initial results from the open data challenge needed further exploration. This included exploring variables from the secondary datasets and examining the SDO distribution with spatial analysis. The geospatial methods support the evaluation of how SDO is distributed across counties and states. This study extended the initial analysis to include more recent datasets to investigate digital opportunities among K-12 students, such as the speed quality information from Ookla and IMLS public library information in each county.

## 3.2 Data Collection

Van Dijk and Hacker (2003) emphasized that data concerning the digital divide are often driven by political interest and subjective perspectives (p. 321). The previous discussion of metrics development has focused on assessing the digital divide across nations (e.g., European Union) or for international comparison (e.g., ITU metrics) which tells us very little about variation within countries. The use of individual survey data is also problematic because sampling procedures may not be transparent. Reusing those small sample data could lead to bias and misinterpretation.

### 3.2.1 Quantitative Data Sources

In this study, I utilized datasets from Microsoft Airband Initiative, U.S. Census Bureau, National Center for Education Statistics, Federal Commission of Communication, Ookla, and Stanford Education Data Archive. The types of data and their sources are displayed as follows:

- Public school characteristics and administrative record (National Center for Education Statistics, 2021)
- Microsoft broadband usage data at ZIP code and county level (Kahan & Lavista Ferres, 2020)
- Broadband availability data (Federal Commission of Communication, 2021b)
- Fixed broadband speed test data (Ookla, 2020)
- ACS 5-year estimates Table B28005: Age by presence of a computer and types of internet subscription in household (U.S. Census Bureau, 2019a)
- Stanford Education Data Archive (SEDA) version 4.1 (Reardon et al., 2021)
- 2019 TIGER County level shapefiles (U.S. Census Bureau, 2019b; Walker, 2021)

Each data file required its own processing; the following sections provide the basic information of each data source but did not address the specific steps in data cleaning. Details of the data processing and cleaning is provided in my GitHub page: https://github.com/caryjim/Digital_Equity

### 3.2.2 Description of Data Sources

This study utilizes secondary open data at the population level with a focus on K-12 public schools and student population. This section provides information on the original data sources and each dataset's characteristics.

#### 3.2.2.1 Microsoft Airband Initiative Broadband Data

The Microsoft broadband usage datasets are available for 2019 and 2020 and were organized into two geographic levels: ZIP code and county level in each state (Kahan & Lavista Ferres, 2020). The 2020 broadband usage data file contains 3,142 counties and 5 variables with the two-letter postal abbreviation of the state, county ID, county name, percentage of broadband availability (per FCC 2019 report), and broadband usage estimates at each county in 2020. There are 20 counties in which broadband availability, or the usage estimates are missing in the data file. These counties are located in Alaska (AK), Nebraska (NE), Oregon (OR), Texas (TX), and Virginia (VA).

The broadband usage data at the ZIP code level contains 32,653 observations and 8 variables after removing duplicated records, where two counties in AK and NM had an incorrect county ID. The columns are delineated as the two-letter abbreviation of the state, county ID, county name, postal code (ZIP code), and broadband usage estimates at each ZIP code along with their error information. In the 2020 ZIP code-level data file, it contains 3,132 counties and 51 states.

#### 3.2.2.2 The U.S. Census American Community Survey (ACS) Table B28005

The ACS Table B28005 is exported with the 5-year estimates: Age by presence of a computer and types of internet subscription in the household (U.S. Census Bureau, 2019a). There are 3,142 county and 33,120 ZIP code in the U.S. Census Table B28005. The data file contains the following predefined columns:

- GEO_ID
- County ID
- County name
- State
- Total number of estimated households

- Total number of estimated households with children under the age of 18, and the following subcategories applied to households with children:
  - Estimated number of households with PC
  - Estimated number of households with PC and Broadband
  - Estimated number households without Internet
  - Estimated number households without PC

The initial table B28005 is further processed to normalize the counts for each county as percentages. Three variables were derived per household with children in each county: Percentage of homes with students, percentage of homes with children and a PC, percentage of homes with children with a PC and broadband subscription. The estimated population count is also exported from U.S. Census ACS at the county level.

#### 3.2.2.3 NCES K-12 Public School Data

The K-12 public school characteristics and administrative information is extracted from the NCES Elementary and Secondary Information System (National Center for Education Statistics, 2021). The 2019-2020 school year dataset is used in this study. There are 99,338 rows of schools found in the initial dataset with 37 variables broken down by location information, locale (urban-centric), Title I information, National School Lunch Program participation, race/ethnicity of the students, pupil-to-teacher ratio, and teacher full-time status. A county-level school data file is available in which the number of schools have been aggregated to each county. The district-level and county-level data files are used to cross-check the information provided in the initial school-level dataset.

#### 3.2.2.4 Broadband Availability Data

The Federal Communications Commission releases an annual broadband progress report.

This report includes information on broadband deployment of fixed terrestrial and mobile networks. The released report in June 2020 provided the percentage of population per each county who have access to 25/3 Mbps fixed terrestrial broadband and 5/1 Mbps mobile LTE services by the end of December 2018 (Federal Commission of Communication, 2021b). In this study, the broadband access variable within the Microsoft Broadband Usage dataset is used to represent the percentage of the population who have access to fixed terrestrial broadband.

#### 3.2.2.5 Ookla Speedtest

The Ookla Speedtest is a trademarked online application that collects speed information generated by internet users globally. The global dataset includes fixed broadband and mobile (cellular) network information. The Speedtest keeps track of the user' location, download speed, upload speed, and latency information. It also contains aggregated information of geo-layers, device counts, and the number of speed tests conducted at a location. The $4^{th}$ quarter of 2020 Ookla Speedtest Open data is used in this study (Ookla, 2020).

#### 3.2.2.6 Stanford Education Data Archive (SEDA) version 4.1

The SEDA version 4.1 (Reardon et al., 2021) dataset contains information of K-12 public school students' academic achievement estimates in Math/Reading from $3^{rd}$ to $8^{th}$ grade across the United States. There are also covariate data files that provide information on the proportion of students by their race/ethnicity, locale, English language learning status, special education status, within-race group comparison, poverty indicators, and socio-economic composite measure.

#### 3.2.2.7 2019 U.S. County Level Shapefiles

To obtain the geo-spatial information needed for the spatial analysis, the shapefiles have been obtained from the U.S. Census TIGER/Line webpage for the year 2019 at the county level (U.S. Census Bureau, 2019b). An R package is also used to retrieve a county-level shapefile (Walker, 2021) within the data processing procedure during data exploration.

#### 3.2.2.8 Public Libraries Survey (PLS)

The public libraries administrative data are collected yearly by the Institute of Museum and Library Services (IMLS) in the United States. The 2019 fiscal year of the public libraries survey data (IMLS, 2022) is used in this study. The PLS collects yearly information from public libraries across the U.S. The dataset included data about the circulation services, staffing, collections information, operating revenues, and expenditures of public libraries. This information was used to help us evaluate the number and sizes of the public libraries located in each county within the U.S.

## 3.3 Key Variables

### 3.3.1 Broadband Internet Availability

Broadband internet is referred to as the high-speed internet services available to general consumers or businesses. The various types of broadband internet services and how the general public can obtain the related technologies have been discussed in Section 2.3.2. The Federal Communications Commission (FCC) determines the definition of broadband internet which is set as a minimum of 25 Mbps download and 3 Mbps upload speeds (Kruger, 2018). The annual report of broadband deployment (FCC, 2020) further distinguished the speed categories by download/upload ranges as 4/1, 10/1, 100/10, 250/25, and 1000/100 Mbps. This study used the 25/3 Mbps speed as the baseline value which any families or children at home would need to participate in various online activities. For this study, the percentage of access to broadband internet in each county was sourced from the FCC annual report where internet service providers

(ISPs) self-report their coverage area across the United States. The fixed terrestrial broadband at the 25/3 Mbps speed range is the basic level of broadband connectivity I used in this study. Therefore, it did not differentiate by connection medium through cables, fiber-optics, or other forms of landline connections for each home. The percentage of broadband access represents the availability of broadband internet services for each county's population.

### 3.3.2 Broadband Speed

The measure of broadband internet speed is based on the speed tests conducted by internet users across the country (Ookla, 2020). The initial dataset was aggregated to the county level and the median download and upload speed were estimated to represent the speed ranges that are commonly experienced by the people in each county. The average latency has been aggregated to the county level for descriptive purposes. Latency is a term to describe the amount of delay in milliseconds for the data to travel in a complete loop between the user's device and a server in a network. It is useful to have information on the typical delays experienced by the users during the data exploration. As expected, there is an inverse relationship between latency and download/upload speed (see Table 2 in Chapter 4 Results). This means that areas that tend to have access to a higher range of download and upload speeds will have less delays in their internet connection.

### 3.3.3 Broadband Internet Usage

Broadband internet usage is a variable that means different things to different people. The U.S. Census Current Population Survey (CPS) in 2019 included a computer and internet use supplement study by sampling different regions within the country. Their questionnaire (U.S. Census Bureau, 2019c) focused on asking about ownership of different digital devices, the users of these devices, and the location of their use. The range of digital devices included desktop computers, laptops or notebooks, tablets, smartphones, smart watches or glasses, smart TVs, and internet connected games systems. Thus, usage is delineated by the type of digital device, types of online activity performed, and location of the activity. Dutton and Reisdorf (2019) described the different typology of internet users and non-users in which most categorizations are based on technology types or patterns of use. Therefore, we lack the knowledge of the actual use of the internet at broadband speed beside self-reported information. The broadband usage dataset across the county level from Microsoft Airband's initiative solved this problem. The usage variable from this dataset is based on telemetry of the machines that use any of the Microsoft services/software and is a better indicator of internet usage across the country compared to sub-sample self-reported survey data.

### 3.3.4 PC Ownership and Broadband Subscription

Within the U.S. Census Bureau American Community Survey (ACS), there are several tables for the "Computer and Internet Use" topic. The *PC ownership and Broadband (Internet) Subscription* variable in the B28005 Census table is used in this study (U.S. Census Bureau, 2019a). This variable represents a suitable digital device for learning (e.g., a personal computer) at home with broadband internet subscription in a household with children. There are two versions of the table in which the 5-year estimates during 2019 are the only datasets that include all counties within the U.S.

### 3.3.5 Locale: Urbanicity and Rurality

As previously described in Section 2.3.3 where geographical characteristics may play a role in the differences of digital infrastructure and connectivity. In the NCES K-12 public school dataset, there is an urban-centric locale variable that provides categorical information of four major groups of locale: urban, suburban, town, and rural. These major categories are further divided into subcategories based on their distance from the nearest urban cluster (see NCES 2019 for more information on locale classifications). The categorical information was used to calculate a percentage of schools by the urban or rural categories. Schools located in urban and suburban areas were counted toward the percentage of urbanicity. Schools located in town and rural locales were counted toward the rurality percentage. For example, a county may have five schools in urban settings and five schools in rural areas out of a total of ten. Then, the calculated percentage for this county will be 50% urbanicity and 50% rurality.

### 3.3.6 Title I Status

In K-12 public schools, Title I status is an indicator that has been used to determine resource allocation and fundings across states and at the federal level. Title I is estimated by the percentage of children from low-income families at each school or educational agency (More details can be found on the U.S. Department of Education, 2018 webpage). While there are other indicators for socioeconomic status or poverty level, they may not represent the student population directly. Therefore, the Title I school status is a reliable indicator of high-need schools and their locations across the county. The original Title I variable from the NCES K-12 public schools contained six categories and due to the aggregation of this information to a county level, a percentage describing the amount of Title I schools in each county was calculated. It means that any schools identified as eligible for Title I funds or any schools that have been certified as Title I schools have been included in the calculation.

## 3.4 Research Questions

The current landscape of digital learning or technology-based teaching depends on a stable digital infrastructure for students and their schools. A comprehensive digital infrastructure for K-12 education includes the physical, sociocultural, economic, and human dimension. For K-

12 students, their basic physical technology requirements are to have access to both a computer and reliable internet at school or at home in order to participate in any form of learning with technology. Currently, almost all software applications have to be connected to the internet for ongoing use, thus internet and equipment access is a combined element for digital connectivity. Digital connectivity is the physical aspect of the digital infrastructure. Other factors add complexity to the picture of digital equity, such as geographical differences, socioeconomic factors, and local policies. The human dimension may include the culture and beliefs of the schools and their instructional practices. While it is impossible to address the human dimension in this work, this aspect can be further broken down into the level of digital skills in a multidimensional measure. Digital skills among students also heavily depend on their age group and maturity as they relate to student ability and knowledge. The current research on digital skill has faced the same challenge of inconsistent definitions and a lack of theoretical background due to the variation in technology programming and devices available in K-12 schools. From a long-term perspective, there is an inherited generational gap in using technology as well, because teachers who graduated from their training 10 to 30 years ago may not be using the same technologies in their current classroom. Those differences are very difficult to capture and account for unless there is systematic data collection to record this information.

In Chapter 2, previous research and established indicators were limited in the way they assess the digital divide among the general population. Currently, there is not a conceptual framework to evaluate digital opportunity among the student population in K-12 education or a quantitative measure to examine how access and use of internet and digital devices are distributed across the U.S. Therefore, the following research questions were developed to investigate the digital connectivity aspect of digital equity among K-12 students:

1. How does students' digital opportunity (SDO) distribute across the United States at the county and state level?
2. Are there any geographical associations with the distribution of the Students' Digital Opportunity (SDO) measure across the United States?

## 3.5 Quantitative Methods

This study used a quantitative approach to construct a conceptual framework to evaluate the digital connectivity aspect of digital equity, called the Students' Digital Opportunity (SDO). This quantitative study contains multiple steps in data analysis to ensure the developed model is robust and applicable to decision-making at the county and state level. The following section will elaborate on the various data analyses conducted in this study.

### 3.5.1 Exploratory Data Analysis

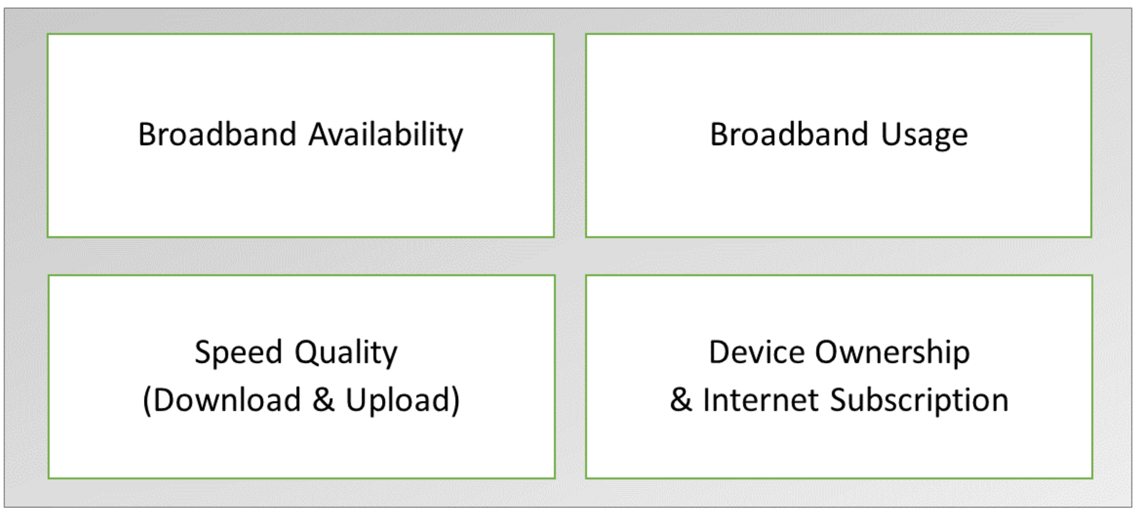
Several techniques have been used to generate an understanding of the different datasets used in this study by examining the patterns and relationships among the variables. A descriptive statistics including correlation analysis provides a summary of the data characteristics. Tables and figures were produced for visual examination of the data characteristics and structure. The visualization of the key variables provided an initial understanding of the multivariate relationship and their distribution. The *psych* package in R was used to conduct the descriptive statistics analysis. An exploratory k-mean clustering analysis with the key variables supported the determination of similar group characteristics by urbanicity, rurality, Title I status, and socioeconomic status with the identified digital connectivity components (broadband, speed, and device ownership). The *cluster* package in R (Maechler et al., 2021) was used for the k-mean clustering analysis. The exploratory data analysis results support decisions of the model design and analytical procedures in this study.

### 3.5.2 Students Digital Opportunity (SDO) Measure

The structural design of the SDO measure was conceived by the last two decades of research (as explained in Chapter 2) where four major components of digital connectivity were examined: access or availability to broadband internet, utilization of broadband service, internet speed, and technology access or computer ownership. Figure 2 displays the conceptual framework of Students Digital Opportunity measure.

Figure 2

*Students' Digital Opportunity Framework*




To my knowledge, there is not a scale nor an instrument similar to what I have proposed to represent the multifaceted components of digital opportunity. Therefore, this work was conducted in an exploratory manner to assess the components within the SDO conceptual model. The SDO measure was conceptualized to represent the relative position of digital opportunity of students across counties in the U.S. with an emphasis on their digital connectivity. The quantitative method used to develop the SDO measure was informed by measurement principles and research (Alwin, 1973; DiStefano et al., 2009; Gorsuch, 1974; Henson & Roberts, 2006; OECD, 2008; Tabachnick & Fidell, 2019; Watkins, 2018) with an intention to support educational decision-making at the county and state level. A factor analysis was used to highlight the relationship between the four identified components to evaluate the underlying structure of the latent variable. During the factor extraction, several strategies were used to determine the number of factors to retain, including the Very Simple Structure (Revelle & Rocklin, 1979) and Horn's parallel analysis (Horn, 1965).

The principal axis factoring approach was used during factor analysis and the Bartlett estimation method (Bartlett, 1937) was used to create a set of standardized score to represent the position of each county. The derived scores are the factor scores of the SDO variable, and the use of Bartlett approach is due to the consideration of minimizing errors (unique factors) to produce an unbiased estimate (DiStefano et al., 2009). During the process of factor analysis, the *psych* package (Revelle, 2021) version 2.1.9 in R programming was used.

### 3.5.3 Reliability and Validity

The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (Kaiser, 1970; 1974) and Bartlett's Test of Sphericity (Bartlett, 1950; 1951; 1954) were used to prescreen the selected variables to determine if they were factorable (both available in *psych* package). After conducting the factor analysis and obtaining the standardized factor scores, the reliability and validity test of the SDO construct was conducted. The Cronbach's alpha method was used to determine the reliability of the SDO construct. The internal consistency and the construct reliability were assessed. The content validity of the SDO construct was examined by checking the factor loadings, communalities, and their uniqueness to determine if the selected component is representative of what was conceptualized.

### 3.5.4 Spatial Analysis

Often, human related events are not random due to their dependence on location.

Lucendo-Monedero et al. (2019) suggested the use of spatial analysis to identify regional spatial patterns in digital divide research. The use of spatial analysis in this study has two purposes. First, applying spatial autocorrelation analysis can help us determine if the distribution of the SDO variables is associated with geographical characteristics. Second, spatial clusters and dispersion from the analysis can be visualized to determine which regions share similar characteristics and how intervention can be applied to address digital equity in terms of disparity of digital connectivity experienced by the K-12 students in different regions. The global and local Moran's *I* statistics are used to model spatial patterns in a descriptive way. The Moran's *I* (Moran, 1950) statistic is a classic measure of spatial dependence of a feature variable with its spatial polygons and points (Chainey & Ratfliffe, 2005). The global Moran's *I* can determine if there is a strong clustering between neighboring areas across an entire map. In other words, it determines if regions follow a similar pattern. A positive spatial autocorrelation coefficient indicates clustering, and a negative value indicates dispersion. Based on our current understanding of digital disparity as a non-random event, we expect to reject the null hypothesis testing of spatial randomness in the global Moran's *I* statistics.

A disadvantage of visual inspection of mapped values is that we cannot determine if the observed clusters are significant. Thus, local Moran's *I* offers additional statistical robustness (Chainey & Ratcliffe, 2005; Rey at al., 2020). The local indicators of spatial associations (LISA) approach proposed by Anselin (1995, 2020) can determine local cluster or repulsion (Brunsdon & Comber, 2019). A local Moran's *I* statistics was obtained for each county with the associated *p*-values. This local level spatial analysis also allowed for exploration of local dependence and its relation to the global Moran's I statistics. Scholars have cautioned against the interpretation of *p*-values from local Moran's statistics in the conventional way (Anselin, 2020; Bivand et al., 2013). The indication of significance is only useful when it is combined with the variable of interest in a Moran scatterplot to allow for classification of the four quadrants. The four quadrants represent High-High and Low-Low regions as spatial clusters versus High-Low and Low-High regions as spatial outliers (Anselin, 2020; Lansley & Cheshire, 2016). While there are follow-up analyses for multiple hypothesis testing (Brunsdon & Comber, 2019) or fitting models with areal data (Bivand al., 2013), these analyses are beyond the scope of this dissertation.

## 3.6 Summary

This chapter described the quantitative research methods and analyses for this study. The description and sources of datasets and the relevant variables are explained. Chapter 4 discusses the data processing needed as part of a data-driven study and the results from the various quantitative data analyses.

# CHAPTER 4

# RESULTS

## 4.1 Introduction

This study utilized open data to evaluate the current state of digital opportunity among K-12 students conceptualized within a digital equity perspective. The key components of equitable digital opportunity are broadband access, usage, speed quality, and device ownership. A composite measure, SDO, was designed to holistically represent the key components and the level of variation among the K-12 students based on their location. Several data products and thematic maps were produced as part of this research and made available to others for replication and use.

## 4.2 Data Extraction and Processing

The programming language R (v 4.0.5) and the software RStudio (v 2021.09.1) were used for the initial data processing and aggregation. The *tidyverse* set of packages (Wickham, 2021) in R programming was used for the data extraction, file processing, and aggregation. The final dataset was aggregated at the county level across the 50 states including the District of Columbia and consists of 3,138 counties with variables associated with school characteristics, socioeconomic status, urbanicity or rurality, Title I status, broadband availability and usage, PC ownership, broadband subscription, and speed quality information.

All K-12 schools located within the 50 states including the District of Columbia (D.C.) are included in the dataset. There are 3,954 school records did not provide their student enrollment information (3.98% of the whole dataset). Schools that are listed as not regular schools, for example, alternative schools or vocational schools are dropped. The final dataset contained 91,182 operating K-12 public schools and approximately 50,031,797 students. The school-level data was further aggregated to a county level for this study. Four counties were omitted in the final dataset as there is not a K-12 public school in operation within their legal boundaries: Loving County, TX; Buffalo County, SD; Issaquena County, MS; and Kalawao County, HI.

There are missing values in the original broadband access and usage data from the Microsoft Airband Initiatives. This is due to a change of county ID and names in the U.S. Census record after 2010 as the boundaries of some towns or rural counties have changed. The missing information in these rural areas (due to boundary changes) are partly due to the use of satellite connection or other non-terrestrial internet connection. A random forest approach (Mayer, 2021) was used to impute the missing values based on the mixed-type data in the final dataset. It is assumed that the counties with missing broadband deployment and usage information have similar patterns with their nearby regions. The distribution of variables with imputed values were checked. A manual step of reviewing the imputed values with nearby counties with similar demographics was completed before data analysis.

After the data analyses, the derived factor scores (SDO), the local Moran's *I* statistics and quadrants information were appended to the output data file at the county level. Data visualization was utilized to further explore distribution of the measure and create thematic maps.

## 4.3 Exploratory Data Analysis

There were several steps in the exploratory data analysis. First, the descriptive statistics of each variable were calculated using base R functions. Then, data visualization of the descriptive statistics were generated with the *tidyverse* and *psych* packages in R. Table 1 displays the summary statistics of each variable in the final dataset. Table 2 displays the Pearson *r* correlation of the key variables. Other relevant variables that were not included in the factor analysis model are the percentage of Title I schools, percentage of urbanicity/rurality of schools, and the socioeconomic measure at each county. These variables are considered the external factors that relate to the components of the SDO measure.

Table 1

*Descriptive Statistics of the Key Variables (Non-Standardized)*

| Variables | Mean | SD | Median | Skewness | Kurtosis |
|---|---|---|---|---|---|
| Urbanicity (Percentage) | .17 | .31 | 0.00 | 1.54 | .74 |
| Rurality (Percentage) | .83 | .31 | 1.00 | -1.54 | .74 |
| Title I Status (Percentage) | .81 | .21 | .88 | -1.13 | .65 |
| Broadband Availability (Percentage) | .84 | .20 | .92 | -1.84 | 3.26 |
| Broadband Usage (Percentage) | .86 | .23 | .37 | .32 | -.89 |
| PC and Broadband at Home (Percentage) | .86 | .09 | .88 | -1.60 | 3.91 |
| Download Speed (Mbps) | 75.0 | 40.62 | 70.40 | .41 | -.68 |
| Upload Speed (Mbps) | 23.52 | 22.42 | 14.87 | 2.19 | 5.52 |
| Latency (milliseconds) | 66.20 | 52.88 | 51.03 | 2.86 | 15.73 |
| Socioeconomic Status (Percentage) | .03 | .74 | .09 | -.64 | 1.01 |

*Note*. There are 3,138 U.S. Counties. SD = standard deviation.

The urbanicity and rurality variables displayed a perfectly inverse relationship ($r$ = -1.0). A moderate positive correlation is observed among broadband availability and use ($r$ = .56), broadband availability and download speed ($r$ = .47), urbanicity with broadband usage ($r$ = .62), socioeconomic status with PC and broadband at home ($r$ = .56), and urbanicity with download speed ($r$ = .64). There are several moderately inverse relationships including rurality to broadband usage ($r$ = -.62), rurality to download speed ($r$ = -.64), and broadband availability to latency ($r$ = -.56).

Table 2

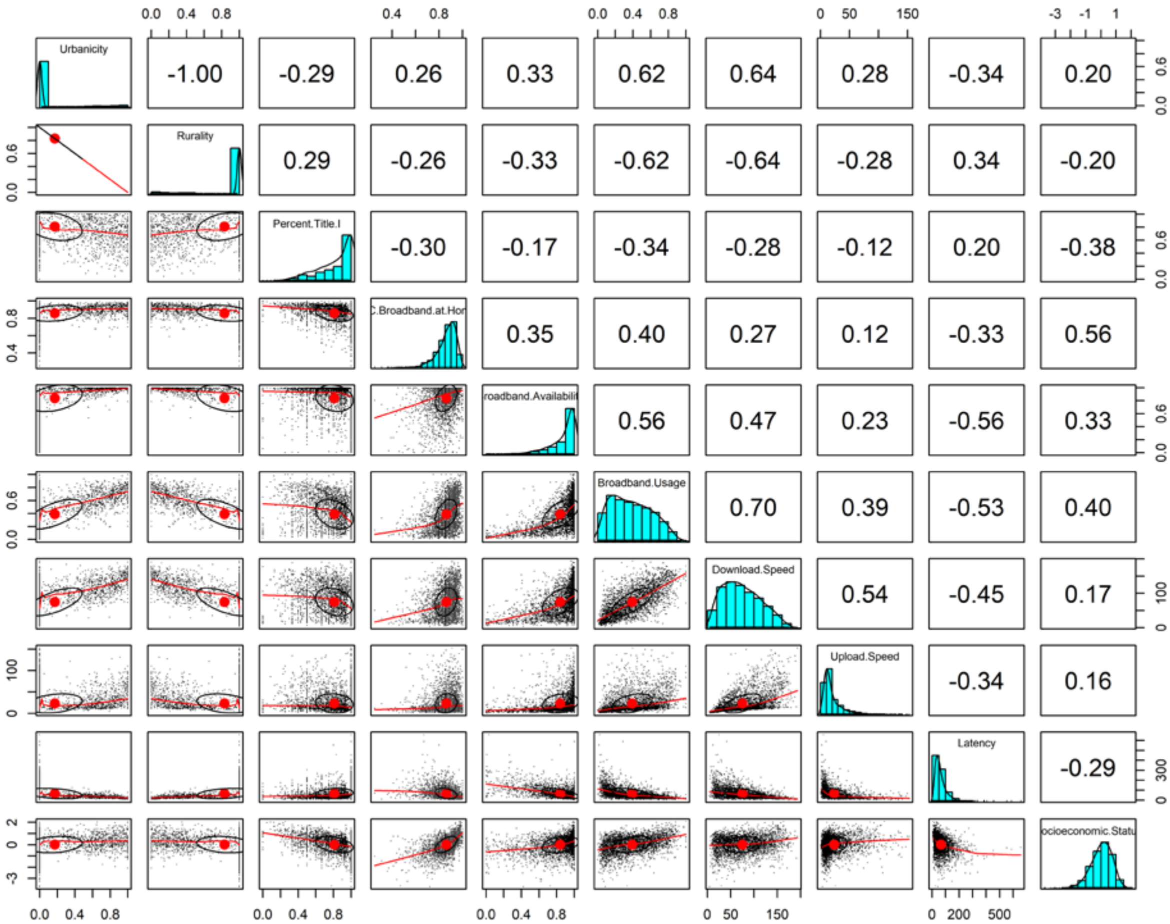
*Pearson* r *Correlation of the Key Variables*

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Urbanicity (Percentage) | 1 | | | | | | | | | |
| 2. Rurality (Percentage) | -1 | 1 | | | | | | | | |
| 3. Title I Status (Percentage) | -.29 | .29 | 1 | | | | | | | |
| 4. Broadband Availability (Percentage) | .33 | -.33 | -.17 | 1 | | | | | | |
| 5. Broadband Usage (Percentage) | .62 | -.62 | -.34 | .56 | 1 | | | | | |
| 6. PC and Broadband at Home (Percentage) | .26 | -.26 | -.30 | .35 | .40 | 1 | | | | |
| 7. Download Speed (Mbps) | .64 | -.64 | -.28 | .47 | .70 | .27 | 1 | | | |
| 8. Upload Speed (Mbps) | .28 | -.28 | -.12 | .23 | .39 | .12 | .54 | 1 | | |
| 9. Latency (milliseconds) | -.34 | 34 | .20 | -.56 | -.53 | -.33 | -.45 | .34 | 1 | |
| 10. Socioeconomic Status (Percentage) | .20 | -.20 | -.38 | .33 | .40 | .56 | .17 | .16 | -.29 | 1 |

*Note*. There are 3,138 U.S. Counties. All variable pairs have significance $p < .001$.

Figure 3

*Scatterplots and Histograms of Key Variables in Pair Panel Format*




A series of scatterplots and histograms were generated to assess linearity, normality, and outliers (Figure 3). The non-normal distribution of several factors can be explained due to the fact there are existing differences of broadband availability, usage, and speed quality across the nation. Some distortion could occur in which variables with a similar level of skewness and kurtosis may form artifactual factors (Bandalos & Finney, 2010). Due to this observation of the skewed distribution of various variables, data transformations were applied (log and square root approaches), however, it made no substantial difference and would have increased the difficulty of later interpretation if implemented. While the current best practices to handle outliers in the case of factor analysis (Tabachnick & Fidell, 2019) were considered, a careful decision was made to keep these cases because there are places in our country that still do not have or have a low level of access to broadband internet. In sum, the distribution of the selected variables and their correlation were checked to determine their likelihood for a common factor before proceeding to factor analysis.

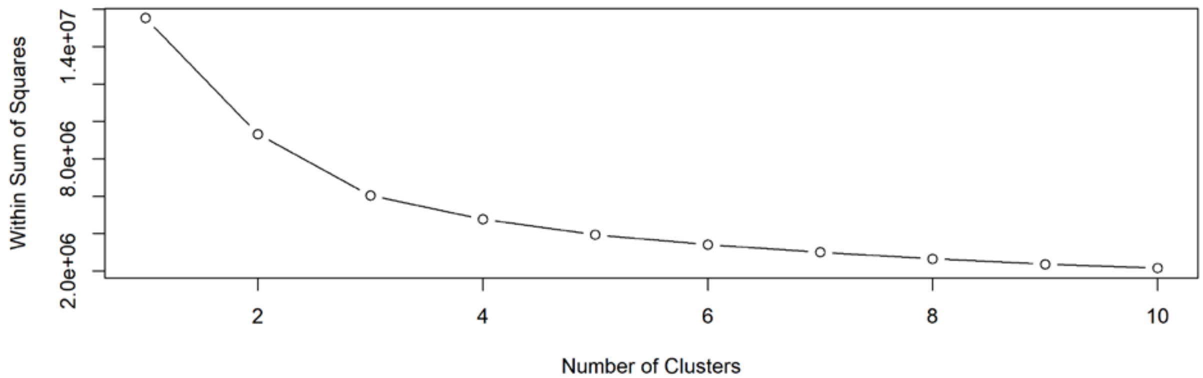
The evaluation of descriptive statistics provided support toward the use of factor analysis (Bandalos & Finney, 2010; Tabachnick & Fidell, 2019) to construct the SDO measure and to derive the factor scores to evaluate the K-12 students digital opportunity.

A decision was made not to include the public libraries dataset in the following analysis. After the initial data exploration, there were only 2,775 counties reported to have an operating library branch or outlets in the year of 2019. Due to the large discrepancy of how public libraries are distributed across the country as public infrastructure (about 363 counties either do not have a library building or it is unreported), it was not suitable to join the public libraries dataset to the K-12 school datasets at the county level. The implication of this decision during data exploration will be further discussed in the next chapter.

Exploratory k-mean clustering was used to explore the similarities of the key variables by their values in the dataset. Two methods were used to determine the optimal number of clusters: the elbow method and the *NbClust* package (Charrad et al., 2014) in R that utilized 25 indexes for comparison. The elbow method was conducted by computing a range of clusters from 1 to 10 with the k-means clustering algorithm (stat package in base R). For each cluster, the within-cluster sum of squares is calculated and plotted by the number of clusters estimated. The "elbow" which is the bend in the plot indicates the appropriate number of clusters in Figure 4.

Figure 4

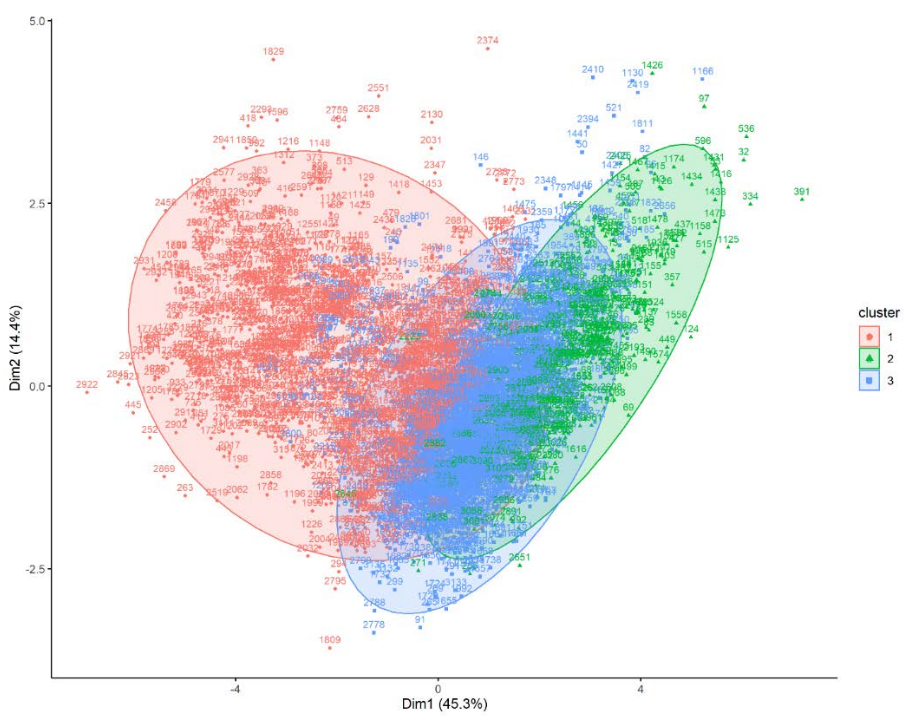
*Optimal Number of Clusters by Elbow Method*




The NbClust indexes suggested three clusters as the best solution in the dataset. Thus, a k-mean clustering solution of a three-cluster model was obtained. There are 1,164 counties identified in the 1st cluster, 381 counties in the 2nd cluster, and 1,593 counties in the 3rd cluster. Figure 5 displays a two-dimensional visualization of the clusters in this analysis.

Figure 5

*Exploratory Clustering Solution - Three Clusters*



The calculated cluster centroids (means) of each variable are displayed in Table 3. Based on this information, it can be determined that Cluster 2 characterizes counties that experienced the lower range of the broadband services including speed quality and usage in mostly rural places with student populations positioned in a below-average socioeconomic status. Cluster 1 represents counties that experienced higher ranges of similar characteristics, and Cluster 3 represents regions that fall between the two clusters in terms of their broadband services and are mostly driven by their rural characteristics.

Table 3

*Cluster Centroids of Key Variables in Each Cluster*

| Variables | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| Urbanicity (Percentage) | .41 | .00 | 0.4 |
| Rurality (Percentage) | .59 | 1.00 | .96 |
| Title I Status (Percentage) | .74 | .88 | .84 |
| Broadband Availability (Percentage) | .94 | .59 | .82 |
| Broadband Usage (Percentage) | .57 | .16 | .31 |
| PC and Broadband at Home (Percentage) | .89 | .80 | .85 |
| Download Speed (Mbps) | 118.16 | 39.68 | 51.90 |
| Upload Speed (Mbps) | 38.45 | 10.91 | 15.62 |
| Latency (milliseconds) | 36.41 | 175.11 | 61.92 |
| Socioeconomic Status (Percentage) | .18 | -.39 | .01 |

## 4.4 Factor Analysis

There were two objectives during the exploratory factor analysis. The first objective was to explore the proposed latent structure of the Students’ Digital Opportunity (SDO) Measure and evaluate each component in a common factor model. Second, it was designed to obtain a score to represent each county for comparison purposes. The goal was to obtain a factor solution that is interpretable and applicable to supporting meaningful decision-making. The observed variables of broadband availability/access, broadband usage, speed quality, and ownership of PC and broadband subscriptions were used to construct the SDO measure. During the exploratory data analysis, the descriptive summary and correlation of the observed variables were examined. The suggested threshold of correlation magnitude between .3 to .7 indicates appropriateness to use the selected variables in factor analysis (Tabachnick & Fidell, 2019). The direction of the correlation was also examined in which all selected variables for the SDO measure were positively correlated. Upload Speed (Mbps) correlated with other variables lower than .3 except with broadband usage (.39) and download speed (.54). Since upload speed is conceptually a part of the speed quality and technically related to download speed, it was included in the factor model analysis. The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (Kaiser, 1970; 1974) and Bartlett's Test of Sphericity (Bartlett, 1950; 1951; 1954) were used to prescreen the variables to determine if they were factorable. A KMO measure of sampling adequacy of .74 indicated sufficient items for each factor and a significant (< .001) value in Bartlett's Test of Sphericity indicated that the correlation matrix is different to the identity matrix significantly. Due to the sensitivity of the Bartlett's Test of Sphericity, a significant result can be expected when a substantial sample size is in place (Tabachnick & Fidell, 2019).

Initial evaluation of the factor structure was performed with Horn's parallel analysis (Horn, 1965) and Very Simple Structure (Revelle & Rocklin, 1979) methods in the *psych* package. The Very Simple Structure (VSS) function utilized three criteria: VSS criterion (Revelle & Rocklin, 1979), Velicer's Minimum Average Partial (MAP) criterion (Velicer, 1976), and the Bayesian Information Criterion (BIC) minimum.

A Horn's parallel analysis produced a scree plot of the observed data compared to the randomly generated data matrix as shown in Figure 6. By examining the scree plot (Figure 6), the line representing the observed data is above the reference line (Eigenvalue of 1) on the y-axis for a one factor model. The random and simulated data were used for comparison and none of their resulting lines are above 1 in the scree plot. Therefore, the Horn's parallel analysis suggested a one-factor solution.

Figure 6

*Scree Plot of Horn's Parallel Analysis*

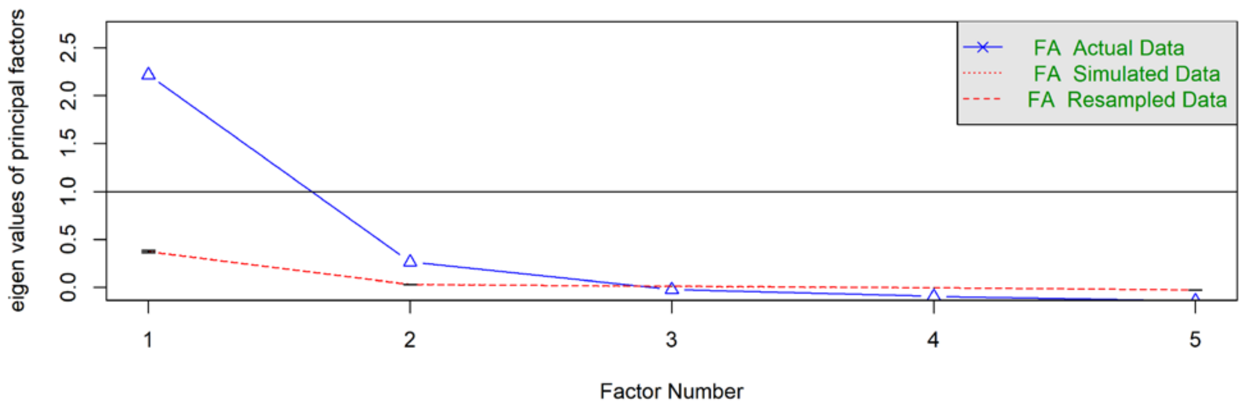
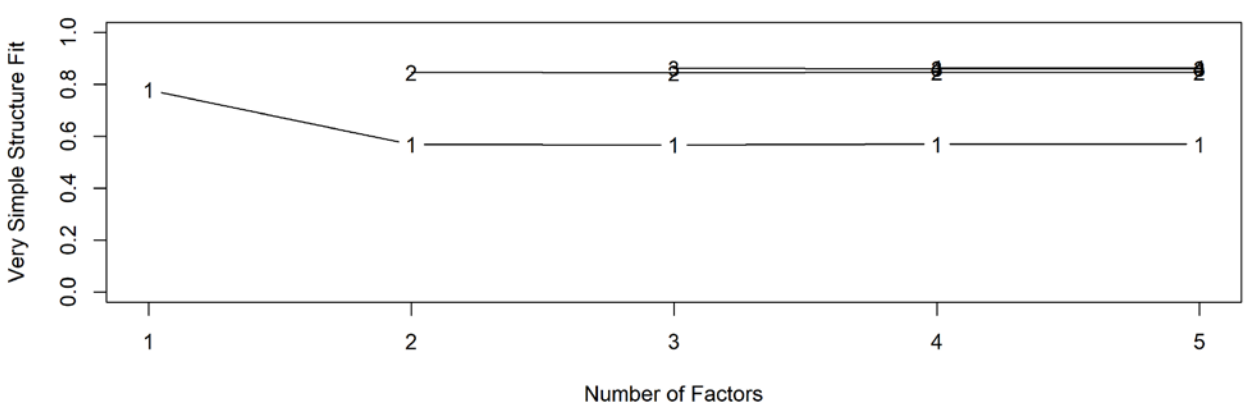



Figure 7

*Very Simple Structure (VSS) Plot*



The VSS method indicated different numbers of factors for different levels of complexity in the factor model. The VSS plot displays the fit results for each level of complexity. The highest value (highest line on the plot) implied an easier model for interpretation. In Figure 7, the VSS plot suggested a two-factor model by the highest lines. The output of the VSS analysis indicated a .78 value with a one-factor solution and .85 value with a two-factor solution. The Velicer's MAP criterion suggested a one-factor model to achieve a minimum of .09 and the BIC (Bayesian Information Criterion) minimum suggested a two-factor model.

Figure 8

*BassAckward Plot of One Factor Solution*

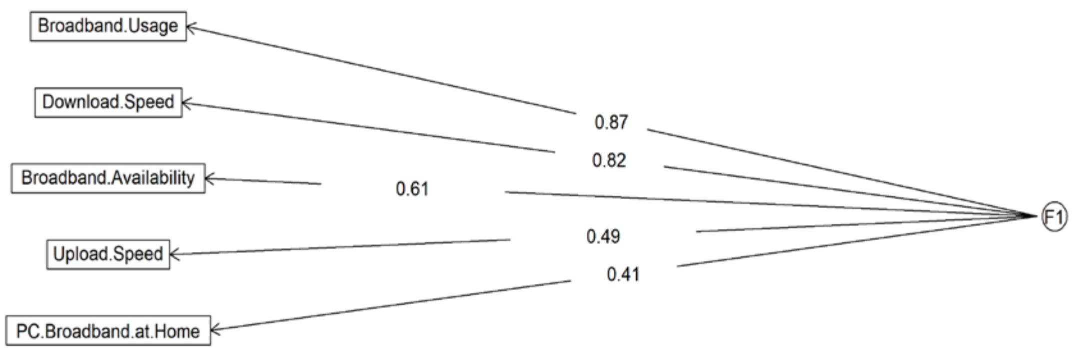
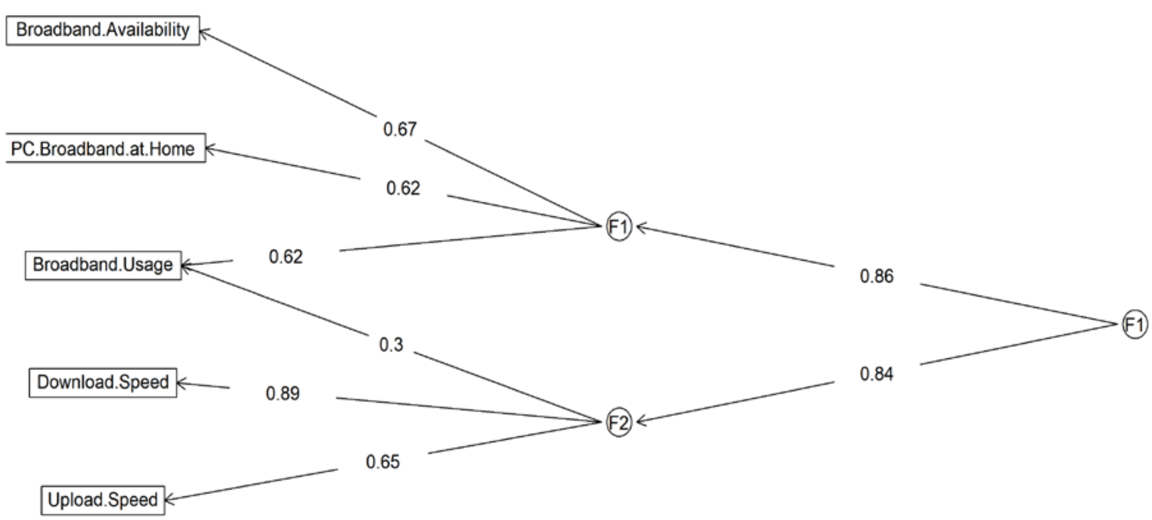



Figure 9

*BassAckward Plot of Two Factor Solution*



Due to the inconsistency in the suggested optimal factor solution at this stage, other approaches were used to inform decision-making in the process of factor analysis. The Bass-Ackward factoring algorithm (Goldberg, 2006) is a top-down technique to evaluate the structure of the selected variables. Two plots were generated to examine the difference between a one-factor model vs. a two-factor model. Figure 8 and 9 displays how the different variables were grouped to form the hypothesized solutions.

While the two-factor solutions may have looked like a hierarchical factor structure, Goldberg (2006) stated that the representation should be viewed as sequential and not hierarchical. By comparing the two structures (Figure 8 and 9), the broadband usage variable appeared to contribute differently to each model. Broadband usage was split by portions in a two-factor model to the speed quality variables (download and upload speeds) versus the broadband availability and device ownership variables. After careful consideration of these suggested solutions, a one-factor solution was ideal for the SDO construct development. The different factor structures suggested by these initial analyses are worthy of further consideration in terms of conceptual or theoretical development of their representation within digital connectivity. This is further discussed in Chapter 5.

Table 4

*Factor Loading, Communality, Uniqueness of the SDO Factor Model*

| Variables | Factor Loading | Communality ($h^2$) | Unique Variance ($u^2$) | Cronbach Coefficient Alpha |
|---|---|---|---|---|
| PC and Broadband at Home | .41 | .17 | .83 | .79 |
| Broadband Availability | .61 | .37 | .63 | .73 |
| Broadband Usage | .87 | .76 | .24 | .66 |
| Download Speed | .83 | .69 | .31 | .67 |
| Upload Speed | .49 | .24 | .76 | .77 |

*Note*. There are 3,138 counties in the sample. Principal axis factoring is used as the extraction method.

The principal axis factoring method was used to obtain a one-factor model. The factors loading, communality, and uniqueness are displayed in Table 4. The one-factor solution has a 44.5% total variance explained, which is the overall effect size of the SDO model. The underlying factor structure is displayed in Figure 10. During the factor extraction process, rotation did not improve the interpretation of the factor structure or factor loading. Figure 11 displays the dimensional plot of factor loading for each variable without rotations.

Figure 10

*Factor Analysis Structure Diagram*

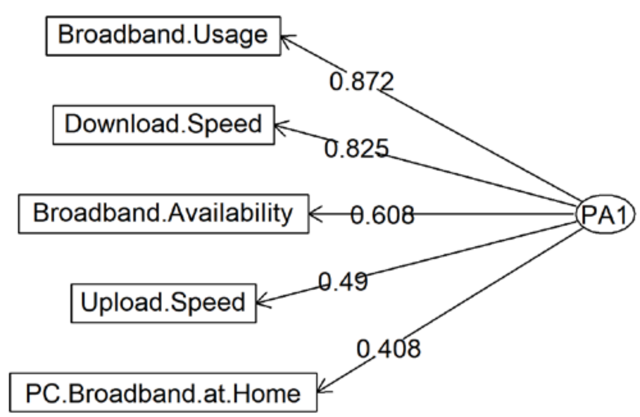
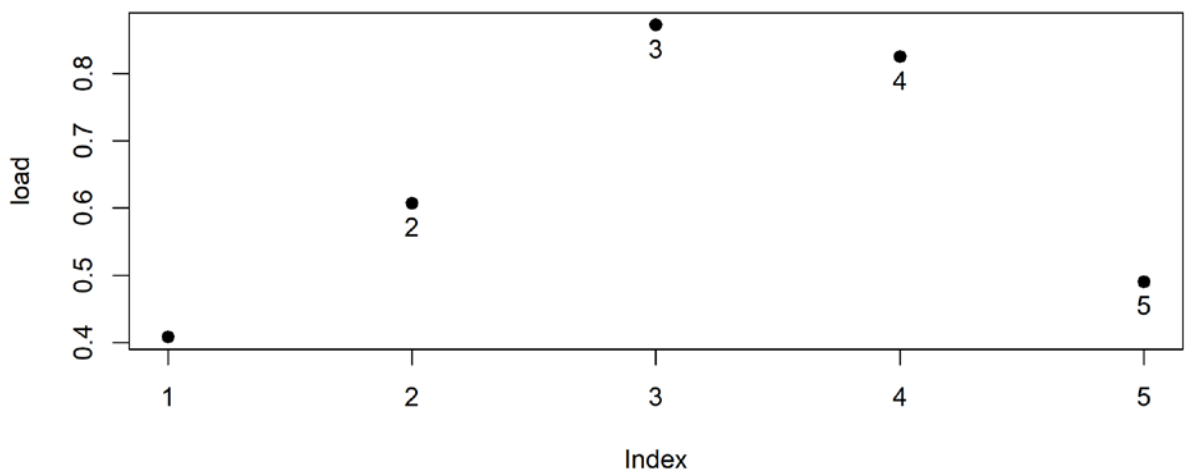



Figure 11

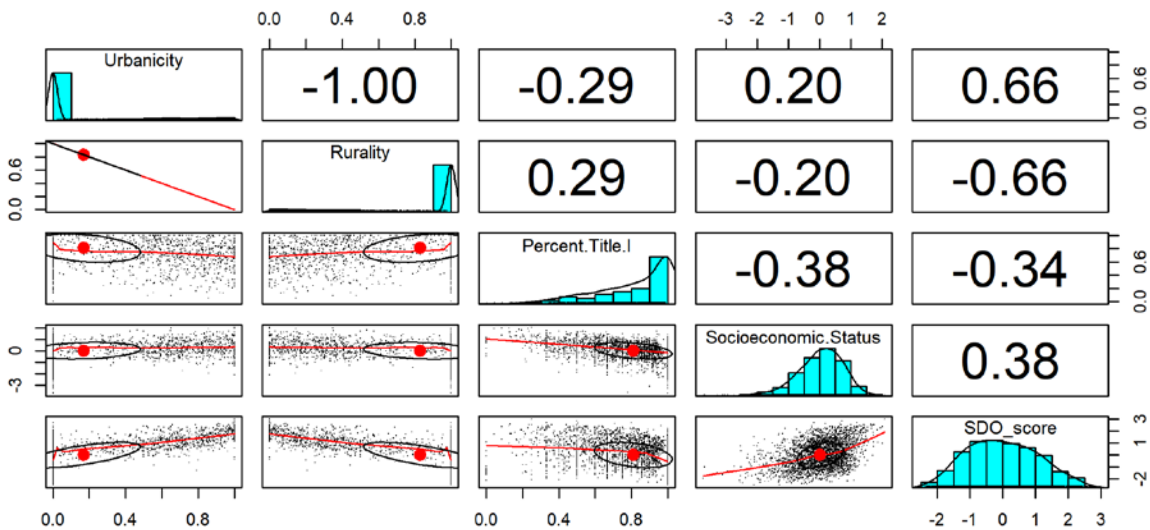
*Dimensional Plot of Factor Loadings*



Reliability analysis was conducted with the alpha function in the *psych* package. The reliability of each variable determined by using the Cronbach Coefficient Alpha method is displayed in Table 4. An alpha coefficient estimate of .77 was obtained for the overall model. Previous guidelines were used to verify if the one-factor solution obtained is adequate (Hogarty et al., 2005, Young & Pearce, 2013), in which case, the factor loadings of each variable should at least be .40. The factor loadings of Upload Speed and PC with Broadband at Home variables are at the borderline, estimated at .49 and .40, respectively. Therefore, further consideration of the nature of these two variables in the SDO model will help to understand the borderline estimated values observed. Their estimated communalities are considered low ($h^2 < .20$) and the unique variances are fairly high ($u^2 > .75$). This indicates their specific variance and error variance of each variable make up a higher portion of its own total variance. The last step was to obtain a standardized factor score from the analysis. The factor scores were derived with the Bartlett method (1937) embedded in the *psych* package *fa* function for each county. The use of the Bartlett method was to minimize the sum of squares for the unique factors across the selected variables (Grice, 2001) in the SDO construct. The standardized factor score provided a numerical value for each county in the SDO composite measure. The SDO measure has a mean of zero and a standard deviation of 1.07 with values ranging from -2.54 to 2.99. To interpret the SDO measure, the mean represents the average digital opportunity for K-12 students across the 3,138 counties. SDO values above the mean indicated an above-average digital opportunity, whereas SDO values below the mean indicated a below-average digital opportunity for students in that region. The estimated SDO scores along with socioeconomic status, urbanicity, rurality, and the Title I status were evaluated by creating a pair panel figure (Figure 12). The SDO measure displays a moderate inverse relationship with rurality ($r$ = - .66) and a positively moderate relationship with urbanicity ($r$ = .66). However, the relationship between socio-economic status (SES) and the SDO values are positively weak ($r$ = .38). Last, the Title I status of schools is inversely correlated with the SDO values ($r$ = -.34) meaning that if a region has a higher percentage of Title I schools, the SDO values tend to decrease. After reviewing all the correlation pairs, the geographic factors indicated by rurality and urbanicity appeared to have a strong relationship than the SES and Title I status variables.

Figure 12

*Correlation and Bivariate Plots of SDO and Other Variables*




This observation is consistent with our current assumption of how digital infrastructure varied largely between urban and rural areas in our country. This will be further investigated in the spatial analysis of this study. An interactive map has been generated with the SDO scores across the country on Tableau (Figure 13). A summary table of the SDO scores for each state is provided in Table 5. Due to the interest in how rurality by school at the county level may be associated with the variation of SDO values across the U.S., an interactive tool for user interaction (Figure 14) allows users to interact with this variable.

Figure 13

*Distribution of SDO Values Across the Main Continent in U.S.*

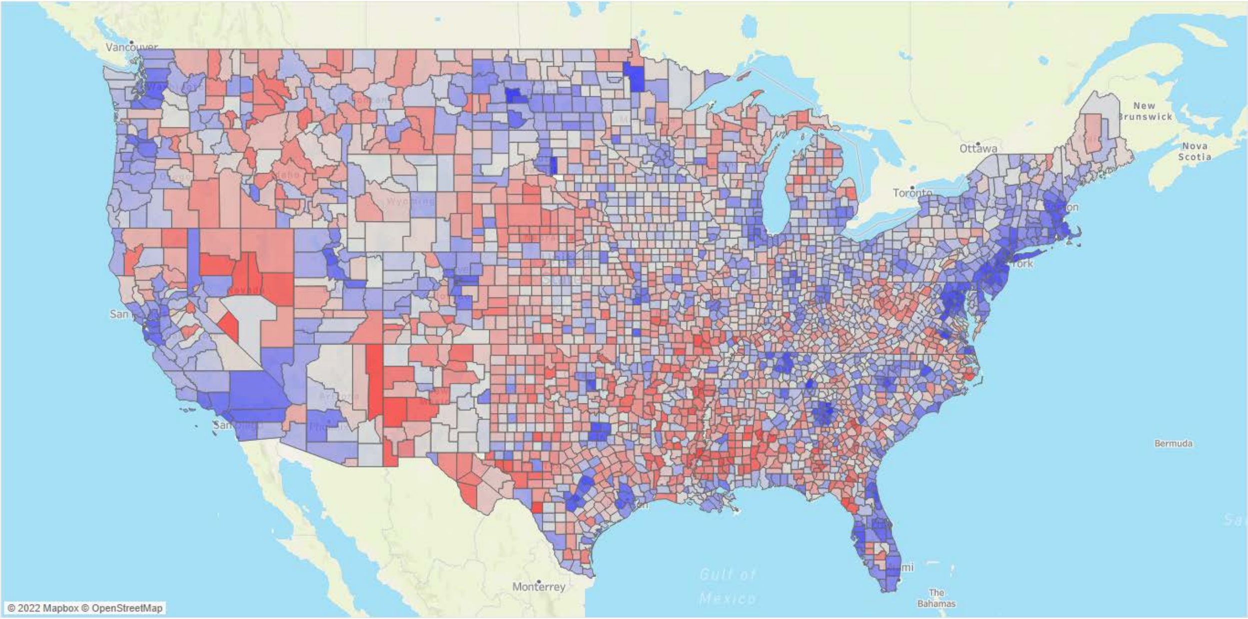



Figure 14

*Interactive Tool to Examine Rurality in Relation to SDO Values*

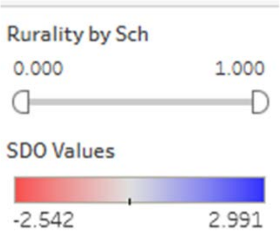



Table 5

*Minimum and Maximum Range of SDO Values by State*

| State Name | Min. of SDO | Max. of SDO | Mean of SDO | SD of SDO |
|---|---|---|---|---|
| Alaska | -2.30 | 1.94 | -0.18 | 2.99 |
| Alabama | -2.36 | 1.93 | -0.22 | 3.04 |
| Arkansas | -2.21 | 1.71 | -0.25 | 2.78 |
| Arizona | -2.31 | 1.58 | -0.37 | 2.76 |

*(table continues)*

| State Name | Min. of SDO | Max. of SDO | Mean of SDO | SD of SDO |
|---|---|---|---|---|
| California | -1.76 | 2.05 | 0.14 | 2.69 |
| Colorado | -1.82 | 2.43 | 0.30 | 3.00 |
| Connecticut | 0.93 | 1.68 | 1.30 | 0.53 |
| District of Columbia* | 1.72 | 1.72 | 1.72 | 0.00 |
| Delaware | 1.41 | 2.11 | 1.76 | 0.50 |
| Florida | -2.27 | 2.33 | 0.03 | 3.25 |
| Georgia | -2.52 | 2.68 | 0.08 | 3.68 |
| Hawaii | 0.82 | 2.04 | 1.43 | 0.86 |
| Iowa | -1.14 | 1.50 | 0.18 | 1.87 |
| Idaho | -1.95 | 1.62 | -0.17 | 2.52 |
| Illinois | -2.31 | 2.03 | -0.14 | 3.07 |
| Indiana | -1.95 | 1.92 | -0.01 | 2.74 |
| Kansas | -1.80 | 2.07 | 0.14 | 2.74 |
| Kentucky | -2.01 | 1.57 | -0.22 | 2.53 |
| Louisiana | -2.54 | 1.69 | -0.43 | 2.99 |
| Massachusetts | 0.78 | 2.29 | 1.53 | 1.07 |
| Maryland | -0.09 | 2.54 | 1.22 | 1.86 |
| Maine | -0.83 | 1.03 | 0.10 | 1.32 |
| Michigan | -1.78 | 1.72 | -0.03 | 2.48 |
| Montana | -1.45 | 2.46 | 0.51 | 2.76 |
| Missouri | -2.34 | 1.85 | -0.24 | 2.96 |
| Mississippi | -2.39 | 1.79 | -0.30 | 2.95 |
| Montana | -1.82 | 1.12 | -0.35 | 2.07 |
| North Carolina | -2.18 | 2.31 | 0.07 | 3.17 |
| North Dakota | -1.32 | 2.69 | 0.69 | 2.83 |
| Nebraska | -1.68 | 1.71 | 0.02 | 2.40 |
| New Hampshire | -0.05 | 2.20 | 1.08 | 1.59 |
| New Jersey | 0.68 | 2.42 | 1.55 | 1.23 |
| New Mexico | -2.26 | 1.50 | -0.38 | 2.65 |
| Nevada | -2.27 | 1.65 | -0.31 | 2.77 |
| New York | -0.36 | 2.42 | 1.03 | 1.96 |
| Ohio | -1.56 | 1.91 | 0.17 | 2.45 |
| Oklahoma | -1.98 | 2.33 | 0.17 | 3.05 |

*(table continues)*

| State Name | Min. of SDO | Max. of SDO | Mean of SDO | SD of SDO |
|---|---|---|---|---|
| Oregon | -1.52 | 1.95 | 0.21 | 2.45 |
| Pennsylvania | -1.41 | 2.38 | 0.49 | 2.68 |
| Rhode Island | 1.27 | 2.05 | 1.66 | 0.56 |
| South Carolina | -1.73 | 1.72 | -0.01 | 2.44 |
| South Dakota | -1.82 | 2.53 | 0.36 | 3.08 |
| Tennessee | -1.29 | 2.64 | 0.67 | 2.78 |
| Texas | -2.45 | 2.41 | -0.02 | 3.44 |
| Utah | -1.42 | 2.07 | 0.33 | 2.47 |
| Virginia | -1.88 | 2.99 | 0.55 | 3.45 |
| Vermont | -1.33 | 1.43 | 0.05 | 1.95 |
| Washington | -1.67 | 1.99 | 0.16 | 2.59 |
| Wisconsin | -1.36 | 1.59 | 0.12 | 2.08 |
| West Virginia | -1.87 | 1.74 | -0.07 | 2.56 |
| Wyoming | -1.37 | 1.11 | -0.13 | 1.75 |

## 4.5 Spatial Analysis

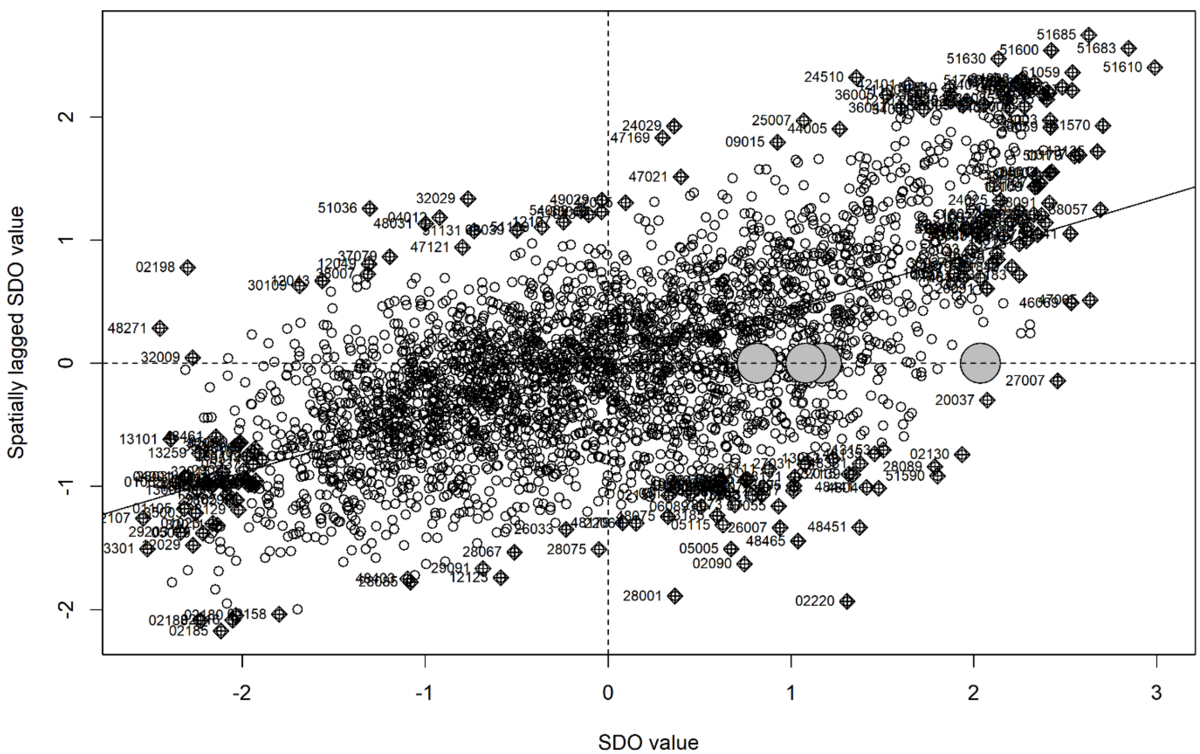
From the previous analysis, a standardized SDO score has been obtained for each county (*n* = 3,138) where K-12 public schools were in operation during the 2019-2020 school year. While it is informative to see how the Students' Digital Opportunity (SDO) scores are distributed across the country (Figure 13) on a map, it is uncertain how geo-spatial characteristics play a role in its (SDO value) variation at the county level. For example, a moderately negative correlation is observed between the SDO values with the degree of rurality. The current assumption is that a majority of rural areas are facing greater challenges in accessing reliable internet services and digital technologies. However, is it true that all rural areas are experiencing the same kinds of disparity in terms of their lack of digital opportunities? Investigating how the SDO values are distributed with their geographic location will help us better understand the challenges in different parts of our country. This information will greatly influence our current understanding of digital equity for the student population in rural areas because their digital opportunities are dependent upon the physical and social infrastructures (i.e. school) in their immediate environment. Spatial autocorrelation with Moran's *I* statistics were used for this part of the analysis.

There were several data processing steps required before conducting the spatial analysis. The use of spatial techniques required the use of shapefiles at the county level. The legal boundaries of each administrative unit that is considered as U.S. County are defined in the U.S. Census shapefiles. The shapefiles contain information about location, geometric information, polygon shapes, and attributes of the geographic features. The U.S. Census TIGER/Line Shapefiles (2019b) was used in this study. The shapefiles were joined to the dataset which included the SDO values estimated from the previous analysis. Several spatial packages in R programming were used for this part of the analysis: *sp* (Pebesma & Bivand, 2021), *sf* (Pebesma, 2021), *tigris* (Walker, 2021), *spdep* (Bivand, 2021), and *choroplethr* (Lamstein, 2020).

During the spatial autocorrelation analysis, several outputs were obtained: global Moran's *I* statistics, local Moran's *I* statistics, *p*-values for both levels, and quadrant information. The spatial autocorrelation of Moran's *I* supported the evaluation of spatial dependence (location dependence) of the estimated SDO scores across all regions. The global Moran's *I* statistical analysis determined if clusters are present across the counties. A significant result ($p$-value of $< .001$) of the global Moran's *I* indicated the presence of clusters. This result is consistent with our assumption that broadband deployment is not a randomized event. A scatterplot was generated with the *moran.plot* function in the *spdep* package. This plot displays the spatial data against its spatially lagged values in a linear fashion (Bivand, 2021). The spatial lag value is a product of multiplying the contiguity-based spatial weight by the average SDO values among neighboring counties. The neighboring counties were those that shared at least one edge with a specific county. The contiguity-based spatial weights in this analysis were estimated by using the queen criterion as recommended by Anselin and Morrison (2019). Figure 15 displays the scatterplot of SDO values and their corresponding spatially lagged values by the four quadrants. The four quadrants represent this association between positive and negative ranges in two dimensions (x-axis for SDO values and y-axis for spatially lagged SDO values) for each county. The top left and bottom right quadrants represent the dispersion counties, or those that are next to counties with dissimilar SDO values. The bottom left and top right quadrants represented the cluster counties with neighboring counties with similar SDO values, e.g., above average SDO values surrounded by above average SDO values (High-High).

Figure 15

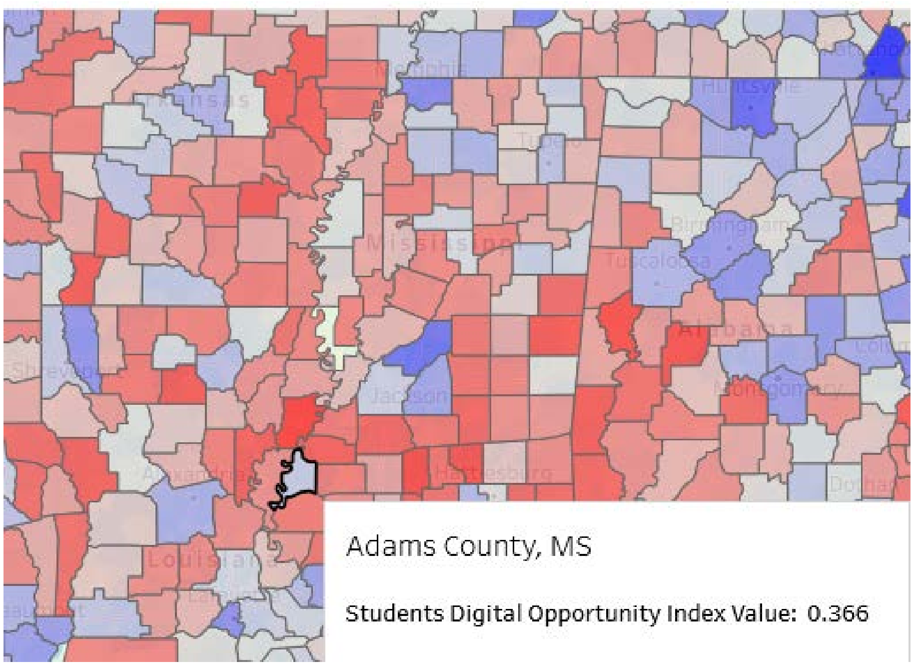
*Association Between SDO Values and Their Spatially Lagged Values*



For example, County ID 28001 (Adams County, MS) is estimated to have an above-average SDO value (Figure 15) and is surrounded by below-average SDO values counties indicated by the below 0 position to the y-axis. This can be visually verified by examining the SDO distribution map (Figure 13). A screenshot (Figure 16) of the regions surrounding Adams County, which is located on the state boundaries between Louisiana and Mississippi is displayed. This figure confirmed the characteristics observed by the autocorrelation calculation  as local Moran's *I* coefficients displayed in the scatterplot quadrants (Figure 15).

Figure 16

*Adams County in Mississippi and Surrounding Counties*




During the estimation of the global Moran's *I* statistics, the local Moran's *I* coefficients were also estimated for each county along with the associated *p*-values. Thus, the global Moran's *I* value is influenced by the variation of the local Moran's *I* values. The local Moran's *I* coefficients obtained fall between a range of -2.19 to 6.31 with a mean of .45 and standard deviation of .87. A positive local Moran's *I* value indicates a positive autocorrelation. This means that a specific county is clustered by counties with similar characteristics, either a High-High or Low-Low SDO combination. The negative local Moran's *I* value indicates a negative autocorrelation where a county has neighbors of dissimilar SDO scores (i.e., High-Low or Low-High). Table 6 provides a summary of the range of local Moran's *I* value for each state.

Table 6

*Minimum and Maximum Range of Local Moran's* I *Coefficients by State*

| State Name | Minimum of Local Moran's *I* Coefficients | Maximum of Local Moran's *I* Coefficients |
|---|---|---|
| Alaska | -2.19 | 4.04 |
| Alabama | -0.62 | 2.57 |
| Arkansas | -0.88 | 3.12 |
| Arizona | -0.94 | 1.85 |
| California | -0.58 | 3.09 |
| Colorado | -0.47 | 3.27 |
| Connecticut | 1.44 | 2.76 |
| District of Columbia | 3.30 | - |
| Delaware | 1.78 | 3.14 |
| Florida | -0.91 | 3.10 |
| Georgia | -0.92 | 4.42 |
| Hawaii | 0.00 | 0.00 |
| Iowa | -0.51 | 0.43 |
| Idaho | -0.83 | 0.99 |
| Illinois | -0.46 | 2.95 |
| Indiana | -0.55 | 1.46 |
| Kansas | -0.70 | 2.23 |
| Kentucky | -0.54 | 1.67 |
| Louisiana | -1.03 | 2.78 |
| Massachusetts | 0.70 | 3.70 |
| Maryland | -0.25 | 4.88 |
| Maine | -0.14 | 0.89 |
| Michigan | -1.09 | 1.73 |

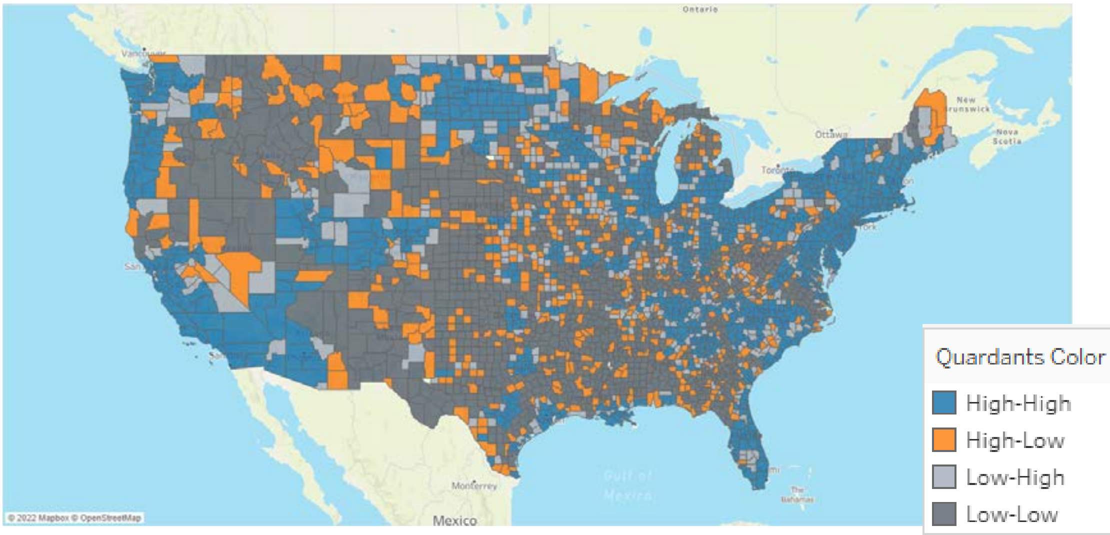
*(table continues)*

| State Name | Minimum of Local Moran's *I* Coefficients | Maximum of Local Moran's *I* Coefficients |
|---|---|---|
| Montana | -0.76 | 1.17 |
| Missouri | -0.51 | 2.83 |
| Mississippi | -1.30 | 3.68 |
| Montana | -0.92 | 1.64 |
| North Carolina | -1.05 | 3.12 |
| North Dakota | -0.83 | 3.10 |
| Nebraska | -0.66 | 2.02 |
| New Hampshire | -0.04 | 3.15 |
| New Jersey | 0.89 | 4.54 |
| New Mexico | -0.84 | 2.38 |
| Nevada | -0.89 | 2.44 |
| New York | -0.21 | 4.06 |
| Ohio | -0.42 | 1.77 |
| Oklahoma | -0.54 | 2.44 |
| Oregon | -0.22 | 1.76 |
| Pennsylvania | -0.42 | 3.75 |
| Rhode Island | 2.09 | 3.15 |
| South Carolina | -0.70 | 1.12 |
| South Dakota | -0.93 | 1.65 |
| Tennessee | -0.65 | 2.47 |
| Texas | -1.59 | 3.14 |
| Utah | -0.35 | 1.96 |
| Virginia | -1.43 | 6.31 |
| Vermont | -0.20 | 0.64 |
| Washington | -0.39 | 1.95 |
| Wisconsin | -0.66 | 1.57 |
| West Virginia | -0.54 | 2.09 |
| Wyoming | -0.16 | 1.07 |

*Note*. There is not a range of values for District of Columbia.

Figure 17

*Quadrants of the SDO Distribution*




However, without using the local Moran's *I* statistics and *p*-values together, it is possible to misinterpret the information as to whether these are true clusters or outliers. Using the *p*-values of less than .05 as a threshold along with the local Moran's *I* information, we can determine if the clustered or dispersed regions are indeed significant. In other words, by using the local indicator of spatial association (local Moran's *I*), we can assess if the cluster patterns observed were not happened by chance. Figure 17 displays the interactive map and Figure 18 displays the clusters and outliers regions after a .05 *p*-value threshold was applied to the interactive map.

Although the map distinguishes the difference between cluster or dispersion regions, there are differences within each cluster. A High-High cluster represents counties with high SDO values such as the Emmons County in North Dakota (Figure 19). A Low-Low cluster represents counties with low SDO values such as the Cherry County in Nebraska (Figure 20). Therefore, a marked map is produced where clusters with High-High or Low-Low values are identified (Figure 21). We can see that there is a high concentration of Low-Low regions in the South and

Southwest parts of the United States along with few cluster regions in the Midwest. The High-High regions tend to cluster around major metropolitan areas in the United States, except the clusters in North Dakota and northern part of South Dakota.

Figure 18

*Clusters and Outliers with* p-*values of .50 or Below (Interactive Map)*

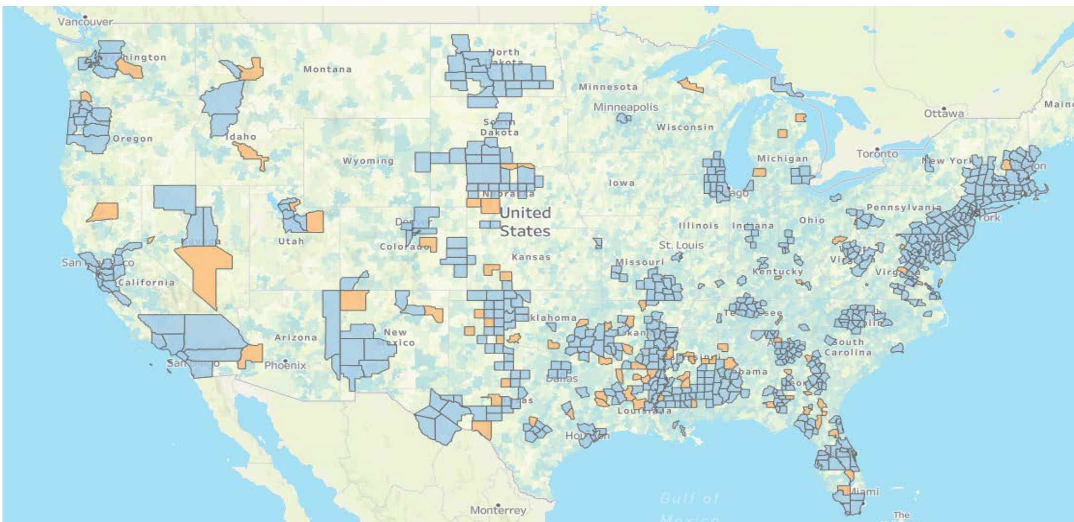



Figure 19

*Example of Clusters in High-High Quadrants, Emmons County in North Dakota*

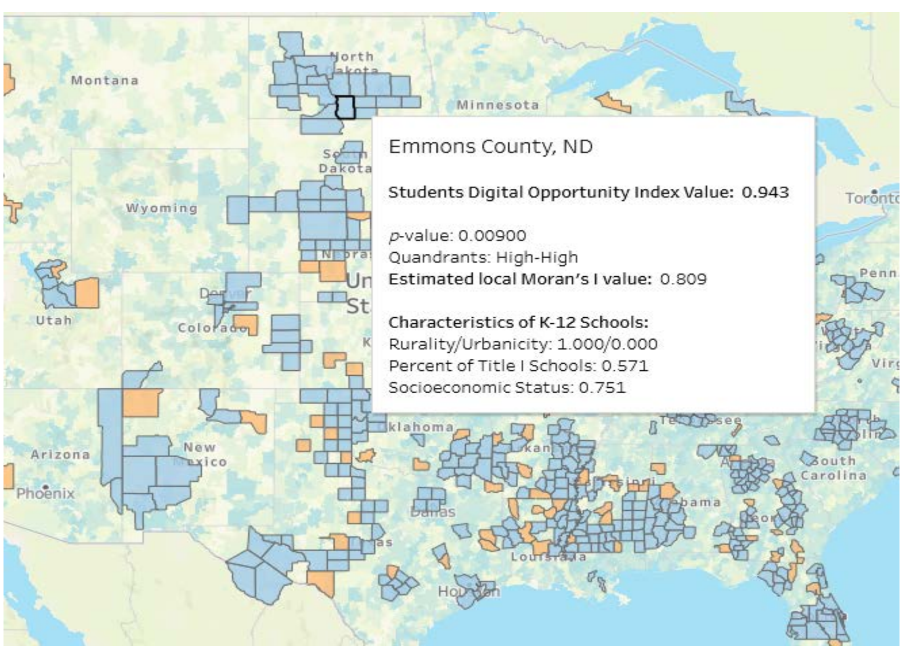


Figure 20

*Example of Clusters in Low-Low Quadrants, Cherry County in Nebraska*

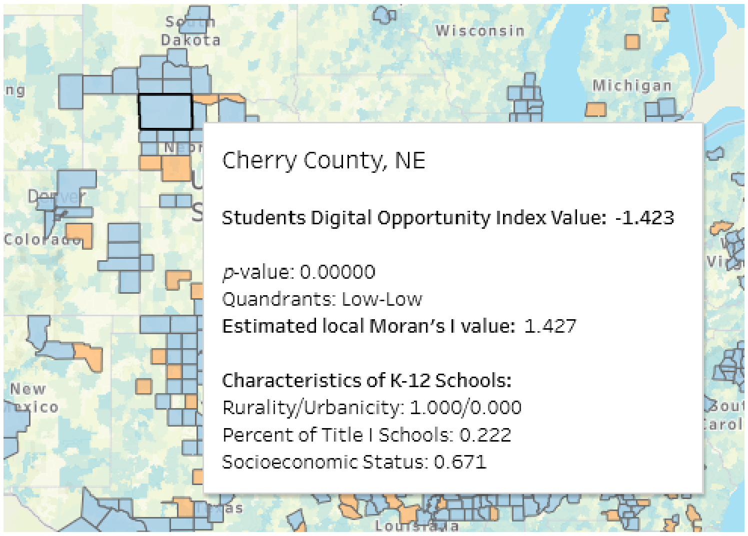



Figure 21

*Significant Clusters of High-High or Low-Low Quadrants*

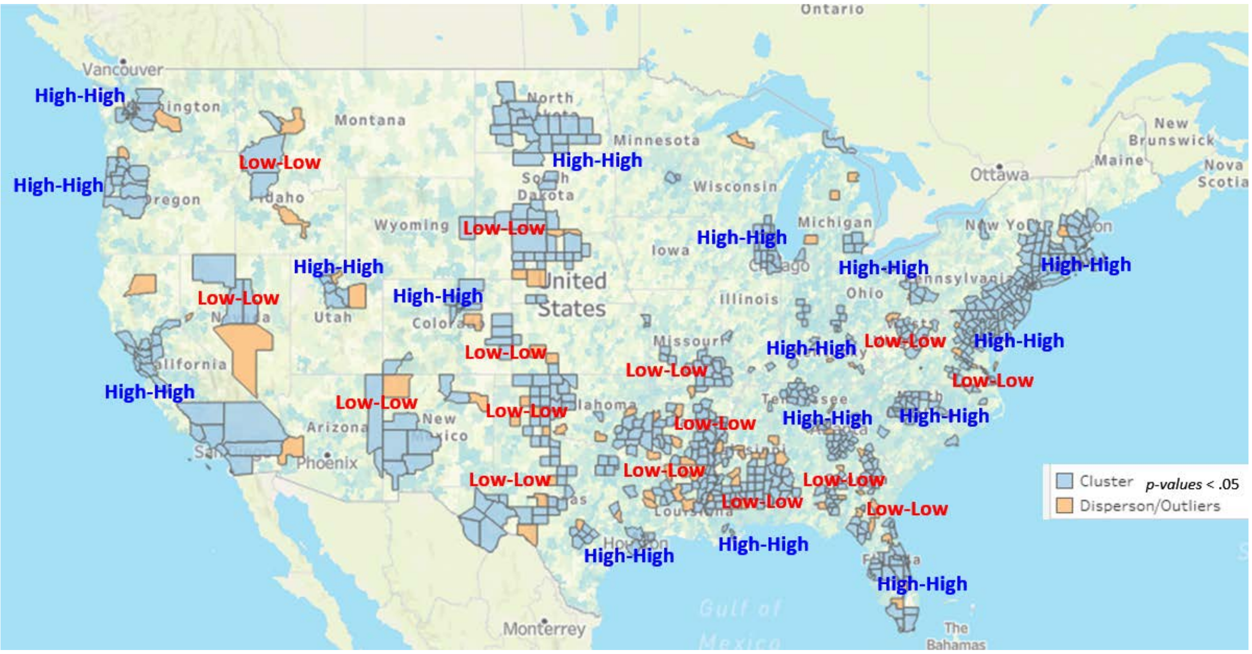



Table 7 was produced to allow a comparison of numerical values with the marked map displays in Figure 21. There are 690 counties across all 50 states and D.C. identified as clusters or outliers by using the less than .05 *p*-values threshold of the local Moran's *I* coefficients. Table

7 displays the number of counties in each state as significant clusters or outliers produced by the spatial analysis.

Table 7

*U.S. Counties as Significant Clusters or Outliers by Count for Each State*

| State Name | No. of Cluster Counties | No. of Outlier Counties | No. of High-High Quadrant Counties | No. of Low-Low Quadrant Counties |
|---|---|---|---|---|
| Alaska | 10 | 2 | 0 | 10 |
| Alabama | 14 | 3 | 0 | 14 |
| Arkansas | 29 | 5 | 0 | 29 |
| Arizona | 2 | 1 | 0 | 2 |
| California | 19 | 1 | 19 | 0 |
| Colorado | 12 | 1 | 6 | 6 |
| Connecticut | 8 | 0 | 8 | 0 |
| District of Columbia | 1 | 0 | 1 | 0 |
| Delaware | 3 | 0 | 3 | 0 |
| Florida | 25 | 4 | 19 | 6 |
| Georgia | 54 | 6 | 23 | 31 |
| Hawaii | 0 | 0 | 0 | 0 |
| Iowa | 0 | 0 | 0 | 0 |
| Idaho | 2 | 1 | 0 | 2 |
| Illinois | 10 | 0 | 9 | 1 |
| Indiana | 5 | 0 | 5 | 0 |
| Kansas | 3 | 2 | 1 | 2 |
| Kentucky | 7 | 2 | 3 | 4 |
| Louisiana | 15 | 6 | 1 | 14 |
| Massachusetts | 12 | 0 | 12 | 0 |
| Maryland | 18 | 0 | 18 | 0 |
| Maine | 1 | 0 | 1 | 0 |
| Michigan | 6 | 4 | 5 | 1 |
| Montana | 2 | 0 | 2 | 0 |
| Missouri | 19 | 0 | 3 | 16 |

*(table continues)*

| State Name | No. of Cluster Counties | No. of Outlier Counties | No. of High-High Quadrant Counties | No. of Low-Low Quadrant Counties |
|---|---|---|---|---|
| Mississippi | 30 | 9 | 0 | 30 |
| Montana | 0 | 1 | 0 | 0 |
| North Carolina | 21 | 0 | 17 | 4 |
| North Dakota | 18 | 0 | 18 | 0 |
| Nebraska | 15 | 3 | 0 | 15 |
| New Hampshire | 6 | 1 | 6 | 0 |
| New Jersey | 21 | 0 | 21 | 0 |
| New Mexico | 7 | 2 | 0 | 7 |
| Nevada | 3 | 2 | 0 | 3 |
| New York | 20 | 0 | 20 | 0 |
| Ohio | 5 | 1 | 5 | 0 |
| Oklahoma | 20 | 3 | 1 | 19 |
| Oregon | 8 | 1 | 8 | 0 |
| Pennsylvania | 17 | 0 | 17 | 0 |
| Rhode Island | 5 | 0 | 5 | 0 |
| South Carolina | 1 | 0 | 1 | 0 |
| South Dakota | 9 | 0 | 2 | 7 |
| Tennessee | 14 | 1 | 13 | 1 |
| Texas | 47 | 11 | 15 | 32 |
| Utah | 5 | 2 | 5 | 0 |
| Virginia | 37 | 4 | 33 | 4 |
| Vermont | 1 | 0 | 1 | 0 |
| Washington | 4 | 0 | 4 | 0 |
| Wisconsin | 6 | 0 | 6 | 0 |
| West Virginia | 12 | 1 | 0 | 12 |
| Wyoming | 1 | 0 | 0 | 1 |

## 4.6 Summary

This chapter provided the quantitative results of the study and the rationale during the data analysis process. Based on the results, it was determined that digital connectivity in terms of digital opportunities varied across the states and counties in the U.S. The geographical influences are more pronounced in certain states, especially regions in the Southern part of the U.S. in comparison to other county-level regions in the country. This evidence is an important resource to further consider how the physical environment, as well as the socioeconomic and demographic factors at play, in regions that are marked as High-High vs. Low-Low clusters in this study. The variation within each cluster could be due to other underlying factors that were not considered in the current model. The following chapter will further discuss the result of this study and the limitations of this work.

# CHAPTER 5

# DISCUSSION AND CONCLUSION

## 5.1 Introduction

This study examined digital opportunity among K-12 public school students in the United States and is the first to evaluate it at the county level across country. A conceptual model was developed to represent the key requirements of broadband internet connection and device ownership before students can participate in digital learning. Based on this conceptual model, a composite measure named the "Students' Digital Opportunity (SDO)" was developed to allow evaluation of this phenomenon. A spatial analysis of the SDO distribution by location further improved the robustness of this measure in providing evidence of how geographical features contribute to the unequal digital opportunities in our country. This work is also intended to serve as a methodological example that differs from the historical way of assessing the digital divide. The use of open data at the population level with a data science approach to investigate a socio-technical phenomenon is a first for this research topic. While the design of the conceptual model was informed by previous literature, the analytical process used a data-driven approach which is another unique feature of this research for the education discipline. The conceptual model of this work is meant to provide an integrated view of the key components of digital equity from the physical dimension of access and use. The SDO measure allow comparison of differences between U.S. counties. This study's research process and data products provide evidence and analytical framework for future research of digital opportunity across locations within the country.

## 5.2 Discussion of the Results

### 5.2.1 Research Question 1

*How does students' digital opportunity (SDO) distribute across the United States at the county and state level?*

To address the first research question, there were several steps in the research process. First, a concept model (SDO) was established with the support of literature and the proposed variables were defined for construct development. Then, the defined components in the SDO model had to align to the proxy variables in the data sources that best represented each component at the county level. Each record in the final dataset represents a county in the United States during the 2019-2020 school year where at least one public school was in operation. There were 3,138 counties included in this study, where four counties (as described in Chapter 4.2) did not have any K-12 schools in operation within their legal boundaries during the 2019-2020 school year. For each county record, the schools' locale categories were converted to a percentage of urbanicity and a percentage of rurality to represent the degree of how urban or rural each county is. Due to the different sizes and population densities of each county, a large county may not necessarily be highly populated. By using percentages of the locale categories as urbanicity or rurality, it reflects the type of area in which the schools are located. The number of Title I schools is also converted to a percentage for each county to represent the proportion of schools with children from low-income families (as defined by Title I) in that area. Currently, any schools with at least 40 percent of children enrolled from low-income families are eligible for the Title I status (U.S. Department of Education, 2018). This designation qualifies the schools to receive certain types of federal aid to support their students' learning and to raise student achievement. The socioeconomic status (SES) composite from the SEDA 4.1 dataset represents the characteristics of the total population at each county, where income, educational attainment, household characteristics, unemployment rate, and social benefits are included in its estimation. The SES composite provides information about the population characteristics at each county including families who do not have K-12 students in their household.

The interactive map of the SDO distribution (Figure 13) displays the county-level variation across the United States. The counties with below-average SDO scores are located mostly in the central part of the country and spread in two directions toward the south and west. By visually examining it at the state level, it is apparent that several states have many counties scoring below average within their boundaries: Mississippi, Arkansas, Oklahoma, Texas, and New Mexico in the South; Idaho, Montana, and Nevada in the West. Table 8 displays the frequency count of counties with below-average SDO scores and above-average SDO scores, and the proportion of below-average counties by percentage within each state. This table is sorted by listing the states with the highest percentage of counties that have below-average SDO scores first.

Table 8

*Distribution of SDO Values by State (Sorted by Percentages)*

| State Abbreviation | No. of Counties by State | No. of Counties with < 0 SDO Values (Below Average) | No. of Counties with > 0 SDO Values (Above Average) | Percentage of Counties Scoring Below Average |
|---|---|---|---|---|
| MS | 81 | 61 | 20 | 75.31 |
| ID | 44 | 33 | 11 | 75.00 |
| AR | 75 | 56 | 19 | 74.67 |
| WV | 55 | 40 | 15 | 72.73 |
| OK | 77 | 54 | 23 | 70.13 |
| NE | 93 | 65 | 28 | 69.89 |
| NM | 33 | 23 | 10 | 69.70 |
| KS | 105 | 71 | 34 | 67.62 |

*(table continues)*

| State Abbreviation | No. of Counties by State | No. of Counties with < 0 SDO Values (Below Average) | No. of Counties with > 0 SDO Values (Above Average) | Percentage of Counties Scoring Below Average |
|---|---|---|---|---|
| MT | 56 | 37 | 19 | 66.07 |
| TX | 253 | 167 | 86 | 66.01 |
| AK | 29 | 19 | 10 | 65.52 |
| MO | 115 | 73 | 42 | 63.48 |
| SD | 65 | 40 | 25 | 61.54 |
| GA | 159 | 95 | 64 | 59.75 |
| KY | 120 | 71 | 49 | 59.17 |
| NV | 17 | 10 | 7 | 58.82 |
| AL | 67 | 39 | 28 | 58.21 |
| IL | 102 | 59 | 43 | 57.84 |
| LA | 64 | 37 | 27 | 57.81 |
| WI | 72 | 41 | 31 | 56.94 |
| WY | 23 | 13 | 10 | 56.52 |
| IA | 99 | 52 | 47 | 52.53 |
| IN | 92 | 48 | 44 | 52.17 |
| CO | 64 | 33 | 31 | 51.56 |
| VT | 14 | 7 | 7 | 50.00 |
| MI | 83 | 39 | 44 | 46.99 |
| AZ | 15 | 7 | 8 | 46.67 |
| SC | 46 | 21 | 25 | 45.65 |
| MN | 87 | 38 | 49 | 43.68 |
| WA | 39 | 16 | 23 | 41.03 |
| TN | 95 | 37 | 58 | 38.95 |
| FL | 67 | 25 | 42 | 37.31 |
| OR | 36 | 13 | 23 | 36.11 |
| UT | 29 | 10 | 19 | 34.48 |
| VA | 133 | 45 | 88 | 33.83 |
| NC | 100 | 33 | 67 | 33.00 |
| PA | 67 | 22 | 45 | 32.84 |
| ME | 16 | 5 | 11 | 31.25 |
| CA | 58 | 17 | 41 | 29.31 |

*(table continues)*

| State Abbreviation | No. of Counties by State | No. of Counties with < 0 SDO Values (Below Average) | No. of Counties with > 0 SDO Values (Above Average) | Percentage of Counties Scoring Below Average |
|---|---|---|---|---|
| OH | 88 | 25 | 63 | 28.41 |
| ND | 53 | 11 | 42 | 20.75 |
| NH | 10 | 2 | 8 | 20.00 |
| NY | 62 | 4 | 58 | 6.45 |
| MD | 24 | 1 | 23 | 4.17 |
| CT | 8 | 0 | 8 | 0.00 |
| DC | 1 | 0 | 1 | 0.00 |
| DE | 3 | 0 | 3 | 0.00 |
| HI | 4 | 0 | 4 | 0.00 |
| MA | 14 | 0 | 14 | 0.00 |
| NJ | 21 | 0 | 21 | 0.00 |
| RI | 5 | 0 | 5 | 0.00 |

*Note*. District of Columbia (DC) is counted as a county and a state.

The observed differences at the state level (Table 10 and 11 in the next section) are similar to other existing report, such as the World Bank Broadband Strategies Toolkit (2022, Figure 6.2) that has identified places with the largest gap of broadband supply and demand in the United States: Mississippi, West Virginia, Alabama, New Mexico, and Arkansas. This indicates possible underlying factors in terms of economic or socio-demographic characteristics that contribute to the differences observed in the general population. The State Educational Technology Directors Association (SETDA) report (Fox, 2019) in 2019 discussed the various funding efforts at the state level, in which Arkansas, Alabama, New Mexico, and West Virginia were mentioned and described. However, there are no details about proposed actions in the state of Mississippi to ensure that their K-12 schools have broadband connectivity and off-campus access for students to continue their learning at home.

#### 5.2.1.1 Interpretation and Use of the SDO Model and Score

Previous to this study, there was not a measurement tool to allow comparison of digital opportunity for the K-12 students population. This has been an ongoing challenge because of the different types of technology or digital infrastructure available at each county or state in the U.S. It is also difficult to compare values from different indicators, such as broadband availability or broadband usage, side-by-side without understanding their interconnected relationships. The students' digital opportunity (SDO) model was created to address these problems and is a holistic way to evaluate all four components in one dimension numerically. The SDO model is also designed with flexibility where the components can be further decomposed at a lower level to assess individual differences when those data are available. The analytical process with factor analysis generated a set of standardized SDO scores (the factor score). The SDO scale can be thought of as a number line where all counties are positioned based on the four components together in the current model. This approach allows comparison on in a unified scale and serves as the base value if comparison over time of the SDO components is desired. The mean of the SDO score represents the current average students' digital opportunity on the SDO scale (mean value is zero). An SDO score below the average (less than zero) in a county indicates that their digital opportunities are lower relative to other places, and these counties should be categorized as priority areas. Counties with above average (higher than zero) SDO score are comparatively in a better position for students digital opportunity. However, the current SDO score does not guarantee that those counties' position are fixed. The current categorization by comparing below- or above- average scores is to provide a baseline in the measurement of the proposed model. Each unit of standard deviation above and below the mean represents the magnitude difference from the average SDO score. The following table (Table 9) displays the mean values of each component in the final factor solution where the standardized factor score is derived. The SDO values is broken down by each standard deviation unit, and the table displays the mean of each component. The value of each variable is unstandardized for easier understanding and interpretation. Figure 22 displays a histogram of the SDO scores distribution (n = 3,138 Counties).

Table 9

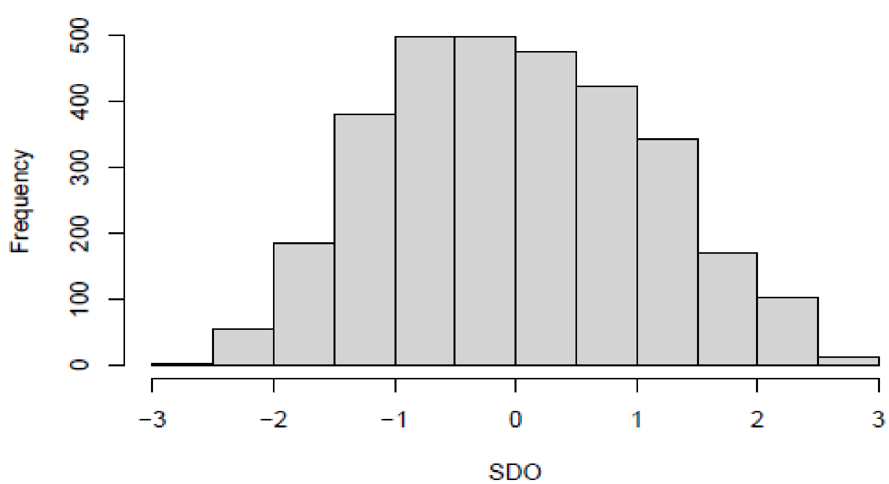
*M. of SDO Model Component by Range of Standard Deviation in SDO Score*

| Variables | Range of Standard Deviations in SDO Score Distribution | | | | | |
|---|---|---|---|---|---|---|
| | < -2 (n = 57) | -2 to -1 (n = 564) | -1 to 0 (n = 994) | 0 to 1 (n = 893) | 1 to 2 (n = 512) | > 2 (n = 118) |
| Broadband Availability (Percentage) | .13 | .63 | .85 | .92 | .96 | .98 |
| Broadband Usage (Percentage) | .05 | .12 | .27 | .49 | .68 | .83 |
| PC and Broadband at Home (Percentage) | .72 | .80 | .85 | .88 | .90 | .93 |
| Download Speed (Mbps) | 14.23 | 30.61 | 56.68 | 88.82 | 123.52 | 155.64 |
| Upload Speed (Mbps) | 5.06 | 9.47 | 17.54 | 27.03 | 36.75 | 65.94 |

*Note*. There are 3,138 U.S. Counties in the sample. Mbps is Megabits per second.

Figure 22

*Distribution of the SDO Scores*



While it is known that K-12 students and schools need access to digital technology and broadband internet, there are no explicit benchmarks or criteria for K-12 schools to set targets for their specific students' needs. By examining the numerical value distribution by each standard deviation in Table 9, we can see that it is necessary to have at least 88 Mbps download speed and 27 Mbps upload speed as the average speed quality. Broadband availability and device ownership would need to be at 90% or above to ensure these are accessible to the K-12 student population. As previously discussed in the factor analysis result (Section 4.4), the broadband usage variable is the major variable that underlies the SDO model. While 49% of broadband usage represents the amount of use by the population in that county, it also represents that this area may utilize broadband internet for various types of activities. For example, if we compare this information to the significant cluster map (Figure 21), by geography, high SDO places are predominately in urban areas. Thus, this reflects that the demand for broadband internet and usage is related to the socio-demographic and economic characteristics in each region. However, the discussion here focuses on K-12 education and not the general population, even though schools and students are affected by their local level environment and resources. So, if we are to ensure all students in K-12 will be provided equal digital opportunities, then we need to strategically increase the four components proposed in the SDO models in places that are currently positioned as below average. In general, the digital challenges in the urban area schools are different from the challenges faced by schools in rural areas. Thus, a one-size-fits all solution will not be appropriate.

If we reflect on how statewide students' performance is measured and evaluated, the SDO model could be applied similarly to allow schools to make comparisons of their digital capacity within themselves. While the goal of evaluating digital opportunity is not related to student performance, this research reflects the need for a reliable and fair approach in determining digital connectivity so equitable solutions and actions can be applied. The reliability of the SDO model is adequate, however, there are some content validity concerns due to the unique variances presented by the upload speed and device ownership components (further discussed in Section 5.3.1). Further understanding of what these components represent and how they vary by different contexts or situations will strengthen the current SDO model and its application.

### 5.2.2 Research Question 2

*Are there any geographical associations with the distribution of the Students' Digital Opportunity (SDO) measure across the United States?*

With the application of spatial autocorrelation, it is apparent that clusters of high SDO and low SDO regions do exist in the United States. In this analysis, as displayed by Figure 18 and Figure 21, most of the significant clusters of high SDO values are present in metropolitan areas across the U.S., especially in major cities along the East Coast and West Coast. A majority of significant clusters of low SDO values are present in the Southern part of the U.S., certain states in the West (Idaho and Nevada), parts of the mid-regions such as Nebraska, and certain parts of Colorado. Table 10 displays the frequency count of the significant clusters and outliers by state, the associated thematic map is displayed in Figure 21.

Table 10

*Significant Clusters and Outliers by State From Spatial Analysis*

| State | No. of Counties with Significant *p*-values ( <.05) | No. of Counties in Clusters of High SDO | No. of Counties in Clusters of Low SDO | No. of Counties as Outliers |
|---|---|---|---|---|
| AK | 12 | 0 | 10 | 2 |
| AL | 17 | 0 | 14 | 3 |

*(table continues)*

| State | No. of Counties with Significant *p*-values ( <.05) | No. of Counties in Clusters of High SDO | No. of Counties in Clusters of Low SDO | No. of Counties as Outliers |
|---|---|---|---|---|
| AR | 34 | 0 | 29 | 5 |
| AZ | 3 | 0 | 2 | 1 |
| CA | 20 | 19 | 0 | 1 |
| CO | 13 | 6 | 6 | 1 |
| CT | 8 | 8 | 0 | 0 |
| DC | 1 | 1 | 0 | 0 |
| DE | 3 | 3 | 0 | 0 |
| FL | 29 | 19 | 6 | 4 |
| GA | 60 | 23 | 31 | 6 |
| HI | 0 | 0 | 0 | 0 |
| IA | 0 | 0 | 0 | 0 |
| ID | 3 | 0 | 2 | 1 |
| IL | 10 | 9 | 1 | 0 |
| IN | 5 | 5 | 0 | 0 |
| KS | 5 | 1 | 2 | 2 |
| KY | 9 | 3 | 4 | 2 |
| LA | 21 | 1 | 14 | 6 |
| MA | 12 | 12 | 0 | 0 |
| MD | 18 | 18 | 0 | 0 |
| ME | 1 | 1 | 0 | 0 |
| MI | 10 | 5 | 1 | 4 |
| MN | 2 | 2 | 0 | 0 |
| MO | 19 | 3 | 16 | 0 |
| MS | 39 | 0 | 30 | 9 |
| MT | 1 | 0 | 0 | 1 |
| NC | 21 | 17 | 4 | 0 |
| ND | 18 | 18 | 0 | 0 |
| NE | 18 | 0 | 15 | 3 |
| NH | 7 | 6 | 0 | 1 |
| NJ | 21 | 21 | 0 | 0 |
| NM | 9 | 0 | 7 | 2 |
| NV | 5 | 0 | 3 | 2 |
| NY | 20 | 20 | 0 | 0 |
| OH | 6 | 5 | 0 | 1 |

*(table continues)*

| State | No. of Counties with Significant *p*-values ( <.05) | No. of Counties in Clusters of High SDO | No. of Counties in Clusters of Low SDO | No. of Counties as Outliers |
|---|---|---|---|---|
| OK | 23 | 1 | 19 | 3 |
| OR | 9 | 8 | 0 | 1 |
| PA | 17 | 17 | 0 | 0 |
| RI | 5 | 5 | 0 | 0 |
| SC | 1 | 1 | 0 | 0 |
| SD | 9 | 2 | 7 | 0 |
| TN | 15 | 13 | 1 | 1 |
| TX | 58 | 15 | 32 | 11 |
| UT | 7 | 5 | 0 | 2 |
| VA | 41 | 33 | 4 | 4 |
| VT | 1 | 1 | 0 | 0 |
| WA | 4 | 4 | 0 | 0 |
| WI | 6 | 6 | 0 | 0 |
| WV | 13 | 0 | 12 | 1 |
| WY | 1 | 0 | 1 | 0 |

A common assumption is the association of rurality with the lack of digital connectivity in the United States. However, by reviewing the thematic map (Figure 21) of significant clusters, not all rural regions are highlighted as low SDO clusters representing a similar level of digital opportunities. For example, North Dakota contains a fair number of rural counties that are not positioned below average in terms of their SDO values. Table 11 displays the significant clusters identified on the thematic map (Figure 21) by degree of rurality from their locale information. This table is sorted by the proportion of rurality in each state. For example, in Arizona, there are two counties displayed as low SDO clusters and one county as an outlier. The percentage of rurality of these three counties on average is 100% (1.00 as a proportion in decimal form in Table 11). While the table only displays the statistically significant regions, keep in mind that each state varies in terms of their geography, infrastructure, and population characteristics in general. Some of these factors may play a role in the observed SDO scores and how they are distributed at the county level in each state.

Table 11

*Significant Clusters with Degree of Rurality by State (Sorted by Rurality)*

| State | No. of Counties with Significant *p*-values (<.05) | Average Rurality of the Significant Clusters (as a Proportion of 1.00) | No. of Counties in Clusters of Low SDO | No. of Counties in Clusters of High SDO | No. of Counties as Outliers |
|---|---|---|---|---|---|
| AZ | 3 | 1.00 | 2 | 0 | 1 |
| ID | 3 | 1.00 | 2 | 0 | 1 |
| NE | 18 | 1.00 | 15 | 0 | 3 |
| NV | 5 | 1.00 | 3 | 0 | 2 |
| SC | 1 | 1.00 | 0 | 1 | 0 |
| SD | 9 | 1.00 | 7 | 2 | 0 |
| VT | 1 | 1.00 | 0 | 1 | 0 |
| WY | 1 | 1.00 | 1 | 0 | 0 |
| MS | 39 | 0.97 | 30 | 0 | 9 |
| WV | 13 | 0.96 | 12 | 0 | 1 |
| OK | 23 | 0.96 | 19 | 1 | 3 |
| NM | 9 | 0.96 | 7 | 0 | 2 |
| AR | 34 | 0.94 | 29 | 0 | 5 |
| ND | 18 | 0.94 | 0 | 18 | 0 |
| AL | 17 | 0.93 | 14 | 0 | 3 |
| AK | 12 | 0.93 | 10 | 0 | 2 |
| LA | 21 | 0.87 | 14 | 1 | 6 |
| MO | 19 | 0.86 | 16 | 3 | 0 |
| KS | 5 | 0.81 | 2 | 1 | 2 |
| TX | 58 | 0.79 | 32 | 15 | 11 |
| TN | 15 | 0.77 | 1 | 13 | 1 |
| GA | 60 | 0.75 | 31 | 23 | 6 |
| KY | 9 | 0.70 | 4 | 3 | 2 |
| OR | 9 | 0.69 | 0 | 8 | 1 |
| NH | 7 | 0.68 | 0 | 6 | 1 |

*(table continues)*

| State | No. of Counties with Significant *p*-values (<.05) | Average Rurality of the Significant Clusters (as a Proportion of 1.00) | No. of Counties in Clusters of Low SDO | No. of Counties in Clusters of High SDO | No. of Counties as Outliers |
|---|---|---|---|---|---|
| NC | 21 | 0.67 | 4 | 17 | 0 |
| ME | 1 | 0.67 | 0 | 1 | 0 |
| MI | 10 | 0.62 | 1 | 5 | 4 |
| CO | 13 | 0.62 | 6 | 6 | 1 |
| IN | 5 | 0.53 | 0 | 5 | 0 |
| FL | 29 | 0.50 | 6 | 19 | 4 |
| UT | 7 | 0.47 | 0 | 5 | 2 |
| MD | 18 | 0.46 | 0 | 18 | 0 |
| MT | 1 | 0.42 | 0 | 0 | 1 |
| DE | 3 | 0.41 | 0 | 3 | 0 |
| VA | 41 | 0.37 | 4 | 33 | 4 |
| CT | 8 | 0.36 | 0 | 8 | 0 |
| OH | 6 | 0.35 | 0 | 5 | 1 |
| NY | 20 | 0.35 | 0 | 20 | 0 |
| WA | 4 | 0.34 | 0 | 4 | 0 |
| WI | 6 | 0.32 | 0 | 6 | 0 |
| IL | 10 | 0.29 | 1 | 9 | 0 |
| PA | 17 | 0.27 | 0 | 17 | 0 |
| MA | 12 | 0.24 | 0 | 12 | 0 |
| NJ | 21 | 0.22 | 0 | 21 | 0 |
| CA | 20 | 0.19 | 0 | 19 | 1 |
| RI | 5 | 0.17 | 0 | 5 | 0 |
| MN | 2 | 0.02 | 0 | 2 | 0 |
| DC | 1 | 0.00 | 0 | 1 | 0 |

*Note*. Hawaii and Iowa are not included in this table as there are no significant clusters or outliers in these two states.

## 5.3 Theoretical Implications

### 5.3.1 Conceptual Model: Students' Digital Opportunity

During the iterative process of exploratory factor analysis, several important insights were drawn as to how we conceptualized the key components and what they represented in the SDO model. Broadband usage and download speed by proportion contributed mostly to the SDO measure by their factor loadings. We can interpret this understanding as students' digital opportunities will be driven mostly by the use of high-speed internet at schools and at home. The download speed indicator helped us understand two things. First, higher download speeds and lower latency values mean that a user can utilize richer digital content without delay or lags in their user experience. Higher speed ranges can also allow more devices to access the internet simultaneously from the same location. While the upload speed did not load consistently with the download speed in the SDO model, it is understandable because upload speeds represent a different type of use. In general, users are accessing information or downloading content to their devices, and therefore a higher download speed range is needed for timely data processing. However, upload speed is associated with the user sending data through their network to the internet. If a user mostly consumes information instead of sharing or sending information, a lower range of upload speeds will not significantly affect their user experience. Thus, for students in a digital learning situation, they may be mostly viewing content online or participating in virtual classes where their download speed will have a greater effect on their experience. However, their upload speed ranges become important during video conferencing or live streaming as their devices are sending data to the internet so that others can receive their information. As technology continues to evolve, the conceptualization of speed as the quality of the broadband connection will also change. For example, Bauer et al. (2010) discussed that the testing methodologies used in assessing speed vary by hardware, network, and software applications. While this technical research is necessary for digital infrastructure, it is difficult to relate this information to K-12 school systems. Therefore, collaborative efforts between technical experts and school administrators is needed to ensure that school and student broadband needs are met.

PC ownership and broadband subscription at home loaded inconsistently with the different factor models during the initial analysis. One limitation of this variable is that this information is collected from the U.S. Census ACS Survey (2019a) where users self-reported their device ownership. There could be a content validity concern because this variable may not truly be representative of the digital devices necessary for K-12 students. From a theoretical perspective, gauging users' digital device ownership and types of broadband services is a challenge. Hilbert (2014) pointed out that as technology is changing, even when a high level of ICT ownership is reported, the ICT tools themselves vary from dated equipment to the latest digital devices. Therefore, the conceptualization of device ownership would need further research to improve our understanding and model of measurement. For example, chrome books or tablets have been promoted to replace traditional PCs or laptops in school because they are low-cost and easy to store. However, to truly become efficient with using a computer for learning and for work, there are a range of skillsets needed to perform a variety of computer-based tasks. These factors are then associated with digital skills or computer literacy which are beyond the physical aspect of digital equity. For example, there are international-level studies, such as the International Computer and Information Literacy Study (ICILS) for 8$^{th}$ grade students (National Center for Education Statistics, n.d.). However, these studies used a country-based sampling approach, and it is inappropriate to use their information to assess how digital skills varied at the county level among K-12 students. In sum, the types of digital device and ownership status will require further research to confirm their relevancy and roles in the measurement of digital equity.

As discussed in the literature review (Chapter 2), several factors play key roles in determining a person's access and use of the internet and digital technologies. The conceptualization of the SDO model is based on an equal opportunity perspective where all students should have the four components described by the SDO model to participate in learning with technology. These are pre-conditions that must be addressed to close the digital opportunity gap shown in this research. From an access perspective within digital equity, the result of this work displayed the disparity of how some K-12 students are not able to participate fully and equitably in association with where they are. The SDO model also represents how students obtain their digital opportunities both at school and in the home. We should no longer separate the discussion of digital equity for K-12 students at school and at home, because as the pandemic has proven to us, learning with digital technology at home is an extension of the current school system implicitly. While the pandemic may end one day, digital technology at home for students to complete homework or practice their knowledge learned from school is an inseparable part of a comprehensive educational experience among the K-12 student population in our country.

### 5.3.2 Beyond K-12 Schools and Students

For the general population, the current viewpoint of "universal access" is idealistic, however as is clearly shown from the exploratory data analysis of this research, there is a discrepancy between access and use (as illustrated in Figure 3). Thus, increasing broadband availability does not directly increase usage or adoption of broadband internet or related services. There are other underlying factors that affect adoption or utilization of the fixed, home-based broadband internet analyzed in this work. If we are to discuss this matter from a digital equity perspective, some suggest that online contents may not be inclusive (Gorski, 2007) for a wide range of audiences. As we observed in the spatial clusters of below-average SDO values, it would be difficult for schools and their communities to change their current situations immediately without purposeful intervention. For counties and states that are receptive to improving their current digital infrastructure from the local level, municipal governments play an important role in forging partnerships between the public and private sectors to close the digital opportunity gap in their regions.

## 5.4 Methodological Implications

One significant contribution of this study to current research is the integration of data science techniques with traditional measurement approaches to evaluate digital opportunity among K-12 students. While previous educational research attempted to analyze digital equity in using technology for learning, most of the work is limited to a local level. There is also a limited understanding of how digital equity among K-12 schools and students is related to geography. The SDO measure provides a baseline in assessing the digital opportunity for K-12 students in different parts of the United States and serves as a feature variable for geographical evaluation.

The spatial methods used in this study also improve our understanding of the phenomenon and challenge the naïve assumptions of how the Internet can transfer information or knowledge across space. Birdsall and Birdsall (2005) emphasized the importance of mapping index measurements by their geographical locations because of the spatial dimension of human development and digital access. This study provided evidence that physical barriers continue to exist even when broadband is made available and does not directly translate into improved digital opportunities. The data product is also a methodological example of using data visualization as part of rigorous research during data exploration and model evaluation.

## 5.5 Practical Implications

Every initiative in education also entails a cost. Currently, the FCC provides support to public schools and libraries in acquiring technology at a discounted rate through the E-rate program. However, it is unclear how much this program could affect the actual adoption and use of technology in K-12 public schools if the persistent gaps of digital access and use are expected. Reardon (2011, p. 92) points out that “We tend to think of the relationship between socioeconomic status and children’s academic achievement as a sociological necessity, rather than a product of a set of social conditions, policy choices, and educational practice.” The digital equity observed in K-12 schools presented similar social conditions as a result of socioeconomic infrastructure at the county level. This work attempts to weave the complex nature of digital disparity with geographical characteristics to shed light on how counties vary within states and across the country. The result of this work can further inform federal and state-level policy as this study evaluated digital equity for the K-12 population at the macro level. Further analysis at the micro level can be applied to determine what the underlying causes are and how they contribute to the observed differences on the thematic maps. The data products of this work are openly accessible and can be combined with other sources of county-level data to further research of digital equity issues.

From the initial data exploration, it was discovered that there are only 2,775 counties that have at least one operating library, branch, or outlet during the year of 2019. The K-12 school information at the county level indicated the presence of schools among 3,138 counties in the U.S. It is assumed that most places in the U.S. have a library outlet that serves the local community. At the time of research, it is uncertain why there are hundreds of counties (about 363) without a library facility or outlet during the year of 2019. This is a major concern if compared to how many counties have at least one school in operation; there are roughly 12% of the counties in our country that do not have a public library reported to be in operation.

This observation challenges the current assumption of how public libraries can be utilized to address the digital divide in their communities. While public libraries serve all age groups in their communities, they would need to offer a variety of programs and services to meet all technological needs. For example, it is often reported that adults use library services to complete online applications for jobs or to obtain other information resources (Ball, 2009). Children are assumed to have direct access to a library in their communities. For example, Brake (2020) suggested schools and libraries use the FCC's E-Rate funding and the recent COVID-19 recovery bill's allocation to purchase loanable equipment for students. However, this type of recommendation is simply not feasible if there is not a library facility in the neighborhood. If the schools in those regions did not have the capacity to obtain these resources in the past, the lack of a coexisting library infrastructure may put the people residing there further behind in terms of their digital opportunities.

Thus, from an infrastructure perspective, should the priority be placed on ensuring that public libraries are established and available first before expanding the existing library services to address the digital divide? The International Federation of Library Associations and Institutions (IFLA) discussed that infrastructure in today's society is beyond roads or bridges from the traditional transport perspective; libraries are arguably a core part of the connectivity infrastructure that connects educational, social, cultural, innovation, and democratic infrastructures (IFLA, 2021).

There is still hope for expanding digital infrastructure outside of traditional institutions. Local initiatives have been implemented to provide broadband internet access and technology services beyond the already burdened anchor institutions like schools or libraries. For example, the 'channelAustin' case study is a model of community participation as a digital intervention in

Texas, in which the municipal government played an important role in forging the partnerships between the city and the local non-profits and media organizations (Fuentes-Bautista, 2014).

## 5.6 Limitations of the Results

This study evaluated the physical dimension of digital equity and did not address the human or socio-cultural dimension of digital equity among K-12 students. Another limitation of the current study is that the time period only reflects the current state of broadband access and use at the county level across the U.S. Without comparable information from the past or the future, this result cannot support a broader evaluation of how digital equity has changed over time. However, longitudinal data collection of the four key components in the SDO model will allow us to track and analyze how digital equity may change or shift from a spatial-temporal perspective.

The estimated SDO measure and scores cannot be used for comparison at other geographical levels. This is a limitation of the current research design in which group differences were analyzed at the county level and visualization of the county-level thematic map (choropleth) was used for a meaningful representation of the country-level information. While the SDO scores represented the average digital opportunities of K-12 students in each county, this value does not reflect the individual differences (heterogeneity) within each county. Others wanting to replicate this methodology will need to aggregate the school-level information to the proper spatial unit and conduct exploratory factor analysis to re-evaluate these factors at the intended geographical levels. The shapefiles from the U.S. Census provided the legal boundaries for each county for spatial analysis and allowed for the production of choropleth maps to display the variation of SDO across counties. However, we should keep in mind that while the legal and administrative units set by the U.S. Census are useful for measurement and evaluation purposes, human related activities may often cross boundaries, depending on how the local infrastructure is organized.

Another limitation concerns the type of educational entities outside of K-12 public schools. There is little understanding of how digital equity varies among children who attend private schools, home schools, schools operated by the Bureau of Indian Education (BIE), or tribally controlled schools (Bureau of Indian Education, n.d.). This research lies in the physical dimension of the pre-conditions students would need to equally participate in digital learning. Digital equity from a perspective of having inclusive and accessible content online within education is not well-known. There is an imbalance of information production versus information utilization due to the varying levels of access and skills children may have based on their surrounding factors (Subramony, 2011). This is the human aspect of digital equity for which we do not have a clear conception in quantitative research. Data sources detailing this aspect of digital equity would be valuable to enhance the current students' digital opportunity construct.

Currently, there are regions that do not have fixed broadband through landlines, and there are other forms of broadband technology that this study did not address. Therefore, we need to collect information and data on how places that do not have landline broadband connections obtain their internet connectivity if universal access is the country's goal in broadband deployment.

## 5.7 Future Research

This research focuses on a construct development for K-12 students' digital opportunity by using secondary datasets. If passive data collection at the micro-level without risking the users' privacy is possible in the future, it will be a game changer for future research of digital equity. These data can provide evidence and possibly explain how people are accessing and using broadband internet and how device types would influence speed quality. There are many more research opportunities in the digital content and skills area, for example, if we can assess how users experience or interact with the technology according to their differences in digital opportunities, it will be groundbreaking to quantitatively capture these nuances at the individual level. By having telemetry monitoring data from software applications, hardware, and school broadband networks, we could further develop guidelines on how much broadband speed and what types of digital devices are appropriate for K-12 students to participate from home and at school. Because student learning does not happen in a vacuum, we can assume that K-12 teachers and staff may also experience similar barriers in obtaining digital connectivity and devices. To effectively address digital equity for K-12 students, future research is needed to assess the school system as a whole, including the digital needs among leadership teams, teachers, instructional staff, and students.

There are efforts at the national level to collect data on K-12 students' digital skills and computer literacy, such as the National Assessment of Education Progress (NAEP) or the International Computer and Information Literacy Study (ICILS). However, digital skills or computer literacy require another conceptual model that needs to be agreed upon by the research communities. There is still a lot of uncertainty about how K-12 students should be trained to meet the future workforce. By changing curriculum and instruction to be digitally oriented, teachers and other instructional staff will need to gain new skills to keep up with the rapid development of technology. Parent involvement in a child's digital experience is valuable as shown by previous work on neighborhood computer clubs and parent computer courses (Chen & Dym, 2003). It is evident that at the local level, school staff, parents, and the neighborhood can come together to strengthen the relationship to further improve a student's digital opportunity.

From a human information behavior perspective, especially in rural regions, the community's perception may play a bigger role in determining their internet utilization. Therefore, user interview and secondary data collection at the local level can provide an understanding of the nuances of the human aspects of digital equity. Due to the large variation in policy and local practice among the 50 states in the U.S., future micro-level studies are important to determine if the current broadband infrastructure and future technology policy are addressing the needs of the U.S. students. We generally assume that increasing access to the internet and technology will have a positive effect on civil engagement and other benefits. However, the results show that high SDO areas tend to have a larger range of socioeconomic groups, and thus any differences due to the high population density in urban areas may be masked. In other words, simply living in a region with a higher level of access and use of digital technologies may not be representative of actual access and use for certain groups, especially for minorities or marginalized communities. Datasets from other sources can be combined to further understand the relationship between the current state of digital connectivity and other demographic characteristics in urban areas. Future development of rural areas is also important. One way to evaluate the chances of a rural area to improve the residents' digital opportunities is by determine how far the rural area is to the nearest urban center. This is based on the assumption that when urban areas continue to grow in the U.S., there may be economic development opportunities for these less populated areas. Therefore, the spillover effect of clusters with high SDO identified in the current model can be analyzed to support future decision-making in resource allocation and strategic planning.

## 5.8 Conclusion

Imagine that a child is entering their first year in our K-12 school system in 2022, their digital opportunity is already at a different place compared to their peers. This image is quite troubling especially as this study has shown the great differences in digital access and use across the U.S. The next generation of children is going to spend their next 10 to 12 years in schools, and we are not sure if the current expansion of digital infrastructure will be fast enough to meet their needs. This work has brought the issue in front of practitioners and policymakers with a hope to raise awareness and interest in beginning dialogues about how to ensure basic requirements are met before students can participate in learning with technologies. Having internet access is only one aspect of the matter; equipment ownership and speed quality need to be present simultaneously to have a complete digital experience. For students who are already lagging behind in terms of their digital opportunities, what can we do to help them make up for lost time and chances? My work did not focus on educational outcomes or achievement as a product of the digital disparity. However, other educational scholars identified evidence that the concentration of disadvantaged students is the critical mechanism influencing educational outcomes and other psychosocial factors (Horgrebe & Tate, 2012). The map products reflect a similar issue where low SDO values are concentrated in certain regions. Our current school funding mechanism mainly focuses on improving student achievement; if necessary instructional resources like digital connectivity and devices are not available, is it fair to expect students to somehow overcome these challenges? This study only begins to address this topic. More questions about how to ensure digital equity among the K-12 school systems in our country are bound to appear as we continue to work on this issue. Using evidence-based tools and models like the SDO measure created in the course of this work, is a first step to making informed decisions and tracking the progress toward addressing the digital disparity in our country.

## REFERENCES


Alwin, D. F. (1973). The use of factor analysis in the construction of linear composites in social research. *Sociological Methods and Research*, *2*(2), 191–214.

American Library Association. (2019, January 14). *Library Funding*. Advocacy, Legislation & Issues. https://www.ala.org/advocacy/library-funding

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis* , *27*(2), 93–115.

Anselin, L., & Morrison, G. (2019, November 17). *Contiguity-based spatial weights*. R Notes. https://spatialanalysis.github.io/lab_tutorials/Contiguity_Spatial_Weights.html

Anselin, L. (2020). *Local Spatial Autocorrelation (1): LISA and Local Moran*. GeoDa: An Introduction to Spatial Data Science. https://geodacenter.github.io/workbook/6a_local_auto/lab6a.html

Attewell, P. (2001). The first and second digital divides. *Sociology of Education*, *74*(3), 252–259. https://doi.org/10.2307/2673277

Aydin, M. (2021). Does the digital divide matter? Factors and conditions that promote ICT literacy. *Telematics and Informatics*, *58,* 1–9. https://doi.org/10.1016/j.tele.2020.101536

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, *193*, 1–17. https://doi.org/10.1016/j.jpubeco.2020.104345

Ball, M. A. (2009). Aggregating broadband demand: Surveying the benefits and challenges for public libraries. *Government Information Quarterly*, *26*(4), 551–558. https://doi.org/10.1016/j.giq.2009.05.004

Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and Confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (1st ed., pp. 93–114). Routledge.

Barack, L. (2005). Gauging the digital divide. *School Library Journal*, *51*(8), 21.

Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology*, *28,* 97–104.

Bartlett, M. S. (1950). Test of significance in factor analysis. *British Journal of Psychology*, *3*(2), 77–85.

Bartlett, M. S. (1951). The effect of standardization on a $X^2$ approximation in factor analysis. *Biometika*, *38*(3/4), 337–344.

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, *16*(2), 296–298.

Bauer, S., Clark, D. D., & Lehr, W. (2010). *Understanding broadband speed measurements.* TPRC 2010. https://papers.ssrn.com/abstract=1988332

Bertot, J. C. (2009). Capacity planning for broadband in public libraries: Issues and strategies. *Library Technology Reports*, *45*(1), 38–42.

Birdsall, S. A., & Birdsall, W. F. (2005). Geography matters: Mapping human development and digital access. *First Monday*, *10*(10). https://firstmonday.org/ojs/index.php/fm/article/view/1281/1201

Bivand, R. S., Pebesma, E. J., Gómez-Rubio, V., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-7618-4

Bivand, R. (2021). *spdep: Spatial dependence: Weighting schemes, statistics* (Version 1.1-12) [R package]. https://cran.r-project.org/web/packages/spdep/index.html

Brake, D. (2020). *Lessons From the Pandemic: Broadband Policy After COVID-19*. Information Technology & Innovation Foundation. https://itif.org/publications/2020/07/13/lessons-pandemic-broadband-policy-after-covid-19

Brandtzæg, P. B., Heim, J., & Karahasanović, A. (2011). Understanding the new digital divide - A typology of Internet users in Europe. *International Journal of Human Computer Studies*, *69*(3), 123–138. https://doi.org/10.1016/j.ijhcs.2010.11.004

Brunsdon, C., & Comber, L. (2019). *An introduction to R for spatial analysis and mapping* (2nd ed.). SAGE Publications Ltd.

Bureau of Indian Education. (n.d.). *Tribally-Controlled Schools*. 25 USC Ch. 27 Tribally Controlled School Grants. https://www.bie.edu/topic-page/tribally-controlled-schools

Calabrese, M., & Nasr, A. (2020). *The online learning equity gap: Innovative solutions to connect all students at home.* New America Open Technology Institute , Wireless Future Project. https://www.newamerica.org/oti/reports/online-learning-equity-gap/

Campos-Castillo, C. (2015). Revisiting the first-level digital divide in the United States: gender and race/ethnicity patterns, 2007–2012. *Social Science Computer Review*, *33*(4), 423–439. https://doi.org/10.1177/0894439314547617

Chainey, S., & Ratcliffe, J. (2005). *GIS and crime mapping* (1st ed.). John Wiley & Sons, Ltd.

Chakraborty, J., & Bosman, M. M. (2005). Measuring the digital divide in the United States: Race, income, and personal computer ownership. *The Professional Geographer*, *57*(3), 395–410. https://doi.org/10.1111/j.0033-0124.2005.00486.x

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). *NbClust: An R package for determining the relevant number of clusters in a data set* (Version 3.0) [R package]. https://www.jstatsoft.org/article/view/v061i06

Chen, J.-Q., & Dym, W. (2003). Using computer technology to bridge school and community. *Phi Delta Kappan*, *85*(3), 232–234. https://www.jstor.org/stable/20441540

Cohron, M. (2015). The continuing digital divide in the United States. *Serials Librarian*, *69*(1), 77–86. https://doi.org/10.1080/0361526X.2015.1036195

Connected Nation - Texas. (2020, December 20). *Texas Municipal Broadband State Laws and Regulations*. https://www.connecttexas.com/internet-service/municipal-broadband-texas-state-regulations

Corrocher, N., & Ordanini, A. (2002). Measuring the Digital Divide: A framework for the analysis of cross-country differences. *Journal of Information Technology*, *17*(1), 9–19. https://doi.org/10.1080/02683960210132061

Cotten, S. R., Hale, T. M., Moroney, M. H., O'Neal, L., & Borch, C. (2011). Using affordable technology to decrease digital inequality. *Information, Communication & Society*, *14*(4), 424–444. https://doi.org/10.1080/1369118X.2011.559266

Council of Economic Advisers. (2015). *Mapping the digital divide.* (Issue Brief July 2015). The White House, Executive Office of the President, Council of Economic Advisers. https://www.whitehouse.gov/sites/default/files/wh_digital_divide_issue_brief.pdf

Council of Economic Advisers. (2016). *The digital divide and economic benefits of broadband access*. (Issue Brief March 2016). The White House, Executive Office of the President, Council of Economic Advisers. https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160308_broadband_cea_issue_brief.pdf

DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation*, *14*(20), 1–11.

Dolan, J. E. (2016). Splicing the divide: A review of research on the evolving digital divide among K-12 students. *Journal of Research on Technology in Education*, *48*(1), 16–37. https://doi.org/10.1080/15391523.2015.1103147

Dolnicar, V., Vehnovar, V., & Sicherl, P. (2004). Benchmarking digital divide: Definitions used and methods applied. *Proceedings of the 26th International Conference on Information Technology Interfaces, ITI 2004*, 684.

Dutton, W. H., & Reisdorf, B. C. (2019). Cultural divides and digital inequalities: attitudes shaping Internet and social media divides. *Information, Communication & Society*, *22*(1), 18–38. https://doi.org/10.1080/1369118X.2017.1353640

Edgerton, A. K., & Cookson, P. W. J. (2020, November 10). *Closing the digital divide: The critical role of the Federal government*. Learning Policy Institute Blog. https://learningpolicyinstitute.org/blog/covid-closing-digital-divide-federal-government

Empirica. (2003). *Measuring the Information Society in the EU, the EU Accession Countries, Switzerland and the US*. SIBIC Statistics and Indicators. http://www.sibis-eu.org/statistics/stat_ind.htm

European Commission. (2001). *Statistical Indicators Benchmarking the Information Society (SIBIS) project*. European Commission, CROS, Research Projects under Framework Programmer, FP5 Projects. http://www.sibis-eu.org/index.htm

Federal Communications Commission. (2015). *2015 Broadband Progress Report*. Broadband Progress Report . https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2015-broadband-progress-report

Federal Communications Commission (FCC). (2021a). *Homework Gap and Connectivity Divide*. FCC Initiatives. https://www.fcc.gov/about-fcc/fcc-initiatives/bridging-digital-divide-all-americans

Federal Communications Commission. (2021b). *Fourteenth broadband deployment report*. https://www.fcc.gov/reports-research/reports/broadband-progress-reports/fourteenth-broadband-deployment-report

Ford, G. S. (2018). Is faster better? Quantifying the relationship between broadband speed and economic growth. *Telecommunications Policy*, *42*(9), 766–777. https://doi.org/10.1016/j.telpol.2018.05.006

Fox, C. (2019). *State K-12 Broadband Leadership 2019: Driving Connectivity, Access and Ensuring Student Success* (ERIC No. ED594505). State Educational Technology Directors Association (SETDA). https://files.eric.ed.gov/fulltext/ED594505.pdf

Fuentes-Bautista, M. (2014). Rethinking localism in the broadband era: A participatory community development approach. *Government Information Quarterly*, *31*(1), 65–77. https://doi.org/10.1016/j.giq.2012.08.007

Gao, N., & Hayes, J. (2021). *The digital divide in education.* Public Policy Institute of California Just the Facts. https://www.ppic.org/publication/the-digital-divide-in-education/

García, E., & Weiss, E. (2020). *COVID-19 and student performance, equity, and U.S. education policy: Lessons from pre-pandemic research to inform relief, recovery, and rebuilding*. Economic Policy Institute Report. https://www.epi.org/publication/the-consequences-of-the-covid-19-pandemic-for-education-performance-and-equity-in-the-united-states-what-can-we-learn-from-pre-pandemic-research-to-inform-relief-recovery-and-rebuilding/

Goldberg, L. R. (2006). Doing it all Bass-Ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality*, *40*(4), 347–358. https://doi.org/10.1016/j.jrp.2006.01.001

Gonzales, A. (2016). The contemporary US digital divide: from initial access to technology maintenance. *Information Communication and Society*, *19*(2), 234–248. https://doi.org/10.1080/1369118X.2015.1050438

Goolsbee, A., & Guryan, J. (2006). The impact of internet subsidies in public schools. *The Review of Economics and Statistics*, *88*(2), 336–347. https://doi.org/jstor.org/stable/40042999

Gorski, P. C. (2007). Digital equity. In G. L. Anderson (Ed.), *Encyclopedia of Activism and Social Justice*. Sage Publications.

Gorsuch, R. L. (1974). *Factor analysis*. W. B. Saunders Company.

Graves, K. F., & Bowers, A. J. (2018). Toward a typology of technology-using teachers in the "New Digital Divide": A latent class analysis of the NCES Fast Response Survey System teachers' use of educational technology in U.S. Public Schools, 2009 (FRSS 95). *Teachers College Record*, *120*(8), 1–42. https://academiccommons.columbia.edu/doi/10.7916/d8-ddmv-y558

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*(3), 430–450. https://doi.org/10.1037/1082-989x.6.4.430

Grubesic, T. H., & Murray, A. T. (2002). Constructing the divide: Spatial disparities in broadband access. *Papers in Regional Science*, *81*(2), 197–221.

Hargittai, E. (2002). Second-Level Digital Divide: Differences in People's Online Skills. *First Monday*, *7*(4). https://doi.org/10.5210/FM.V7I4.942

Helsper, E. J. (2017). The social relativity of digital exclusion: Applying relative deprivation theory to digital inequalities. *Communication Theory*, *27*(3), 223–242. https://doi.org/10.1111/comt.12110

Henderson, H. (2017). Broadband. In *Encyclopedia of Computer Science and Technology*. Retrieved July 31, 2021, from https://search.credoreference.com

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research. *Educational and Psychological Measurement*, *66*(3), 393–416. https://doi.org/10.1177/0013164405282485

Hilbert, M. (2014). Technological information inequality as an incessantly moving target: The redistribution of information and communication capacities between 1986 and 2010. *Journal of the Association for Information Science and Technology*, *65*(4), 821–835. https://doi.org/10.1002/asi.23020

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Perron, J. M., & Mumford, A. K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, *65*(2), 202–226. https://doi.org/10.1177/0013164404267287

Hogrebe, M. C., & Tate, W. F. (2012). Place, Poverty, and Algebra: A statewide comparative spatial analysis of variable relationships. *Journal of Mathematics Education at Teachers College*, *3*(2), 12–24. https://journals.library.columbia.edu/index.php/jmetc/issue/view/25

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185.

Horrigan, J. B. (2010). Broadband Adoption and Use in America. In *FCC Omnibus Broadband Initiative Working Paper Series* (OBI Working Paper Series No. 1, Issue 1). http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-296442A1.pdf

Horrigan, J. B. (2015). *The numbers behind the broadband "homework gap"*. Pew Research Center. [FactTank News, Research Topics: Internet Connectivity]. https://www.pewresearch.org/fact-tank/2015/04/20/the-numbers-behind-the-broadband-homework-gap/

Howard, P. N., Busch, L., & Sheets, P. (2010). Comparing digital divides: Internet access and social inequality in Canada and the United States. *Canadian Journal of Communication*, *35*(1), 109–128.

Hsieh, J. J. P. A., Rai, A., & Keil, M. (2008). Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. *MIS Quarterly: Management Information Systems*, *32*(1), 97–126. https://doi.org/10.2307/25148830

Hüsing, T., & Selhofer, H. (2002). The digital divide index - A measure of social inequalities in the adoption of ICT. *European Conference on Information Systems (ECIS) Conference Proceedings*, *35, June 6-8*, 1273–1286. http://www.ifiptc8.org/asp/aspecis/20020042.pdf

Hüsing, T., & Selhofer, H. (2004). DIDIX: A digital divide index for measuring inequality in IT diffusion. *IT & Society*, *1*(7), 21–38. https://www.semanticscholar.org/paper/DiDix.-A-Digital-Divide-Index-for-Measuring-in-IT-Selhofer/55fa8c6a7c91b3de8f3f0143a7146e96f408a60b

IFLA, International Federation of Library Associations and Institutions . (2021, March 21). Libraries at the Heart of Educational, Social, Cultural, Innovation and Democratic Infrastructure. *Library Policy and Advocacy Blog.* https://blogs.ifla.org/lpa/2021/03/21/libraries-at-the-heart-of-educational-social-cultural-innovation-and-democratic-infrastructure/

IMLS, Institute of Museum and Library Services. (2022). *Public Libraries Survey (FY 2019)*. https://www.imls.gov/research-evaluation/data-collection/public-libraries-survey

International Telecommunication Union (ITU). (2003). *ITU Digital Access Index: World's First Global ICT Ranking* [Press Release]. https://www.itu.int/newsarchive/press_releases/2003/30.html

Jayakar, K. (2011). Promoting universal broadband through middle mile institutions: A legislative agenda. *Journal of Information Policy*, *1*, 102–124.

Jim, C., Evans, S. A., & Grant, A. (2021a, July 12-16). *Multidimensional approaches to illustrate the digital divide among K-12 students* [Paper presentation]. The 49th Annual Conference of the International Association of School Librarianship (IASL), Virtual, Denton, TX, United States. https://doi.org/10.29173/iasl8282

Jim, C. K., Grant, A., & Ladd, B. T. (2021b, October 13-15). *Visualization of the digital divide among K-12 Students: Open data, quantitative measures, and policy implications* [Paper presentation]. 2021 Northeastern Educational Research Association (NERA) Conference, Virtual, United States. https://opencommons.uconn.edu/nera-2021/

Kahan, J., & Lavista Ferres, J. M. (2020). *Microsoft United States broadband usage percentages dataset* [Data set].https://github.com/microsoft/USBroadbandUsagePercentages

Kaiser, H. F. (1970). A second-generation little jiffy. *Psychometrika*, *35*, 401–415.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36.

Katz, V. S., & Gonzalez, C. (2016). Toward meaningful connectivity: using multilevel communication research to reframe digital inequality. *Journal of Communication*, *66*(2), 236–249. https://doi.org/10.1111/JCOM.12214

KewalRamani, A., Zhang, J., Wang, X., Rathbun, A., Corcoran, L., Diliberti, M., & Zhang, J. (2018). *Student access to digital learning resources outside of the classroom (NCES 2017-098)*. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017098

Kinney, B. (2010). The internet, public libraries, and the digital divide. *Public Library Quarterly*, *29*(2), 104–161. https://doi.org/10.1080/01616841003779718

Kruger, L. G. (2018). *Defining broadband : Minimum threshold speeds and broadband policy (CRS Report R45039 - Version 6)*. Congressional Research Service. https://crsreports.congress.gov/product/details?prodcode=R45039

Kuttan, A., & Peters, L. (2003). *From digital divide to digital opportunity* (1st ed.). Scarecrow Press, Inc.

Lamstein, A. (2020). *choroplethr: Simplify the creation of choropleth maps in R* (Version 3.7.0) [ R packages]. https://cran.r-project.org/web/packages/choroplethr/index.html

Lansley, G., & Cheshire, J. (2016). *An introduction to spatial data analysis and visualisation in R*. https://data.cdrc.ac.uk/tutorial/an-introduction-to-spatial-data-analysis-and-visualisation-in-r

Liu, Y. H., Prince, J., & Wallsten, S. (2018). Distinguishing bandwidth and latency in households' willingness-to-pay for broadband internet speed. *Information Economics and Policy*, *45*, 1–15. https://doi.org/10.1016/j.infoecopol.2018.07.001

Looker, E. D., & Thiessen, V. (2003). Beyond the digital divide in Canadian schools: From access to competency in the use of information technology. *Social Science Computer Review*, *21*(4), 475–490. https://doi.org/10.1177/0894439303256536

Lucendo-Monedero, A. L., Ruiz-Rodríguez, F., & González-Relaño, R. (2019). Measuring the digital divide at regional level. A spatial analysis of the inequalities in digital development of households and individuals in Europe. *Telematics and Informatics*, *41*(April), 197–217. https://doi.org/10.1016/j.tele.2019.05.002

Maechler, M., Rousseeuw, P., Struyf, A., & Hubert, M. (2021). *cluster: Finding groups in data* (Version 2.1.2) [R package]. https://cran.r-project.org/web/packages/cluster/index.html

Mandel, L. H., Bishop, B. W., McClure, C. R., Bertot, J. C., & Jaeger, P. T. (2010). Broadband for public libraries: Importance, issues, and research needs. *Government Information Quarterly*, *27*(3), 280–291. https://doi.org/10.1016/j.giq.2010.02.004

Mayer, M. (2021). *missRanger: Fast imputation of missing values* (Version 2.1.3) [R package]. https://cran.r-project.org/web/packages/missRanger/index.html

Middleton, K. L., & Chambers, V. (2010). Approaching digital equity: is wifi the new leveler? *Information, Technology & People*, *23*(1), 4–22. https://doi.org/10.1108/09593841011022528

Mills, B. F., & Whitacre, B. E. (2003). Understanding the non-metropolitan-metropolitan digital divide. *Growth and Change*, *34*(2), 219–243. https://doi.org/10.1111/1468-2257.00215

Monroe, B. (2004). *Crossing the digital divide: race, writing, and technology in the classroom*. Teachers College Press.

Mori, C. K. (2011). “Digital inclusion”: Are we all talking about the same thing? In J. Steyn & G. Johanson (Eds.), *ICTs and Sustainable Solutions for the Digital Divide: Theory and Perspectives* (pp. 45–64). IGI Global. https://doi.org/10.4018/978-1-61520-799-2.ch003

Mossberger, K., Tolbert, C. J., & Stansbury, M. (2003). *Virtual inequality: Beyond the digital divide*. Georgetown University Press.

National Center for Education Statistics (NCES). (n.d.). *International computer and information literacy study (ICILS)*. About ICILS - Overview. https://nces.ed.gov/surveys/icils/overview.asp

National Center for Education Statistics (NCES). (2019). *Locale Classifications*. Education Demographic and Geographic Estimates (EDGE). https://nces.ed.gov/programs/edge/Geographic/LocaleBoundaries

National Center for Education Statistics (NCES). (2021). *Elementary/Secondary Information System* (No. 2019–20 v. 1a). Common Core of Data (CCD). https://nces.ed.gov/ccd/elsi/

National Telecommunications and Information Administration (NTIA). (1995). *Falling Through the Net: A Survey of the "Have Nots" in Rural and Urban America*. https://www.ntia.doc.gov/ntiahome/fallingthru.html

National Telecommunications and Information Administration (NTIA). (2014, February 25). *Working to Close the Digital Divide in Silicon Valley*. NTIA Blog. https://www.ntia.doc.gov/blog/2014/working-close-digital-divide-silicon-valley

Nishida, T., Pick, J. B., & Sarkar, A. (2014). Japan's prefectural digital divide: A multivariate and spatial analysis. *Telecommunications Policy*, *38*(11), 992–1010. https://doi.org/10.1016/j.telpol.2014.05.004

Noam, E. (2011). Let them eat cellphones: Why mobile wireless is no solution for broadband. *Journal of Information Policy*, *1*, 470–485. https://www.jstor.org/stable/10.5325/jinfopoli.1.2011.0470

OECD, Organisation for Economic Co-operation and Development. (2008). *Handbook on constructing composite indicators: methodology and user guide*. https://www.oecd.org/sdd/42495745.pdf

OECD, Organisation for Economic Co-operation and Development. (2019). *Enhancing access and connectivity to harness digital transformation*. https://www.oecd.org/going-digital/topics/digital-infrastructure/

Ookla. (2020). *Ookla open data initiative: Speed test by Ookla global fixed and mobile network performance maps* [Data set].Open Data Registry on AWS. https://www.ookla.com/ookla-for-good/open-data

Pebesma, E. (2021). *sf: Simple features for R* (Version 1.0-4) [R package]. https://cran.r-project.org/web/packages/sf/index.html

Pebesma, E., & Bivand, R. (2021). *sp: Classes and methods for spatial data* (Version 1.4-6) [R package]. https://cran.r-project.org/web/packages/sp/index.html

Peck, C., Hewitt, K. K., Mullen, C. A., Lashley, C. A., Eldridge, J. A., & Douglas, T.-R. M. O. (2015). Digital youth in brick-and-mortar schools: Examining the complex interplay of students, technology, education, and change. *Teachers College Record*, *117*(5), 1–40. https://journals.sagepub.com/doi/abs/10.1177/016146811511700505

Pérez-Amaral, T., Valarezo, A., López, R., & Garín-Muñoz, T. (2021). Digital divides across consumers of internet services in Spain using panel data 2007–2019. Narrowing or not? *Telecommunications Policy*, *45*(2), 1–17. https://doi.org/10.1016/j.telpol.2020.102093

Philip, L., Cottrill, C., Farrington, J., Williams, F., & Ashmore, F. (2017). The digital divide: Patterns, policy and scenarios for connecting the 'final few' in rural communities across Great Britain. *Journal of Rural Studies*, *54*, 386–398. https://doi.org/10.1016/j.jrurstud.2016.12.002

Philip, L., & Williams, F. (2019). Remote rural home based businesses and digital inequalities: Understanding needs and expectations in a digitally underserved community. *Journal of Rural Studies*, *68*, 306–318. https://doi.org/10.1016/j.jrurstud.2018.09.011

Pick, J. B., & Sarkar, A. (2015). *The Global Digital Divides: Explaining Change*. Springer-Verlag.

Prescott, S. (2020, April 20). What we don't know about teachers' home internet access. *New America Education Policy Blog.* https://www.newamerica.org/education-policy/edcentral/what-we-dont-know-about-teachers-home-internet-access/

Rafalow, M. H. (2021a). Digital equality requires more than access. *Kappan*, *102*(6), 26–29.

Rafalow, M. H. (2021b). *Digital divisions: How schools create inequality in the tech era*. The University of Chicago Press.

Ragnedda, M. (2019). Conceptualising the digital divide. In B. Mutsvairo & M. Radnedda (Eds.), *Mapping the Digital Divide in Africa: A Mediated Analysis* (pp. 27–43). Amsterdam University Press. https://doi.org/10.5117/9789462986855

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. j. Murnane (Eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation.

Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Jang, H., & Chavez, B. (2021). *Stanford Education Data Archive (Version 4.1)* [Data set]. https://edopportunity.org/get-the-data/

Reisdorf, B. C., Dutton, W. H., Triwibowo, W., & Nelson, M. E. (2017). The unexplored history of operationalising digital divides: a pilot study. *Internet Histories*, *1*(1–2), 106–118. https://doi.org/10.1080/24701475.2017.1311165

Reisdorf, B. C., Yankelevich, A., Shapiro, M., & Dutton, W. H. (2019). Wirelessly bridging the homework gap: Technical options and social challenges in getting broadband to disconnected students. *Education and Information Technologies*, *24*(6), 3803–3821. https://doi.org/10.1007/s10639-019-09953-9

Revelle, W. (2021). *psych: Procedures for psychological, psychometric, and personality research* (Version 2.1.9) [R package]. https://cran.r-project.org/web/packages/psych/index.html

Revelle, W., & Rocklin, T. (1979). Very Simple Structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, *14*(4), 403–414.

Rey, S. J., Arribas-Bel, D., & Wolf, L. J. (2020). *Geographic data science with PySAL and the PyData Stack*. Geographic Data Science with Python. https://geographicdata.science/book/intro.html#

Reynolds, R., & Chiu, M. M. (2016). Reducing digital divide effects through student engagement in coordinated game design, online resource use, and social computing activities in school. *Journal of the Association for Information Science and Technology*, *67*(8), 1822–1835. https://doi.org/10.1002/asi.23504

Rice, R. E., & Katz, J. E. (2002). Comparing internet and mobile phone digital divides. *Proceedings of the American Society for Information Science and Technology*, *39*(1), 92–98. https://doi.org/10.1002/meet.1450390110

Riddlesden, D., & Singleton, A. D. (2014). Broadband speed equity: A new digital divide? *Applied Geography*, *52*, 25–33. https://doi.org/10.1016/j.apgeog.2014.04.008

Robinson, L., Schulz, J., Dodel, M., Correa, T., Villanueva-Mansilla, E., Leal, S., Magallanes-Blanco, C., Rodriguez-Medina, L., Dunn, H. S., Levine, L., McMahon, R., & Khilnani, A. (2020). Digital inclusion across the Americas and the Caribbean. *Social Inclusion*, *8*(2), 244–259. https://doi.org/10.17645/si.v8i2.2632

Rogers, S. E. (2016). Bridging the 21st century digital divide. *TechTrends*, *60*(3), 197–199. https://doi.org/10.1007/s11528-016-0057-0

Salemink, K., Strijker, D., & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. *Journal of Rural Studies*, *54*, 360–371. https://doi.org/10.1016/j.jrurstud.2015.09.001

Scheerder, A., van Deursen, A., & van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. *Telematics and Informatics*, *34*(8), 1607–1624. https://doi.org/10.1016/j.tele.2017.07.007

Schmidt, D., & Power, S. A. (2021). Offline World: The Internet as Social Infrastructure among the Unconnected in Quasi-Rural Illinois. *Integrative Psychological and Behavioral Science*, *55*(2), 371–385. https://doi.org/10.1007/s12124-020-09574-9

Schweik, C., Smith, J., & Meyer, C. (2018). World librarians: A peer-to-peer commons for closing the global digital divide. *Journal of Librarianship and Scholarly Communication*, *6* (Special Issue: The Role of Scholarly Communication in a Democratic Society), eeP2201. https://doi.org/https://doi.org/10.7710/2162-3309.2249

Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, *6*(3), 341–362. https://doi.org/10.1177/1461444804042519

Selwyn, N., Gorard, S., & Williams, S. (2001). Digital divide or digital opportunity? The role of technology in overcoming social exclusion in U.S. education. *Educational Policy*, *15*(2), 258–277. https://doi.org/10.1177/0895904801015002002

Servon, L. J. (2002). *Bridging the digital divide: Technology, community, and public policy*. Blackwell Publishers Ltd.

Shenglin, B., Simonelli, F., Zhang, R., Bosc, R., & Li, W. (2017). *Digital Infrastructure: Overcoming the digital divide in China and the European Union*. https://www.g20-insights.org/policy_briefs/digital-infrastructure-overcoming-digital-divide-emerging-economies/

Sicherl, P. (2019). Different statistical measures create different perceptions of the digital divide. *The Information Society*, *35*(3), 143–157. https://doi.org/10.1080/01972243.2019.1582568

Strusani, D., & Houngbonon, G. V. (2020). *What COVID-19 means for digital infrastructure in emerging markets* (IFC Thought Leadership, World Bank Group, EM Compass Note 83). https://www.ifc.org/wps/wcm/connect/publications_ext_content/ifc_external_publication_site/publications_listing_page/what+covid+19+means+for+digital+infrastructure+in+emerging+markets

Subramony, D. (2011). Socio-cultural issues in educational technology integration. *Colleagues*, *6*(1), 18–20. http://scholarworks.gvsu.edu/colleagueshttp://scholarworks.gvsu.edu/colleagues/vol6/iss1/10

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (4th ed.). Pearson Education, Inc.

U. S. Census Bureau. (2019a). *Age by Presence of a Computer and Types of Internet Subscription in Household (Table 28005)*. https://data.census.gov/cedsci/table?q=ACSDT1Y2019.B28005&tid=ACSDT5Y2019.B28005&hidePreview=true

U.S. Census Bureau. (2019b). *TIGER/Line Shapefiles*. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2019.html

U.S. Census Bureau. (2019c). *Current Population Survey, Computer and Internet Use Supplement Technical Documentation* (Issue November). https://www2.census.gov/programs-surveys/cps/techdocs/cpsnov19.pdf

U.S. Department of Education. (2018, October 24). *Title I. Improving Basic Programs Operated by Local Education Agencies.* https://www2.ed.gov/programs/titleiparta/index.html

U.S. Department of Education. (2021, June 15). *Federal Role in Education*. https://www2.ed.gov/about/overview/fed/role.html

United Nations Statistics Division. (2021). *SDG Indicator Goal 4*. The Sustainable Development Goals Report. https://sdgs.un.org/goals/goal4 van Deursen, A. J. A. M. J., & van Dijk, J. A. A. G. M. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New Media and Society*, *21*(2), 354–375. https://doi.org/10.1177/1461444818797082

van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *Information Society*, *19*(4), 315–326. https://doi.org/10.1080/01972240309487

van Ness, C., & Varn, J. (2021). *Governors prioritize expanding internet access for K-12 students*. National Governors Association Commentary. https://www.nga.org/news/commentary/governors-prioritize-expanding-internet-access-for-k-12-students/

Vehovar, V., Sicherl, P., Hüsing, T., & Dolnicar, V. (2006). Methodological challenges of digital divide measurements. *The Information Society*, *22*(5), 279–290. https://doi.org/10.1080/01972240600904076

Velicer, W. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327.

Walker, K. (2021). *tigris: Load Census TIGER/Line shapefiles* (version 1.5). [R Package]. https://cran.r-project.org/web/packages/tigris/index.html

Warf, B. (2013). Contemporary digital divides in the United States. *Tijdschrift Voor Economische En Sociale Geografie*, *104*(1), 1–17. https://doi.org/10.1111/j.1467-9663.2012.00720.x

Warschauer, M. (2003). Demystifying the Digital Divide. *Scientific American*, *289*(2), 42–47. https://www.jstor.org/stable/26060401

Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Whitacre, B. E., & Mills, B. F. (2007). Infrastructure and the rural—urban divide in high-speed residential internet access. *International Regional Science Review*, *30*(3), 249–273. https://doi.org/10.1177/0160017607301606

Whitacre, B., & Rhinesmith, C. (2015). Public libraries and residential broadband adoption: Do more computers lead to higher rates? *Government Information Quarterly*, *32*(2), 164–171. https://doi.org/10.1016/j.giq.2015.02.007

Wickham, H. (2021). *tidyverse: Easily install and load the 'Tidyverse'* (Version 1.3.1) [R Package]. https://cran.r-project.org/web/packages/tidyverse/index.html

World Bank Group. (2022). *6.2 Assessing the broadband demand gap*. Broadband Strategies Toolkit Module 6 Driving Demand. https://ddtoolkits.worldbankgroup.org/broadband-strategies/driving-demand/assessing-broadband-demand-gap

World Economic Forum. (2014). *Delivering digital infrastructure: Advancing the internet economy*. http://reports.weforum.org/delivering-digital-infrastructure/introduction-the-digital-infrastructure-imperative/?doing_wp_cron=1610818305.8569428920745849609375

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Yu, L. (2006). Understanding information inequality: Making sense of the literature of the information and digital divides. *Journal of Librarianship and Information Science*, *38*(4), 229–252. https://doi.org/10.1177/0961000606070600