Proceedings of the 56th Hawaii International Conference on System Sciences | 2023

Spatial and Socioeconomic Analysis of Purposeful Mobile Internet Use in US States

Avijit Sarkar
University of Redlands
avijit_sarkar@redlands.edu

James B. Pick
University of Redlands
james_pick@redlands.edu

Geeta Bhat
University of Redlands
geeta_bhat@redlands.edu

Abstract

The digital divide in the United States has received renewed attention during the COVID-19 pandemic. As achievement of digital equity remains a high priority, this study examines spatial patterns and socioeconomic determinants of the purposeful use of mobile internet for personal and business needs in US states. Agglomerations of mobile internet use are identified using K-means clustering and the extent of agglomeration is measured using spatial autocorrelation analysis. Regression analysis reveals that mobile internet use is associated with employment in management, business, science, and arts occupations, affordability, age structure, and the extent of freedom in US states. Spatial randomness of regression residuals shows the effectiveness of the conceptual model to account for spatial bias. Implications of these findings are discussed.

Keywords: Digital divide, Mobile internet, Regression, Clustering, Spatial Autocorrelation

1. Introduction

The COVID-19 pandemic has renewed attention to the digital divide in the United States. Over the past two years, the pandemic has highlighted that high-speed internet access is a necessity for Americans in all strata of society to communicate with each other, maintain social interactions, attend work and school online, conduct online shopping, access healthcare, entertainment, amid myriad forms of personal, social, and business-related activities.

During the pandemic, mobile internet (MI) adoption has grown worldwide. To stay connected with friends and family, MI has been used for instant messaging, making or receiving calls including video calls, and social networking. MI has enabled people to read the news, search for information online, watch video, play games, listen to music, among various forms of hedonic activities. MI also enabled users to order, purchase, and sell goods online, access online banking, pay bills, access mobile money among various forms of business-related activities. MI has brought health information, job-related information, and government services to the fingertips of internet users. Studies have shown that the frequency of engaging in such activities at least once a day, week, or month such increased for every one of these activities from 2019 to 2020 [9], showing the relevance and necessity of MI spurred by the COVID-19 pandemic.

Globally, 51% of the world's population or 4 billion people used MI by the end of 2020 [9]. In the United States (US), the base of smartphone users expanded from 80 million to over 220 million between 2011 and 2021. Since 2015, the proportion of US households that used a mobile data plan grew from 64% to over 74% (98 million households) at the end of 2021. This growth in the base of smartphone users and mobile data plan users has however not been equitable, with disparities across demographic groups, socioeconomic status, as well as geographic location. Disparities have been reported in latest NTIA data [18] based on age, gender, education, race and ethnicity, and population density. Globally, key barriers to MI adoption and use are lack of access to devices, networks, and services, knowledge and skills deficits, affordability, and relevance of content [9].

Despite these barriers, people have used MI for a broad spectrum of online activities. In this context of purposeful use of MI, it is essential to analyze the disparities and gaps between "individuals, households, businesses and geographic areas at different socioeconomic levels with regard both to their opportunities to access information and communication technologies and to their use of the (mobile) Internet for a wide variety of activities [19, p.5]". In recent years, a handful of digital divide studies have started to focus on the use of the internet for various activities and disparities therein [17, 27, 30, 37]. Apart from these studies, there is a gap in prior work in systematic examination of MI based purposeful activities in the US. This study aims to fill this gap. The objective of this study is to analyze spatial patterns of MI usage for

URI: https://hdl.handle.net/10125/103258
978-0-9981331-6-4

personal and business activities and determine the influences of demographic and socioeconomic factors on MI use.

The research questions are: (1) What are the spatial patterns and agglomerations of purposeful MI use for personal and business needs in US states, as measured by spatial autocorrelation? (2) How is purposeful use of MI clustered in US states and what are the demographic and socioeconomic attributes of such clusters? (3) What are the influences of demographic, occupational, affordability, social capital, innovation, and societal openness factors on MI use in US states? Analysis of spatial patterns and agglomerations is essential to understand the influence of geography on the digital divide overall, specifically for MI use. Without accounting for such agglomerations, the determination of demographic and socioeconomic influences on MI use is likely to be affected by spatial bias. Finally, the study is conducted at the state level for the US. As evident from NTIA reports over many years, the state unit of analysis provides meaningful insights about the digital divide that can help policymakers shape their own state's telecommunications and digital initiatives, priorities, and policy.

As digital equity and inclusion issues increasingly come to the forefront, their spatial underpinnings are often ignored by digital divide researchers. In addition, there is very little research on purposeful use of mobile internet within the digital divide literature. This paper contributes to the growing body of literature on actual use of the internet, particularly MI, rather than its access. In doing so, it comprehensively examines spatial patterns and disparities in MI use. These are the distinguishing features of this paper. The remainder of this paper is organized into sections on literature review of technology use, particularly the internet, conceptual model of purposeful MI use for personal and business purposes, spatial patterns of MI use, regression findings, policy implications, limitations, and conclusions.

2. Literature Review

There has been growing attention to research on the digital divide and digital inequalities. The persistence of unequal purposeful MI use has become even more visible during the covid-19 pandemic. The review groups prior research into subsections on technology use at the state level, then county and individual levels, purposeful use at the county and individual levels, and studies of racial and ethnic differentials on the divide.

2.1 Technology use at the state level

For US states, a study [20] based on data from 2007-2010 was analyzed based on the Spatially Aware Technology Utilization Theory (SATUM) theory, which is discussed later on in this paper. The study analyzed eight dependent variables of ICT use, namely desktop/laptop, internet access, broadband adoption, cellphone-only use, high-speed wireless devices, fixed phone use, and Facebook and Twitter use. Based on independent variables of demographics, race/ethnicity, economy, education, innovation, societal openness, and social capital, there were strong positive effects from college education, social capital, and mixed effects for race & ethnicity factors, and the dependent variables were spatially agglomerated except for social media.

A study based on the US Census American Community Survey and centered on the household level in 2018, provided insights that there was nearly full access of households to the internet, while broadband or smartphone presence was 85 percent, while there were substantial differentials by age (inverse), urban versus rural, higher income, and varied means of accessing the internet [13]. The study emphasizes the huge US expansion in computer and internet use from 9% in 1984 to near saturation of computer use in 2018.

2.2 Technology use at county & individual levels

Research on determinants of technology use at the county and individual levels in the US has revealed considerable geographic agglomeration and determinants of county technology uses [11, 29]. Based on a large US sample of geo-referenced tweets and photos, a study mapped the patterns of use nationwide and conducted a more intensive case study and examined correlates of social media use in California [11]. Results showed differing geographic concentration for the two social media entities, with tweets heavily concentrated in urban areas of the northeast, mid-Atlantic and south, whereas high photo density occurred in areas of tourism and technology such as West Coast cities, Lake Tahoe, Yosemite, Austin, Orlando, Ann Arbor, and Boston. For California, tweets were associated with education, income, and graduate study in professional-science-arts, while photos were related to white and Asian ethnicity and graduate study in management, science, and arts.

At the county level, studies have provided insights into geographic patterns and determinants of use for ICT use and mobile ICT use [29]. There was generally concentration of high use in the counties of the megalopolis from Boston to Washington, the West Coast, western Colorado and Utah, with low-usage in the mid to lower South, with the exception of the Atlanta metropolitan area and most of Florida. Regression analysis indicated determinants of ICT use to be demographic attributes, urban location, service occupation, and mixed findings for ethnicities.

2.3 Purposeful Use: County & Individual Levels

As technology uses approach saturation in the US, gaps still loom gaps in the purposeful uses of ICT. For instance, an underserved community may receive broadband throughout its area from federal investment, but community users may not have the training, education, or financial resources to engage in purposeful uses, for example to access scientific journals, financial services, etc. Purposeful technology use in the US has recently become an area of research interest, stimulated by the online needs of covid-19 pandemic.

Studies of use of e-entertainment services in US counties [28] and of social media uses in US counties [23] examined specialized purposeful uses, based on the SATUM model, with data drawn from nationwide surveys and government sources. For the e-entertainment purposeful variables of "obtained the latest news," "adding video to website," "watching a movie online" and "ordering iTunes from a website," there were fairly consistent determinants including young dependency ratio, education, working age population, service occupations, and Asian ethnicity.

Research on the purposeful uses of the internet across a sample of middle and low level economies was based on extensive surveys by GSMA [5], The study indicated, for 198 nations, that although the "coverage gap," i.e. gap in extent of MI broadband coverage, indicated gap narrowed from 24% to 6% for 2014-2020, while the "usage gap", i.e. gap in use of mobile services, remained even at 43%. This underscores the justification in the present study to focus on purposeful uses, which ultimately can account for narrowing of the usage gap. The report [5] analyzed leading purposeful uses for survey respondents, finding substantial ranges of adoption of services, for example, whereas 79% of MI users looked for information in the search bar or app, while 61% changed settings of data usage limit. GSMA's survey of MI use among internet users globally is another example of a study focused on MI usage [9].

2.4 Influence of Racial and Ethnic Differences on digital divides

Prior US digital divide studies have included race and ethnicity variables, but only a few have drilled down to try to gain deeper and detailed understanding of why the differences exist. In a survey study of individuals in Chicago in 2013, findings showed that the MI use is strongly associated with political and economic activities online, with the largest effects for African-American and Latino respondents, particularly for Latino respondents who live in Latino neighborhoods [17]. Using multilevel analysis, the study confirmed that mobile access led to 75-80% increases in the probable number of economic and civic activities for Latinos, compared to 40-45 percent increase for Blacks [17].

Another study examined whether the internet use of older population in urban and rural areas varied by racial and ethnic group [4]. The study found that older Black and Hispanic individuals had lower odds of using the internet, and further that rural living lowered the probability of internet use more for Blacks than for Whites. It suggests targeted programs to reduce the digital divide especially of older Black people living in rural areas. These relatively rare studies of racial and ethnic internet use differentials can be helpful in explaining findings in the present study.

3. Conceptual Model: Purposeful Use of MI

The digital divide on the theoretical side has had several conceptual frameworks developed, but there is not presently a standard framework. Among the conceptual models that were considered for the present study are Adoption-Diffusion Theory [24], Van Dijk's theories [34,35], the Unified Theory of Acceptance and Use of Technology (UTAUT) [26], and SATUM [21]. Among its pluses are that the conceptual model was developed specifically for quantitative digital divide research, and it posits to analyze, for Internet and ICT dependent variables, the combined associations on the dependent variables of a variety of independent variables including demographic, occupational, economic, educational, innovation, affordability, freedom, and social capital influences. It also includes geographical mapping, cluster analysis, the results of which can be spatially rendered, and spatial autocorrelation. It is more adaptable than some of the other theories, such as Adoption-Diffusion and van Dijk's, to make use of data for governmental spatial units of analysis [22]. For these reasons, the SATUM theory is adopted of the present investigation. Due to limited space, readers are referred to [22] for a more detailed comparison of the pluses and minuses of SATUM compared to the other theories mentioned.

The remainder of this section justifies the independent variables based on induction from prior studies or reasoning by the authors. This process of inducing from the literature and reasoning to formulate factors to include in the model is justified for an exploratory study, such as the present one [31]. The current SATUM model is conceptualized in Figure 1. The independent variables are grouped into factors on the left. The model justification is given next for these independent variables and dependent variables, while the SATUM methods, in the center of the figure, are presented in Section 4.

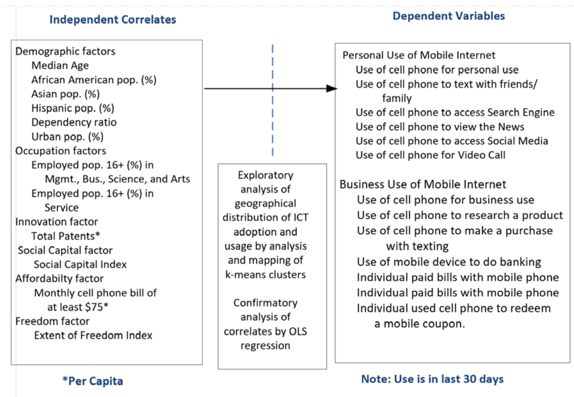

Figure 1. Conceptual Model based on SATUM.

3.1 Justification for independent variables

Demographic variables

- Median Age: This variable has been utilized frequently in digital divide research and generally is found to have inverse effect [10,13,23,26,28].
- % of Pop. that identifies as African-American, Asian, and Hispanic or Latino: These racial and ethnic variables have been included in prior studies and often been significant in the findings [4,13,17,23,26,29]. Generally, % Asian is a positive correlate, while African American and Hispanic/Latino variables are usually mixed in direction of effect.
- Age Dependency Ratio (% of Pop under 19 and 65+): Dependency ratio for young age only has appeared in several studies and been an important factor [23,28,29].
- Percent of Urban Population: Percent urban has consistently been a positive factor in several prior studies [4,13,20,26,28, 29].

Occupational variables

- Percent of the Employed Pop Age 16+ in the Management, Business, Science, and Arts (MBSA) Occupations: The association of workforce in these “creative” jobs with higher tech cities has been put forward and justified by the work of Richard Florida [7]. Variables for scientific and technical workforce have been studied in the U.S. and sometimes found to be significant positive correlates of technology use [2].
- Percent of the Employed Pop Age 16+ in Service Occupations: Employment in service occupations has occasionally been included in the literature, and often has positive and significant effects when included [23,26,28,29]. It can be reasoned that service occupations include job requirements for computer skills and use, often intensive use.

Innovation – Total Patents divided by the Total Pop. of the State: Innovation is reasoned to stimulate technology and computer use both in technology companies, universities and research centers, but also for communities interacting with these technology hubs.

Social Capital variable: Social capital has been an important variable in some U.S. digital divide studies. It represents the linkages and ties for social groups within a population through physical and communication means. It was a major factor in a survey sample of technology use by individuals in the US [3], as well as in a nationwide study of the decisions of people to go online, including from influences of peers [1].

Affordability – Monthly Cell Bill is $75+: Affordability has been shown sometimes to impact adoption to technology, and tends to be more influential in developing countries or economically weaker segments of developed nations, such as India and Japan [21, chapters 5,6]. We reason that pricing of internet services can have an inverse influence on individuals, especially in lower income households [18].

Freedom – Extent of freedom in US states: Freedom has been shown to be important in a limited number of studies. It was seen the Arab Spring, an opening in many Arab countries in 2010 and 2022, in which feelings of freedom were associated with increased internet communications. It was also shown to be important in Africa for a study of 51 countries [21, chapter 9], in which freedom was represented by the proxy of laws related to ICT.

3.2 Justification for Dependent Variables

The dependent variables are not individually justified by literature since there is very little quantitative/geographical prior research on purposeful use of indicators. The justification for including them goes back to the saturation in recent years in simple use of technologies in the US and the need to spring forward and conduct research on purposeful use and impacts [35]. The need to segue to purposeful use was initiated by some research a decade ago [37]. Today, given the rising needs of purposeful use during the covid-19 pandemic [14], we feel justified in reasoning that purposeful use is becoming the new digital divide in the US. This explains the exploratory inclusion of a set of purposeful personal uses and purposeful business uses. Personal Use of MI dependent variables include the following. Individual uses Cell Phone for: (1) personal use (2) to text message friends/family (3) to access a search engine (4) view the News on cell phone (5) access Social Media, and (6) to make a video call. Business Use of MI dependent variables include the following. Individual used cell phone (7) for business use, (8) to research product, (9) to make purchase with text, (10) for online banking, (11) to pay bills online, and (12) to redeem a mobile coupon.

4. Methodology

Descriptive statistics were computed for all 12 dependent indicators of purposeful MI use, as well as 12 independent variables for a sample of n = 49 (lower-48 states plus Washington D.C.). Bivariate correlations were computed for each pair of independent variables as a preliminary screening for multicollinearity and variables such as median household income and proportion of population with Bachelors education were eliminated due to statistically significant high correlation with MBSA occupation. Next, each of the 12 dependent variables were mapped using a Geographical Information System (GIS). GIS mapping provides visual cues about spatial distributions of MI usage and point to similarities and differences between states and well as between the variables themselves. Due to space limitations, maps of MI use are not provided but spatial patterns are discussed in Section 6.

Following diagnosis of spatial autocorrelation, K-means clustering is deployed to determine two separate clusters of MI use for (i) personal and (ii) business purposes. In each case, k-means is applied for the dependent variables, with k=5 and k=6. K=6 resulted in more meaningful clusters. States are assigned to clusters based on the levels of MI use, from high to moderate to low. Cluster centers are used to determine the ratio of MI use between the highest and lowest use clusters. This ratio indicates the extent of disparity in MI use between US states. Clusters are characterized based on the demographic and socioeconomic attributes and similarities and differences are observed. The clusters are also mapped and descriptively reveal spatial agglomerations of MI use.

To diagnose whether purposeful MI use in US states show statistically significant patterns of agglomeration of high and low values, or if such patterns are spatially randomly distributed, the Moran's I test statistic [15] is computed as follows:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j}\, z_i\, z_j}{\sum_{i=1}^{n} z_i^2}$$

where z_i is the deviation of an attribute for feature *i* from its mean ($x_i - \bar{x}$), $w_{i,j}$ is the spatial wight between features *i* and *j*, *n* is the total number of features, and S_0 is the sum of all spatial weights. Moran's I measures the extent of spatial autocorrelation of MI use in US states. The Moran's I test is inferential; the null hypothesis is that the values of a variable are randomly distributed spatially. The test statistic ranges in value between –1 and +1. Moran's I statistic value close to 0 for a dependent variable (MI use) would indicate spatial randomness while values close to –1 and +1 indicate the presence of spatial bias for a dependent variable that needs to be accounted for while examining associations of independent variables with the dependent variable in question. Interpretation of Moran's I is performed using the p value for statistical significance (if p is not significant, the variable is randomly distributed spatially). Further, if the Z score is positive, the values of a variable are more geographically agglomerated (high values located near high ones and low values near low ones). If it is negative, the spatial pattern resembles a "checkerboard" pattern, in which high values are surrounded by low ones and vice versa [15,16]

Finally, Ordinary Least Squares (OLS) regressions are employed to test posited associations between the 12 dependent variables and 12 independent variables. OLS regressions were conducted in stepwise fashion allowing in variables with significance levels of equal or less than 0.05. The Variance Inflation Factor (VIF) was computed as an additional test of multicollinearity, and a cutoff of 5.0 for VIFs was used. Each of the 12 OLS regressions had a VIF lower than 5.0 and no multicollinearity problems were detected. Moran's I of regression residuals were also computed to check if the model accounted for any spatial bias present in the spatial distribution of the dependent variables.

5. Data

Data on all 12 dependent indicators of MI use were sourced from Esri's 2021 Market Potential database [6] which measures the likely demand for products and services among US consumers as well as consumer attitudes on media consumption, internet activities, cell phones and service, etc. Market potential is estimated by Esri based on data collected from various surveys such as the MRI Survey of the American Consumer and Doublebase 2020 Survey from MRI-Simmons. It is important to note that Esri's 2021 Market Potential database documents changes in consumer demand and behavior due to the COVID-19 pandemic. The 12 dependent variables include six indicators of MI use for personal purposes and six indicators for business purposes.

Data on 7 out of the 12 independent variables are sourced from the U.S. Census Bureau's American Community Survey. 5-year estimates (2016-2020) centered on 2018 were used for these independent variables. Data on urban population was sourced from the U.S. Census Bureau's 2010 Census. While it would have been ideal to capture urban population data from the 2020 Census, it is still unavailable. Data on patents issued to residents in U.S. states in 2021 were sourced from the Performance and Accountability Report of the U.S. Patent and Trademark Office (USPTO) and was used as an indicator of innovation. Data on freedom in U.S. states was obtained from the Freedom in the 50 states project report [25]. Data on social capital in U.S. states were obtained from a report commissioned by the Joint Economic Committee of the U.S. Congress (US Senate, 2018). Details about the social capital index and its research design methodology can be found in the report [33]. Finally, data on monthly cellphone bill exceeding $75 in U.S. households was compiled at the state level from Esri's 2021 Market Potential database [6] as an indicator of affordability.

The variable estimates were normalized based on 2018 population in U.S. states. This ensures time simultaneity since the dependent variables are for the year 2021. Finally, Alaska and Hawaii are excluded from the study due to missing data. Variable descriptions and descriptive statistics are in Table 1.

Table 1. Variable definitions & descriptive stats.

Description of Variable (n=49)	Min	Max	Mean	SD
Dependent Variables: Personal Use				
Individual uses Cell Ph for personal use	0.5307	0.6102	0.5648	0.0165
Individual used cell ph to text message friends/family	0.5709	0.6562	0.6144	0.018
Individual used cell ph to access a search engine	0.4272	0.5597	0.4899	0.0246
Individual has viewed the News on cell ph	0.3528	0.4751	0.4044	0.0261
Individual used cell ph to access Social Media	0.358	0.4413	0.3916	0.0145
Individual used cell ph for video call	0.2403	0.3485	0.2738	0.017
Dependent Variables: Business Use				
Individual uses Cell Ph for business use	0.1257	0.1742	0.1511	0.0102
Individual Used cell ph to Research Product	0.3201	0.4122	0.3652	0.0187
Individual used cell ph to make purchase w/ text	0.26	0.3438	0.2986	0.0159
Individual used mobile device to do banking	0.2171	0.3004	0.2586	0.0179
Individual paid bills with mobile phone	0.1736	0.2136	0.1908	0.0085
Individual used cell ph to redeem a mobile coupon	0.1414	0.1917	0.1581	0.0105
Independent Variables				
Median Age	31.1	44.8	38.6224	2.3717
% of Population - African-American	0.0056	0.4539	0.1159	0.1064
% of Population - Asian	0.0079	0.1483	0.0359	0.0278
% of Population - Hispanic or Latino	0.0159	0.492	0.1219	0.1049
Age Dependency Ratio, % of Pop under 19 & 65 +	0.4975	0.785	0.7086	0.0472
Urban Pop (%)	0.387	1	0.739	0.149
Pop 16+ in the MBSA Occupn (%)	0.1381	0.3565	0.1911	0.0359
Pop 16+ in Service Occupn (%)	0.0699	0.1182	0.0827	0.0076
Innovation: Patents per capita	0.0001	0.0024	0.0008	0.0005
Social Capital index	0.0001	4.2333	2.1514	1.0178
Affordability: Monthly Cell Bill is >$75	0.5248	0.6164	0.5699	0.0199
Index of freedom	0	1.381	0.9083	0.2606

6. Spatial Patterns and Clusters of MI Use

In 2021, personal use of MI varied between 53% to 61% of the population in US states, with states in the Northeast and Pacific Northwest leading the nation. Surprisingly, West Virginia (WV) in the Appalachian region is a leading state in MI use for personal purposes. Due to the remoteness of the region and infrastructural malaise, WV has often been a laggard in internet access and use [18]. Personal use of MI was also found to be reasonably high in states along with Atlantic Coast, a handful of prairie states, and also in the Rocky Mountain state of Colorado. Conversely, MI use for personal purposes was low in the South and Southwest in states such as Texas, Louisiana, Mississippi, Georgia, and also in California and Arizona. Low-moderate levels of use was also found in some states in the Midwest and also in the South. For MI use for business purposes, the leading states were in the Northeast, Pacific Northwest (Washington state), and also Colorado among Rocky Mountain states. Several of the prairie and midwestern states has moderate levels of use for business purposes, while states in the south and west had low to very low levels of MI use for business needs. It is pertinent to note that overall, business use of MI (ranging between 12.6-17.4%) significantly lags personal use in US states in 2021.

K-means clusters (k=6) of MI use for personal purposes shows that Washington D.C. as the sole member of cluster 1 (Table 2), representing highest use, followed by 9 states in cluster 2, which represents high use. Six of these states are in the Northeast in the Boston-Washington megalopolitan area, in the Pacific Northwest (Washington), Rocky Mountains (Colorado), and upper Midwest (Minnesota). A total of 20 states, 5 in cluster 3 and 15 in cluster 4 comprise moderate use of MI for personal purposes. These states are agglomerated all over the country including the Northeast, West, prairies, Midwestern rust belt, and the South. Cluster 5 is comprised of 17 low use states located in the Midwest, Appalachia region, and South. Finally, Louisiana and Mississippi are in cluster 6, representing the lowest levels of MI use for personal purposes. The ratio of highest to lowest use varies between 1.11 to 1.40 indicating that that the extent of disparity between leaders and laggards is not too high.

K-means clusters (k=6) of MI use for business purposes (Figure 2) show strong similarity with the corresponding clusters for personal purposes. Highest use is in Washington D.C. with a 100% urban population in cluster 1, followed by six states in cluster 2. Moderate use of MI for online banking, bill payment, mobile-based purchase, etc. is found in 12 states in cluster 3, followed by moderate-low use in 20 states in cluster 4. Cluster 3 states are predominantly in the North and Northwest, Upper Midwest, and the Northeast.

Cluster 4 states are found all over the country in the Midwest, South and Southeast, and also in the West. Low to very low use states (10 total) in clusters 5 and 6 are in the rural Appalachia, deep South, and West (New Mexico). The ratio of highest to lowest use cluster centers varies between 1.21 to 1.39 indicating that the extent of disparity between the leading and lagging states is not too high.

Table 2. K-means clusters of MI use, personal purposes, 2021.

Cluster Centers - MI Use, Personal Use, K=6									
	1 Highest MI Use	2	3	4	5	6 Lowest MI Use	MAX	MIN	Ratio = Highest/Lowest
Cell Phone for personal use	0.595	0.570	0.565	0.563	0.558	0.538	0.595	0.538	1.105
News on cell phone	0.475	0.439	0.425	0.403	0.382	0.363	0.475	0.363	1.310
Social Media on cell phone	0.441	0.407	0.402	0.391	0.381	0.365	0.441	0.365	1.210
Access search engine on cell	0.560	0.518	0.512	0.489	0.471	0.439	0.560	0.439	1.274
Text friends & family	0.644	0.628	0.644	0.611	0.604	0.577	0.644	0.577	1.116
Make video call	0.349	0.290	0.276	0.276	0.261	0.249	0.349	0.249	1.402
# of States	1	9	5	15	17	2			
	District of Columbia	Colorado	Delaware	Arizona	Alabama	Louisiana			
		Connecticut	Maine	California	Arkansas	Mississippi			
		Maryland	New Hamps	Florida	Indiana				
		Massachusett	Oregon	Georgia	Iowa				
		Minnesota	Vermont	Idaho	Kansas				
		New Jersey		Illinois	Kentucky				
		Rhode Island		Nevada	Michigan				
		Virginia		New York	Missouri				
		Washington		North Caroli	Montana				
				North Dakot	Nebraska				
				Pennsylvani	New Mexico				
				Texas	Ohio				
				Utah	Oklahoma				
				Wisconsin	South Carolina				
				Wyoming	South Dakota				
					Tennessee				
					West Virginia				

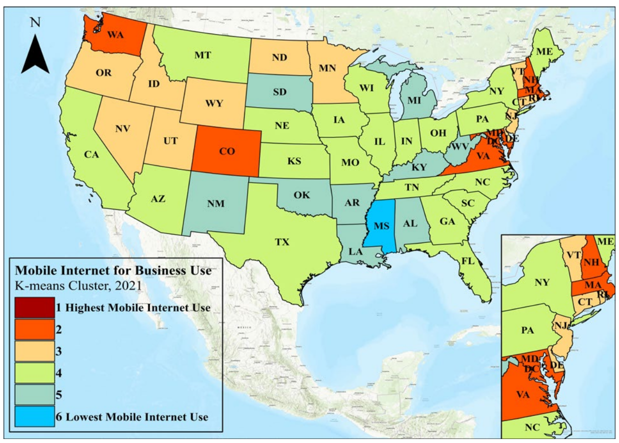

Figure 2. K-means clusters of MI use, business purposes, 2021.

There are a few noticeable differences between the clusters of MI use for personal and business purposes. States in the moderate and low use clusters are comparatively less urban than highest and high use clusters with an almost 17% difference in urban population per capita for business use. On the occupational front, the proportion of population engaged in MBSA occupations is highest is clusters 1 and 2 at 36% and 22% respectively, but progressively decreases to 15% in the lowest cluster 6, for personal use. Similar gaps are found for business use as well. Some other differences such as patents issued per capita are also observed between the high and low use clusters. Overall, cluster analysis findings point to some demographic, occupational, and innovation-related underlying factors in MI use differences for both personal and business purposes.

It is important to note that the clusters of both personal and business use show states that are spatially contiguous. This spatial contiguity of states within clusters is in alignment with the first law of geography, which states "Everything is related to everything else, but near things are more related than distant things [32]." This spatial arrangement has implications; the agglomeration of states that are similar to each other in terms of their MI use points to the presence of spatial bias, which manifests itself in the form of spatial autocorrelation (Longley et al., 2015). Spatial autocorrelation analysis shows moderate to high levels of agglomeration of each of the 12 dependent variables, with Moran's I ranging from 0.416 to 0.653 (the bottom of Table 3), significant at the .001 level. This confirms the presence of spatial bias which needs to be accounted for during regression analysis.

7. Determinants of Mobile Internet Use

OLS regression results (in Table 3) reveal that the dominant correlates of MI use are MBSA occupations (MBSA) among the occupational variables, affordability (measured by the proportion of population whose monthly cell phone bill exceeds $75), and median age and age dependency ratio among the demographic variables. MBSA occupation is found to be positively associated at the .01 level or lower with 10 out of the 12 dependent variables. MBSA include a broad set of occupations including management, business, and finance, computer and mathematical, engineering, architecture, healthcare, and life, physical, and social sciences. Those engaged in such occupations are often highly educated and high earners (Pearson correlation coefficients of MSBA with Bachelors education of .958, and with median household income of .846, both significant at the .001 level). It is likely that such individuals are likely to be more skilled internet users and their professional and personal needs spur MI use for both personal and business purposes.

Affordability is positively associated with 11 out of the 12 dependent variables (at .001 level) indicating that as the proportion of population in US states with monthly cell phone bill exceeding $75 increases, MI use tends to increase. Since higher cell phone bills are often associated with high volume data plans, it is likely that those with such plans are more likely to engage in a broad spectrum of activities, particularly those that consume higher volumes of data such as watching and accessing the news, making video calls, and researching products and services online. Age dependency ratio is inversely associated with 10 out of the 12 dependent variables indicating that as the proportion of those younger than 19 years old and older than 65 years old increases, MI use tends to decrease. This combined with the positive association of median age with the dependent variables points to the dominance of cohorts in age groups 20-64 years with MI use. Viewed in another way, this cohort, also referred to as the working age population appears to drive MI use for both personal and business purposes.

Median age is positively associated with 6 out of the 12 dependent variables, of which 4 are in personal use category. This finding is somewhat surprising since higher age has often been perceived to lower technology adoption and usage. However, each of the six dependent variables – using cell for personal use, viewing the news, accessing a search engine, text messaging, mobile banking, and researching products online have become ubiquitous and dominant forms of MI use during the pandemic across age groups [9].

The three dominant correlates are followed by the extent of freedom, which is found to be positively associated with 5 out of the 12 dependent variables. The index of freedom, comprised of personal, fiscal, and regulatory components points to the overall importance of societal openness in relation to MI based activities. This is a novel finding for the US digital divide, particularly in light of personal freedom oriented societal discourse in the United States during the COVID-19 pandemic. Apart from freedom, race and ethnicity is found to have inverse association with a small set of dependent variables, particularly the indicators of business use. Such inverse associations of African American and Hispanic segments of the population are consistent with lower levels of technology (internet) access, adoption, and usage among these race and ethnic groups. That said, increases in internet connectivity in minority race and ethnic groups have been reported between 2019-2021 (Goldberg, 2022). While the urban-rural digital divide in the United States has been documented in several contexts, urbanization is a not a dominant correlate for MI use. OLS regressions also reveal limited association of social capital with the dependent variables pointing to the limited role of social bonding and transfer of skills between cell phones users using MI. Lastly, service occupation and patents issued per capita are found to have no association with any of the dependent variables.

The Variance Inflation Factor (VIF) for each of the regressions is lower than a cutoff of 5.0 indicating that multicollinearity is not of concern. Regression diagnostics (Joint-Wald statistic is significant while Koenker and Jarque-Bera statistics are not significant) indicate that regression assumptions have been met. Finally, our conceptual model of MI use explains 66.2-93.7% of the variation in the dependent variables, showing the robustness of the proposed model. Moran's I of the regression residuals (bottom of Table 3) are lower than the Moran's I for each of the 12 dependent variables, indicating spatial randomness of the residuals. In fact, absolute values of Moran's I are reduced by 41%-86%, with an average reduction of 57.13%. Based on this evidence, it can be concluded that the conceptual model is able to account for the spatial bias present in the dependent variables.

Overall, OLS regressions findings shed light on the importance of economic (MBSA occupation), affordability, and demographic (age-related) factors for purposeful MI use. Regressions findings also reveal that societal openness also influences MI's purposeful use. The limited influences of traditional demographic factors such as race/ethnicity and urbanization are also revealed. The findings are also largely similar for both types of purposeful use of MI – personal and business.

8. Discussion and Implications

There is overall contrast between personal use and business use of the MI. Although there are similarities in the generally much higher relative use of the internet in the north of the US compared to the South, there are some fine spatial differences that are revealing. Consider first that the range of personal use is lower than for business use. This can be observed by looking at the legends for the personal compared to the business use. The high category for personal use, compared to the low category, is about 10 percent higher, whereas for business use the high category is only 30 percent higher. This implies that personal use of cell phones is more even in percentage of use throughout the nation, which might relate to the near saturation of cell phones in the US population, as seen in the 85% presence in households of internet subscriptions in 2018 [13], a figure that is expected to be higher in 2021 due to greater need to use the internet during the pandemic [14]. On the other hand the greater percentage variation in business use nationally might be the result of a broad range of pandemic impacts on businesses, many of which scaled back on workforce and capital expenditures.

K-means cluster analysis for personal use and business use resemble each other more closely than the contrast just discussed for the individual personal and business variables. In both cluster analyses, the very highest singleton cluster is Washington DC, which reflects the very high dependency on MI there. From the standpoint of explanation, Washington DC has high education, is 100% urban and has high concentrations of MBSA occupations. It is interesting also this DC has a high proportion of African American population, which reflects improvements in racial/ethnic MI usage. At the low end of clusters is Louisiana and Mississippi for personal MI use and Mississippi for business MI. These states are known to have very low technology and internet utilization from other studies [20]. There are mostly strong geographic agglomerations following Tobler's Law. This includes at the high end (cluster 2) the agglomerated areas in the Boston to Washington megalopolis, although for business use, cluster 2 splits into two pieces. For both cluster maps there are several isolated cluster 2 states, in particular, for personal use, Washington state, Colorado, and Minnesota, and for business use, Washington and Colorado.

The regression findings reveal as determinants many aspects that are heretofore unreported. This includes the positive effect of median age, which is opposite to many other studies, but may reflect pandemic influences on evening out age differences. The strong effect of MBSA occupations may relate to its close correlation with education and it also confirms the results on creative occupations being tied to technology-based cities [7]. The strong association with monthly cell bill of at least $75 has rarely been reported in US studies and points to the need to have future research to determine why it has appeared so prominently in 2021. The lack of association for urban population for both personal and business dependent variables reflects a leveling of the geographic spread of the MI, while also possibly is explained by pandemic-related movement of people and household out of urban areas. Again, the explanation points to need for further research emphasizing fine points of urban geographies.

9. Conclusions and Future Research

This study has analyzed MI use for US states using a variety of survey and government data. The dependent variables are for year 2021, which reflects the influence of the covid-19 pandemic. The conceptual frame is SATUM, which includes spatial and multivariate analytics. The dependent variables reflect purposeful use of mobile devices and groups the usage into five variables each for personal and business usage. The independent variables are either induced from the digital divide literature or reasoned by the investigators, an approach appropriate for exploratory research.

The findings indicate distinctive and, in some ways, new geographical patterning of dependent variables and k-means clusters. There are new patterns that emerge such as the states with big metro areas being classified in the lowest cluster category for personal use and the distinctive cluster leadership of Washington DC in MI purposeful use. Some of the mapping subtleties are not immediately interpretable and need further research.

The OLS regression analysis is surprising in the positive effect of median age, and the strong effects of MBSA occupation, monthly cell phone bill of at least $75, inverse effect of dependency ratio, and impact of the Freedom Index, especially for personal use, while the common-place effect of urban has little effect. The study has the limitation of using the state geographic unit, which might obscure more detailed within-state variations. It also does not do a direct comparison of 2021 dependent variables with the same set of pre-pandemic ones. It is hoped that this exploratory study will stimulate further research to explain some outcomes and associations that are unclear in interpretation based on the research literature. It is also hoped that future research will expand the range of purposeful uses and begin to sort out patterns and groupings of the many purposeful uses available.

10. References

[1] Agarwal, R, Animesh, A., and Prasad, K. 2009. Social interactions and the digital divide: Explaining variations in Internet use. Info. Systems Research 20:277-294.

[2] Azari, R., & Pick, J. B. (2005). Technology and society: socioeconomic influences on technological sectors for United States counties. International Journal of Information Management, 25(1), 21-37.

[3] Chen, W. 2013. The implications of social capital for the digital divides in America. The Information Society 29:13-25.

[4] Choi, E.Y., Kanthawala, S., Kim, Y.S., and Lee, H.Y. 2022. Urban/rural digital divide exists in older adults: does it vary by racial/ethnic groups? Journal of Applied Gerontology 41(5):1348-1356.

[5] Delaporte, A., and Bahia, K. 2021. The State of Mobile Internet Connectivity, 2021. London, England, GSMA. Retrieved from https://www.gsma.com/r/wp-content/uploads/2021/09/The-State-of-Mobile-Internet-Connectivity-Report-2021.pdf

[6] Esri. 2021. Esri Market Potential Methodology Statement. https://storymaps.arcgis.com/stories/4c91b3b9228f4036862aad3ea4187bfa.

[7] Florida, R. 2012. The Rise of the Creative Class Revisited. New York: Basic Books.

[8] Goldberg, R. 2022. New NTIA data show enduring barriers to closing the digital divide, achieving digital equity. Available at https://ntia.gov/blog/2022/new-ntia-data-show-enduring-barriers-closing-digital-divide-achieving-digital-equity.

[9] GSMA. 2021. Connected society: The state of Mobile internet connectivity 2021. Retrieved from https://www.gsma.com/r/wp-content/uploads/2021/09/The-State-of-Mobile-Internet-Connectivity-Report-2021.pdf.

[10] Lenhart, A., Purcell, K., Smith, A., and Zickuhr, K. 2010. Social media and mobile internet use among teens and young adults. Pew Research Center. Washington, DC: Pew Research Center Report. Retrieved from http://pweinternet.org/Reports/2010/Social-Media-and-Young-Adults.aspx.

[11] Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, 40(2), 61-77.

[12] Longley, P.A, Goodchild, M.F., Maguire, D.J, and Rhind, D.W. (2015). Geographic Information Systems and Science, 4th ed., New York, NY: John Wiley and Sons.

[13] Martin, M. 2021.Computer and internet use in the United States: 2018. American Community Survey Reports ACS-49. Washington, D.C.: U.S. Census Bureau.

[14] McClain, C., Vogels, E.A., Perrin, A., Sechopoulos, S., and Rainie, L. 2021. The Internet and the pandemic. Report. Washington, DC: Pew Research Center.

[15] Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. Biometrika (37:1), pp.17-23.

[16] Openshaw, S. (1984). The Modifiable Areal Unit Problem. Geobooks, Norwich, England.

[17] Mossberger, K., Tolbert, C. J., & Anderson, C. (2017). The mobile Internet and digital citizenship in African-American and Latino communities. Information, Communication & Society, 20(10), 1587-1606.

[18] NTIA. 2022. Digital Nation Data Explorer. National Telecom. & Info. Admin. US Dept. of Commerce. Available at https://www.ntia.doc.gov/data/explorer.

[19] OECD. 2001. Understanding the digital divide. Org. For Economic Cooperation and Development. Retrieved from https://www.oecd.org/sti/1888451.pdf.

[20] Pick, J. B., Sarkar, A., & Johnson, J. (2015). United States digital divide: State level analysis of spatial clustering and multivariate determinants of ICT utilization. Socio-Economic Planning Sciences, 49, 16-32.

[21] Pick, J. B., & Sarkar, A. (2015). The global digital divides: Explaining change. Berlin: Springer.

[22] Pick, James, and Avijit Sarkar. "Theories of the digital divide: Critical comparison." 2016 49th Hawaii Int. Conference on System Sciences (HICSS). IEEE, 2016.

[23] Pick, J., Sarkar, A., & Rosales, J. (2019). Social media use in American counties: geography and determinants. ISPRS Int. Journal of Geo-Information, 8(9), 424.

[24] Rogers, E. 2003. Diffusion of Innovations, 5th edition, Free Press, New York, NY.

[25] Rugers, W.P., and Sorens, J. 2021. Freedom in the 50 States, 6th ed. Cato Institute.

[26] Sarkar, A., Pick, J.B., & Rosales, J. (2017). ICT-enabled E-Entertainment Services in United States Counties: Socio-Economic Determinants and Geographic Patterns, in Choudrie J., Kurnia, S. and Tsatsou, P. (Eds.), Social Inclusion and Usability of ICT-enabled services (115-146). London, England: Routledge.

[27] Sarkar, A., Pick, J. B., & Parrish, E. (2017). Geographic Patterns and Socio-Economic Influences on Internet Use in US States: A Spatial and Multivariate Analysis. Available at SSRN 2942292.

[28] Pick, J.B., & Sarkar, A. (2018). Global Digital Divides. In Warf, B. (Ed.), The SAGE Encyclopedia of the Internet (217-225). Thousand Oaks, California: SAGE.

[29] Author. 2016.

[30] Smith, A., Anderson, M., & Caiazza, T. (2018). Social media use in 2018. Washington, DC: Pew Research Center.

[31] Stebbins, R.A. 2001. Exploratory Research in the Social Sciences. Qualitative research methods series 48. Thousand Oaks, CA: SAGE Publishing.

[32] Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. Economic Geography, 46(sup1), 234-240.

[33] U.S. Senate. 2018. The Geography of Social Capital in America. SCP REPORT NO. 1-18.

[34] Van Dijk, J.A.G.M. 2005. The Deepening Divide: Inequality in the Information Society. Thousand Oaks, CA: SAGE Publications.

[35] Van Dijk, J. 2021. The Digital Divide. Cambridge, UK: Polity Press.

[36] Venkatesh, V., M.G. Morris, G.B. Davis, F.D. Davis. 2003. User Acceptance of Information Technology: Toward a Unified View, MIS Quarterly, (27: 3), 425-478.

[37] Warschauer, M., and Matuchniak, T. (2010). New technology and digital worlds: analyzing evidence of equity in access, use, and outcomes. Review of Research in Education, 34(1), 179-22.

Table 3. OLS Regression Results.

	Personal Use						Business Use					
Ind. Vars.	Indv. used Cell Ph for personal use	Indv. has viewed the News on cell ph	Indv. used cell ph to access Social Media	Indv. used cell ph to access a search engine	Indv. used cell ph to text message friends/family	Indv. used cell ph For Video Call	Indv. used cell ph for business use	Indv. used cell ph for online banking	Indv. used cell ph to make purchase w/ text	Indv. used cell ph to redeem a mobile coupon	Indv. paid bills with cell phone	Used cell ph to Research Product
Median Age	0.677***	0.274***		0.212***	0.577***			0.183**				0.179***
African-Am Pop.							-0.216***				-0.316**	
Asian Pop.			-0.327***						-0.222***		-0.499***	
Hispanic Pop.							-0.356***					
Age Dependency Ratio		-0.282***	-0.492***	-0.201**		-0.360***	-0.175*	-0.166*	-0.310***	-0.526***	-0.709***	-0.253***
Urban Pop	-0.208*	0.128*				0.081						
MBSA Occupn	0.328**	0.425***	0.353***	0.533***	0.541***	0.471***	0.514***	0.467***	0.423***	0.208*	0.111	0.459***
Service Occupn			0.135*						0.157**			
Patents												
Social Capital	0.219*					-0.187*						
Affordability: Monthly Cell Bill is >$75		0.414***	0.478***	0.462***	0.310***	0.325***	0.482***	0.550***	0.556***	0.441***	0.479***	0.501***
Freedom	0.206*			0.147**	0.301***			0.151**				0.123*
MAX VIF	1.657	2.707	3.199	2.645	1.251	4.272	3.055	2.65	3.199	2.636***	3.637***	2.645
Joint Wald Statistic	252.519***	1699.833***	530.896***	847.792***	246.788***	458.578**	937.083***	721.742***	838.257***	587.780***	183.386***	824.275***
Koenker (BP)	3.242	7.263	3.778	3.825	2.474	2.181	10.853	6.38	4.371	6.162	2.521	4.386
Jarque-Bera	1.600	4.135	3.447	4.158	2.451	10.254	1.087	1.66	2.155	1.076	9.177	0.707
Spatial Autocorrelation												
Moran's I: Dep. Var.	0.416***	0.653***	0.450***	0.585***	0.509***	0.417***	0.562***	0.591***	0.575***	0.540***	0.538***	0.531***
Moran's I: Std. Residual	0.118	0.254*	0.162	0.227*	0.167	-0.055	-0.084	0.308**	0.258*	0.216*	0.312**	0.076
Adjusted R^2	0.731***	0.937***	0.846***	0.903***	0.718***	0.911***	0.919***	0.870***	0.910***	0.857***	0.662***	0.904***
	n = 49			* p < 0.05, ** p < 0.01, *** p < 0.001								