Does Bigger Screen Lead to More Cellular Data Usage?

Hailiang Chen
Department of Information Systems
College of Business
City University of Hong Kong
Kowloon, Hong Kong
hailchen@cityu.edu.hk

Baojun Ma
School of Economics and Management
Beijing University of Posts and Telecommunications
Beijing, China
mabaojun@bupt.edu.cn

Yu Pan
School of Business and Management
Shanghai International Studies University
Shanghai, China
13311887777@163.com

September 2016

ABSTRACT

This study utilizes a unique dataset from a telecommunications company to examine the relationship between screen size and cellular data consumption for smartphone users. We first find an overall positive correlation based on population-level cross-sectional regressions, but after adopting a quasi-experimental design, there is no significant relationship between screen size and cellular data consumption measured by the amount of data transmitted. Motivated by the theory of task-technology fit, our additional analyses of the QQ browser mobile app suggest that screen size has a positive effect on mobile web browsing. Our study provides important implications for mobile network operators in understanding consumer behavior and promoting mobile data services to users with different devices.

Keywords: task-technology fit, telecommunications, screen size, data usage, smartphone, mobile web

1. Introduction

The use of mobile devices such as smartphones and tablets is continuously increasing over the past few years. Together with that, we also see a rapid growth in the mobile data traffic, which reached 3.7 exabytes (i.e., 3.7 million terabytes) per month at the end of 2015, according to the Cisco Visual Networking Index.[1] Although the mobile data market generated more than US$100 billion of revenues in 2011, it is still in the early stage of development, and it is important for network operators to understand different aspects of customer behavior in order to further grow this market (Informa Telecoms & Media 2012). Screen size is one major consideration factor when consumers purchase a smartphone (Fowler 2014; Pandey and Nakra 2014). Prior studies such as Chae and Kim (2004) have shown that the reduced screen size of traditional mobile phones compared with that of desktop computers and laptops significantly affects mobile users' behaviors and perceptions. However, little academic research has directly examined or quantified the effect of screen size on consumers' mobile data consumption on smartphones. There have been a few industry initiatives that try to address this issue. One study by The NPD Group reports that monthly Wi-Fi and cellular data consumption on smartphones with screens 4.5 inches and larger is significantly higher than on smartphones with screens below 4.5 inches (The NPD Group 2013). Another report published by OpenSignal suggests that data usage over Wi-Fi networks increases with screen size but there is only a weak correlation between screen size and data usage over cellular networks, based on data collected from Android phones (OpenSignal 2013). The purpose of this study is to fill this research gap and assess the causal relationship between screen size and cellular data usage on smartphones.

The screen size of smartphones has experienced dramatic changes over time in the past decade.[2] The first iPhone launched by Apple in 2007 was "only" 3.5 inches, which at that time was considered to

[1] The Cisco Visual Networking Index also estimates that mobile data traffic was fewer than 10 gigabytes per month in 2000 and less than 1 petabyte (=1,000 terabytes) per month in 2005. (Retrieved from on September 16, 2016, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html)

[2] According to data compiled by phoneArena.com, the average smartphone screen size was 2.59 inches back in 2007; this number increased to 4.86 inches in 2014. (Retrieved from on September 16, 2016,

be much bigger than the average size of traditional feature phones (Barredo 2014; Taylor 2014). Later models of iPhones (until 5C/5S released in 2013) are either 3.5 or 4 inches. In 2011, Samsung created the "Phablet" market by launching the Galaxy Note with a 5.3-inch screen (Paczkowski 2011) and has since then introduced even bigger phones with screens of 6 inches or more. To compete against rivals offering larger-screen smartphones, Apple introduced the 4.7-inch iPhone 6 and the 5.5-inch iPhone 6 Plus in September 2014. On the other hand, although the general trend in the smartphone industry is that screen size keeps increasing over time, some consumers still prefer smartphones with smaller screens. To satisfy the demand of such consumers, Apple released the iPhone SE model with a 4-inch display in March 2016.

Since the screen size of smartphones is increasing simultaneously with mobile data usage over time, it is unclear whether and how much the increase in the screen size contributes to the increase in people's mobile data usage. To investigate this problem, we conduct two empirical studies to assess the impact of screen size on overall cellular data usage and mobile browser usage, respectively. We obtain a unique dataset from a telecommunications company that includes detailed records of cellular data usage in April 2014 for all subscribers in a large city in China. Besides the information about every user's data streaming activities at the transaction level (i.e., every data connection with the cellular network), the company also keeps track of the devices used by subscribers to connect to the network. A diverse set of phone models is represented in the dataset. We manually collect the public information on the screen sizes of 2,906 phones produced by 109 manufacturers appearing in our dataset from the Internet.

In our first study, we first conduct regression analyses at the individual level to explore the relationship between screen size and monthly cellular data usage. The intensity level of cellular data usage behavior is measured by the total amount of data transmitted in megabytes (MBs) in a month. The main independent variable of interest is the diagonal screen size of the phone used by the user. We include a wide variety of control variables: phone characteristics such as screen resolution, weight, and retail price, individual characteristics such as gender, age, and membership tier (Basic, Silver, Gold, or

---

http://www.phonearena.com/news/Did-you-know-that-smartphone-screens-nearly-doubled-in-size-since-2007_id52067)

Diamond), and the amount of data quota included in the subscribed service plan. Overall, the regression analyses on the population of users at the individual level seem to suggest that screen size is positively associated with cellular data usage on smartphones, which is also consistent with the industry reports based on correlation analysis (The NPD Group 2013; OpenSignal 2013).

The exploratory regression analyses provide some interesting insights. However, the regression analyses may suffer from problems such as omitted variable bias or selection bias and thus produce misleading results. In an attempt to infer a causal relationship between screen size and cellular data usage, we utilize the quasi-experimental design and employ propensity score matching (Rosenbaum and Rubin 1983) to address selection bias that may lead to the potentially overestimated effects in the regression analyses. We randomly select two groups of 1,500 users whose phones have the screen size of 4 inches and 5 inches, respectively. We select the 4-inch phone users as the control group, as 4 inches is the most popular screen size in our dataset. The other 5-inch group serves as the treatment group that receives a treatment of one inch more in screen size. We use observed individual characteristics (gender, age, membership tier) and observed behavior characteristics pertaining to a user's text messaging and voice call services, to predict each user's propensity to join the treatment group. In order to conduct the comparison in a regression framework, we adopt the one-to-one Nearest Neighbor matching method (Becker and Ichino 2002; Leuven and Sianesi 2003) to find matched pairs of treated and untreated users who have a similar tendency to adopt a larger screen but differ from each other in the actual screen choice. Our results show that there is no significant difference in cellular data usage between the 4-inch phone users and the 5-inch phone users. We also have the same findings for comparisons between other screen sizes (results available upon request). We conclude that a bigger screen does not lead to more cellular data usage on smartphones.

In our second study, we investigate whether screen size has a positive effect on the mobile web browsing behavior. We collect usage data of the QQ browser mobile app in February 2015 for the matched sample of 2,118 users in the first study. 1,347 users installed this app and used it at least once in the study period. Our dataset for this study is a user-day panel. The dependent variable is the number of data connections with the mobile network by a user in a day. As the dependent variable is now a count variable, we employ the negative binomial panel regression model to conduct our analysis. The main independent variable is still the screen size of a user's smartphone. We also control for phone characteristics, individual characteristics, data quota limit, and daily dummies in the model. Our result suggests that there is a positive effect of screen size on mobile web browsing.

As the first academic research that aims to infer a causal relationship between screen size and cellular data usage, this paper serves as an initial attempt of utilizing data analytics to understand consumer behavior and generate insights to support decision-making in the telecommunications industry. Prior studies about telecommunications in the IS and marketing fields have primarily focused on the pricing strategies of different services. Text messages, voice call, and data services are three main revenue streams for network operators. How to design various pricing models to optimize the revenue or profit is a central question for service providers.[3] This paper is related to the growing literature of studies that examine various kinds of consumer behavior, such as adoption of mobile services and content generation/consumption, in the mobile era. For example, Hong and Tam (2006) investigates the factors that determine the adoption of mobile data services in non-work settings; Venkatesh et al. (2012) extends the unified theory of acceptance and use of technology (UTAUT) to propose the UTAUT2 in a consumer context, and the authors conduct a two-stage online survey to explain the acceptance and use of mobile Internet technology; Ghose and Han (2011) empirically analyzes the interdependence between content generation and content consumption behavior on the mobile Internet; Xu et al. (2014) studies how the adoption of a mobile news app may affect the traffic to a corresponding mobile web site. As mobile

[3] One stream of studies focuses on examining the profit optimization and welfare implications of pricing models for one type of service. For instance, the three-part tariff plan or "fixed-up-to" plan is increasingly becoming a popular pricing strategy. This pricing plan requires a basic access fee, a free usage allowance, and a variable charge for usage beyond the free allowance. Masuda and Whang (2006), Lambrecht et al. (2007), and Iyengar et al. (2008), to name a few, have studied this multi-part tariff plan for voice call or internet access service under different contexts. Another stream of studies attempts to investigate the consumer demand and pricing strategies of multiple services at the same time. For instance, Kim et al. (2010) proposes a structural model to empirically evaluate the own- and cross-price elasticity demand of short message service (SMS) and voice calls; Niculescu and Whang (2012) proposes a model to examine the adoption of voice and data services for different types of consumers and applies it to the Japanese wireless services market; Lahiri et al. (2013) compares the profit and social welfare of two different pricing schemes for wireless services: one is based on service type (e.g., voice, text messaging, data, etc.) and the other is based on traffic (e.g., bytes) regardless of service type.

devices are becoming ubiquitous and more data are made available to researchers, we expect there will be more research utilizing data analytics to study consumer behavior in the telecommunications industry.

Our study provides important practical implications for both network operators and smartphone manufacturers. Although there is no significant relationship between the screen size of the device connected to the mobile network and a user's overall cellular data usage, screen size does have a significant impact on the data traffic from the mobile web. As network operators cooperate with mobile app developers to offer customized mobile data services specific to certain apps (e.g., exempt an app from counting toward a user's data plan), device characteristics such as screen size can be an important factor in pricing such data services. Our results also suggest that consumers do not necessarily strictly prefer larger screen smartphones in terms of staying connected through the mobile network, so smartphone manufacturers should closely monitor changes in consumer tastes and preferences to inform their product design decisions.

2. Study 1: Screen Size and Cellular Data Usage

Intuitively, it seems reasonable to expect that a bigger screen size is likely to lead to more cellular data usage. Similar as desktop/laptop computers, it is easier to read, touch, and interact with a bigger screen on smartphones, so the bigger screen may be perceived to be easier to use. However, a smartphone is also a mobile device. A smaller screen makes it easier to hold and carry (e.g., fit into the pocket), and more importantly, it can also be easily used with one hand. In that regard, a smaller screen may be perceived to be easier to use. Prior research in IS suggests that perceived ease of use is one key determinant of information technology acceptance and actual usage (e.g., Davis 1989, Davis et al. 1989, Venkatesh and Davis 2000). According to Davis (1989), perceived ease of use is defined as "the degree to which a person believes that using a particular system would be free of effort" (page 320). Although the theories of technology acceptance and use (e.g., Venkatesh et al. 2003) are developed under organizational contexts, they are also extended and tested in consumer contexts (e.g., Venkatesh et al. 2012).

It is ex ante difficult to assess the net impact of screen size on cellular data usage; hence we argue that bigger screen may lead to either more or less mobile data usage. We introduce our unique dataset in Section 2.1. To empirically evaluate the relationship between screen size and cellular data usage, we first conduct exploratory regression analyses on a large sample of mobile users (Section 2.2). Then in Section 2.3, we attempt to infer a causal relationship by adopting a quasi-experimental design based on propensity score matching.

2.1 Data

We first describe our dataset and introduce the main variables of interest. Our dependent variable, cellular data usage, is constructed from detailed transactions of mobile data services. Our main independent variable, diagonal screen size, and other phone characteristics are constructed based on information collected from various online resources. We also have information on some non-sensitive individual characteristics and the details of users' service plans.

Cellular Data Usage

Our main dataset is obtained from a leading telecommunications company in China. We have access to information about usage behavior of all 3G data users in a large city in April 2014. All personally identifiable information is anonymized. The number of active 3G users is over 1 million. For each user, we observe each data transmission between the user and a base transceiver station, which is regarded as a transaction in this study. On average, there are 930 transactions for a phone user in one month (30.5 per day). For each transaction, we have information on the amount of data transmitted in bytes.

We aggregate information at the monthly level to conduct a cross-sectional study. $DataUsage_i$ is our main dependent variable, which denotes the total amount of data transmitted in megabytes (MBs) for all transactions related to the user i in this month.

Screen Size

The company keeps track of the terminal models of phones that are connected to the network. In our dataset, there are 2,906 phone models (including traditional feature phones) produced by 109 manufacturers. Apple, Huawei, Lenovo, Nokia, Samsung, and Xiaomi are the main manufacturers (or brands). Later in our analyses, we include brand dummies to control for similar usage patterns across different phone models of the same brand (Hui 2004). We search on the Internet for the screen size information for all the brands and terminal models of phones in our dataset. Some manufacturers provide detailed technical specifications information on their official websites. Many e-Commerce websites selling electronics products also provide information about screen sizes. We first try to collect as much information as possible from manufacturers' official websites and various e-Commerce websites. Then we use the Google search engine to conduct individual searches on the remaining brands and terminal models to look for the screen size information. Overall, this is a very labor-intensive data collection process. We define *Screen* as the variable that measures the diagonal screen size of a device in inches.

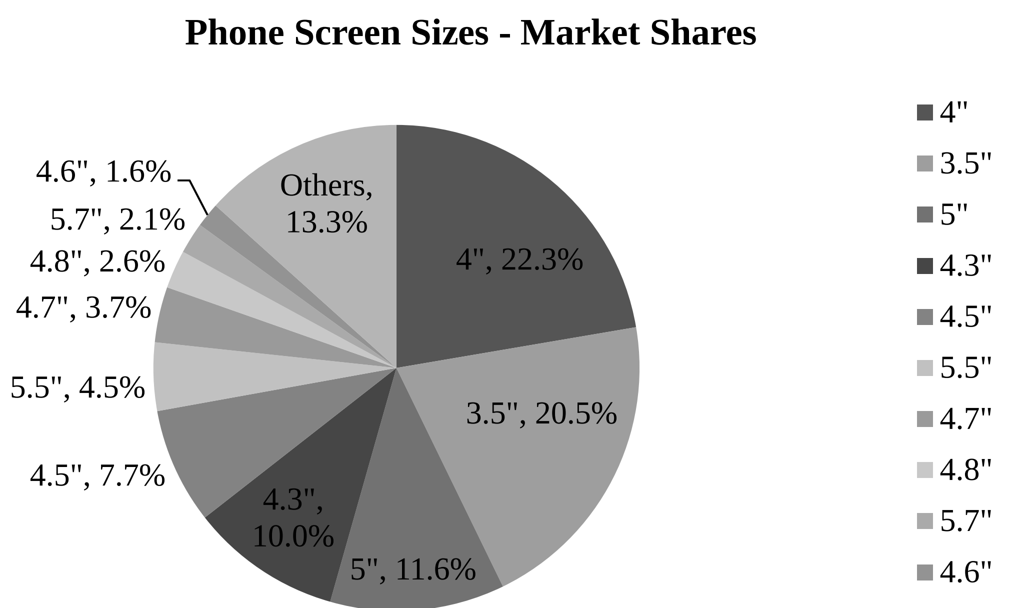

Figure 1. Market Shares of Different Phone Screen Sizes

Figure 1 presents the market shares of phones with different screen sizes measured by the total number of users in our dataset. Many device models from different manufacturers may share the same screen size. For the phone market at the time of this study, major models include Apple iPhone 5 series (4 inches), iPhone 4 series (3.5 inches), Samsung Galaxy S4 (5 inches), Xiaomi's MI 2 and MI 2S (4.3

inches), MI 2A (4.5 inches), Huawei Ascend (4.5 inches), Samsung Galaxy Note 2 (5.5 inches), and so on. Although traditional feature phones with screen sizes below 3 inches are also present in our dataset, the number of their users is very small and only 2.5% of the overall population. Because these phones are not designed for mobile data services, we exclude them from our analyses and focus only on smartphones with screen sizes above 3 inches.

Phone Characteristics

In case other technical specifications of a phone may also affect cellular data usage, we try to collect such information for the phone models in our dataset. Due to the difficulty in finding such information, especially for non-popular phones, we manually collect the information on screen resolution, weight, and retail price for the most popular 170 phone models. We consider only the top 170 phone models, because the number of users associated with them (862,482 users) is about 80% of the overall population (1,067,842 users in total). $Resolution_i$ is the number of pixels per inch (PPI) for the smartphone used by user i. $Weight_i$ is the weight of user i's smartphone in grams. $Price_i$ is the retail price in February 2015 for user i's smartphone on Tmall.com, which is one of the most popular B2C websites in China.

Individual Characteristics

We also observe some individual characteristics including gender, age, and membership tier. $Female_i$ is 1 if the gender of user i is female, and 0 otherwise. Age_i is user i's age as of 2014. The company offers four membership tiers: Basic, Silver, Gold, and Diamond. Non-Basic membership tiers enjoy different levels of benefits, e.g., free replacement SIM cards, exclusive shopping discounts, car rental discounts, a dedicated relationship manager, and so on. Most users are basic members. A non-basic membership tier can often be attained if a certain spending threshold is reached in the past few months. For each of the four membership tiers, we define a dummy variable to denote whether user i belongs to that tier. These four dummy variables are $Basic_i$, $Silver_i$, $Gold_i$, and $Diamond_i$. Membership tier information is available for all users, but gender and age information are missing for some users, which is assumed to be missing at random according to company practice.

Service Plan Characteristics

We also obtain information on the service plan subscribed by each user, including the quota limits for text messages, voice call, and data services, and the cost of the service plan in Chinese Yuan (CNY). The amount of free data included in a service plan should have a significant impact on the actual data usage, although we do observe some users go over the limit. We define $DataPlan_i$ as the number of MBs included in the monthly plan.

Table 1. Summary Statistics

Variable	# Obs.	Mean	Std. Dev.	Min	Median	Max
DataUsage	862,482	424.99	513.48	0.00	299.64	55,850.45
Screen	862,482	4.30	0.67	3.20	4.30	6.30
Resolution	862,482	303.51	62.12	146	326	471
Weight	862,482	139.10	20.08	93	140	233
Price	862,482	1,404.99	974.06	128	1,135	3,520
Female	862,482	0.22	0.41	0	0	1
Age	862,482	30.60	8.00	10	29	94
Diamond	862,482	0.002	0.05	0	0	1
Gold	862,482	0.02	0.13	0	0	1
Silver	862,482	0.14	0.34	0	0	1
DataPlan	862,482	291.06	290.27	40	300	4,096

Table 1 presents the summary statistics of the variables for the sample of 862,482 users who use the 170 most popular smartphones in our dataset. The average monthly data usage is 424.99 MBs and the median is smaller at 299.64 MBs. The mean and median screen sizes are both 4.3 inches, while the minimum and maximum screen sizes are 3.2 and 6.3 inches, respectively. The phone resolution ranges from 146 to 471 PPI, and the weight ranges from 93 to 233 grams. The average phone price is about 1,405 CNY (or 225 USD) and the maximum price is 3,520 CNY. Around 22% of users are female, and the average user age is 30.6 years old. Only a small portion of users is either Diamond or Gold, and 14% of users are Silver. The mean and median data quota limits are 291.06 MBs and 300 MBs, respectively.

2.2 Population-Level Regressions

We first conduct cross-sectional regression analyses to assess the correlation between screen size and cellular data usage at the individual level after controlling for the effects of various characteristics pertaining to the phone, individual, and service plan. We specify the following regression model:

$$\begin{aligned} Log(DataUsage_i) &= Screen_i + Log(Resolution_i) + Log(Weight_i) + Log(Price_i) + Female_i \\ &+ Log(Age_i) + Diamond_i + Gold_i + Silver_i + Log(DataPlan_i) \\ &+ BrandDummies + \epsilon \end{aligned} \tag{1}$$

The dependent variable in Equation (1) is the natural logarithm of $DataUsage_i$, where i denotes a user. The main independent variable is $Screen_i$, which is the diagonal screen size of the phone owned by user i. We control for the screen resolution, weight, and retail price of user i's phone. We also control for $Female_i$, Age_i, three membership-tier dummies ($Diamond_i$, $Gold_i$, and $Silver_i$), and the amount of data quota limit, which are all defined in the data section. The basic membership tier would serve as the benchmark case. Brand dummy variables are included to control for any manufacturer specific effects. Standard errors are clustered by brands to account for the residual correlation across devices produced by the same manufacturer.

The results of individual-level regressions are presented in Table 2. Column (1) contains the results for the regression on the overall population without controlling for other phone characteristics, and Column (2) contains the results for the regression on the sample of top 170 phone models whose screen resolution, weight, and retail price are controlled for. The results in these two columns are qualitatively similar, although without controlling for other phone characteristics in Column (1) seems to over-estimate the effect of screen size on cellular data usage. In Column (1), the coefficient estimate on $Screen_i$ is 0.401 and statistically significant at the 1% level, which implies that at the individual level an increase of 1 inch in screen size is associated with a 40.1% increase in cellular data usage on smartphones. In Column (2), we also find a positive and statistically significant relationship between screen size and cellular data usage, but the magnitude is much smaller (0.161), which implies that an increase of 1 inch in screen size is associated with a 16.1% increase in cellular data usage. This is equivalent to an increase of roughly 68 MBs in a month at the average level (425 MBs). Overall, these two regressions suggest a positive correlation between screen size and the amount of data transmitted on smartphones at the individual level, which is consistent with the findings of industry reports such as The NPD Group (2013) and OpenSignal (2013).

Table 2. Population-level Regressions

	(1)	(2)
$Screen_i$	0.401** (0.046)	0.161** (0.055)
$Log(Resolution_i)$		0.411** (0.122)
$Log(Weight_i)$		0.074 (0.164)
$Log(Price_i)$		0.177** (0.059)
$Female_i$	-0.165** (0.034)	-0.160** (0.036)
$Log(Age_i)$	-1.242** (0.074)	-1.159** (0.085)
$Diamond_i$	0.361* (0.148)	0.455** (0.091)
$Gold_i$	0.259** (0.059)	0.316** (0.043)
$Silver_i$	0.436** (0.034)	0.452** (0.038)
$Log(DataPlan_i)$	0.519** (0.036)	0.459** (0.026)
#Brand Dummies	105	20
Obs	1,067,842	862,482
Adj. R^2	0.206	0.202

Note: * p-value<0.05, ** p-value<0.01.

The coefficient estimates on other control variables in Table 2 are largely consistent with expectations. First, among the other phone characteristics, we find a positive and significant correlation between screen resolution and cellular data usage. Specifically, a 10% increase in screen resolution is associated with a 4.11% increase in cellular data usage. However, we do not find a significant effect of phone weight on cellular data usage. This is not surprising given that most smartphones are very light despite some variations across different phone models. We also find a positive and significant correlation between phone price and cellular data usage, although the coefficient estimate is much smaller (0.177). This implies that mobile users who own more expensive smartphones on average tend to consume more cellular data. Second, all the individual characteristics are significantly associated with cellular data usage. The coefficient estimate on $Female_i$ is negative and statistically significant at the 1% level in both regressions, suggesting that in general females use 16% less cellular data than males do. The coefficient estimate on Age_i is also negative and statistically significant at the 1% level in both regressions in Table 2, which imply that younger people are heavier data users than older people. The coefficient estimates on the three membership-tier dummies are all positive and statistically significant at the 1% level, which imply that non-Basic members consume more data than Basic members. Finally, the coefficient estimate on Log($DataPlan_i$) is positive and statistically significant at the 1% level in both regressions. Roughly, a 10% increase in the monthly data quota limit is associated with a 5% increase in actual data usage.

2.3 Quasi-experimental Design

The exploratory regression analyses in the previous subsection can provide some interesting insights, but these analyses may potentially suffer from omitted variable bias or selection bias. Although we have included many control variables in the regressions, we may still not be able to infer causality as we do not observe all individual characteristics, or more importantly, selection bias could potentially become a serious concern. For instance, heavy users of cellular data services may self-select to purchase smartphone models that have larger screens; without accounting for such selection bias, our analysis could possibly result in an over-estimation of the effect of screen size on cellular data usage.

In order to address these additional concerns and infer a causal relationship, we further examine the problem at hand following a quasi-experimental design. Specifically, we adopt the Propensity Score Matching technique (Rosenbaum and Rubin 1983) to find a control group and a treatment group that are observationally identical, but users in these two groups own devices of different screen sizes. We then compare the cellular data usage of these two groups and see if there is any significant difference. The main shortcoming of this approach is that we can compare only two groups of mobile users related to two screen sizes at a time, instead of considering all screen sizes in a cross-sectional regression.

To illustrate the detailed steps of our analyses, we use the comparison between 4 inches and 5 inches as an example. As shown in Figure 1, 4 inches is the most popular phone screen size in our dataset, while 5 inches is the third most popular one. In general, 4 inches is considered as neither too big nor too small; it is also very close to the median screen size (4.3 inches) of all phones in our dataset. Our results remain the same for comparisons between other screen sizes. We randomly select two groups of 1,500 users whose phones have the screen size of 4 and 5 inches, respectively. Many phone manufacturers and models are represented in these random samples. We identify 31 phone models from 13 manufacturers in the random sample of 1,500 users whose phones fall under the 4-inch category; there are 27 phone models from 9 manufacturers represented in the 5-inch category. Table 3 provides the list of top 10 phone models for each screen size based on the number of users in our random samples. Having a diverse mix of manufacturers and models helps alleviate confounding concerns related with manufacturer- or model-specific characteristics that would arise if the random sampling is based on phone models only, for example, Apple iPhone 5 (4 inches) and Samsung Galaxy S4 (5 inches). The group of users with a 4-inch smartphone is the control group and the group with a 5-inch smartphone is the treatment group. The purpose of the propensity score matching technique is to find matched users in the 4-inch group (control group) for each of the users in the 5-inch group (treatment group).

Table 3. Top 10 Phone Models with Different Screen Sizes

4-inch		5-inch	
Apple	iPhone 5S	Samsung	Galaxy S4 I9500
Apple	iPhone 5	Samsung	Galaxy Grand Duos I9082
Apple	iPhone 5C	Samsung	Galaxy S4 I9502
Samsung	Galaxy S Duos	Huawei	G610-U00
Samsung	Galaxy Trend II Duos	Xiaomi	MI 3W
Samsung	Galaxy Trend Duos	Huawei	Honor 3C
Xiaomi	MI-ONE S	Samsung	Galaxy Grand Duos I9082I
Xiaomi	MI-ONE Plus	Coolpad	7295
Nokia	Lumia 520	Sony	Xperia C
Huawei	Ascend G330D	OPPO	Find5 X909

We run a Probit regression to evaluate how different observed characteristics affect the user's probability of joining the treatment group. The dependent variable of this Probit regression is $Treatment_i$, which is 1 if the user is in the 5-inch group and 0 if the user is in the 4-inch group. The independent variables include individual characteristics (gender, age, and membership status) and service usage characteristics pertaining to text messaging and voice calls. Since only a few users are Diamond or Gold, to ensure that a sufficient number of observations is available in each membership category for matching, we group all three non-basic members into one category and create a variable called $PremierMember_i$, which is 1 if user *i* has a Diamond, Gold, or Silver membership and 0 otherwise. We do not use any mobile data usage information to predict a user's propensity to adopt a larger screen in case the outcome variable (mobile data usage) is simultaneously influenced by this screen choice. Instead, we rely on the observed text messaging and voice calling behavior in the same month to help predict a user's propensity of being treated. These two types of service usage behavior are indicative of how active a user is in her social network, which in turn could be helpful in predicting a user's choice in screen size. SMS_i is the number of text messages sent and received by user *i*. $Voice_i$ is the number of voice calls made and received by user *i*. $VoiceDuration_i$ is the number of hours spent on all voice calls by user *i*.

Table 4. Propensity Score Matching – Probit Regression

	(1)	(2)
$Female_i$	-0.250**	0.051
	(0.057)	(0.067)
$Log(Age_i)$	0.518**	-0.047
	(0.099)	(0.118)
$PremierMember_i$	-0.616**	-0.069
	(0.066)	(0.083)
$Log(SMS_i)$	0.123**	-0.007
	(0.022)	(0.027)
$Log(Voice_i)$	-0.179**	-0.001
	(0.045)	(0.054)
$Log(VoiceDuration_i)$	0.012	0.017
	(0.057)	(0.067)
Constant	-1.172**	0.158
	(0.368)	(0.435)
Obs	3,000	2,118
Log Likelihood	-1,976.4	-1,467.1

Note: * p-value<0.05, ** p-value<0.01.

Table 5. Relationship between Screen Size and Cellular Data Usage

	(1)	(2)
$Screen_i$	0.493*	0.504
	(0.213)	(0.264)
$Log(Resolution_i)$	0.332	0.271
	(0.157)	(0.180)
$Log(Weight_i)$	-0.423	-0.358
	(0.493)	(0.730)
$Log(Price_i)$	-0.153	-0.127
	(0.196)	(0.244)
$Female_i$	-0.137	-0.163
	(0.080)	(0.092)
$Log(Age_i)$	-0.947**	-0.888**
	(0.118)	(0.121)
$Diamond_i$	0.683**	0.704**
	(0.088)	(0.094)
$Gold_i$	-0.132	-0.178
	(0.101)	(0.109)
$Silver_i$	0.271**	0.268**
	(0.038)	(0.033)
$Log(DataPlan_i)$	0.516**	0.524**
	(0.061)	(0.049)
# Brand Dummies	16	16
Obs	3,000	2,118
Adj. R^2	0.275	0.275

Note: * p-value<0.05, ** p-value<0.01.

Table 4 presents the results of the Probit regression. Column (1) is the regression before matching on the sample of 3,000 users. We can see that gender, age, membership status, number of text messages, and number of voice calls are all good predictors of a user's choice of phone screen size except the duration of voice calls. Using these coefficient estimates, we then calculate the predicted propensity score for each user. In order to compare the control and treatment group in a regression framework, we adopt the one-to-one Nearest Neighbor matching method (Becker and Ichino 2002; Leuven and Sianesi 2003) to find matched pairs of users in the control and treatment groups. 2,118 users in our sample are matched. In other words, each of the 1,059 users from the control group appears to be observationally identical to a corresponding user from the treatment group, and they differ from each other only in the choice of screen size. Column (2) of Table 4 presents the result of the Probit regression on the final matched sample of 2,118 users. The purpose of this regression is to verify if the matched users from two groups are observationally identical. As shown in Column (2), all predictors are statistically insignificant at the 5% level, suggesting that our matching procedure is able to achieve the desired purpose.

In Table 5, we re-assess the relationship between screen size and cellular data usage based on the regression of cellular data usage on screen size and other control variables (Equation 1) on the matched sample of 2,118 users (Column 2). The regressions in Table 5 are the same as in Table 2 except on a smaller sample. Column (1) of Table 5 presents the regression on the random sample of 3,000 users. The coefficient estimate on $Screen_i$ is 0.493 and statistically significant at the 5% level, which suggests that after controlling for other factors a one-inch increase in screen size is associated with 49.3% more cellular data usage. The effect size in this random sample is much larger than that in Column (2) of Table 2, but its significance level decreases due to the smaller sample size. The larger effect size could be due to the presence of only two screen sizes (4 and 5 inches). Among the control variables, some coefficient estimates become insignificant and others remain similar as in Table 2. However, on the matched sample of 2,118 users, the coefficient estimate on $Screen_i$ is still positive (0.504) but no longer statistically significant at the 5% level. The results for coefficient estimates on other variables remain qualitatively similar as in Column (1). By adopting the quasi-experimental design, we find that there is no significant effect of screen size on cellular data usage.

3. Study 2: Screen Size and Mobile Browser Usage

Mobile users can perform all kinds of tasks such as web browsing, watching videos, and playing games on smartphones. For some tasks such as making phone calls, screen size makes no difference for users. For other tasks, screen size could play an either positive or negative role. Due to this complexity, it may be difficult to predict how screen size affects the total mobile data usage. However, given a specific task in which screen size plays a significant role, the theory of task-technology fit can help us predict how screen size may affect mobile data usage.

Goodhue and Thompson (1995) define task-technology fit as “the degree to which a technology assists an individual in performing his or her portfolio of tasks” (page 216). Among the various tasks that can be performed on smartphones, web browsing is one basic and fundamental task for mobile devices. More importantly, screen size is one key characteristic of the technology required for performing the web browsing task because screen size directly impacts how much information can be presented and how exactly it is presented (Jones et al. 1999). A few recent studies in IS have investigated the differences between mobile Internet and desktop Internet that arise due to the smaller screen size of mobile devices than that of personal computers. Ghose et al. (2013) explores how Internet browsing behavior differs on mobile phones and personal computers. The authors show that search costs are higher on mobile phones because the screen is smaller, so links at the top of the screen are more likely to be clicked. Adipat et al. (2011) adopts the design science research framework and proposes an approach to adapt the presentation of web pages for mobile handheld devices; one of the important considerations in the study is the small screen size. Therefore, a larger screen is preferred for mobile web browsing holding everything else equal. Based on the results from prior studies in both IS and Human-Computer Interaction (HCI), we predict that there should be a positive effect of screen size on mobile browser usage.

In this section, we conduct a second study to examine the role of screen size in affecting the intensity of web browsing on mobile devices. We collect usage data of the QQ Browser app in February 2015 for the matched sample of 2,118 in Study 1. According to China Internet Watch (2015), QQ browser mobile app is one of the most popular mobile web browsers in China. Among 2,118 users in our sample, 1,347 users installed the QQ browser app and used it at least once in February 2015.

Our dataset is a panel data at the user/day level. Table 6 summarizes the descriptive statistics of the main variables for this study. The number of observations is 37,716 (1,347 users and 28 days). Most of the variables are the same as in Study 1 and already described above. The primary difference is that we have a different dependent variable, $BrowserUsage_{it}$, which is a count variable and denotes the number of data connections with the mobile network by user i in day t. The mean and median of this variable are 2.67 and 0, respectively, but the maximum can reach 1,032. The mean of *Screen* is 4.70 inches as only two screen sizes (4 and 5 inches) are present in the sample. The summary statistics for other control variables are similar as in Table 1.

Table 6. Summary Statistics

Variable	# Obs.	Mean	Std. Dev.	Min	Median	Max
BrowserUsage	37,716	2.67	9.63	0	0	1,032
Screen	37,716	4.70	0.46	4	5	5
Resolution	37,716	322.11	97.63	187	326	441
Weight	37,716	140.19	18.89	112	140	180
Price	37,716	1,524.55	943.09	215	1,536	3,520
Female	37,716	0.22	0.41	0	0	1
Age	37,716	30.83	7.81	16	29	74
Diamond	37,716	0.001	0.03	0	0	1
Gold	37,716	0.02	0.13	0	0	1
Silver	37,716	0.12	0.33	0	0	1
DataPlan	37,716	296.57	294.00	40	300	1,624

Since the dependent variable, *BrowserUsage*$_{it}$, is a count variable and there is evidence of over-dispersion, we estimate a negative binomial panel regression model (Hausman et al. 1984, Hilbe 2011):

$$f(Y_{it}|X_{it}) = \frac{e^{-\mu_{it}}\mu_{it}^{Y_{it}}}{Y_{it}!}, Y_{it} = 0,1,2,3,\ldots \quad (2)$$

where Y_{it} is the number of data connections by user *i* in day *t*; $E(Y_{it}|X_{it}) = \mu_{it} = \exp(X_{it}\beta + \varepsilon_{it})$ is the conditional mean; ε_{it} is the unobserved heterogeneity and is assumed to follow a log-gamma distribution, with $\varepsilon_{it} \sim \Gamma(\theta, \theta)$ (Greene 2002). The explanatory variables X_{it} include our main variable of interest, *Screen*$_{it}$, other phone characteristics (resolution, weight, retail price), individual characteristics (gender, age, membership status), service plan's data limit, and daily dummies that control for time effects such as trends and seasonality.

In Table 7, we report the results from the population averaged negative binomial estimators for panel models (Hilbe 2011) to exploit the panel nature of the data and account for any potential within-panel correlation. Alternatively, one could also employ random effects estimators for the panel model, which would produce the subject specific estimates, i.e., what would happen to a particular user in terms of her mobile browser usage if her screen size changes. Here we prefer the population averaged panel model because we are more interested in the impact of screen size on mobile browser usage for the general population, or what would happen to an average person if her screen size changes.[4]

The coefficient estimate on $Screen_{it}$ in Table 7 is positive and statistically significant at the 1% level. The coefficient estimate of 0.906 in Column (3) means that an increase of one inch in screen size leads to a roughly 147.4% (i.e., $e^{0.906}$-1=1.474) increase in the number of mobile data connections from the QQ browser app. This shows that there is a positive relationship between screen size and mobile browser usage. Together with the results from the first study, we conclude that screen size may not influence the overall cellular data usage, but when a larger screen size facilitates a task on smartphone, it is likely to improve the performance of this task and thus leads to more usage. This is consistent with the theory of task-technology fit.

The coefficient estimates on the control variables in Table 7 also yield some interesting insights. First, among the phone characteristics, screen resolution is not associated with the usage of mobile browser. This makes sense because text information is still probably the dominant type of information presented in web browsers and a higher screen resolution does not make much difference for web browsing. Consistent with the regression of overall cellular data usage in Equation (1), phone weight also does not affect mobile browser usage. However, we find a negative association between the retail price of a phone and mobile browser usage, indicating that people who own more expensive smartphones tend to use mobile browsers less. Second, among the individual characteristics, we find mostly negative associations for gender, age, and two of the premier membership statuses (Diamond and Gold). Overall, our results suggest that female, younger, and higher-end users tend to use mobile browsers less. Finally, we also find a positive association between data quota limit and mobile browser usage.

---

[4] Note that the population averaged estimates and subject specific estimates are equivalent for linear models but not for nonlinear models such as negative binomial regressions. In negative binomial panel models, the population averaged estimator β_{PA} is derived from $E(Y_{it}) = \mu_{it} = \exp(X_{it}\beta_{PA})$ and the random effects estimator β_{RE} is derived from $E(Y_{it}|\alpha_i) = \mu_{it} = \exp(\alpha_i + X_{it}\beta_{RE})$. Taking expectation with respect to α_i in the later equation does not lead to the former one due to the nonlinear nature of the model.

Table 7. Relationship between Screen Size and Mobile Browser Usage

	(1)
$Screen_{it}$	0.906** (0.257)
$Log(Resolution_{it})$	0.080 (0.262)
$Log(Weight_{it})$	-1.045 (1.052)
$Log(Price_{it})$	-0.498** (0.112)
$Female_{it}$	-0.284* (0.111)
$Log(Age_{it})$	-1.121** (0.208)
$Diamond_{it}$	-1.295** (0.162)
$Gold_{it}$	-0.398* (0.274)
$Silver_{it}$	0.082 (0.145)
$Log(DataPlan_{it})$	0.322** (0.063)
Daily Dummies	Yes
#Users	1,347
Obs	37,716

Note: * p-value<0.05, ** p-value<0.01.

4. Conclusion

To build an in-depth understanding of consumer behavior is essential for increasing the average revenue per user (ARPU) for network operators. As the screen size of smartphones is becoming bigger and bigger, whether a bigger screen would induce more cellular data usage is an interesting and important question. Motivated by this question, we utilize a unique dataset to empirically test the relationship between screen size and cellular data usage on smartphones. Individual-level regressions on the population seem to suggest that there is a positive correlation between screen size and cellular data usage on phones, as shown in the industry reports (The NPD Group 2013; OpenSignal 2013). However, after we adopt the quasi-experimental design and employ the propensity score matching technique to address potential concerns due to omitted variable bias or selection bias, we find that a bigger screen does not lead to more cellular data usage on smartphones.

To shed more light on this matter, we further test the effect of screen size on the intensity of the web-browsing task on smartphones. We find that a bigger screen is associated with more mobile web browsing after controlling for various factors related with technical specifications of the phone, individual characteristics, and mobile data limit. Our result is in line with the Mobile Benchmark Report by Adobe (2014), which shows that consumers prefer to browse on phones with 5 inches or larger screens. The report also notes that the market share of Apple's Safari is continuously declining from December 2013 to May 2014 in the mobile browser market, possibly due to the fact that all iPhone models have either 3.5-inch or 4-inch screens before September 2014.

Although various information are available in our dataset, we unfortunately do not observe users' data usage behavior over Wi-Fi networks. Statistics from different data sources suggest that people may consume more data over various Wi-Fi networks, either public or private (Informa Telecoms & Media 2012). Future research can examine how screen size plays a role in affecting users' Wi-Fi and overall mobile data consumption. The availability of Wi-Fi networks and the choice between free and paid data could potentially complicate the matter and require different kinds of analysis.

REFERENCES

Adipat, B., Zhang, D., and Zhou, L. 2011. "The Effects of Tree-View Based Presentation Adaptation on Mobile Web Browsing," *MIS Quarterly* (35:1), pp. 99-122.

Adobe. 2014. "2014 U.S. Mobile Benchmark Report." Retrieved from http://www.cmo.com/content/dam/CMO_Other/ADI/ADI_Mobile_Report_2014/2014_US_Mobile_Benchmark_Report.pdf

Barredo, A. 2014. "A Comprehensive Look at Smartphone Screen Size Statistics and Trends." Retrieved from https://medium.com/@somospostpc/a-comprehensive-look-at-smartphone-screen-size-statistics-and-trends-e61d77001ebe

Becker, S. O., and Ichino, A. 2002. "Estimation of Average Treatment Effects Based on Propensity Scores," *The Stata Journal* (2:4), pp. 358-377.

China Internet Watch. 2015. "Top 3 China Mobile Browsers in Q4 2014." Retrieved from https://www.chinainternetwatch.com/12329/top-3-china-mobile-browsers-in-q4-2014/

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp. 319-342.

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. 1989. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science* (35:8), pp. 982-1002.

Fowler, G. A. 2014. "Find the Best Phone-Screen Size for You," Retrieved from http://www.wsj.com/articles/how-to-find-the-phone-that-fits-your-hand-1395795606

Ghose, A., Goldfarb, A., and Han, S. P. 2013. "How Is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research* (24:3), pp. 613-631.

Ghose, A., and Han, S. P. 2011. "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science* (57:9), pp. 1671-1691.

Goodhue, D.L., and Thompson, R.L. 1995. “Task-Technology Fit and Individual Performance,” *MIS Quarterly* (19:2), pp. 213-236.

Greene, W.H. 2002. *Econometric Analysis*, Prentice Hall, New York, NY.

Hausman, J., Hall, B. H., and Griliches, Z. 1984. “Econometric Models for Count Data with an Application to the Patents-R&D Relationship,” *Eonometrica* (52:4), pp. 909-938.

Hilbe, J. M. 2011. *Negative Binomial Regression*. Cambridge, UK: Cambridge University Press.

Hong, S.-J., and Tam, K. Y. 2006. “Understanding the Adoption of Multipurpose Information Appliances: The Case of Mobile Data Services,” *Information Systems Research* (17:2), pp. 162-179.

Hui, K. L. 2004. “Product Variety Under Brand Influence: An Empirical Investigation of Personal Computer Demand,” *Management Science* (50:5), pp. 686-700.

Informa Telecoms & Media. 2012. “Understanding Today's Smartphone User: Demystifying Data Usage Trends on Cellular & Wi-Fi Networks.” Retrieved from http://www.informatandm.com/wp-content/uploads/2012/02/Mobidia_final.pdf

Jones, M., Marsden, G., Mohd-Nasir, N., Boone, K., and Buchanan, G. 1999. “Improving Web interaction on small displays.” In Proceedings of the eighth international conference on World Wide Web (WWW '99), Toronto, Canada.

Iyengar, R., Jedidi, K., and Kohli, R. 2008. “A Conjoint Approach to Multipart Pricing,” *Journal of Marketing Research (JMR)* (45:2), pp. 195-210.

Kim, Y., Telang, R., Vogt, W. B., and Krishnan, R. 2010. “An Empirical Analysis of Mobile Voice Service and Sms: A Structural Model,” *Management Science* (56:2), pp. 234-252.

Lahiri, A., Dewan, R. M., and Freimer, M. 2013. “Pricing of Wireless Services: Service Pricing Vs. Traffic Pricing,” *Information Systems Research* (24:2), pp. 418-435.

Lambrecht, A., Seim, K., and Skiera, B. 2007. “Does Uncertainty Matter? Consumer Behavior under Three-Part Tariffs,” *Marketing Science* (26:5), pp. 698-710.

Leuven, E., and Sianesi, B. 2003. “PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing,” https://ideas.repec.org/c/boc/bocode/s432001.html

Masuda, Y., and Whang, S. 2006. “On the Optimality of Fixed-up-to Tariff for Telecommunications Service,” *Information Systems Research* (17:3), pp. 247-253.

Niculescu, M. F., and Whang, S. 2012. “Research Note—Codiffusion of Wireless Voice and Data Services: An Empirical Analysis of the Japanese Mobile Telecommunications Market,” *Information Systems Research* (23:1), pp. 260-279.

OpenSignal. 2013. “The Relationship between Screen Size and Data Usage.” Retrieved from http://opensignal.com/reports/data-usage-by-screen-size/

Paczkowski, J. 2011. “Note Is Creating 'Phablet' Market, Says Samsung.” Retrieved from http://allthingsd.com/20111229/samsungs-note-creates-phablet-market/

Pandey, M., and Nakra, N. 2014. “Consumer Preference towards Smartphone Brands, with Special Reference to Android Operating System,” *IUP Journal of Marketing Management* (13:4), pp. 7-22.

Rosenbaum, P. R., and Rubin, D. B. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika* (70:1), pp. 41-55.

Taylor, B. 2014. “Why Smartphone Screens Are Getting Bigger: Specs Reveal a Surprising Story.” Retrieved from http://www.pcworld.com/article/2455169/why-smartphone-screens-are-getting-bigger-specs-reveal-a-surprising-story.html

The NPD Group. 2013. “Smartphone Data Consumption Is 44 Percent Greater on Larger Screen Phones, According to Npd.” Retrieved from http://www.connected-intelligence.com/about-us/press-releases/smartphone-data-consumption-44-percent-greater-larger-screen-phones

Thompson, R.L., Higgins, C.A., and Howell, J.M. 1991. “Personal Computing: Towards a Conceptual Model of Utilization,” *MIS Quarterly* (15:1), pp. 125-143.

Venkatesh, V., and Davis, F. D. 2000. “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies,” *Management Science* (45:2) pp. 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. 2003. “User Acceptance of Information Technology: Toward a Unified View,” *MIS Quarterly* (27:3), pp. 425-478.

Venkatesh, V., Thong, J. Y. L., and Xu, X. 2012. “Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology,” *MIS Quarterly* (36:1), pp. 157-178.

Xu, J., Forman, C., Kim, J. B., and Van Ittersum, K. 2014. “News Media Channels: Complements or Substitutes? Evidence from Mobile Phone Usage,” *Journal of Marketing* (78:4), pp. 97-112.