American Economic Journal: Applied Economics 2022, 14(3): 488–526
https://doi.org/10.1257/app.20200367

Peer Effects in Product Adoption†

By Michael Bailey, Drew Johnston, Theresa Kuchler, Johannes Stroebel, and Arlene Wong*

We use de-identified data from Facebook to study the nature of peer effects in the market for cell phones. To identify peer effects, we exploit variation in friends' new phone acquisitions resulting from random phone losses. A new phone purchase by a friend has a large and persistent effect on an individual's own demand for phones of the same brand. While peer effects increase the overall demand for phones, a friend's purchase of a particular phone brand can reduce an individual's own demand for phones from competing brands, in particular if they are running on a different operating system. (*JEL* C45, D12, L63, M31, Z13)

Peer effects in consumption are pervasive. For example, an individual's choice of which car to purchase is likely influenced by the recent car purchasing decisions of her friends. Such peer effects have important implications for firms and policymakers. For instance, in the presence of peer effects, the elasticity of aggregate demand may be larger than the elasticity of individual demand since any direct incremental sales in response to a price reduction may lead to further extra sales through peer effects. Similarly, from a macro perspective, such peer effects in consumption suggest that the effects of stimulus policies on aggregate demand are larger than those estimated from directly affected individuals.

Despite the economic importance of peer effects in consumption and product adoption decisions, there is limited evidence on their exact nature and the resulting implications. For example, peer effects may lead someone to buy a new phone when her friend gets a new phone, but the effect of this purchase on firm profits

* Bailey: Facebook (email: mcbailey@fb.com); Johnston: Harvard University (email: drewmjohnston1@gmail.com); Kuchler: New York University, Stern School of Business, NBER, and CEPR (email: tkuchler@stern.nyu.edu); Stroebel: New York University, Stern School of Business, NBER, and CEPR (email: johannes.stroebel@nyu.edu); Wong: Princeton University and NBER (email: arlenewong@princeton.edu). Camille Landais was coeditor for this article. For their helpful comments, we thank several anonymous referees, as well as Tim Armstrong, John Campbell, Dean Eckles, Liran Einav, Matt Gentzkow, Ed Glaeser, Paul Goldsmith-Pinkam, Matthew Jackson, Michal Kolesar, Matthew Notowidigdo, Alex Peysakhovich, Luigi Pistaferri, Paulo Somaini, and Chris Tonetti, as well as seminar and conference participants at Baruch College, Carnegie Mellon University, Facebook, Frankfurt School of Finance and Management, Ludwig Maximilians University Munich, New York University, University of Rochester, Stanford University, Stockholm University, University of North Carolina Chapel Hill, University of Bonn, University of British Columbia, University of Mannheim, University of Texas at Austin, Wharton School of the University of Pennsylvania, Yale University, the Empirical Macro Workshop in Las Vegas, and the NBER Summer Institute. We thank the Center for Global Economy and Business at NYU Stern for generous research support. We also thank Abhinav Gupta, Sung Lee, and Hongbum Lee for excellent research assistance. This research was facilitated through a research consulting agreement between some of the academic authors (Johnston, Kuchler, and Stroebel) and Facebook. Bailey is an employee at Facebook.
† Go to https://doi.org/10.1257/app.20200367 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

depends on whether it represents incremental demand or the retiming of an already planned purchase. The implications of such peer effects for the competitive dynamics between firms also depend on whether any changes in demand are restricted to the brand purchased by the peer or whether there are positive or negative demand spillovers to competing brands.

In this paper, we explore the nature of peer effects in the US cell phone market. We find that peer effects are large, heterogeneous, and long-lasting and that they generate substantial incremental demand. Positive peer effects are largest for the brand purchased by the peer, but the size of the peer effect on same-brand demand often exceeds the effect on total phone demand. This finding suggests that some incremental same-brand purchases come at the expense of purchases from competing brands, in particular those on different operating systems.

We work with de-identified data from Facebook, the world's largest online social networking site. At the end of our sample period in May 2016, Facebook had around 226 million active users in the United States and Canada (Facebook 2016). In this dataset, we observe individuals' social networks as represented by their Facebook friends, which have been shown to provide a fair representation of real-world US friendship networks. For mobile active users, we also observe data on the device model used to log into their Facebook accounts, allowing us to identify the timing of new phone acquisitions.

We use this data to explore how phone purchases by a user's friends influence the user's own phone-purchasing behavior. To identify peer effects separately from common shocks or common preferences within friendship groups, we exploit quasi-random variation in friends' phone purchases. Useful sources of variation need to shift a friend's probability of acquiring a new phone in a given week, without affecting the probability of a user herself purchasing a new phone through any channel other than peer effects. We use two separate sources of variation that fit these requirements.

First, we use the number of friends who break or lose their phones in a given week to instrument for how many friends purchase a new phone in that week. The identifying assumption is that the number of friends who break or lose their phones in a given week is conditionally random and unrelated to a user's own propensity to buy a new phone in that week. We provide various pieces of evidence in support of this assumption. We identify individuals who randomly break or lose their phones by applying natural language processing and machine learning techniques to the universe of public posts on Facebook. This approach allows us to detect posts, such as "Phone broken … Ordered a new one but if anyone needs me urgently, call Joe," that signal the random phone loss by a peer. We show that people are substantially more likely to buy a new phone in the week after posting such messages. Our second instrument for the number of friends who obtain a new phone in a given week is the number of peers who are likely up for a contract renewal, which is often aligned with an upgrade to a new device.

We improve the power of these instruments by exploiting variation not only in *how many* friends experience the conditionally random event but also in *which* friends do so. Specifically, for both instruments, we use neural networks to estimate the probability that each individual would obtain a new phone conditional on the event, exploiting, for example, that older individuals are more likely to buy a new phone immediately after breaking their old device. The eventual instrument, then, is the sum of these estimated propensities across all individuals who experience the event, controlling for the distribution of these propensities in the overall pool of friends. This research design allows us to control, for example, for the average age in a person's friendship network and only identify off variation in whether it is the person's old or young friends who happen to break their phones in a given week.

Across both instruments, we obtain peer effect estimates of similar magnitude. Having 1 additional friend who purchases a new phone in a given week increases an individual's own probability of buying a new phone in the following week by 0.040 and 0.022 percentage points, estimates obtained using the random phone loss instrument and the contract renewal instrument, respectively. These estimated effects are large relative to the weekly probability of buying a new phone of about 1 percentage point. We argue that much of the communication between friends about the new phone purchase that drives the observed peer effect occurs off the Facebook platform and to a substantial extent through real-world interactions. Consistent with this interpretation, we show that peer effects from geographically proximate friends are larger than peer effects from friends who live further away.

In addition to exploring the immediate response of an individual's own purchasing behavior to new phone acquisitions by her friends, we also analyze the extent to which this situation generates new purchases instead of pulling forward already-planned future purchases. We find that a random phone loss by an individual has a positive effect on the total number of phones purchased by her friends in each of the following ten months, though the magnitude of this effect starts to decline after about three months. Peer effects thus cause an increase in the total number of phone purchases, at least over intermediate horizons. Quantitatively, having 1 extra friend purchase a new phone increases an individual's own probability of purchasing a new phone over the next 4 months by 0.6 percentage points, relative to a baseline probability of buying a new phone over this horizon of about 14.8 percent.

In the next step, we explore heterogeneities in peer effects along characteristics of potential influencers and potentially influenced individuals. We focus on heterogeneities in the local average treatment effects of the random phone loss instrument, which has the most power in the baseline specification, but find similar patterns of heterogeneity in the corresponding OLS estimates. We observe that close friends on Facebook exert a larger influence on one another than friends with weaker tie strength. We also find large heterogeneities in the peer effects exerted by different demographic groups but little variation in individual susceptibility to influence along the same demographic characteristics. For example, less-educated individuals have the largest effects on their friends' purchasing behaviors, but these individuals are no more likely to be influenced by phone purchases of their friends. These heterogeneities in peer influence have important implications for understanding the effectiveness of seed marketing campaigns, which target a small set of early adopters who can generate follow-on demand through peer effects. We also find that those individuals who exert larger peer effects are generally more price sensitive, measured as the effect of a price cut for a phone model on the probability of purchasing that model. This result suggests that the difference between the elasticities of aggregate and individual demand induced by peer effects is even larger than implied by the average peer effect.

In the second part of the paper, we explore whether peer effects are limited to the brand purchased by the peer or whether there are demand spillovers to other brands. To do so, we first predict the probability that each individual would purchase a phone in each of three broad brand categories: iPhone, Galaxy, and "other." We then exploit variation in this probability among friends who randomly break their phones in a given week (conditional on the average of this probability among all friends) to instrument for the number of friends who purchase phones of that particular brand. The identification assumption is similar to before: conditional on the characteristics of all friends and other controls, it is random whether, in a given week, the friends who happen to lose their phones are those who are likely to replace it with a new iPhone or those who are likely to purchase a new Samsung Galaxy.

There are three key takeaways from the cross-brand analysis. First, for all three brand categories, positive peer effects are largest for phones in the same category as that purchased by the peer. Second, these same-brand peer effects are largest for less well-known but cheaper "other" phones, and they are smallest for the expensive and well-known iPhones. These facts suggest that social learning is an important part of the explanation for these peer effects since social learning should be more important for lesser-known brands.[1] The third main takeaway relates to across-brand demand spillovers. Specifically, we find that when a friend buys a new phone, this purchase increases a person's own propensity of buying a phone from competing brands on the same operating system, while reducing their propensity of buying a phone from competing brands on different operating systems. In other words, while some of the observed positive same-brand peer effects arise by generating entirely new demand, others come from pulling demand away from rival firms with competing operating systems. Importantly, these demand spillovers across operating systems could have easily been positive. For example, a user who buys a Galaxy might have caused her friends to desire a new phone—of any type, including iPhones—through a "keeping up" effect. The observed across-brand demand spillovers are thus again consistent with an important social learning component: when your friends use a certain operating system, you are more likely to learn about that system. This would increase your demand for all phones using that operating system (even those produced by a different manufacturer), in part at the expense of phones using competing operating systems.[2]

[1] Our results do not allow us to rule out that "keeping up" effects (which are likely to be larger for more expensive brands) also contribute to the observed peer effects; instead, our findings suggest that such effects cannot be the entire story.

[2] In addition to social learning, network externalities provide a second mechanism that might explain some of the patterns of across-brand demand spillovers. Such network externalities would arise if having more friends use a certain operating system would increase a user's own value of using that same operating system. In the context of cell phones, network externalities may primarily come from the use of the FaceTime video messaging app, which is only available on the iOS operating system. However, while we cannot rule out that such network externalities play some role, such externalities cannot explain a number of the patterns we document in this paper, all of which would naturally follow from a social learning story (e.g., the fact that peer effects are largest for the same model rather than equally spread across all phones of the same operating system or the fact that we find peer effects to decline in time since model release). As a result, our findings are only consistent with a story in which there is at

The observed across-brand demand spillovers highlight that peer effects have important competitive implications for firms: losing a customer to a competitor does not only mean missing out on positive peer effects that this customer could have had but may also lead to future losses of other customers through competitive peer effects. These implications of peer effects for the demand of competitors' brands complement a large literature that has explored similar spillover effects of advertising (e.g., Sahni 2016; Shapiro 2018; Sinkinson and Starc 2019). In that literature, researchers regularly find positive demand spillovers to nonadvertised competitor brands. Our finding of negative across-brand demand spillovers highlights that the implications of peer effects for the competitive dynamics between firms can be qualitatively different to those from the spillover effects of marketing activities.

Our paper contributes to a literature that has studied the role of peer effects in a wide range of economic and financial decisions. Peers have been shown to influence consumption choices (e.g., Goolsbee and Klenow 2002; Mobius, Niehaus, and Rosenblat 2005; Kuhn et al. 2011; Moretti 2011; Aral and Walker 2012; Gilchrist and Sands 2016; De Giorgi, Frederiksen, and Pistaferri 2016; Han, Hirshleifer, and Walden 2019) as well as a variety of household financial decisions (e.g., Duflo and Saez 2003; Hong, Kubik, and Stein 2004; Bursztyn et al. 2014; Beshears et al. 2015; Ouimet and Tate 2017; Kuchler and Stroebel 2020), housing market decisions (e.g., Bailey et al. 2019; Bailey, Cao, Kuchler, and Stroebel 2018), and charitable giving (e.g., DellaVigna, List, and Malmendier 2012). Peer effects also play an important role in explaining education decisions (e.g., Hoxby 2000; Sacerdote 2001, 2011), program participation (Dahl, Løken, and Mogstad 2014), labor market outcomes (Mas and Moretti 2009), mutual fund investments (Kuchler et al. 2020), international trade flows (Bailey, Gupta et al. 2021), and the spread of and response to COVID-19 (Bailey, Johnston et al. 2021; Kuchler, Russel, and Stroebel 2020). Prior work has studied peer effects in product and technology adoption decisions; one focus of this literature has been how social learning can help the diffusion of new technologies in developing countries (e.g., Foster and Rosenzweig 1995; Conley and Udry 2010; Oster and Thornton 2012; Kremer and Miguel 2007; Björkegren 2019). In the developed world, peer effects have been shown to affect the adoption of new technologies such as solar panels (e.g., Bollinger and Gillingham 2012; Allcott and Kessler 2019). Within the literature that has studied peer effects in product adoption decisions, we are the first, to our knowledge, to identify important competitive spillovers to other models and brands. Our setting and research design also allow us to expand our understanding of peer effects along other important dimensions. For example, we are able to document that peer effects can generate additional demand rather than just a retiming of demand. We can also identify characteristics of influential individuals as well as the correlation of peer influence with price sensitivity.

least a substantial social learning component to peer effects (in addition to potential other components coming from "keeping up" desires or network externalities).

I. Data Description

A central challenge for studying peer effects in product adoption decisions is the need to observe both social networks and product adoption behavior within the same dataset. We overcome this measurement challenge by exploring peer effects in phone purchasing decisions using de-identified data from Facebook, the world's largest online social networking site. In the United States, Facebook primarily serves as a platform for real-world friends and acquaintances to interact online, and people usually only add connections to individuals on Facebook whom they know in the real world (Jones et al. 2013). As a result, friendships on Facebook provide a good approximation of real-world friendship networks (see Bailey, Cao, Kuchler, Stroebel, and Wong 2018; Bailey, Farrell et al. 2020).

For each Facebook user, we observe basic demographic information such as their date of birth, gender, and county location as well as the set of individuals that they are connected to (Facebook 2020).[3] Using the language adopted by the Facebook community, we call these connections "friends." The vast majority of Facebook users regularly access their Facebook accounts from their cell phones.[4] For these mobile active users, we observe data on the cell phone carrier and the phone model used to access the Facebook app. We use these data to identify when a user obtains a new phone.[5] Since we can only observe a new phone model when the user logs into the Facebook app for the first time from the new device, we can generally pinpoint the timing of the purchase to roughly the week that a new device is acquired. Our unit of observation is therefore the purchasing behavior of a user in a given week.

In our analysis, we focus on US-based Facebook users between 18 and 65 years of age who have between 100 and 1,000 friends on Facebook. We also require users to access Facebook on their phones across two consecutive weeks in order to be able to observe the timing of potential phone purchases. Our primary sample covers the purchasing behavior of these individuals across four consecutive weeks in May 2016. These weeks were chosen to be relatively far away from both major phone release dates and major shopping holidays (such as Black Friday or Labor Day), which could confound our estimates. We are left with about 329 million user-weeks as our baseline estimation sample.

Table 1 provides summary statistics on our sample. The average user in our sample is 35 years old, with a tenth to ninetieth percentile age range of 21 years to 53 years. Roughly 58 percent of users in our sample are male. Fifty-five percent of the users have an iPhone, and 27 percent have a Samsung Galaxy; the rest of the users are relatively fragmented across many other phone models. The average user has a

[3] Facebook is unable to retain for replication purposes a "stable" version of the raw data that do not change over time.

[4] Facebook reports in its July 26, 2018, 10-Q filing, "Substantially all of our daily and monthly active users […] access Facebook on mobile devices."

[5] The process of determining when a user obtains a new phone involves a number of steps, including the removal of likely work phones or phones borrowed from a friend as well as dropping temporary phones with only a few log-ins. Because Facebook only records the device model but no unique device identifier, we are unable to detect switches between two devices of the same model. The overwhelming majority of switches that we detect are to phones released no more than nine months prior to the start of our sample, suggesting they are new purchases rather than hand-downs from friends and family.

Table 1—Summary Statistics

	Mean	Standard deviation	P10	P25	P50	P75	P90
User characteristics							
Age (years)	35.3	12.1	21	25	33	44	53
Male	0.58	0.5	0	0	1	1	1
Phone age (days)	388.8	327.3	63	152	317	544	777
Buys phone (percent)	0.93	9.59	0	0	0	0	0
Has iPhone	0.55	0.50	0	0	1	1	1
Has Galaxy	0.27	0.44	0	0	0	1	1
Friend characteristics							
Friends in sample	322.4	202.8	124	165	258	424	631
Friends with phone purchases	3.00	2.87	0	1	2	4	7
Friends with public statuses	59.5	53.8	17	26	43	73	120
Friends posting about breaking/losing phone	0.26	0.64	0	0	0	0	1
Friends at phone age threshold	1.83	1.84	0	0	1	3	4

Notes: Table presents summary statistics for our baseline panel. The unit of observation is a user-week, and our data consist of approximately 329 million such user-weeks. For each characteristic, we present the mean, standard deviation, and the tenth, twenty-fifth, fiftieth, seventy-fifth, and ninetieth percentiles of the distribution.

phone that is 389 days old, while the median user has a phone that is just over 10 months old. The tenth to ninetieth percentile range of phone age is between 63 days and 777 days. About 0.93 percent of all users acquire a new phone in a given week. The average user has 323 friends in the sample as well as about 3 new phone purchases among friends in a given week.

II. Research Design

We next outline how we use the data described above to identify peer effects in cell phone–purchasing behavior. Our most basic specification seeks to understand a Facebook user's decision to buy a new phone in a given week as a function of the prior or contemporaneous purchases of her friends. The challenge for identifying such peer effects is that individuals tend to be friends with others who are similar to them across many dimensions (McPherson, Smith-Lovin, and Cook 2001; Bailey, Cao, Kuchler, Stroebel, and Wong 2018; Bailey, Cao, Kuchler, and Stroebel 2018). For example, in the context of our study, an Apple enthusiast may primarily be friends with other Apple enthusiasts. Even in the absence of peer effects, these friends may thus have similar phone-purchasing behaviors, such as buying a new iPhone around its release date. As a result, observing a correlation in purchasing behavior within friendship groups does not necessarily provide evidence for peer effects (see Manski 1993 for an extended discussion).

Our approach to solving this identification challenge is to develop instrumental variables for the purchasing behavior of a person's friends. A successful instrument should shift the purchasing behavior of a person's friends without affecting the purchasing behavior of that person through any channel other than peer effects. We propose two instruments that meet this exclusion restriction: first, the number of a user's friends who randomly lose their phones and second, the number of friends

Figure 1. Sample Posts about Randomly Lost Phones

who have owned their phones for exactly two years and whose contract is thus likely up for renewal. We next discuss both of these instruments in more detail.

A. *Random Phone Loss Instrument*

Our first instrument is based on the idea that individuals are substantially more likely to buy a new phone in a week in which they lose or break their current phone. As a result, an individual who has more friends randomly losing their phones in a given week is likely to have more friends buying a new phone in that week. Provided that a random phone loss of a friend only influences the probability that a user herself purchases a new phone through peer effects from any replacement purchase by the friend, the number of friends who experience a random phone loss can then be used to instrument for the number of friends who purchase new phones.

The first step in constructing this instrument is to determine which individuals randomly break or lose their phones in a given week. We do so by analyzing public posts on Facebook that relate to such events. Figure 1 provides examples of such posts, which were relatively common during our sample period, since users regularly posted on Facebook to explain to their friends why they were not returning calls or text messages.

We use a machine learning–based approach to classify the universe of public Facebook posts in a given week, allowing us to assign an indicator $\mathbf{1}(RandomPhoneLoss_{i,t})$ to individuals who post about a random phone loss in that week.[6] Specifically, we use two tools from the natural language processing literature: word embeddings and convolutional neural networks. We will provide a

[6] We only have access to posts from individuals who have set their privacy settings for that specific post to "public" at the time of the analysis, rendering the post visible to any individual with the URL. Table 1 shows that while the average person in our sample has about 322 friends in total, only about 60 of those friends have set their statuses as public.

brief overview of these tools here; a longer explanation of our methodology is available in online Appendix A.1.

Our approach relies on word embeddings, which are low-dimensional vectors that provide a geometric representation of the meaning of the corresponding word. Words with similar meanings will be represented by similar vectors, and the spatial relationships between vectors will capture complex relationships between the corresponding words (see Mikolov, Yih, and Zweig 2013 for details). For instance, after converting words to their embeddings, the embedding most similar to $\left(\overrightarrow{King} - \overrightarrow{Man} + \overrightarrow{Woman}\right)$ is $\overrightarrow{Queen}$. In our application, we use 200-dimensional word embeddings that were trained using all articles on the English edition of Wikipedia. Using these vectors, we can represent each public Facebook post as a matrix, consisting of the stacked vectors of its constituent words.

After generating this numerical representation of each public post, we next use a convolutional neural network (CNN) to determine which posts describe a user breaking or losing her phone. CNNs were originally developed for applications in computer vision, and they expand upon traditional neural networks by transforming the underlying data to make use of its spatial configuration. In the case of image data, CNNs account for relationships between nearby areas of the image; in natural language applications, CNNs make use of the order of words within a passage. This allows us to distinguish between sentences such as "I broke my phone when I was with my friend John" and "I just saw my friend John break his phone." We train the CNN on a large sample of manually classified posts using 10-fold cross-validation and then use it to classify all public posts in our sample. The resulting model performs quite well on unseen posts, identifying many idiosyncratic examples such as "R.I.P phone. You will be missed" that would be difficult to capture with regular expression searches.[7] Online Appendix A.1 includes further details on the training process and the model's performance. In total, we identify around 65,000 public posts about broken or lost phones per week. Table 1 shows that, in a given week, the average person has 0.26 friends who publicly post about breaking or losing their phones.[8]

Panel A of Figure 2 visualizes the first stage of the random phone loss instrument. It shows the probability of purchasing a new phone in each week, splitting individuals according to their posting behavior in week 0. The green-triangle line corresponds to individuals who publicly post about a random phone loss in week 0. The orange-circle line corresponds to individuals with a public post that was not about a random phone loss, and the blue-square line corresponds to individuals without a public post in week 0. In the weeks prior to posting about a random phone loss, the purchasing behavior of individuals who post about such a phone loss in week 0 has a broadly similar trend to that of other individuals, although it has a somewhat higher level. (As we describe below, our research design will account for this higher level.) In week 0, those individuals who posted about a random phone loss have a

[7] It is likely that the CNN identified this particular post after observing hand-classified posts such as "My phone is dead" in the training sample, combined with the fact that "dead" and "R.I.P" occupy similar positions in the embedding space.

[8] We have also implemented a model using a regular expression-based classifier, which produced an instrument that had less power but found largely similar results as our baseline analysis. This simpler classifier is used to reinforce our main model in an approach inspired by ensemble classifiers. See the discussion in online Appendix A.1.

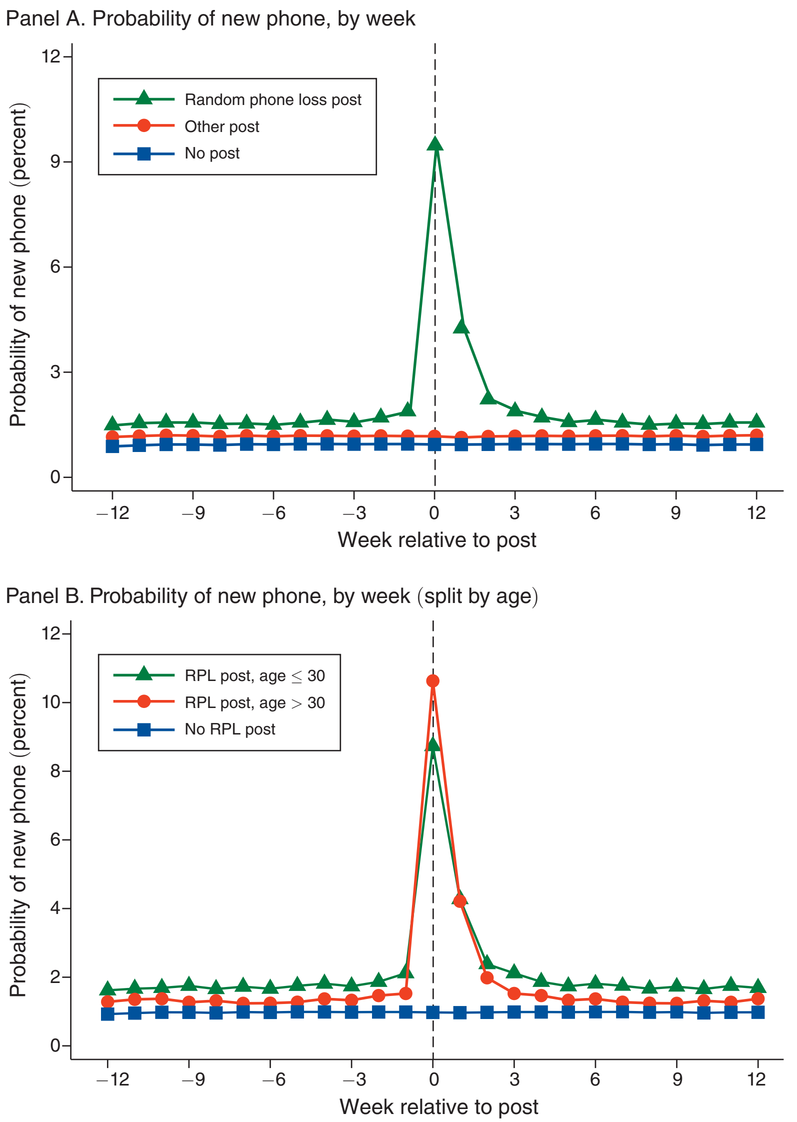

Figure 2. Random Phone Loss Instrument

Notes: Panel A shows the probability of purchasing a new phone in a given week, splitting users by their posting behavior in week 0. The line Random Phone Loss Post (green triangles) shows the behavior of users who have a public post in week 0 that relates to a random phone loss. The line Other Post (orange circles) captures the behavior of those who have a public post in week 0 that does not relate to a random phone loss, while the line for No Post (blue squares) tracks the behavior of those individuals without a public post in week 0. Panel B shows the probability that a user of each age group buys a phone in the weeks after posting about randomly losing or breaking her phone (RPL = Random Phone Loss).

substantial increase in the probability of acquiring a new phone. Specifically, about 9 percent of individuals with a post identified by our classifier get a new phone in the week of posting about losing their phone. The probability of purchasing a new phone remains slightly elevated in the week following the post about the random phone loss before returning to its baseline rate.

While the probability of getting a new phone spikes in the week of the post and remains elevated in the following week, the sum of these probabilities is far below 100 percent, meaning that we do not observe a new phone purchase for every individual whom we identify as having posted about a random phone loss. There are several reasons for this result. First, our classifier is likely to include some "false positive" posts that we incorrectly identify as indicating a random phone loss. For example, our classifier cannot perfectly separate posts that mention that someone's "phone is dead" into those that talk about a dead battery and those that talk about a permanently broken phone.[9] A second explanation is that some users may continue to use a phone with a broken screen or damage of another type. Users may also be able to repair broken phones or recover lost or stolen phones. Finally, our data do not allow us to identify individuals who replace a broken phone with a new phone of the exact same model. In these instances, however, peer effects are likely to be small, and not observing these switches is unlikely to substantially bias our results.

Based on this classification of a random phone loss, a basic identification strategy would instrument for the number of friends who purchase a new phone in a given week with the number of friends who publicly post about randomly breaking or losing their phones in that week. The associated identifying assumption would be that the number of friends losing or breaking their phones in a given week is conditionally random. To strengthen the validity of this exclusion restriction, we include a number of controls in specifications using this first instrument. One possible concern is that the purchasing behavior of individuals with friends who are more likely to lose or break their phone, or with friends who are more likely to post about it publicly, may be fundamentally different. To address such concerns, we directly control for the number of friends who have posted publicly about losing or breaking their phones in the previous year as well as for the number of friends who have public statuses by default.[10]

While posting about breaking or losing one's phone leads to a sizable increase in the average probability of obtaining a new phone, there is substantial heterogeneity in the size of this increase across individuals with different characteristics. For example, panel B of Figure 2 shows that, among individuals who publicly post about losing their phones in week 0, the probability of getting a new phone in that week is 11 percent for individuals over the age of 30, while it is only about 9 percent for individuals under 30 years of age. How many friends purchase a phone in a given

[9] Properly weighting "false positives" and "false negatives" was an important consideration when constructing our classifier, and we chose a threshold that balanced the number of the posts found with the conditional probability of switching of the posters. We also trained an alternative classifier that was better at rejecting false positives and gave a conditional $\Pr\left(BuysPhone_{i,t} \mid \mathbf{1}\left(RandomPhoneLoss_{i,t}\right)\right)$ of 13.4 percent, although the number of posts found decreased by 85 percent. This associated decrease in the number of true positives thus weakened our instrument.

[10] Additionally, it is important that having friends lose or break their phones in a given week is not correlated with individuals losing or breaking their own phones in that week. One reason for such a correlation could be common experiences that are correlated with breaking or losing a phone (e.g., a bachelor party, a trip to the beach, or time spent in a high-crime area). To assess whether phone loss events are temporally correlated across friends, we perform a series of tests on users who post about losing or breaking their phones in week t, calculating the probability that one of their friends posts about losing or breaking their phones in each week from $t-5$ to $t+5$. We were unable to find evidence that users lose or break their phones at the same time as their friends (see online Appendix B). Even though such concerns seem to be minor, we include a control indicating whether the user has posted about a random phone loss in all regressions that make use of this instrument.

week is therefore not only affected by *how many* friends lose their phones in that week but also by *which* friends lose their phones. Under our assumption that phone loss is a conditionally random event, which friends lose their phones is also plausibly random. We use this insight to further improve the power of our instrument.

Specifically, we exploit small-sample variation in whether those friends who randomly lose their phones in a given week are more or less likely to purchase a new phone, conditional on the distribution of this propensity among all friends. For example, one could use the average age among people posting about a random phone loss as an instrument, controlling for the average age among all friends. Many other demographic characteristics are also correlated with a user's conditional probability of buying a new phone, and all of these characteristics (and their interactions) could serve as potential instruments. However, using many of these potentially weak instruments would risk overfitting the first stage, therefore biasing our instrumental variables estimates toward the OLS estimates. Since fitting the first stage is a prediction exercise, recent literature suggests using machine learning tools to optimally fit the first stage when there are a large number of possible instruments (e.g., Belloni, Chernozhukov, and Hansen 2014; Mullainathan and Spiess 2017; Peysakhovich and Eckles 2017; Athey 2019; Chernozhukov et al. 2018). We build on the ideas in this work and use a neural network to create a single propensity score from the large space of possible instruments.

$$(1) \qquad ProbBuyRandomPhoneLoss_{i,t} = Prob\Big(\mathbf{1}\big(BuysPhone_{i,t}\big)\,\big|\,X_{i,t}, \mathbf{1}\big(RandomPhoneLoss_{i,t}\big) = 1\Big).$$

The vector $X_{i,t}$ collects a large number of observable characteristics of user i at time t.[11] We train the neural network using data from a separate sample of weeks, 2016–2015 to 2016–2017 and 2016–2023 to 2016–2025. This approach, which is similar to the jackknife IV approach in Angrist, Imbens, and Krueger (1999), allows us to avoid overfitting in-sample noise, thus ensuring that we obtain unbiased estimates when building our instruments based on $ProbBuyRandomPhoneLoss_{i,t}$. Online Appendix A.2 provides details on the design and the performance of the neural network used to estimate the propensity score.

We then construct the first instrument for the number of friends of person i who purchase a phone in week t by summing these propensities among user i's friends who post about a random phone loss:

$$(2) \qquad Instrument_{i,t}^{Lose} = \sum_{j \in Fr(i)} \mathbf{1}\big(RandomPhoneLoss_{j,t}\big) \cdot ProbBuyRandomPhoneLoss_{j,t},$$

[11] We use the following characteristics as features when training our neural networks: current phone age, current phone model, carrier, user age, user gender, user browser, Instagram usage flag, user education level, US state, friend count, activity flags, account age, profile picture flag, number of friendships initiated, and area-level average income.

where $Fr(i)$ is the set of all users who are friends with user *i*. As discussed above, we add controls for the average of $ProbBuyRandomPhoneLoss_{j,t}$ among all of a user's friends in the IV regressions with this instrument. This step allows us to exploit small-sample variation in the probability of replacing a lost phone of the friends who randomly lose their phones in a given week, without capturing a possible direct relationship between the average conditional probability among a user's friends and that user's own probability of purchasing a new phone in that week.[12]

While the exclusion restriction is inherently untestable, we verify its plausibility by exploring whether our instrument is conditionally related to important observable user characteristics. In particular, for each user, we first calculate the unconditional probability that she purchases a phone in a given week, $ProbBuyUncond_{i,t}$, based on observable characteristics of the user (see online Appendix A.2 for details). In Figure 3, we then show the correlation between our instruments and the predicted probability that the user purchases a phone in a given week. In the top row, we explore the random phone loss instrument (equation (2)). Panel A shows unconditional relationships. We find that $Instrument_{i,t}^{Lose}$ is correlated with a user's own predicted probability of buying a new phone, probably due to homophily. However, panel B shows that after controlling for the characteristics of a user's overall group of friends—which are also included as controls in our IV specifications—there is no residual relationship between $Instrument_{i,t}^{Lose}$ and the estimated probability that an individual herself purchases a new phone. This lack of conditional correlation between our instrument and observable user characteristics that influence purchasing decisions supports the credibility of our identifying assumption that no such correlation exists with unobservable user characteristics, either.

It is important to point out that the set of compliers in a specification using $Instrument_{i,t}^{Lose}$ to instrument for the total number of phone purchases by friends is likely different to the set of compliers when using the total number of friends who break their phones. To the extent that these compliers differ in the strength of the peer effects they exert, the two approaches may therefore estimate different local average treatment effects (LATEs), though it is unclear whether either one of these LATEs would be preferable in terms of being more representative of a population average treatment effect.

B. *Phone Age Instrument*

Our second instrument is based on the observation that during the period of our study, there were two main contract structures in the US cell phone market. The first

[12] We also explore the possibility that the group of friends who would ever publicly post about a random phone loss is a selected subset of all of a user's total friends. In this case, controlling for the average conditional probability among all of a user's friends may not suffice to eliminate a possible direct relationship between the instrument and the errors in the second stage. To address this possibility, we also control for the average conditional probability of purchasing among a user's friends for whom $\mathbf{1}(RandomPhoneLoss_{i,t}) = 1$ at any point in the year prior to our sample period. In the case of a user having no such friends, we set their average probability to a value outside the normal range of the data (in our case, to -1), and we include a binary control for missing data. This procedure allows us to avoid dropping observations when the user had no friends who had $\mathbf{1}(RandomPhoneLoss_{i,t}) = 1$ in the prior 12 months.

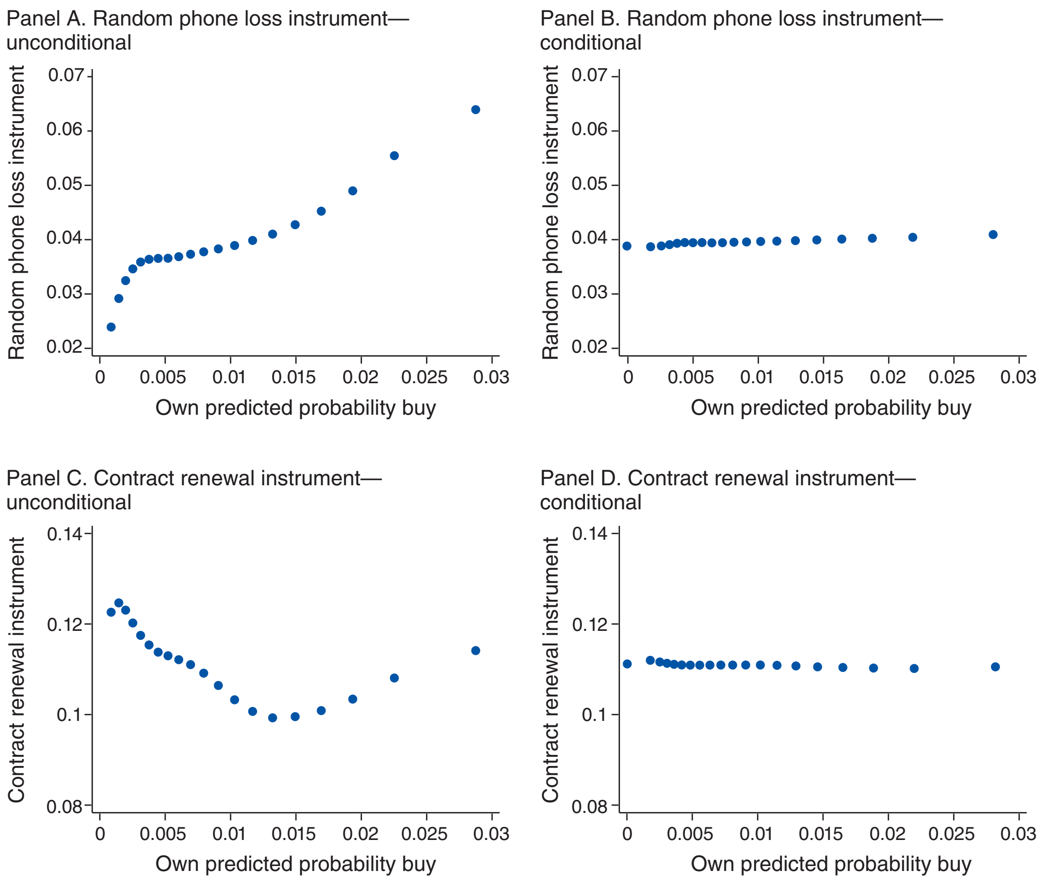

Figure 3. Conditional Independence of Baseline Instruments

Notes: Panel A shows the unconditional relationship between a user's own predicted probability to buy a new phone, $ProbBuyUncond_{i,t}$, on the horizontal axis and the random phone loss instrument, $Instrument^{Lose}_{i,t}$, on the vertical axis. Panel B shows the same relationship but conditions on the controls included in equation (6), with the exception of $ProbBuyRandomPhoneLoss_{j,t}$, the horizontal axis variable. Panels C and D in the bottom row show the analogous relationships for the contract renewal instrument.

involved month-to-month contracts in which a user would purchase her own phone. This type of contract was offered primarily by T-Mobile, AT&T, and MetroPCS. The second contract structure involved carriers subsidizing customers' phone purchases in exchange for a two-year service commitment at a set price. Service of this kind was offered primarily by Sprint and Verizon during that time.

Figure 4 shows the weekly probability of a user obtaining a new phone by the age of their current phone. Panel A shows that this probability is generally increasing in phone age, but it spikes when phones cross the two-year age threshold (the dark gray area). Panel B, which shows this probability separately by carrier, highlights that this spike is concentrated among customers whose service is provided by Verizon or Sprint.

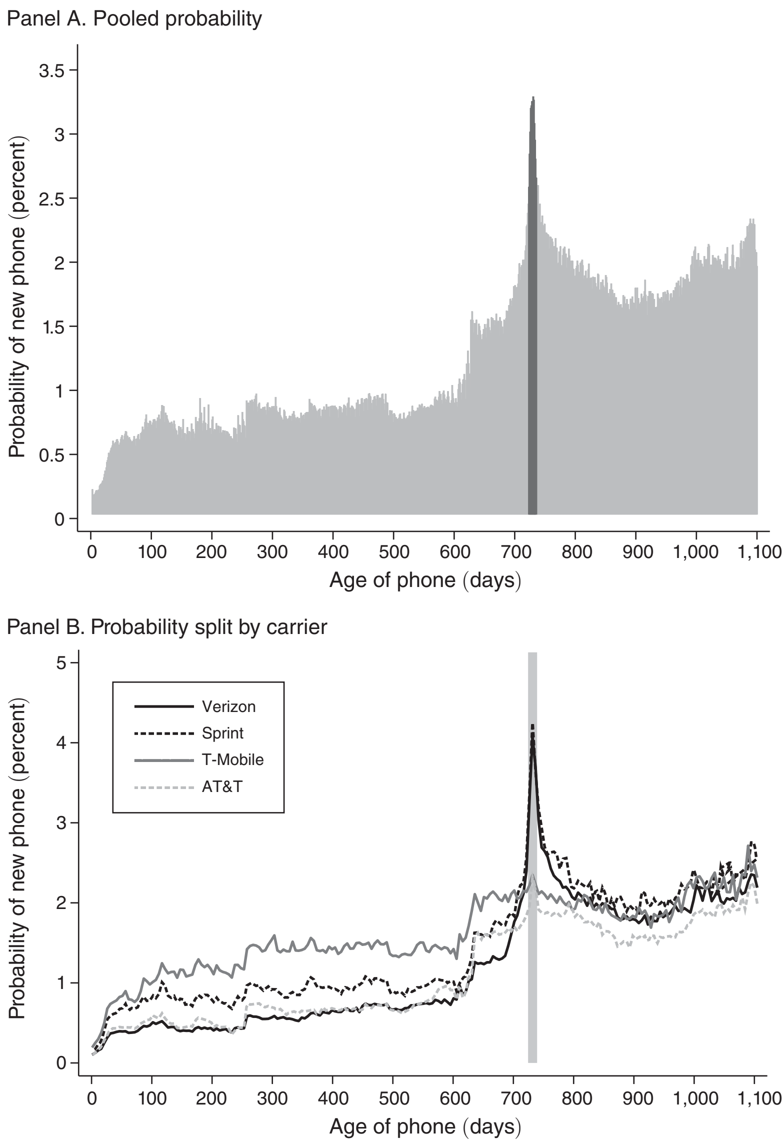

Figure 4. Probability of New Phone by Phone Age

Notes: Panel A shows how a user's probability of getting a new phone varies with the age of their current phone. Panel B shows the same split by user carrier.

As before, we use a neural network to estimate, for each consumer, the probability of buying a new phone in the week when his current phone is two years old:

$$(3) \qquad ProbBuy2y_{i,t} \,=\, Prob\Big(\mathbf{1}\big(BuysPhone_{i,t}\big)\,|\,X_{i,t}, \mathbf{1}\big(Phone2yOld_{i,t}\big) \,=\, 1\Big),$$

where $\mathbf{1}\big(Phone2yOld_{i,t}\big) \,=\, 1$ is an indicator that is set to 1 for individuals whose phones are between 721 and 735 days old. As suggested by panel B of Figure 4, a key predictor here is a user's current carrier, but other demographic characteristics included in $X_{i,t}$ also influence this conditional probability. We then instrument for the number of friends who get a new phone with the sum of $ProbBuy2y_{j,t}$ across all friends who are at the two-year phone age threshold in a given week:

$$Instrument_{i,t}^{2y} = \sum_{j \in Fr(i)} \mathbf{1}\left(Phone2yOld_{j,t}\right) \cdot ProbBuy2y_{j,t}. \tag{4}$$

Since individuals who have more friends with older phones are plausibly different from individuals with friends who have younger phones, we directly control for the number of friends whose phones are between 721 and 735 days old. We also add controls for the number of friends who were at the 2-year phone age threshold in the 12 months prior to our sample, as well as the average value of $ProbBuy2y_{j,t}$ among those people, in addition to the average value of $ProbBuy2y_{j,t}$ among all friends. By including these controls, we are effectively using only small-sample variation in the conditional probabilities of a user's friends who are at the contract renewal threshold in a given week, without using variation in the number of these friends. The bottom row of Figure 3 shows that after including these controls, there is no relationship between $Instrument_{i,t}^{2y}$ and a user's own estimated probability of purchasing a phone in a given week, $ProbBuyUncond_{i,t}$.

C. *Empirical Specification and Inference*

Using these instruments, we estimate instrumental variables regressions to measure peer effects in the cell phone market. The first and second stages of the IV regression, respectively, are

$$FriendsBuyPhone_{i,(t-1,t)} = \delta Instrument_{i,t-1} + \omega X_{i,t} + e_{i,t} \tag{5}$$

$$\mathbf{1}\left(BuysPhone_{i,t}\right) = \beta \widehat{FriendsBuyPhone}_{i,(t-1,t)} + \gamma X_{i,t} + \epsilon_{i,t}. \tag{6}$$

The key dependent variable in the second stage, $\mathbf{1}\left(BuysPhone_{i,t}\right)$, is an indicator of whether individual i purchases a new phone in week t. The vector $X_{i,t}$ represents a rich set of fixed effects and linear controls based on characteristics of the users and their friends. In addition to the controls we already discussed above, we include fully interacted fixed effects for user characteristics (age bucket $\times$ gender $\times$ education $\times$ state $\times$ week), device characteristics (device $\times$ carrier $\times$ phone age buckets $\times$ week), and friend characteristics (number of friends $\times$ number of friends switching phones in the last 6 months $\times$ week). We also control for the predicted (unconditional) probability that a user purchases a phone in that week, $ProbBuyUncond_{i,t}$. In the online Appendix, we show that our baseline results are robust to different specifications of the controls and fixed effects.

Our instrument in the first-stage regression is based on shocks to friends in week $t-1$ (e.g., the number and characteristics of friends who broke their phones in that week). The IV estimate β corresponds to the total user purchases in week t that were induced by the instrument, scaled by the first-stage estimate δ of how many relevant friend purchases were induced by the instrument. This scaling should account for all friend purchases caused by the instrument that occurred prior to the user's purchasing decision in week t and that could thus have influenced that purchasing decision. As mentioned above, our data do not allow us to precisely pinpoint the timing of purchases, and Figure 2 shows that friends who randomly lose their phones in week $t-1$ have a somewhat elevated purchasing probability in week t. An analogous, though weaker, increase in purchasing in week t occurs when a user reaches the contract renewal threshold in week $t-1$. We therefore include all friend purchases in weeks t and $t-1$ in our endogenous variable, $FriendsBuyPhone_{i,(t-1,t)}$:

$$(7) \quad FriendsBuyPhone_{i,(t-1,t)} = \sum_{j \in Fr(i)} \mathbf{1}\left(BuysPhone\right)_{j,t-1} + \sum_{j \in Fr(i)} \mathbf{1}\left(BuysPhone\right)_{j,t}.$$

This approach potentially overcounts the relevant number of instrument-induced purchases of new phones by friends since it can include some friend purchases in week t that occurred after the user has already purchased a phone in that week; as a result, the second-stage coefficient estimates of β provide a conservative measure of the magnitude of peer effects.[13]

Inference.—In any setting where peer effects might be important (whether or not they are the focus of the analysis), these peer effects can introduce a correlation in the error terms across individuals. Such a correlation would invalidate the independence assumptions used to derive the asymptotic properties of standard estimators. In a world with non-overlapping network communities, one can account for this possible across-observation dependence due to peer effects by clustering standard errors at the level of the community. For complete networks like the one we are studying, statistical inference remains a relatively open area of research, and our vast sample size limits our ability to use the social graph to fully model the structure of the variance-covariance matrix (similar issues arise in a literature that explores the use of cluster-robust estimators when working with spatially dependent data) (see Bester, Conley, and Hansen 2011). We therefore follow a number of recent papers to explore the robustness of our statistical inference to various approaches of constructing standard errors. In particular, Eckles, Kizilcec, and Bakshy (2016) and Zacchia (2020) propose to partition the social graph into a number of communities with limited cross-community dependence and to then cluster the standard errors at the community level.[14] Even though the presence of some across-cluster

[13] Using only friend purchases in week $t-1$ as the endogenous variable would instead undercount the relevant friend purchases induced by the instrument since it would miss purchases that occurred early in week t (before the user's own purchasing decision in that week). It would thus understate the first stage (and overstate the second stage), providing an upper bound on the magnitude of peer effects rather than a lower bound, as our baseline specification does.

[14] Operationally, we start with a dataset that uses a distributed variant of the Kernighan-Lin algorithm to divide the global Facebook social graph into about 21,000 distinct communities. Individuals in our sample are assigned to their communities created by this graph. The 0.2 percent of our sample assigned to communities with fewer than 100 other members of our sample are grouped into a "residual" community (these individuals are likely to be recent

friendship links implies that there remains the potential for across-cluster correlation in the error terms, this clustering approach substantially reduces potential biases in standard errors from such dependencies. Online Appendix A.3 shows that our standard errors are essentially unaffected when moving from heteroskedasticity-robust standard errors to community-robust standard errors. This suggests that in our setting, statistical inference is not substantially affected by residual across-individual dependencies in error terms.

III. Peer Effects in Phone Purchasing

We next explore how a user's propensity to purchase a new phone is affected by the phone purchases of her friends. We begin by presenting the baseline estimates of peer effects. Section IIIA then explores the timing of these peer effects, showing that an individual acquiring a new phone increases the aggregate propensity that her friends purchase a new phone for at least several months. In Section IIIB, we explore heterogeneities in both influence and susceptibility to influence across demographic characteristics.

Baseline Results.—Column 1 of Table 2 presents OLS estimates from regression (6). The results suggest that having one more friend purchase a phone in weeks t or $t-1$ increases a person's own propensity to buy a phone in week t by 0.032 percentage points. This estimate is large relative to a baseline probability of purchasing a new phone of just under 1 percentage point per week. However, as discussed above, this OLS estimate might also pick up the effects of common shocks or preferences in addition to any peer effects. The rest of Table 2 therefore presents causal effects from IV estimations. Columns 2 and 3 show the reduced forms from the random phone loss instrument and the contract renewal instrument, respectively, while columns 4 and 5 show the corresponding second-stage estimates.

Both second-stage IV estimates are similar in magnitude to the OLS estimate: the IV estimate is slightly larger than the OLS estimate when using the random phone loss instrument, and it is slightly smaller than the OLS estimate when using the contract renewal instrument; neither of these differences is statistically significant. This similarity in estimated peer effects across OLS and IV specifications is perhaps surprising since one might have expected that common shocks or common preferences would lead to a substantial upward bias in the OLS estimates. In contrast, our result here suggests that—after controlling for observable characteristics of individuals and their friends—correlated unobservable shocks or preferences induce at most a small bias to our OLS estimates, at least when analyzing the effect of peer purchases on the near-contemporaneous purchasing behavior of individuals.

In terms of magnitudes, a simple back-of-the-envelope calculation suggests that a new phone purchase by an individual in one week leads to an additional 0.08 phone

immigrants, who are members of communities where most members are outside the United States). Overall, the 81 million users in our primary sample are assigned to 5,140 distinct communities with an average size of 15,910. The average user in our sample has 53.4 percent of her friends within the same community; at the tenth/fiftieth/ninetieth percentile of our sample, this number is 21%/54%/84%.

Table 2—All Instruments—All Phones

	OLS	Reduced form: Broken phone	Reduced form: Contract renewal	Second-stage: Broken phone	Second-stage: Contract renewal
	(1)	(2)	(3)	(4)	(5)
Number of friends buying (t − 1 and t)	0.032			0.040	0.022
	(0.000)			(0.005)	(0.013)
Instrument		0.046	0.024		
		(0.007)	(0.014)		
Controls + Fixed Effects	Y	Y	Y	Y	Y
Mean dependent variable	0.93	0.93	0.93	0.93	0.93
Number of observations	329m	329m	329m	329m	329m
Effective *F*-statistic				4,627	878

Notes: Table shows estimates of regression (6). Column 1 presents the OLS estimate, columns 2 and 3 present reduced-form estimates using our two instruments, and columns 4 and 5 present the corresponding second-stage IV estimates. The dependent variable in all specifications is an indicator for whether user i purchases a new phone in week t. All coefficients reported are multiplied by 100 to ease interpretability. We include interacted fixed effects for individual i's demographics (age bucket × state × gender × education), individual i's beginning-of-week device (current phone × current phone age in buckets of 50 days × carrier), and individual i's friends (total friends × number of friends switching phones in the previous 6 months). We control linearly for the user's unconditional probability of buying a new phone, estimated as described in online Appendix A.2 and for the average conditional purchase probability among a user's friends. In columns 2 and 4, we additionally control for individual and friend posting behavior (the number of friends with public statuses, the number of friends posting in a given week, the number of friends who post about random phone loss in the 12 months prior to our sample, the average conditional probability of buying a new phone among friends who posted about random phone loss in the prior 12 months, and a dummy for whether the user herself posted about a random phone loss in the given week). In columns 3 and 5, we additionally control linearly for the number of friends whose phones are between 721 and 735 days old, the number of friends who have had phones of this age in the 12 months prior to our sample, and the average conditional probability of buying a new phone among those friends. We report Olea and Pflueger (2013) effective *F*-statistics. Standard errors are clustered at the level of the community (see the discussion in Section IIC and online Appendix A.3).

purchases through peer effects in the following week.[15] Put differently, a little less than one in ten phone purchases causes a follow-on purchase in the subsequent week through peer effects.[16]

One interesting question is whether the estimated treatment effects are the result of individuals hearing about their friends' new phone purchase through Facebook or through off-line interactions. We think that at most a small part of the overall

[15] The average peer of people in the sample has 258 friends (which is lower than the average number of friends of people in the sample, which was restricted to only include individuals with at least 100 friends), and a new purchase by these peers increases the probability of each friend purchasing a new phone the following week by about 0.032 percentage points (the average of the two IV estimates). A simple back-of-the-envelope estimate of the overall effect is thus $258 \times 0.00032 \approx 0.08$.

[16] We rule out two possible alternative explanations for the patterns in Table 2. First, we explore if they might primarily capture the correlated behavior of family members as a result of contract incentives such as "Buy One, Get One Free" offers that are sometimes available for members of the same family plan. When we repeat our analysis after excluding each user's family members from their friends (where we identify family members through a combination of self-reports and model-based imputations), we find baseline estimates of very similar magnitude. In addition, while "Buy One, Get One Free" offers might in principle explain correlated purchases that are close in time, they could not explain the long-lasting patterns we show in Section IIIA. We also find that our estimates are not driven by Facebook disproportionately advertising cell phones to people whose friends recently experienced a random phone loss or whose friends' contracts were up for renewal. To show this, we repeat our baseline regressions only for users who did not see any cell phone ads on Facebook during our sample period. The peer effects we estimate in this sample are near identical to those in the full sample. The finding is consistent with our institutional understanding of the scope of ad targeting.

observed peer effect comes from interactions on Facebook—indeed, in this setting, we view Facebook primarily as a tool to measure phone purchases and social networks, instead of the primary medium for information flow. There are a number of reasons for this. First, only about 2.3 percent of individuals who post about losing their current phone actually post about purchasing a new phone in the following weeks, and even then, Facebook posts are usually seen by only about a quarter of an individuals' friends (Bernstein et al. 2013). Second, we highlight below that peer effects from geographically proximate friends are substantially larger, suggesting an important role of in-person interactions in propagating information about new phone purchases. Third, we show below that the effect of a friend's phone purchase on own purchasing behavior is strong for a number of months following the friends' purchase. We think it is much more plausible that this effect comes through hearing about the friend's purchase over time (as well as through second-order peer effects), instead of the delayed effect capturing the purchase of a new phone many months after viewing a social media message—in particular given the evidence that individuals remember only a fraction of social media content even immediately after viewing it (Counts and Fisher 2011).

The difference in magnitude across the two IV estimates in columns 4 and 5 of Table 2 highlights that the local average treatment effects we capture using each of these instruments may differ from the average treatment effect in the population. Specifically, our first instrument captures the average peer effects of individuals who post publicly about losing their phones (and who then quickly purchase a new one) on those individuals' friends. Our finding suggests that the peer effects exerted by these individuals may be somewhat larger than the average peer effects in the population, perhaps because individuals who quickly replace a (partially) broken phone care a lot about phones and are therefore more likely to influence their friends. In addition, due to homophily, the users who are friends with these people may themselves be more interested in phones, so their own purchasing behavior may be more affected by peer effects than that of the average person. In contrast, the IV coefficient estimated using the contract renewal instrument identifies the average peer effects from individuals who keep the same phone for two years before replacing it. As can be seen from Table 1, a two-year-old phone is in the right tail of the phone age distribution. This result suggests that users who wait that long to replace their phones may be less interested in up-to-date technology than the average user, perhaps explaining why eventual purchases by these individuals have a below-average effect on the purchasing behavior of their peers.

These differences in local average treatment effects raise the possibility for substantial heterogeneities in peer effects, both along characteristics of the potential influencers and characteristics of the individuals who are potentially influenced.[17] We explore these heterogeneities, which have important implications for firms' marketing strategies and price-setting behaviors, in Section IIIB.

[17] The differences in LATEs across instruments also suggest a potential alternative interpretation of the observation that OLS and IV estimates have similar magnitudes. In particular, it could still be the case that the OLS estimate presents a substantially upward-biased estimate of the true average peer effect in the population, and at the same time that the IV estimates both correspond to LATEs capturing the peer effects from relatively influential individuals, with the two effects approximately offsetting each other.

A. *Peer Effects at Longer Horizons*

The specifications reported in Table 2 analyze the effects on an individual's phone-purchasing behavior immediately following a new phone acquisition by a peer. In this section, we explore two related questions. First, for how long does the purchase of a phone by a peer influence an individual's own purchasing behavior? Second, do these peer effects primarily represent the retiming of already-planned purchases, or do they generate purchases that would not have happened otherwise?

To address these questions, we expand the horizon over which we measure a user's phone purchasing behavior to include up to 43 weeks following the initial phone purchase by a peer. Specifically, we construct dependent variables of the form $\mathbf{1}\left(BuysPhone_{i,(t,t+3)}\right)$, $\mathbf{1}\left(BuysPhone_{i,(t+4,t+7)}\right)$, and so on, to capture whether a user purchases a new phone during a number of four-week periods. In Figure 5, we report the β-coefficients from using these variables as dependent variables in regression (6). Though these regressions are similar to our baseline specification reported in Table 2, the interpretation of the longer-horizon effects is somewhat more complicated. In particular, since individuals and their friends often have many friends in common, second-degree peer effects become increasingly relevant at longer time scales: a friend's purchase in week t may influence a common friend's purchase in week $t+1$, which in turn affects the user's own purchasing decision in week $t+2$. The coefficients presented in Figure 5 provide the LATEs associated with a friend purchasing a new phone in weeks t or $t+1$ on the user purchasing at various horizons, capturing both the direct effect of the initial friend purchase and any higher-order effect of purchases by common friends that were caused by the initial purchase.

A number of patterns emerge from the IV coefficients in Figure 5. First, having an extra friend purchase a new phone in response to a random phone loss is not associated with an elevated probability of a user herself purchasing a new phone in the weeks prior to the random phone loss by the friend (this probability is even marginally lower in the month prior to the friend's random phone loss, though the effect is barely significant and tiny in magnitude). This finding provides further support for the exclusion restriction associated with the random phone loss instrument, which requires that individuals with and without a randomly induced phone purchase by a friend would behave conditionally similarly in the absence of the random friend purchase.

Second, Figure 5 shows that the effect on user purchasing of having an extra friend randomly buy a new phone in week $t = 0$ is roughly as large over the first four weeks following the friends' purchase as it is over each of the subsequent three months. After that, the aggregate effect declines and generally stabilizes. During the period that we observe, the aggregate effect on own purchasing behavior in response to a friend replacing a lost phone does not show signs of a reversal. This finding implies that peer effects induce an increase in the total level of phone purchases and not merely a shift in the timing of a fixed number of purchases.[18] The observed

[18]This result does not mean that no individuals have their purchases pulled forward through peer effects. Indeed, in all weeks $t' > 1$, there are two countervailing forces that determine the aggregate effect of a random phone purchase in week $t = 0$ and $t = 1$ on the total purchases by all the person's friends. Firstly, there are potentially negative effects on the purchasing probability of people who had their purchases pulled to previous weeks

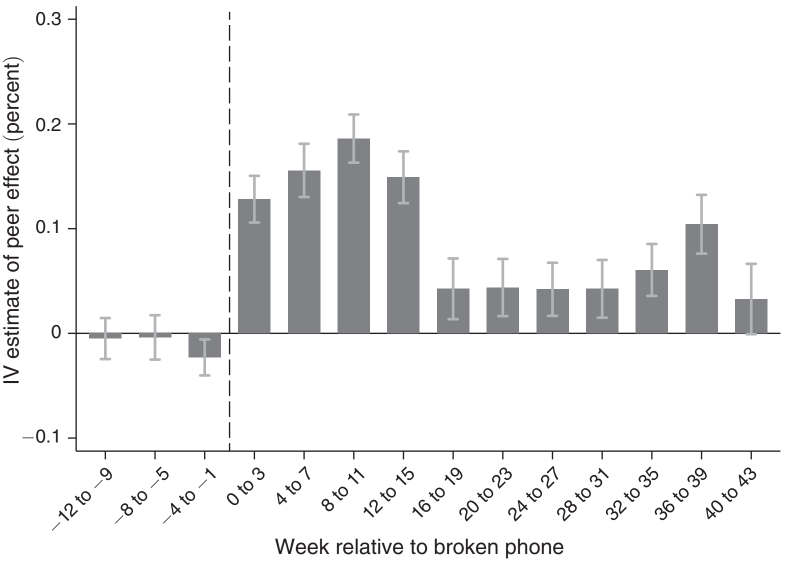

Figure 5. Peer Effects at Alternative Horizons

Notes: Figure shows estimates from IV regression (6) at various horizons. The dependent variables are indicator variables for whether a user purchases a new phone in the given four-week period. The IV coefficients capture the total effect of friend purchases in weeks $t = 0$ or $t = 1$, induced by a random phone loss in week $t = 0$. Error bars show 95 percent confidence intervals.

cumulative increase in purchasing probability is economically meaningful: having an additional friend who purchases a phone in week t increases the chance that a user purchases a phone between weeks t and $t + 15$ by 0.6 percentage points. In our sample, the average chance that a given user purchases a cell phone over this period is 14.8 percent, so a friend's purchase increases the user's own probability of buying a new phone in the next 4 months by about 4 percent of the baseline probability.

B. *Heterogeneities in Treatment Effects*

The previous observation that our two instruments identified LATEs of different magnitudes hinted at the presence of substantial heterogeneities in peer effects. In particular, it suggested that those friends whose behavior was shifted by each of our instruments might be differentially influential on average. To further explore such heterogeneities, we next analyze how peer effects vary with the observable characteristics of users and their friends. These heterogeneities are estimated with IV regressions using the random phone loss instrument, which has the most power; online Appendix A.4 provides the exact regression specifications. Directionally, the patterns of heterogeneity in the resulting LATEs are generally similar to the patterns of heterogeneity in the corresponding OLS estimates, suggesting that they are not only a feature of the local average treatment effects identified by our random

$0 < t < t'$. However, any such effects are more than offset by positive effects on the number of total purchases through delayed or higher-order peer effects.

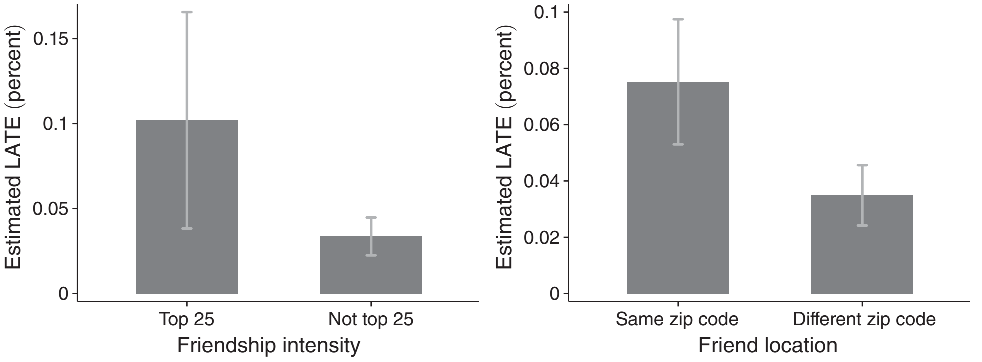

Figure 6. Peer Effects Heterogeneity by Relationship Strength and Geographic Proximity

Notes: Figure shows IV estimates of equation (6) using the random phone loss instrument. In the left panel, we split each user's friends into those inside and outside the top 25 using a model of friendship intensity. In the right panel, we split all friends into those living in the same predicted zip code and those living in a different predicted zip code as the user. Error bars show 95 percent confidence intervals.

phone loss instrument but also of the (potentially biased) average treatment effects obtained through OLS analysis.

Heterogeneity by Relationship Strength and Geographic Proximity.—We first explore whether the magnitude of the peer effects we observe is affected by the strength of the relationship of the user-friend pair. To measure the closeness of friendship links, we rank a user's friendships according to a model of tie strength based on characteristics such as mutual friends and interaction frequency, similar to Gilbert and Karahalios (2009). The left panel of Figure 6 shows that the estimated peer effect from a friend in the top 25 closest friendships is more than twice as large as the peer effect from a friend who is not in the top 25 (we choose this cutoff since tie strength declines much less strongly across ranks beyond the top 25 friends). It is reassuring that peer effects from closer friends are larger. In fact, there are a number of possible explanations that are consistent with this finding. First, purchases by these friends may be more salient to a user, perhaps because she is more likely to interact with these friends. Second, it is likely that individuals are more willing to trust information that they receive from closer peers. Third, the desire to keep up with closer friends may be higher than the desire to keep up with friends who are less close.

We also explore whether peer effects from geographically proximate friends are larger than those from friends who live further away. The right panel of Figure 6 shows that the estimated peer effect from a friend who lives in the same predicted zip code is more than twice as large as the peer effect from a friend who lives in a different predicted zip code. This evidence is highly consistent with our previous interpretation that much of the observed peer effects are the result of in-person interactions between individuals, which are more likely to occur when two individuals live close to each other than when they live further apart.

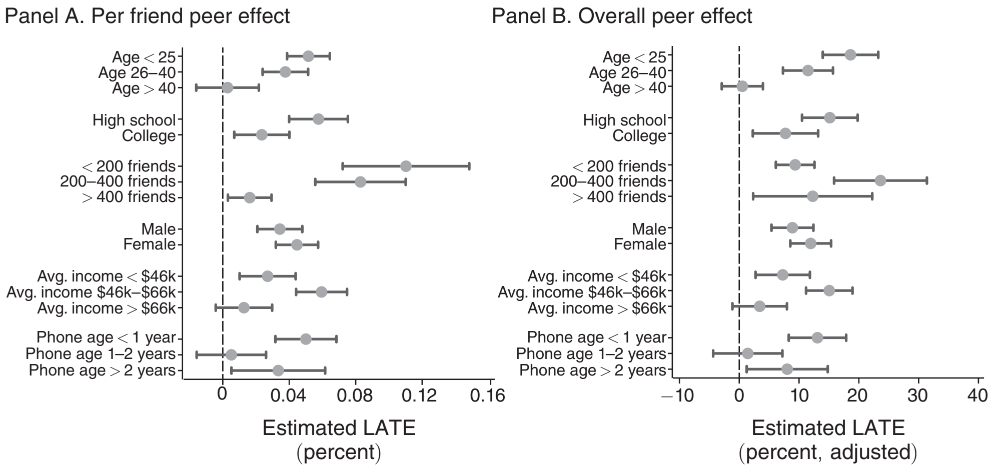

Figure 7. Peer Effect Heterogeneity by Friend Characteristics

Notes: Figure shows IV estimates of equation (6) using the random phone loss instrument. Estimated peer effects are split by characteristics of the peer with the random phone loss. Panel A shows the mean peer effect a user in each group exerts on each of her friends. Panel B reports the average total influence of a user in each group on all of her friends, computed by multiplying the coefficients found in panel A with the average number of friends in each demographic group. We report the full specifications in the online Appendix. Error bars show 95 percent confidence intervals.

Heterogeneity by Friend Characteristics.—We next explore heterogeneities in the magnitude of peer effects exerted by different individuals. Identifying characteristics of socially influential individuals is an important exercise for marketing researchers and practitioners, and "influencer campaigns" are now an integral part of most consumer marketing strategies (see Ferguson 2008; Tucker 2008; Bakshy et al. 2011; Aral and Walker 2012). Here, we contribute to this research effort by documenting demographic characteristics that are indicative of large social influence and by exploring how social influence and price sensitivity are correlated across demographic groups. We discuss that the latter correlation has important implications for firms' dynamic price-setting behavior.

Figure 7 documents heterogeneity in peer effects along peer demographic characteristics. Panel A shows the "per friend" peer effect, corresponding to the causal effect of a purchase of a new phone by a person with those characteristics on average on each of their friends. Panel B measures the "overall" peer effect, which adjusts the per friend peer effect by the fact that different demographic groups have differentially many friends. This second category is of particular interest for designing influencer-based marketing campaigns. We find that younger individuals exert larger peer effects on each of their friends. Combined with the fact that these individuals have more friends on average, we find that the overall peer effect exerted by individuals declines substantially in age. This finding suggests that acquiring younger customers is more valuable to firms than acquiring older customers, at least in the phone market, since younger customers will generate more follow-on demand through peer effects.

We also find that the peer effects exerted by individuals who report high school as their highest education level are larger than the peer effects exerted by individuals who report having gone to college. In addition, we find that the per friend peer effect exerted by individuals is declining in the number of friends they have, perhaps because the marginal friend is less close. However, despite the declining influence on each friend, the overall peer effects do not follow a similarly monotonic pattern. Users with between 200 and 400 friends seem to have the most influence in aggregate, having a large per friend peer effect and relatively many friends. We find that women are somewhat more influential than men, although these differences are relatively small. We also find that users from middle-income areas are more influential than users from richer or poorer areas. Users who lose a phone that is less than one year old have the largest influence on the purchasing behavior of their friends (recall from Table 1 that the median phone age in our sample was 317 days). In turn, individuals who do not regularly replace their phones—and who are therefore likely to not value new technology as much—exert smaller peer effects on their friends. The peer effects of these people may be lower both because they are less likely to talk to their friends about having a new phone and because they may be perceived as less valuable sources of information when they do talk to their friends.[19]

Peer Influence versus Price Sensitivity.—One important implication of peer effects is that the aggregate demand curves faced by firms are more elastic than individual demand curves (see Glaeser, Scheinkman, and Sacerdote 2003). The magnitude of this difference depends in part on the correlation between individuals' price elasticities and the magnitude of the peer effects they exert. Specifically, if a price cut primarily attracts additional demand from individuals who exert only small peer effects, the difference between the individual and aggregate demand curves will be substantially smaller than when a price cut primarily increases the demand of individuals who exert large peer effects.

In our data, we do not have individual-level estimates of price sensitivity. To explore whether the most influential individuals are likely to have relatively high or relatively low price sensitivity, we split individuals into 8 mutually exclusive groups along the interacted dimension of user age (above or below 35 years), user phone age (above or below 1 year), and user gender. We estimate the per friend influence and the total influence for each of these eight groups using instrumental variables specifications similar to the ones described above. We also measure the price sensitivity of each group by calculating the percentage increase in the number of users in each group who purchase an iPhone 6s or iPhone 6s Plus in the week before and after a major price cut in September 2016.[20]

[19] When comparing the 16 IV coefficients presented in the left panel of Figure 7 to the corresponding coefficients from an OLS specification, we obtain a correlation of 0.81, suggesting that our conclusions regarding the relative strength of peer effects of different individuals may generalize beyond the specific LATE studied here. The main difference is in the heterogeneity by age, where there are fewer differences across age groups in the OLS specification than in the IV specification.

[20] On September 7, 2016, Apple announced an immediate price cut of $100 for the iPhone 6s and the iPhone 6s Plus. We use purchasing data from one week on each side of this date to measure price sensitivity, but our findings are robust to comparisons that use several weeks on either side of the price cut to determine the price sensitivity of each group. In the week following this price cut, we observe a 4 percent increase in the number of iPhones registered, with heterogeneity in the size of this jump across demographic groups.

We next explore the correlation between peer influence and price sensitivity across the eight groups. We find the correlation with per friend influence to be 0.89 and the correlation with total influence to be 0.90.[21] This result suggests that price cuts disproportionately attract extra demand from individuals who are relatively influential and that the deviations between individual and aggregate demand curves in this market are thus likely to be large. The higher implied price elasticity of aggregate demand will push firms toward setting lower prices than they would in the absence of peer effects.[22]

The positive correlation between price sensitivity and peer influence may also provide an explanation for the sometimes puzzling observation that many markets clear through queuing rather than through price adjustments. If higher prices disproportionately reduce demand from those individuals with large peer effects on their friends, then an optimal dynamic pricing strategy might be willing to trade off lower revenues today in return for additional sales generated through peer effects in future periods. In other words, while increasing the price would increase revenues today, it might reduce overall long-run revenues due to substantially lower peer effects going forward. In scenarios in which demand exceeds supply and firms do not want to increase prices to avoid selling to less influential individuals, an alternative assignment mechanism is required. Assignment via queuing is likely to disproportionately select individuals who might exert the largest peer effects among those willing to buy at the low price. This mechanism can help rationalize, for example, why Apple does not increase the price for its iPhones, despite the large queues outside its stores around device release dates. Similar mechanisms might be at work in other settings where limited supply is assigned through queuing that can help select individuals who will exert particularly large peer effects and thus generate subsequent sales (e.g., new sneakers, new restaurants, or the famous Cronuts).

Heterogeneity by User Characteristics.—Figure 8 explores heterogeneities in the susceptibility to influence of different individuals, separating users along the same demographic characteristics as in Figure 7. There are only small differences in susceptibility to influence across most demographic groups. The exception is that a user's number of friends is a major determinant of their susceptibility to influence from the average friend. The findings are consistent with the marginal friend being less close and therefore less influential for a user's purchasing behavior.

[21] We obtain similar correlations when we estimate peer effects with OLS regressions (acknowledging the potential biases in these specifications), suggesting the patterns may be generalizable beyond the specific LATE considered here. We also expand this exercise by further splitting each group into those with more or fewer than 300 friends, providing us with estimates of peer influence and price sensitivity for 16 mutually exclusive groups. Despite the fact that the estimates for peer influence are substantially noisier, the correlations across these objects are 0.66 and 0.22 for the per friend and total peer effects, respectively. Running IV regressions with more endogenous variables is not computationally feasible, preventing us from extending our analysis to consider the correlation between price sensitivity and peer effects at finer demographic splits.

[22] Through this channel, peer effects are a force that lowers markups and improves consumer welfare and allocative efficiency in this market. Online Appendix A.5 presents a simple model that formally explores this relationship between the correlation of peer influence and price sensitivity, the aggregate demand elasticity, and price markups. It is possible that peer effects may affect optimal price-setting through other channels (see Easley and Kleinberg 2010; Campbell 2013; Garcia and Shelegia 2018), and the overall effect of peer effects on prices depends on the relative importance of these various channels.

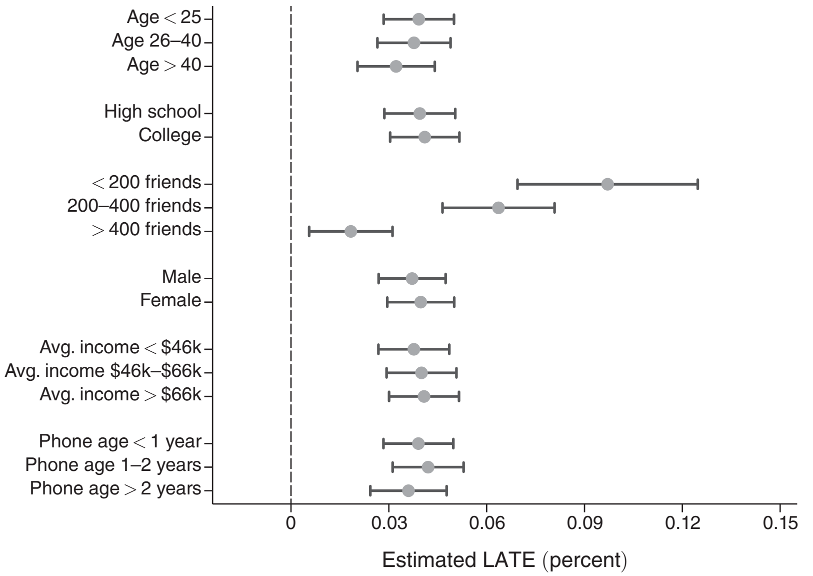

Figure 8. Peer Effect Heterogeneity by User Characteristics

Notes: Figure shows IV estimates of equation (6) using the random phone loss instrument. Estimated peer effects are split by user characteristics. We report the full specifications in the online Appendix. Error bars show 95 percent confidence intervals.

Heterogeneity by Peer and User Characteristics.—In the final set of heterogeneity analyses, we explore peer effects along characteristics of both the user and the peer. For example, we explore whether all individuals are primarily influenced by peers who are similar on observable characteristics or whether all individuals are most influenced by the same types of peers, regardless of their own characteristics.

Panel A of Figure 9 shows the cross-heterogeneity of peer effects by area-level income. Across all user income groups, friends from middle-income areas tend to be the most influential. Panel B shows that, for both high school–educated and college-educated users, high school–educated friends have the largest peer effect. Panel C shows that men and women are both more influenced by female friends than by male friends, though this effect is somewhat larger for female users. Panel D shows that younger users generally have the largest peer effects on their friends, with friends aged 25 years or less having particularly large effects on users older than 40 years. The only exception is the large effect of friends over 40 years old on users below 25 years old, although these peer effect estimates are not very precise, and they could be capturing correlated purchasing behavior between parents and children. Overall, these results suggest that individuals who are more influential on average are, in general, more influential on all users, not just those who are similar to them on demographic characteristics.

IV. Peer Effects for Specific Phone Purchases

In the previous section, we explored how a user's decision to purchase any new phone is affected by whether her friends recently acquired a new phone. In this

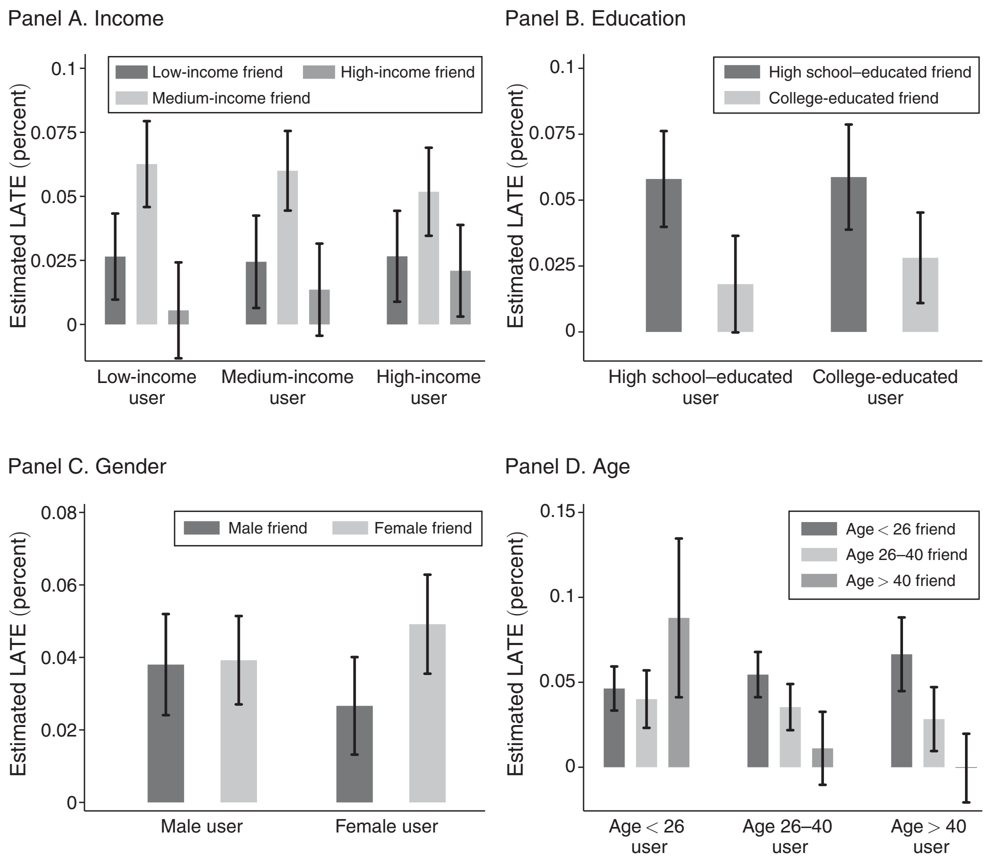

Figure 9. Peer Effect Heterogeneity by Pairwise Characteristics

Notes: Figure shows instrumental variables estimates of equation (6) using the random phone loss instrument. Estimated peer effects are split by user and peer characteristics. We report the full specifications in the online Appendix. Error bars show 95 percent confidence intervals.

section, we study whether the observed peer effects are specific to the phone brand purchased by the peer and explore whether there are positive or negative spillovers to competing brands.

We focus on the two major cell phone lines, Apple's iPhones and Samsung's Galaxy phones, which are used by 55 percent and 27 percent of users in our data, respectively. We pool the other highly fragmented brands into a residual category, which includes a variety of phones operating largely on the Android system. The set of brand categories we consider is thus given by $C = \{iPhone, Galaxy, Other\}$. We are then interested in understanding how a friend's purchase of a phone in brand category $c \in C$ affects a user's probability of buying a phone in the same brand category as well as their probability of buying phones in a different category. This investigation allows us to explore, for example, whether a friend's purchase of an iPhone increases a user's own demand for all phones, including those of iPhone competitor Galaxy, or whether it primarily pulls demand away from Galaxys and toward iPhones.

Identification Challenge and Empirical Approach.—To give a concrete example of the challenge with identifying peer effects in phone brand choice, imagine that there are two individuals, Amy and Bob, both of whom have five friends. Among Amy's friends, four would buy an iPhone if they were to replace their current phones, while only one of Bob's friends would buy an iPhone. In addition, homophily on characteristics such as tech-savviness imply that both Amy and Bob are similar to their friends in terms of phone preferences: even in the absence of peer effects, Amy would likely buy an iPhone, while Bob would probably buy a different phone. As a result, standard OLS specifications that regress whether people buy a certain phone brand on whether their friends buy that same brand would not necessarily identify peer effects since correlated preferences (and correlated shocks) would induce similar purchasing behavior even in the absence of any peer effects.

To document the role of peer effects in determining the purchases of specific phone brands, we thus adapt the IV strategy described above. To conceptualize our approach, imagine now that there is a third person, Carl, who is very similar to Amy. Carl also has five friends, of whom four would purchase an iPhone if they were to replace their phones, and Carl's own propensity to purchase an iPhone is also very similar to that of Amy. Now imagine that, in a given week, both Amy and Carl have one of their friends break their phones. By chance, it happens that Amy's unlucky friend is one who is likely to replace her broken phone with an iPhone, while Carl's unlucky friend is likely to replace it with a Galaxy. Importantly, this variation in the phone brands bought by Amy's and Carl's friends is not driven by differences in the composition of their friends—our thought experiment holds this composition constant by construction. Instead, the brands purchased by Amy's and Carl's friends are determined by which of their friends randomly break their phones in a given week, something that should not be correlated with Amy's and Carl's normal purchasing preferences after controlling for the brand preferences of all friends. As a result, any difference in Amy's and Carl's probabilities of buying different phone brands in the weeks following their friends' random phone losses (and subsequent phone replacements) is informative about the causal role of peer effects.

To operationalize this research design, we first construct a measure of each individual's probability of purchasing a phone of each brand. Specifically, for each phone brand c, we fit a neural network to predict the propensity that individual i will purchase a phone of brand c in week t, based on observable characteristics of that individual. We predict both the unconditional probability of buying a phone of brand c, $ProbBuyUncond^{c}_{i,t}$, and the propensity of such a purchase conditional on posting about a random phone loss, $ProbBuyRandomPhoneLoss^{c}_{i,t}$ (the conditional and unconditional propensities are highly correlated across individuals). We estimate these propensities using information on individuals' demographics and current phones (see online Appendix A.2 for details). For instance, we find that older users prefer iPhones, while all users are more likely to buy a phone of the same brand as their current device. As before, neural networks allow us to uncover nonlinear and interactive relationships between the various observable characteristics.

Our research design then proposes to use the sum of these predicted propensities of the friends who randomly lose their phones as instruments for the number of friends buying a phone of the respective brand, controlling for the average propensities among all friends. As described above, this is based on the assumption that, conditional on the characteristics of all of a user's friends, it is random whether, in a given week, the friends who lose their phones are those who are more likely to purchase iPhones, Galaxys, or other phones. Formally, we calculate, for each individual and each phone brand c, the average conditional probabilities of purchasing phones in each brand category among all her friends as given by equation (8); this will be the central variable to control for the composition of different people's friends, which could be correlated with those people's own phone preferences.

$$(8)\quad AllFriendsAvgProbBuyRPL_{i,t}^{c} = \frac{1}{|Fr(i)|}\sum_{j\in Fr(i)} ProbBuyRandomPhoneLoss_{j,t}^{c}$$

We also sum up this probability among her friends who randomly lose their phones in a given week as given by equation (9); this will be our instrument for the number of friends purchasing a phone of a particular brand:

$$(9)\quad LossFriendsSumProbBuyRPL_{i,t}^{c} = \sum_{j\in Fr(i)} \mathbf{1}\left(RandomPhoneLoss\right)_{j,t} \cdot ProbBuyRandomPhoneLoss_{j,t}^{c}.$$

We indeed find homophily in the propensities to buy phones of a certain brand. The left column of Figure 10 plots, for each phone brand c, a user's own $ProbBuyUncond_{i,t}^{c}$ on the horizontal axis and our instrument for friend purchases of category c, $LossFriendsSumProbBuyRPL_{i,t}^{c}$, on the vertical axis. We find that individuals who themselves are more likely to purchase a certain brand usually have friends who are also more likely to buy that brand, although the relationships are not always monotonic.

The right column of Figure 10 shows the same relationship as the left column but conditions on a number of control variables also included in our regressions, the most important of which is $AllFriendsAvgProbBuyRPL_{i,t}^{c}$. Conditional on the brand preferences in the overall friend population, the brand preferences of those friends who randomly lose their phones in a given week are essentially uncorrelated with the brand preferences of person i, at least to the extent that those preferences are captured by observable characteristics such as demographics and current phone brand. This finding makes it more plausible that they are also uncorrelated with brand preferences based on unobservable characteristics of person i, an assumption that is at the heart of our identification strategy.[23]

[23] As described above, it is possible that the sample of users who post about losing or breaking their phones is a selected subsample of a user's friends. If this were the case, controlling for the average probability among all friends may not accurately capture the distribution from which the randomly shocked friends are drawn. We address these concerns by also controlling for the average value of $ProbBuyRandomPhoneLoss_{j,t}^{c}$ among a user's friends who posted about losing or breaking their phones in the 12 months prior to our sample. Our results are unaffected by the inclusion of these controls.

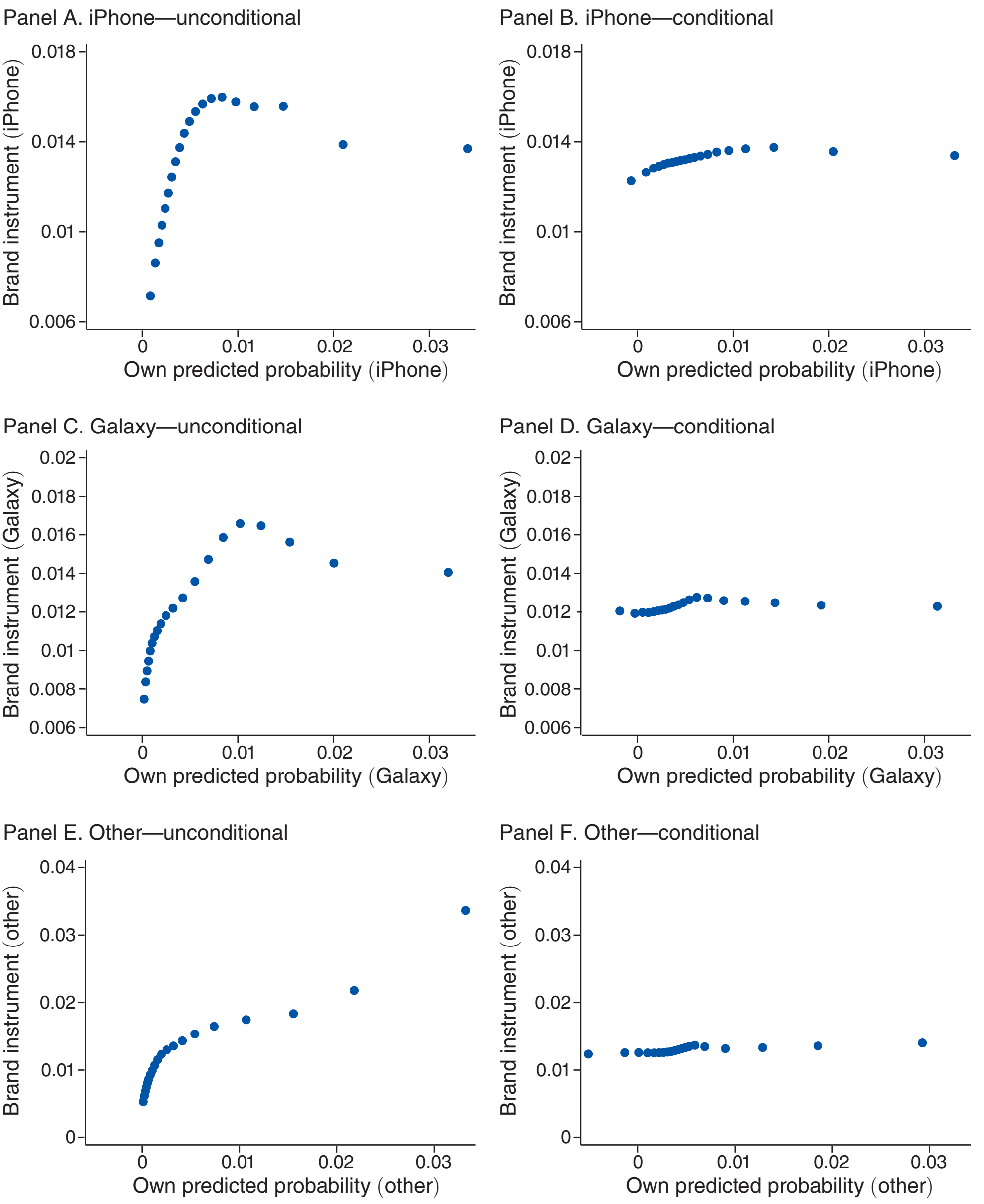

Figure 10. Conditional Independence of Brand Instruments

Notes: Figure shows the relationship between a user's own predicted probability to buy a specific new phone of brand category c, $ProbBuyUncond^{c}_{i,t}$, on the horizontal axis and the instrument, $LossFriendsSumProbBuyRPL^{c}_{i,t}$, on the vertical axis. The first row shows this relationship for $c = iPhone$, the middle row for $c = Galaxy$, and the bottom row for $c = Other$. The left column shows the unconditional relationship. The right column shows the same relationship but conditions on the controls included in equation (11), with the exception of $ProbBuyUncond^{c}_{i,t}$, the variable plotted on the horizontal axis.

Regression Specification.—To study peer effects at the brand level, we perform three instrumental variables regressions, one for each $c'' \in C$. We fit three first stages for each regression, i.e., one for each of the three brand categories $c' \in C$ that a friend could have bought:

$$(10)\quad FriendsBuyPhone^{c'}_{i,(t-1,t)} = \sum_{c \in C} \delta^{c'}_{c} LossFriendsSumProbBuyRPL^{c}_{i,t} + \sum_{c \in C} \phi^{c'}_{c} AllFriendsAvgProbBuyRPL^{c}_{i,t} + \omega X_{i,t} + e_{i,t}.$$

Our three second stages (one for each $c'' \in C$) are of the form

$$(11)\quad \mathbf{1}\left(BuysPhone\right)^{c''}_{i,t} = \sum_{c' \in C} \beta^{c''}_{c'} \widehat{FriendsBuyPhone^{c'}_{i,(t-1,t)}} + \sum_{c' \in C} \Phi^{c''}_{c'} AllFriendsAvgProbBuyRPL^{c'}_{i,t} + \gamma X_{i,t} + \epsilon_{i,t}.$$

The indicator variables $\mathbf{1}\left(BuysPhone\right)^{c''}_{i,t}$ capture whether person i purchased a phone of brand category c'' in week t. The coefficients of interest are comprised by the series of $\beta^{c''}_{c'}$, which capture the effects of a friend purchasing a phone in category c' on an individual purchasing a phone in category c''. The central control variable in both the first and second stages of the regression is the average conditional probability of buying a phone of each brand across all of individual i's friends, $AllFriendsAvgProbBuyRPL^{c}_{i,t}$. The vector $X_{i,t}$ includes the controls and fixed effects described in Section IIC as well as controls for the unconditional probability that user i buys a phone of each type $c \in C$ in week t, given by $ProbBuyUncond^{c}_{i,t}$, and the average of these propensities among the user's friends. We also estimate a fourth specification with $\mathbf{1}\left(BuysPhone\right)_{i,t}$ as the dependent variable, which allows us to examine whether friend purchases of certain brands led to more overall user purchases.

Since some of the (positive or negative) spillovers across brands would likely materialize only over time, we also study the effects of a friend purchase on the cumulative probabilities of phone purchases in different brand categories over the subsequent weeks and months. We take an approach similar to that outlined in Section IIIA, constructing dependent variables of the form $\mathbf{1}\left(BuysPhone\right)^{c''}_{i,(t,t+24)}$. We then perform a second set of instrumental variables regressions of the form outlined in equation (11), replacing the original dependent variables with these multiperiod cumulative purchase indicators. As discussed in Section IIIA, the coefficient estimates in these longer-horizon regressions should be interpreted as the "total" peer effect caused by a friend purchasing a phone of brand c at time $t-1$ or t, including the higher-order peer effects through purchases of common friends that were induced by this initial purchase.

Estimates of Brand-Level Peer Effects.—Table 3 shows results from regression (11). Columns 1–4 analyze a user's purchasing behavior in the week after the friend's random phone loss, analogous to the baseline specification in Table 2, while columns 5–8 analyze the cumulative purchasing behavior in the 24 weeks following

Table 3—Peer Effects in Phone Purchasing—Category-Level Analysis

	Dependent variable: Buys between t and t + 1				Dependent variable: Buys between t and t + 24			
	iPhone (1)	Galaxy (2)	Other (3)	Any phone (4)	iPhone (5)	Galaxy (6)	Other (7)	Any phone (8)
Friends buy iPhone	0.027	−0.002	−0.007	0.018	0.340	−0.006	−0.172	0.162
	(0.005)	(0.004)	(0.004)	(0.007)	(0.069)	(0.022)	(0.043)	(0.059)
Friends buy Galaxy	−0.002	0.047	0.019	0.065	−0.335	0.658	0.521	0.844
	(0.009)	(0.009)	(0.009)	(0.016)	(0.058)	(0.047)	(0.058)	(0.087)
Friends buy other	−0.016	−0.012	0.074	0.046	−0.368	0.043	1.229	0.904
	(0.007)	(0.007)	(0.009)	(0.013)	(0.051)	(0.038)	(0.064)	(0.079)
Controls + Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Mean dependent variable	0.38	0.29	0.25	0.93	11.74	6.44	5.79	23.97
Number of observations	329m	329m	329m	329m	329m	329m	329m	329m

Notes: Table shows estimates of regression (11). In columns 1–4, the dependent variables measure purchasing probabilities between weeks t and $t + 1$; in columns 5–8, the dependent variables measure cumulative purchasing probabilities between weeks t and $t + 24$. We include interacted fixed effects for individual i's demographics (age bucket $\times$ state $\times$ gender $\times$ education), individual i's device (current phone $\times$ current phone age in buckets of 50 days $\times$ carrier), and for individual i's friends (total friends $\times$ number of friends switching phones in the previous 6 months). We control linearly for the users' unconditional probabilities of buying a new phone in each category c and for the average conditional and unconditional probabilities of purchasing a phone in each category among the users' friends. We additionally control for individual and friend posting behavior (the number of friends with public statuses, the number of friends posting in a given week, the number of friends who post about random phone loss in the 12 months prior to our sample, the average conditional probability of buying a phone of each type c among friends who posted in the prior 12, and a dummy for whether the user herself posted about a random phone loss in the given week). Standard errors are clustered at the level of the community (see the discussion in Section IIC and online Appendix A.3).

the friends' random phone loss. Columns 1 and 5 show the effects on an individual's probability of purchasing an iPhone, columns 2 and 6 display the effects on an individual's probability of purchasing a Galaxy, while columns 3 and 7 show the effects on an individual's probability of purchasing a phone in the "Other" category. Columns 4 and 8 show the effects on the individual's probability of purchasing any new phone.

We find that friend purchases in each of our three brand categories lead a user to increase their overall probability of purchasing a new phone (see columns 4 and 8). In all categories, the same-brand peer effects are positive and larger than any across-brand peer effects. For instance, a friend purchasing a Samsung Galaxy primarily increases an individual's own probability of also purchasing a Galaxy—both in the period immediately following the friend's purchase and over longer horizons. In terms of magnitude, the same-category peer effects are largest for devices in the "Other" category and are smallest for iPhones. These findings are consistent with a substantial part of the observed peer effects being the result of information acquisition through social learning. In particular, during our sample period, iPhones were the most well-established brand, suggesting a smaller role for information acquiring through peers; on the other hand, social learning would likely have been most important for the more obscure phones in the fragmented "Other" category.

In addition to these large and positive same-brand peer effects, we also find heterogeneous across-brand demand spillovers. Specifically, we find large positive spillovers from purchases of Samsung Galaxy phones to purchases of phones in the "Other" category; these two brand categories share the Android operating system. This positive demand spillover is also consistent with an important role played by social learning: while most of the learning from a friend's phone purchase is about the precise brand bought by the friend, an individual may also learn about features of the Android operating system, making her more likely to buy any type of Android phone. There are fewer spillovers in the other direction, and the small negative spillover from purchases of phones in the "Other" category to purchases of Samsung Galaxy phones is not statistically significant.

On the other hand, demand spillovers tend to be negative across brands that use different operating systems. Friend purchases of phones in the Galaxy or "Other" categories (which largely use the Android operating system) decrease user purchases of iPhones, which use the competing iOS software. Similarly, friend purchases of iPhones tend to have a negative spillover effect to a user's demand for Galaxy phones and phones in the Other category. It is important to note that these demand spillovers across operating systems could have easily been positive. First, it could have been that a user who buys a Galaxy causes her friends to desire more expensive phones—of any type, including iPhones—through a "keeping up" effect. Second, positive across-brand spillovers could have emerged, even across competing operating systems, through the salience channel documented in a marketing literature that shows how advertising can increase sales of (nonadvertised) options by reminding people of their existence (e.g., Shapiro 2018; Sinkinson and Starc 2019). Third, positive demand spillovers to other brands using different operating systems could have resulted from perception transfers across competing brands (see Roehm and Tybout 2006 for related work in the marketing literature). Our finding of substantial negative demand spillovers to competing brands using different operating systems therefore helps researchers understand the implications of peer effects on the competitive dynamics between firms and distinguish them from the spillover effects of marketing activities.

Summary of Brand-Level Findings.—There are four key takeaways from the cross-brand analysis. First, for all three brand categories, there exist large positive peer effects for same-brand purchases. Second, these same-brand peer effects are largest for the lesser-known but cheaper phones in the "Other" category, and they are smallest for the expensive and well-known iPhones. Third, we generally find positive different-brand demand spillovers for brands sharing an operating system and negative different-brand spillovers for brands on competing operating systems. Fourth, positive different-brand, same-operating-system spillovers are smaller than the positive same-brand effects. These findings point toward social learning as a substantial contributor to the observed peer effects: when a friend purchases a new phone, individuals learn about that phone brand and, to a lesser extent, about other phones using the same operating system. As a result, demand should increase the most for the specific brand purchased by the friend; it should increase somewhat less for competing brands that share the same operating system. The importance of this social learning is largest for the least-well-known brands. Some of the incremental same-brand purchases from peer effects correspond to newly generated demand, and some correspond to a shifting of demand from other brands on competing operating systems.

Peer Effects at the Model Level.—In online Appendix A.6, we also study peer effects at the device model level and explore the presence of same-brand, different-model peer effects. Specifically, we analyze whether having a friend buy an iPhone 6s primarily increases a person's own probability of also purchasing an iPhone 6s or whether it increases the individual's probability of purchasing an iPhone in general. For this analysis, we cannot use an instrumental variables research design as we do in the main body of the paper: while observable characteristics allow us to predict whether a given individual would purchase an iPhone or a Galaxy, it is much harder to predict whether an individual would buy an iPhone 6 or an iPhone 6s. We therefore run OLS specifications that regress an individual's probability of purchasing a specific phone model on the phone model purchases of her friends. While the absolute magnitudes of the estimates should thus be interpreted with caution, some interesting patterns emerge about the relative size of effects for different phone models. First, same-model peer effects are more than an order of magnitude larger than different-model peer effects. Second, these same-model peer effects do not vary with the cost of the model, but they are decreasing in the time since the model release, providing further evidence for an important social learning channel behind the peer effects. Third, same-brand, different-model peer effects are more than twice as large as different-brand peer effects. The spillovers of peer effects to other models of the same brand are largest for Apple, which co-brands all of its devices under the iPhone brand, and smallest for LG, which does not do so.

V. Conclusion

In this paper, we document that new phone purchases by friends have substantial, positive, and long-lasting effects on an individual's own demand for phones of the same brand. Our research design cannot precisely identify the channel behind the observed peer effects, but our results are most consistent with an important role of social learning in explaining the observed peer effects. While peer effects expand the overall market for phones, there can be substantial negative demand spillovers to competitor brands on different operating systems as a result of a phone purchase by a friend. These negative across-brand demand spillovers have important implications for firms: losing a customer to a rival firm does not only mean missing out on positive peer effects that this customer could have had but will also lead to future losses of other customers through competitive peer effects. These findings emphasize how a customer's value to a firm exceeds the direct effect that this customer has on the firm's profits.

An interesting question for future work concerns the generalizability of our findings to understanding the decision to adopt products other than cell phones. Indeed, we hope that future research will further broaden our understanding of the importance of peer effects in product adoption decisions across a wider range of product categories (see, for example, related work by Kuchler, Stroebel, and Wong 2021). In this light, our research emphasizes the increasingly important role of data from online services—such as Facebook, LinkedIn, Twitter, eBay, Mint, Trulia, and Zillow—in overcoming important measurement challenges across the social sciences (see, for example, Baker 2018; Giglio et al. 2015; Einav et al. 2015; Piazzesi, Schneider, and Stroebel 2019). Specifically, we hope that the increasing availability of social network data, such as the Social Connectedness Index described in Bailey, Cao, Kuchler, Stroebel, and Wong (2018); Bailey, Farrell et al. (2020); Bailey, Gupta et al. (2021); Bailey, Johnston et al. (2020), will help to improve our understanding of the effects of social interactions on social, political, financial, and economic outcomes.

REFERENCES

Allcott, Hunt, and Judd B. Kessler. 2019. "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." *American Economic Journal: Applied Economics* 11 (1): 236–76.

Angrist, J.D., G.W. Imbens, and A.B. Krueger. 1999. "Jackknife Instrumental Variables Estimation." *Journal of Applied Econometrics* 14 (1): 57–67.

Aral, Sinan, and Dylan Walker. 2012. "Identifying Influential and Susceptible Members of Social Networks." *Science* 337 (6092): 337–41.

Athey, Susan. 2019. "The Impact of Machine Learning on Economics." In *Economics of Artificial Intelligence: An Agenda*, edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 507–47. Chicago, IL: University of Chicago Press.

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel. 2021. "International Trade and Social Connectedness." *Journal of International Economics*, 129: 103418.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel. 2018. "The Economic Effects of Social Networks: Evidence from the Housing Market." *Journal of Political Economy* 126 (6): 2224–76.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* 32 (3): 259–80.

Bailey, Michael, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel. 2019. "House Price Beliefs and Mortgage Leverage Choice." *Review of Economic Studies* 86 (6): 2403–52.

Bailey, Michael, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel. 2020. "Social Connectedness in Urban Areas." *Journal of Urban Economics* 118: 103264.

Bailey, Michael, Drew M. Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, and Johannes Stroebel. 2021. "Social Networks Shape Beliefs and Behavior: Evidence from Social Distancing during the COVID-19 Pandemic." NBER Working Paper 28234.

Bailey, Michael, Drew Johnston, Theresa Kuchler, Dominic Russel, Bogdan State, and Johannes Stroebel. 2020. "The Determinants of Social Connectedness in Europe." In *12th International Conference on Social Informatics*, ed. Samin Aref, Kalina Bontcheva, Marco Braghieri, Frank Dignum, Fosca Giannotti, Francesco Grisolia, and Dino Pedreschi, 1–14. Cham, Switzerland: Springer Nature.

Bailey, Michael, Drew Johnston, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2022. "Replication data for: Peer Effects in Product Adoption." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E130686V1.

Baker, Scott R. 2018. "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data." *Journal of Political Economy* 126 (4): 1504–57.

Bakshy, Eytan, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. 2011. "Everyone's an Influencer: Quantifying Influence on Twitter." In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, 65–74. New York, NY: Association for Computing Machinery.

Balachander, Subramanian, and Sanjoy Ghose. 2003. "Reciprocal Spillover Effects: A Strategic Benefit of Brand Extensions." *Journal of Marketing* 67 (1): 4–13.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives* 28 (2): 29–50.

Bernstein, Michael S., Eytan Bakshy, Moira Burke, and Brian Karrer. 2013. "Quantifying the Invisible Audience in Social Networks." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ed. Susanne Bødker, Steven Brewster, Patrick Baudisch, Michel Beaudouin-Lafon, and Wendy E. Mackay, 21–30. New York, NY: Association for Computing Machinery.

Beshears, John, James J. Choi, David Laibson, Brigitte C. Madrian, and Katherine L. Milkman. 2015. "The Effect of Providing Peer Information on Retirement Savings Decisions." *Journal of Finance* 70 (3): 1161–1201.

Bester, C. Alan, Timothy G. Conley, and Christian B. Hansen. 2011. "Inference with Dependent Data Using Cluster Covariance Estimators." *Journal of Econometrics* 165 (2): 137–51.

Björkegren, Daniel. 2019. "The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda." *Review of Economic Studies* 86 (3): 1033–60.

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. "Enriching Word Vectors with Subword Information." *arXiv*: 1607.04606.

Bollinger, Bryan, and Kenneth Gillingham. 2012. "Peer Effects in the Diffusion of Solar Photovoltaic Panels." *Marketing Science* 31 (6): 900–912.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2014. "Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions." *Econometrica* 82 (4): 1273–1301.

Campbell, Arthur. 2013. "Word-of-Mouth Communication and Percolation in Social Networks." *American Economic Review* 103 (6): 2466–98.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *Econometrics Journal* 21 (1): C1–68.

Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35–69.

Counts, Scott, and Kristie Fisher. 2011. "Taking it All In? Visual Attention in Microblog Consumption." In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 97–104. Palo Alto, CA: AAAI Press.

Dahl, Gordon B., Katrine V. Løken, and Magne Mogstad. 2014. "Peer Effects in Program Participation." *American Economic Review* 104 (7): 2049–74.

De Giorgi, Giacomo, Anders Frederiksen, and Luigi Pistaferri. 2016. "Consumption Network Effects." NBER Working paper 22357.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *Quarterly Journal of Economics* 127 (1): 1–56.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118 (3): 815–42.

Easley, David, and Jon Kleinberg. 2010. *Networks, Crowds, and Markets*. New York, NY: Cambridge University Press.

Eckles, Dean, René F. Kizilcec, and Eytan Bakshy. 2016. "Estimating Peer Effects in Networks with Peer Encouragement Designs." *Proceedings of the National Academy of Sciences* 113 (27): 7316–22.

Einav, Liran, Theresa Kuchler, Jonathan Levin, and Neel Sundaresan. 2015. "Assessing Sale Strategies in Online Markets Using Matched Listings." *American Economic Journal: Microeconomics* 7 (2): 215–47.

Erdem, Tülin. 1998. "An Empirical Analysis of Umbrella Branding." *Journal of Marketing Research* 35 (3): 339–51.

Facebook. 2016. "Facebook Form 10-Q, Quarterly Report, Q2."

Facebook. 2020. "Various Datasets."

Ferguson, Rick. 2008. "Word of Mouth and Viral Marketing: Taking the Temperature of the Hottest Trends in Marketing." *Journal of Consumer Marketing* 25 (3): 179–82.

Foster, Andrew D., and Mark R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103 (6): 1176–1209.

Garcia, Daniel, and Sandro Shelegia. 2018. "Consumer Search with Observational Learning." *RAND Journal of Economics* 49 (1): 224–53.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Andreas Weber. 2015. "Climate Change and Long-Run Discount Rates: Evidence from Real Estate." NBER Working Paper 21767.

Gilbert, Eric, and Karrie Karahalios. 2009. "Predicting Tie Strength with Social Media." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ed. Saul Greenberg, Scott E. Hudson, Ken Hinckley, Meredith Ringel Morris, and Dan R. Olson, Jr., 211–20. New York, NY: Association for Computing Machinery.

Gilchrist, Duncan Sheppard, and Emily Glassberg Sands. 2016. "Something to Talk About: Social Spillovers in Movie Consumption." *Journal of Political Economy* 124 (5): 1339–82.

Glaeser, Edward L., Jose A. Scheinkman, and Bruce I. Sacerdote. 2003. "The Social Multiplier." *Journal of the European Economic Association* 1 (2–3): 345–53.

Goolsbee, Austan, and Peter J. Klenow. 2002. "Evidence on Learning and Network Externalities in the Diffusion of Home Computers." *Journal of Law and Economics* 45 (2): 317–43.

Han, Bing, David Hirshleifer, and Johan Walden. 2016. "Visibility Bias in the Transmission of Consumption Beliefs and Undersaving." Unpublished.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein. 2004. "Social Interaction and Stock-Market Participation." *Journal of Finance* 59 (1): 137–63.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867.

Janakiraman, Ramkumar, Catarina Sismeiro, and Shantanu Dutta. 2009. "Perception Spillovers Across Competing Brands: A Disaggregate Model of How and When." *Journal of Marketing Research* 46 (4): 467–81.

Jones, Jason J., Jaime E. Settle, Robert M. Bond, Christopher J. Fariss, Cameron Marlow, and James H. Fowler. 2013. "Inferring Tie Strength from Online Directed Behavior." *PloS ONE* 8 (1): e52168.

Kim, Yoon. 2014. "Convolutional Neural Networks for Sentence Classification." *arXiv*: 1408.5882.

Kremer, Michael, and Edward Miguel. 2007. "The Illusion of Sustainability." *Quarterly Journal of Economics* 122 (3): 1007–65.

Kuchler, Theresa, Yan Li, Lin Peng, Johannes Stroebel, and Dexin Zhou. 2020. "Social Proximity to Capital: Implications for Investors and Firms." NBER Working Paper 27299.

Kuchler, Theresa, Dominic Russel, and Johannes Stroebel. 2020. "The Geographic Spread of COVID-19 Correlates with the Structure of Social Networks as Measured by Facebook." NBER Working Paper 26990.

Kuchler, Theresa, and Johannes Stroebel. 2020. "Social Finance." NBER Working Paper 27973.

Kuchler, Theresa, Johannes Stroebel, and Arlene Wong. 2021. "Peer Effects and Markups in Product Markets." Working Paper.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn. 2011. "The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery." *American Economic Review* 101 (5): 2226–47.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60 (3): 531–42.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review* 99 (1): 112–45.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415–44.

Mikolov, Tomas, Wen-tau Yih, and Geoffrey Zweig. 2013. *Linguistic Regularities in Continuous Space Word Representations.* Redmond, WA: Microsoft Research.

Mobius, Markus M., Paul Niehaus, and Tanya S. Rosenblat. 2005. "Social Learning and Consumer Demand." Unpublished.

Moretti, Enrico. 2011. "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales." *Review of Economic Studies* 78 (1): 356–93.

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives* 31 (2): 87–106.

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business and Economic Statistics* 31 (3): 358–69.

Oster, Emily, and Rebecca Thornton. 2012. "Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up." *Journal of the European Economic Association* 10 (6): 1263–93.

Ouimet, Paige, and Geoffrey Tate. 2017. "Learning from Coworkers: Peer Effects on Individual Investment Decisions." NBER Working Paper 24058.

Peysakhovich, Alexander, and Dean Eckles. 2017. "Learning Causal Effects from Many Randomized Experiments using Regularized Instrumental Variables." *arXiv*: 1701.01140.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel. 2019. "Segmented Housing Search." NBER Working Paper 20823.

Roehm, Michelle L., and Alice M. Tybout. 2006. "When Will a Brand Scandal Spill Over, and How Should Competitors Respond?" *Journal of Marketing Research* 43 (3): 366–73.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116 (2): 681–704.

Sacerdote, Bruce. 2011. "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In *Handbook of the Economics of Education*, Vol. 3, ed. Eric A. Hanushek, Stephen Manchin, and Ludger Woessmann, 249–77. Amsterdam: Elsevier.

Sahni, Navdeep S. 2016. "Advertising Spillovers: Evidence from Online Field Experiments and Implications for Returns on Advertising." *Journal of Marketing Research* 53 (4): 459–78.

Shapiro, Bradley T. 2018. "Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants." *Journal of Political Economy* 126 (1): 381–437.

Sinkinson, Michael, and Amanda Starc. 2019. "Ask Your Doctor? Direct-to-Consumer Advertising of Pharmaceuticals." *Review of Economic Studies* 86 (2): 836–81.

Tucker, Catherine. 2008. "Identifying Formal and Informal Influence in Technology Adoption with Network Externalities." *Management Science* 54 (12): 2024–38.

Zacchia, Paolo. 2020. "Knowledge Spillovers through Networks of Scientists." *Review of Economic Studies* 87 (4): 1989–2018.