NBER WORKING PAPER SERIES

THE EMPIRICAL ECONOMICS OF ONLINE ATTENTION

Andre Boik
Shane Greenstein
Jeffrey Prince

Working Paper 22427
http://www.nber.org/papers/w22427

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2016

We thank the Kelley School of Business and the Harvard Business School for funding. We thank Scott Savage and Mo Xiao for excellent suggestions. We thank seminar audiences at Georgetown, Harvard and Northwestern, and conference participants at the American Economic Association Annual Meetings, the International Industrial Organization Conference, and Silicon Flatirons. Philip Marx provided excellent research assistance, and Kate Adams provided excellent editorial assistance. We are responsible for all errors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w22427.ack

The Empirical Economics of Online Attention
Andre Boik, Shane Greenstein, and Jeffrey Prince
NBER Working Paper No. 22427
July 2016
JEL No. D12,L81,L86

ABSTRACT

In several markets, firms compete not for consumer expenditure but instead for consumer attention. We model and characterize how households allocate their scarce attention in arguably the largest market for attention: the Internet. Our characterization of household attention allocation operates along three dimensions: how much attention is allocated, where that attention is allocated, and how that attention is allocated. Using click-stream data for thousands of U.S. households, we assess if and how attention allocation on each dimension changed between 2008 and 2013, a time of large increases in online offerings. We identify vast and expected changes in where households allocate their attention (away from chat and news towards video and social media), and yet we simultaneously identify remarkable stability in how much attention is allocated and how it is allocated. Specifically, we identify (i) persistence in the elasticity of attention according to income and (ii) complete stability in the dispersion of attention across sites and in the intensity of attention within sites. We illustrate how this finding is difficult to reconcile with standard models of optimal attention allocation and suggest alternatives that may be more suitable. We conclude that increasingly valuable offerings change where households go online, but not their general online attention patterns. This conclusion has important implications for competition and welfare in other markets for attention.

Andre Boik
Department of Economics
UC Davis
1107 SSH
1 Shields Avenue
Davis, CA 95616
aboik@ucdavis.edu

Shane Greenstein
Technology Operation and Management
Morgan Hall 439
Harvard Business School
Soldiers Field
Boston, MA 02138
and NBER
sgreenstein@hbs.edu

Jeffrey Prince
Kelley School of Business
Indiana University
1309 E. Tenth St.
Bloomington, IN 47405
jeffprin@indiana.edu

An online appendix is available at http://www.nber.org/data-appendix/w22427

1. Introduction

> "…[I]n an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it." (Simon, 1971).

Herb Simon brought attention to the economic importance of attention, first articulated about information systems, which applies to any situation with abundant information. The observation remains relevant today, even more so for the information supplied by the commercial Internet. A scarce resource, users' attention, must be allocated across the Internet's vast supply of web sites. Firms compete for user attention.

At first glance, competition among Internet sites has much in common with other competitive settings. Users make choices about where to allocate their time, and in any household there is only a finite amount of such time to allocate, which translates into a finite budget of time for which firms compete. In some cases (e.g., electronic commerce), the firms try to convert that attention into sales of products. At over $360 billion per year, e-commerce comprises eight percent of total US sales in 2016.[1] In other cases (e.g., most media), firms try to convert that attention into advertising sales, which amounts to $67 billion of spending.[2] Firms compete for users by investing in web page design, in internal search functions, and in other aspects such as the speed at which relevant information loads. Over time, new firms enter with new offerings, and users can respond by making new choices, potentially substituting one source of supply for another.

However, first impressions mislead. Competition among web sites lacks one of the standard hallmarks of competition. Relative prices largely do not determine user choice among

[1] US Census, 2016. https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf.
[2] E-marketer, 2016. http://www.emarketer.com/Article/US-Digital-Display-Ad-Spending-Surpass-Search-Ad-Spending-2016/1013442.

options, nor do prices determine competitive outcomes. Most households pay for monthly service, then allocate online time among endless options without further expenditure. Unless a household faces a binding cap on usage, no price shapes any other marginal decision. Instead, choice depends on the non-monetary frictions and the gains of the next best choice. Present evidence suggests only a small fraction of users face the shadow of monetary constraints while using online resources (Nevo, Turner, Williams, 2015). Relatedly, subscription services also play little role. As we will show below, only one of the top twenty sites (Netflix) is a subscription service, i.e., where the price of a web site plays an explicit role in decision making.

In this study, we use extensive microdata on user online choice to help us characterize demand for the services offered online. The demand for services by a household is the supply of attention for which firms compete. The study characterizes household heterogeneity in allocation of attention *at any point in time*, and how households substitute between sources of supply *over time*. We ground the analysis in a specific time period, the allocation of US household attention in the years 2008 and 2013, which was a time of enormous change in the supply of online options for the more than 70% of US households with broadband connections to the Internet. During this five-year period, US households experienced a massive expansion in online video offerings, social media, and points of contact (e.g., tablets, smartphones), among other changes.

Our dataset contains information for more than forty thousand primary home computers, or "home devices," at US households in 2008 and more than thirty thousand in 2013. These data come from ComScore, a firm that tracks households over an entire year, recording all of the web sites visited, as well as some key demographics. The unit of observation is a week's worth of choices made by households. We calculate the weekly market for online attention (total time), its concentration (in terms of time) for sites (our measure of breadth, or "focus"), and the weekly fraction of site visits that lasted at least 10 minutes (our measure of depth, or "dwelling"). In addition, we measure shares of attention for different site categories (e.g., social media). Using these measures of online attention, we analyze how they vary both horizontally (across demographics) and vertically (over time, 2008-2013).

We find that demand is comprised of a surprising mix of discretionary and inflexible behavior. First, we find strong evidence that income plays an important role in determining the allocation of time to the Internet. This finding reconfirms an earlier estimate of a relationship between income and extent of Internet use (Goldfarb and Prince, 2008), but does so using a more expansive and detailed dataset, and for later years when broadband access is more prevalent. We find that higher income households spend less total time online per week. Households making $25,000-$35,000 a year spend ninety-two *more* minutes a week online than households making $100,000 or more a year in income, and differences vary monotonically over intermediate income levels. Relatedly, we also find that the amount of time on the home device only slightly changes with increases in the number of available web sites and other devices – it slightly declines between 2008 and 2013 – despite large increases in online activity via smartphones and tablets over this time. Finally, the monotonic negative relationship between income and total time suggests online attention is an inferior good, and we find that this relationship remains *stable*, exhibiting a similar slope of sensitivity to income. We call this property *persistent attention inferiority*. There is a generally similar decline in total time across all income groups, which is consistent with a simple hypothesis that the allocation of time online at a personal computer declines in response to the introduction of new devices.

We also examine how breadth and depth changed with the massive changes in supply (i.e., video proliferation and Internet points of contact) between 2008 and 2013. Our casual expectation was that depth would increase, and more tentatively, that breadth would increase as well, but the findings do not conform to such expectations. Rather, breadth and depth have remained remarkably *stable* over the five years. While there is a statistical difference in the joint distribution of breadth and depth, it is just that – statistical and driven by our large sample. The size of the difference is remarkably small, with little implied economic consequence. We call this property *persistent attention distribution*. Despite the evidence that income and other economic variables affect total time online, demographics – perhaps surprisingly – predict little of the variation in breadth and depth. For one, breadth and depth are not well-predicted by income and there is only a limited role played by major demographics, such as family education, household size, age of head of household, and presence of children.

This stability of breadth and depth contrasts with substantial volatility in the types of sites households visit. Between 2008 and 2013, households substitute online categories such as chat and news for social media and video. In addition, demographics again are predictive of the outcome – household characteristics such as income strongly predict the category of sites that are visited. For example, higher income households prefer services that examine credit history, offer educational services, support games, provide news, support online banking, offer online shopping, provide online sports services, and supply online video services. To summarize: new offerings *did* alter where households went online, only mildly altered how much total time they spent on their machines, and *did not* meaningfully alter their general breadth and depth – as if the determinants of total time, and which sites to visit, are distinct from the determinants of breadth and depth.

These findings have important implications for competition for online attention. Our results imply that reallocation of online attention takes place in the presence of inflexibility of breadth/depth decisions. Reallocation of online attention comes almost entirely in the form of changes in how households select from a portfolio of different web sites, but not in the form of changes in total time or breadth and depth. Altogether, these findings are inconsistent with some models of attention allocation, especially those models lacking any frictions, such as setup costs. They are consistent with a theory that is behavioral in its foundations. This model suggests households are endowed with a fixed set of "slots" of attention to allocate to sites, as if households typically have fixed amounts of time. These amounts of time do not vary but are switched between different categories of web sites. As discussed below, these observations lead to many open questions about online competition.

1.1.Contribution to prior literature

The commercial Internet supports enormous amounts of economic activity, and it has experienced increases in online offerings throughout its short existence. Starting from modest beginnings in the mid-1990s, this sector of the US economy today supports tens of billions of dollars of advertising revenue and trillions in revenue from online sales. Not surprisingly, that phenomenon has spawned an extensive literature, and it has grown so much that it merits handbooks to cover the research (Peitz and Waldfogel, 2012). These handbooks organize the literature around many sub-topics, such as the supply and demand for infrastructure, online and offline competition (Lieber and Syverson, 2012), and the supply and demand for online advertising (Anderson, 2012).

One theme cuts across many of these topics: all households get their time from some other non-Internet leisure activity, and different online activities compete with each other in the household's budget for time. While researchers recognize that users pay an opportunity cost during online time by withdrawing from other leisure activity or household production activity (Webster, 2014, Wallsten, 2013), the household's time for, and attention to, its online activities remains incompletely characterized. No work has characterized the three basic types of online attention measurements – how much attention is used, how is it allocated, and where is it allocated? Hence, there is no widely accepted baseline model of aggregate demand for online activity (and supply of attention) built from a common understanding of online behavior.

Such a characterization can inform research about the economic allocation of time in general. Below we will present a standard economic model of time allocation, which follows the prior literature (Hauser et al. 1993, Ratchford et al. 2003, Savage and Waldman 2009) and finds its roots in Becker (1965). Prior research has used this approach to demonstrate the demand for, and market value of, for example, speed in broadband access, which users spread over a vast array of content (Rosston, Savage, and Waldman, 2010, Hitt and Tambe, 2007). We take this approach in a different direction, highlighting theoretical ambiguities regarding predicted changes in online attention with increased online offerings, ambiguities which highlight the role of frictions in user allocations. We create novel measures of online attention allocation designed to capture the total time allocated to online offerings and the breadth and depth of a household's online attention, and then ask whether user patterns of online behavior are consistent with the predictions of a basic theoretical model of the allocation of time without frictions.

This new direction will also have implications for prior work about the consumer surplus generated by online activity. Prior research has, again, taken the standard model of time allocation in a frictionless labor/leisure framework and estimated a specification for the parameters characterizing demand for time on all households (Goolsbee and Klenow, 2006, Brynjolfsson and Oh, 2012). In contrast, because we can see more about the user's allocation of time, we can use that additional information to characterize the entire time spent online, and the distribution of online time. That will focus on behavior inconsistent with a frictionless model of the labor/leisure tradeoff.

This theme also can inform research into disputes, which, until now, leave aside examination of how the specific dispute fits into the larger household allocation decision. For example, search engine competition has motivated some studies on competition for attention (Athey, Calvano and Gans, 2013, Gabaix, 2014). In addition, there has been some formal statistical work on the competition for attention in the context of conflicts for very specific applications, such as, for example, conflicts between news aggregators and news sites (Chiou and Tucker, 2015, Athey and Mobius, 2012), and conflict between different search instruments (Baye et al. 2016). Each of these disputes contrasts implications from settings in which frictions are a large or small factor in user choice. Our results will be consistent with models that stress the transaction costs of user online activity.

The focus of this study contrasts with the typical focus in the marketing literature on online advertising. As the Internet ecosystem increases the availability of online offerings, consumers can adjust their online attention to gain value in several ways. Specifically, consumers can: 1) Increase the total amount of attention they allocate to the Internet, 2) Re-allocate their ad-viewing attention to better targeted ads, and/or 3) Re-allocate their attention to more and/or higher value sites. Much of the prior work pertaining to online advertising has focused on #2, namely, the principles of targeting ads. This is largely driven by firms tapping into “big data” and extensive information about users’ private behavior, which was previously unobserved and merits study for marketing purposes. The marketing literature on targeting tends not to focus on why behavior changes by consumers as supply changes. In contrast, our analysis centers on the reaction of households to changes in supply, which focuses on the determinants of #1 and #3, which are generally under the control of the consumer, and as of this writing, have been less studied and are less understood. This leads to a different conceptualization about competition for attention.

As we conducted this study, we were surprised to learn that the findings (partially) overlap with conclusions drawn from field work conducted by economic anthropologists and researchers on user-machine design. That line of research also collects microdata and uses it to characterize features of demand. It has documented the periodic – or “bursty” – use of many online sources, consistent with some of our findings concerning breadth (Lindley, Meek, Sellen, Harper, 2012, Kawsaw and Brush, 2013). It also documents the “plasticity” of online attention, as an activity that arises from the midst of household activities as a “filler” activity (Rattenbury, Nafus and Anderson, 2008, Adar, Teevan, Dumais, 2009), which provides an explanation for the consistency of breadth and depth patterns within a household in spite of large changes in the available options. We make these links in the discussion of the findings. Hence, we view our work as a bridge between economic analysis and conversations within other sites of social science.

2. Dynamics of the Internet Ecosystem: 2008-2013

The era we examine is one characterized by rapid technical advance and widespread adoption of new devices. Continuing patterns seen since the commercialization of the Internet in the 1990s (Greenstein, 2015), new technical invention enabled the opportunity for new types of online activity and new devices. For example, the cost of building an engaging web site declined each year as software tools improved, the effectiveness of advertising improved, and the cost of microprocessors declined. In addition, the cost of sending larger amounts of data to a user declined each year as broadband network capacity increased. By the beginning of our sample many online suppliers and startups had begun experimenting with applications that made extensive use of data-intensive video.

The start of our time period is near the end of the first diffusion of broadband networks. By 2007, close to sixty-two million US households had adopted broadband access for their household Internet needs, while by 2013 the numbers were seventy-three million. The earlier year also marked a very early point in the deployment of smart phones, streaming services, and social media. The first generation of the iPhone was released in June 2007, and it is widely credited with catalyzing entry of Android-based phones the following year. By 2013, more than half of US households had a smartphone. Tablets and related devices did not begin to diffuse until 2010, catalyzed, once again, by the release of an Apple product – in this case, the iPad in April, 2010.

Also relevant to our setting are the big changes in online software. Streaming services had begun to grow at this time, with YouTube entering in February, 2005, and purchased by Google in October 2006. Netflix and Hulu both began offering streaming services in 2008.

Social media was also quite young. For example, Twitter launched in March 2006, while Facebook launched in February 2004, and offered widespread public access in September 2006. By 2013, social media had become a mainstream online application, and, as our data will show, was widely used. In summary, the supply of options for users changed dramatically over the time period we examine.

3. A Model of Online Attention

In this section we present a standard model of attention allocation applied to households' online attention allocation decisions. Subsequently, we use the model to examine the predicted effects of two shocks and evaluate the assumptions needed for the model to rationalize our empirical findings.

3.1. The Standard Model with Setup Costs

As pertains to attention allocation, we could propose a number of different models. Our model of online attention is standard and follows the basic structure of the seminal work by Becker (1965) on the allocation of time, which has been adapted by others in various ways to examine household demand for broadband (e.g. Savage and Waldman 2009). Alternatives include a number of different search models (e.g., Gabaix 2014) that develop intuition for length and variety in the allocation of time. In either case, these models fix ideas and help guide the dimensions of analysis, but yield only simple predictions. We develop the standard approach and then make references to a behavioral approach with similar predictions.

Critical to any model is that visits to online sites do not carry a price; rather, the cost of a site visit is the opportunity cost of that attention which could be allocated elsewhere. Further, we suppose that there is a setup cost to visiting each site. The setup cost can be interpreted as either a necessary minimum time cost to absorb the information at a site, a cognitive cost of switching sites, a time cost of waiting for a new site to load, or so on. The existence of any such cost will generate continuous visits to sites that end only when the marginal utility from additional time spent at the site falls below the marginal utility of visiting some other site (or choosing some offline activity) net of the switching cost.

In this setting, household *i* chooses the amount of time to spend at each Internet site (t_{ij}) on its "home device" to maximize its standard continuous, differentiable utility function net of setup costs:

$$(1)\ max_{t_{i1},\dots,t_{iJ}} U\left(t_{i1}, \dots, t_{iJ}, T_i - \left(t_{i1} + \cdots + t_{iJ}\right); \overrightarrow{W}\right) - \sum_j^J 1(t_{ij} > 0)F$$
$$\text{s.t. } t_{i1} \geq 0, \dots, t_{iJ} \geq 0, T_i \geq (t_{i1} + \cdots + t_{iJ})$$

where F is the setup cost of visiting a site. In equation (1), $\overrightarrow{W}$ represents all relevant features (i.e., content, subscription fee – if any, etc.) for the available web sites. Further, T_i represents all time available to household *i* in, say, a week, and the final argument of *U(.)* is the equivalent of a composite good; in this case, it represents all other activities for which household *i* could be using its time (e.g., sleep, work, exercise, and time on other devices). Hence, this formulation implicitly assumes household *i* fully exhausts all of its available time.

For the moment, we place no structure on the utility function, so we define $t_{ij}^* = argmax\,(1)$ as the attention allocation function that solves this problem. A natural way of characterizing this function is in terms of total time, and the breadth and depth of that allocation of time online. We start with total time on the device over a "representative" period. For illustrative purposes, think of this as a week of time.[3] The model produces the following identity for time online for household *i* (TO_i) when there are *J* sites:

$$(2)\ TO_i = \sum_j t_{ij}^*$$

Next, we consider measures for breadth and depth of online time allocation. That is, how is attention allocated *across* sites, and how intensely is it allocated *within* a site? Our measure of breadth stems from the classic literature in industrial organization. Specifically, we measure

[3] In the data section below, we have experimented considerably with alternative units of analysis, such as a day, week, month and year. Consistent with many available measures of the Internet and, more broadly, leisure time (e.g., Wallsten, 2013), we have found considerable variability in household online use day to day, and hour to hour. However, in preliminary work not shown here, we have found considerable stability in weekly patterns of online behavior, and that the same households differ from one another in much the same way week after week. Hence, in this study, we focus exclusively on characterizing one "representative" week for a household.

breadth using a Herfindahl-Herschman index for time spent at sites visited by household *i*, denoted C_i. We define C_i as:

$$(3)\ C_i = \sum_j^J \frac{t_{ij}^{*2}}{(t_{i1}^{*} + \cdots + t_{iN}^{*})}$$

Defined this way, our measure of breadth captures the level of concentration (in terms of time at sites) household *i* exhibits in its site visits. This measure works equally well in the cross-section and over time. At any point in time it measures heterogeneity across households: a high value for C_i indicates a breadth of visits that is highly concentrated at a small number of sites, whereas a low value for C_i indicates a breadth of visits that is unconcentrated, i.e., spread out across relatively many sites. It also can measure changes over time: C_i gets larger as a household substitutes a larger fraction of its time into fewer web sites.

Our measure of depth takes inspiration from an early constraint on YouTube, specifically the cap on video length of ten minutes, which lasted until mid-2010. We measure depth as the fraction of site visits by household *i* that lasted at least ten minutes, denoted L_i. If the setup cost is strictly positive, the standard model suggests households spend all of their time at each site continuously. Hence, the depth of households' visits can be summarized by the fraction that exceed a given threshold of time, $\bar{t}$:

$$L_i = \frac{\sum_j^J 1(t_{ij} > \bar{t})}{\sum_j^J 1(t_{ij} > 0)}$$

To calculate L_i in practice, we must decompose the optimal time spent at each site during the given time period (e.g., a week). To see this, suppose $t_{i1}^{*} = 30$. Hence, time spent at site #1 during the observed week was thirty minutes. However, this measurement does not distinguish between a thirty-minute block that consists of six separate visits lasting five minutes each and one visit lasting thirty minutes. Our measure of depth would account for such a difference.

In order to construct L_i, we first define $\overrightarrow{S_{ij}}$ as the vector of session lengths (i.e., segments of continuous time) at site *j* for household *i*. Hence, the length of $\overrightarrow{S_{ij}}$ is the number of separate visits made by household *i* to site *j*. Next, let t_{ijk}^{*} be the optimal time spent by household *i* at site *j* during session *k*; therefore, t^*_{ijk} is simply the k^{th} entry in $\overrightarrow{S_{ij}}$, and $\sum_k t^*_{ijk} = t^*_{ij}$. Given these additional definitions, we define L_i as:

$$(4)\ L_i = \frac{\sum_j \sum_k 1(t^*_{ijk} > 10)}{\sum_j \sum_k 1(t^*_{ijk} > 0)}$$

As defined, L_i is the proportion of total site visits that lasted more than ten minutes for household *i*. Again, this measure works equally well in the cross-section and over time. At any point in time it measures heterogeneity across households in the fraction of time spent in longer sessions, with higher *L* indicating a higher fraction. It also measures changes over time in a household, with an increase in *L* indicating that a household has substituted some of its time into longer sessions.

An illustration can help build intuition for how these measures characterize cross sectional heterogeneity in online attention. We consider our first metric (C_i) to be a measure of focus – households with a high value for C_i focus their attention on a relatively small number of sites, and vice versa for households with a low values for C_i. We consider our second metric (L_i) to be a measure of a household's propensity to dwell at the sites it visits – households with a high value for L_i tend to dwell at sites, while households with a low value for L_i behave more like a tourist, visiting for a brief stint. Building on this intuition, we envision the very simple 2x2 classification of households using these two metrics in Table 1 as a conceptual benchmark of heterogeneity across households.

[Table 1 about here]

Now that we have detailed our measures of online attention in terms of "how much?" and "how is it allocated?," we consider one last measure: "where is it allocated?" For this measure, we calculate shares of total time online on the home device for different site categories (we list the specific categories for our analysis below). Thus, we define TS_c as the share of total time across all households spent at sites in category *c*. Formally, we have:

$$(5)\ TS_c = \frac{\sum_i \sum_{j \in c} t^*_{ij}}{\sum_i TO_i}$$

Again, this measure works equally well for characterizing heterogeneity at a point in time, and changes in a household over time. That said, we think this measure suggests one approach to measuring changes in the extent of competition among sites. We expect new site entry to lead to turnover when users direct their attention to new categories of web sites. One measure of competition is the fraction of total attention that moves to these new categories.

Section 3.2. Effects of Two Model Shocks

Over the time period of our data, two important shocks occurred. First, a wave of new sites entered the worldwide web, and many of these new sites offered large amounts of video content. For example, Netflix and Hulu both began offering streaming online video during the earliest year of our data, and YouTube began allowing videos longer than ten minutes within the span of our data. While some sites exited during the time we analyze, the net change in sites was positive, with a notable increase in online video available. This influx of sites manifests as an increase in J to J^* and a change in the full list of sites – and their characteristics – comprising the J^* total sites.

The second shock to our model was due to the release of a new batch of connected devices – in particular, tablets and smartphones. Because our model and data focus on the primary personal computer at the home, this shock essentially altered the composition of the composite good within the model.

An increase in the number of sites from J to J^* and the introduction of alternative devices affects the household utility maximizing problem as follows.

(6) $max_{t_{i1},\dots,t_{iJ^*},t_{i1}^{dev},\dots,t_{iJ^*}^{dev}} U\left(t_{i1},\dots,t_{iJ^*},t_{i1}^{dev},\dots,t_{iJ^*}^{dev},T_i-\left(t_{i1}+\cdots+t_{iJ^*}^{dev}\right);\vec{W}\right)$ - $\sum_j^J 1(t_{ij} > 0)F$

$$\text{s.t. } t_{i1} \geq 0, \dots, t_{iJ^*}^{dev} \geq 0, T_i \geq (t_{i1}+\dots+t_{iJ^*}^{dev})$$

The household faces more site choices and the option to consume them on an alternative device. We assume setup costs affect the alternate device as they do the home device, which suggests the model's insight about the household allocation closely mirrors that without additional sites or an additional device. We ask how these two changes impact three key outcomes within our model: total time, breadth, and depth. That is, we ask how these changes impact how much time and how it is allocated.

Without more information about the utility function and size of setup costs, the model could predict either an increase or decrease in the household's total time online and its breadth and depth of browsing on the home device. Here we place some structure on the household's maximization problem to generate simple predictions about the response of households' attention allocation decisions to the two shocks.

If the utility function is symmetric among sites, quasilinear in an unchanging offline outside option, and the setup costs are small, then an increase in the number of sites weakly increases the total amount of time online, and decreases the concentration of time spent across sites on the home device.[4] The standard model with small setup costs does not make a prediction about the depth of browsing because, without setup costs, a given amount of time spent at a site can be split in any way and still yield the same total utility. The introduction of an alternative device is predicted to weakly decrease the total amount of time spent on the home device, and to have no effect on the breadth of browsing on the home device. With small setup costs, the model again does not make a prediction about the depth of browsing.

When setup costs are large, then the household may have already been constrained to visit fewer than J sites before the shock and will continue to visit the same number of sites after the shock, so that the concentration of time across sites is unchanged. If the household was not constrained by setup costs before the shock, then concentration of time across sites will fall. Additionally, the marginal effect of the introduction of an alternative device is to weakly increase concentration: site visits on the alternative device increase the time share of the sites viewed on the home device.

[Table 2 about here]

Table 2 summarizes the effect of the two shocks on the household's time online (TO_i), breadth of browsing (C_i), and depth of browsing (L_i) under the standard model with small and

[4] The appendix contains the details of the microeconomics behind this prediction and those that follow.

large positive setup costs. The standard model predicts an ambiguous effect on TO_i whether setup costs are small or large, while the model predicts a decrease in C_i, if setup costs are small, and an ambiguous change in C_i, if setup costs are large. The predicted effect on L_i is 0, if setup costs are large and there is no prediction for small setup costs. However, it is worth noting that the standard model with setup costs and symmetric utility suggests the level of L_i is either 0 or 1: all sessions are the same length in equilibrium, so they all are either above or below any specified threshold. Since we do not explicitly model different categories of sites, our model is silent with respect to how households will reallocate attention across different types of site categories in response to the two model shocks. This limitation also constrains our ability to generate a predicted response to the growth in video and social media sites in a formal sense, although informally, the high time demands of such sites suggest a predicted increase in L_i.

A behavioral approach to the same problem leads to similar forecasts. For example, Gabaix (2014) considers an agent facing a decision problem which requires the agent to incorporate information from a large number of variables. In practice, an agent cannot pay attention to all variables; Gabaix formalizes this notion by requiring the agent to incur a “psychic cost” for each variable to which the agent chooses to pay attention. In this setting, Gabaix’s agent will decide to pay attention to only a subset of variables that deliver a marginal benefit (of importance) greater than the psychic marginal cost and will ignore all others. Gabaix’s environment and theoretical predictions parallel ours: a user simply cannot pay attention to all sites and continues to visit additional sites until the marginal benefit of an additional site (net of the opportunity cost of offline activities) reaches the marginal setup cost.

In the following sections, we take our measures of households’ depth and breadth of online browsing to the data to examine how these measures changed over our sample period and to evaluate the standard model’s predictions. We will not provide standard economic measures of substitution because there are no prices with which to measure cross-price elasticities and related values. Instead, we use our measures of “how much,” “how is it allocated,” and “where is it allocated” with regard to online attention on the home device – as defined in equations 2 through 5. By doing so, we can observe if households altered their behavior with respect to these outcomes over the timespan of our data and, if so, how.

3.3. Hypothesis development

Our hypotheses distinguish between distinct determinants originating at the supply-side and demand-side in the attention economy. We postulate that supply determines the menu of available choices, and a different set of factors, such as household characteristics, determines the final allocation.

What determines the shock to the menu of choices available to users? Since these inventions become available to all market participants, such technical advance induces three responses of relevance to competition for attention: (1) Existing web sites improve their offerings in a bid for user attention; (2) entrepreneurial firms conceive of new services to offer online in a bid for user attention; and (3) new devices enter to attract user attention. Collectively, these determine the "supply" of web sites bidding for the attention of users in time *t*, which we summarize as $S_{t.}$

As for demand, we further postulate every household *i* in time *t* has a set of demographic characteristics – education and income – that allocate their attention among the available menu of options. We call these variables X_i. Together with supply, an allocation for a household can be characterized as three relationships:

Total time: $TO_{it} = \mathrm{TO}(\mathrm{S_t}, \mathrm{X_{it}})$

Concentration (breadth): $C_{it} = \mathrm{C}(\mathrm{S_t}, \mathrm{X_{it}})$

Length (depth): $L_{it} = \mathrm{L}(\mathrm{S_t}, \mathrm{X_{it}})$

What are the properties of this allocation? Goldfarb and Prince (2008) have shown that households with high income are more likely to adopt the dial-up Internet, but they do not use the Internet as intensively as those with a lower income. They hypothesize that this is due to the outside option value of leisure/work time. In this setting, if X_{it} is income, the Goldfarb-Prince effect would treat online time as an inferior good, and appear as:

H1. $\mathrm{TO_x}(\mathrm{S_t}, \mathrm{X_{it}}) < 0$.

We seek to learn whether this income effect holds in our measures of online attention, and on very different data at a much later period, when broadband dominates connections. A further question is whether time online on the home device has changed over time. That is, has the improvement in devices attracted user attention away from the improving web sites on PCs, or were web site improvements substantial enough to increase time on the home device, despite the advent of alternative devices? The null hypothesis specifies no change in total time:

H2. $TO(S_t, X_{it}) - TO(S_{t-1}, X_{it-1}) = 0$.

The alternative to the null hypothesis could be a decrease in total time if households substitute into other devices. If we reject H2, then an interesting question focuses on whether the income effect has changed over time. That is, despite declines in the *level* of total time online, has the *rate* of the relationship between income and time online remained the same? Again, the null is no change:

H3. $TO_x(S_t, X_{it}) – TO_x(S_{t-1}, X_{it-1}) = 0$.

We can also ask whether greater online time leads to greater breadth and depth? If so, then – once again, assuming X is income – we would expect a larger X to lead to a lower total time and less breadth and less depth. Initially we seek to test the null hypothesis in a one tail test, where the null is:

H4. $C_x(S_t, X_{it}) = 0$ and $L_x(S_t, X_{it}) = 0$, and the alternative is:

H4A. $C_x(S_t, X_{it}) < 0$ and $L_x(S_t, X_{it}) < 0$.

Once again, and parallel to the discussion for H2 and H3, if we reject H4 for H4A, then the next question concerns changes to the determinants of breadth and depth.

We also can test the reaction of households to growth in supply of options. As has been widely reported, social networking applications and streaming have become more available over time. We expect users to substitute some of their time to these new applications. Did this substitution change the measured breadth and depth? We expect new sources of supply to increase depth and breadth, so we set up a test to reject the null, where the null is for no change, expressed as:

H5. $C(S_t, X_{it}) - C(S_{t-1}, X_{it-1}) = 0$, and $L(S_t, X_{it}) - L(S_{t-1}, X_{it-1}) = 0$.

Similar to the above discussion about H2 and H3, after testing H5, we can further test whether breadth and depth are sensitive to demographics.

We stress that the longer the time period between *t* and *t-1*, the more likely the null hypothesis is rejected. That is because the null defines household stability in the allocation of breadth and depth in spite of changes in options available to households. In addition, options grow much larger with the passage of longer time. In our case, five years is a substantial amount of time for changes in Internet supply. Substitution can arise from a vast array of endless possibilities, either splitting up a large unit of time into many smaller units of time, or from taking many small units and putting them together into one long unit. After years of dramatic changes in supply we would not expect similar patterns to arise.

4. Data

We obtained household machine-level browsing data from Comscore for the years 2008 and 2013. We observe one machine for each household for the entire year, either all of 2008 or all of 2013. Here, the machine should be interpreted as the household's primary home computer. The information collected includes the sites visited on the machine, how much time was spent at each site, and the number of pages visited within the site. We also observe several corresponding household demographic measures including income, education, age, household size, and the presence of children. For simplicity we consider only the first four weeks of a month and do not consider partial fifth weeks, so the maximum number of weeks for a household cannot exceed forty-eight. Importantly, we delete households that have fewer than six months of at least five hours of monthly browsing. We also delete the very few households with more than the 10,080 maximum number of minutes online per week, the result of a defective tracking device. For 2008, we are left with 40,590 out of 57,708 households, and for 2013 we are left with 32,750 out of 46,926 households. In both years, this amounts to over one million machine-week observations. We observe an average of 42.1 and 41.5 (medians 45 and 44) machine weeks per household (s.d. = 6.9 both years) for 2008 and 2013.

Summary statistics of our demographic measures are presented in Table 3. These demographics include household income categories, educational attainment of the head of the household, household size, the age of the head of the household, and an indicator for the presence of children. Comscore's sampling of households is known to target towards higher income households, and we observe that those income levels are comparable across the 2008 and 2013 data. Unfortunately, the education identifiers are mostly missing in 2008, and only available for roughly half of all households in 2013. While there do not appear to be any major differences in the sample composition across years, the 2013 heads of households are mildly younger. In addition, Comscore provides no information on the speed of the broadband connection except to indicate that virtually all of them are not dial-up.

[Table 3 about here]

Table 4 presents summary statistics, such as the concentration of time across sites and the fraction of sessions that exceed ten minutes. If a household is online in a given week, it spends roughly fifteen hours online per week on average in 2008 and fourteen hours online in 2013. Perhaps surprisingly, our measures of browsing behavior are virtually identical across years, with 75% of sessions lasting over ten minutes and households' allocation of time across sites being quite concentrated with an HHI of approximately 2,900. We discuss these similarities in greater detail in the next two sections after associating the variance in them with demographic characteristics of households.

[Table 4 about here]

We face two concerns with the measurement of household time online, biasing the measurement in opposite directions – one upward and the other downward. First, Comscore does not know if a user is watching calmly or left the room with the browser open, possibly biasing our measure upward. It ends the timing for sessions after a fixed period of inactivity, but not until additional time has been added to a session. While we observe some correlates with this mismeasurement (such as multitasking) – and that will help in testing for its importance – there is no practical way to fully eliminate it. In contrast, another factor can bias the length of sessions downward. Many content firms typically use a Content Delivery Network (CDN) to put content at the edge of the network, which reduces the delay experienced by users (e.g., Akamai provides such services to many content firms). In some cases, when the user is switched to a CDN, the name of the CDN (e.g., Akamai) will show up in the browser (e.g., instead of the name of the web site which hired Akamai to host the content on its CDN), even though the user has not discontinued their session. In this case a session will appear shorter than it actually is. Most firms try to retain their brand name in the URL so as not to confuse users, but close inspection of the data shows that this is not always the case. Again, there is no foolproof strategy to detect this bias.

Our approach will recognize these biases, treat them as a source of error, and compensate our analysis when feasible. Fortunately, Comscore did not fundamentally alter its data collection processes between 2008 and 2013 in these dimensions, so our analysis assumes a similar distribution of the biases in the two years of the data. We also do not expect these to differ systematically across application or demographics except in a few instances which we describe below. We also will test this assumption in the analysis when we examine the sensitivity of inferences to multitasking.

5. Empirical Analysis

In this section, we present three types of results that shed light on three corresponding basic questions pertaining to online attention: how, how much and where? In the first subsection, we present findings concerning total time online (how much). In the second subsection, we present findings concerning our measures of fundamental browsing behavior (how). In the third subsection, we present findings on the shares of attention garnered by different online content categories (where). For each of these sets of findings, we analyze how they vary both horizontally (across demographics) and vertically (over time, 2008-2013). We discuss key insights from these comparisons in Section 6.

5.1 Total Time Online

Our first set of analyses concern total time online on the PC. We are limited in our ability to draw conclusions about the total time spent online by a household across all devices, and we possess no information about which household member spent time on the PC in multi-dweller households.

First, our summary statistics show that the average household spends approximately two hours per day on the Internet. When considering vertical changes in total time, our theory predicted that time on the PC could go up or down over the years. We see, in fact, that total time online on the primary home device declined by approximately 5% between 2008 and 2013, which rejects the null on H2. If we assume total time online across all devices increased during this time (see Allen 2015, which supports this assumption), this suggests at least a minimal amount of substitution of online attention across devices. Nonetheless, the decline we observe is rather small, suggesting that much of the increased online attention on tablets and smartphones is in addition to, and not in place of, online attention on the home PC. We will come back to this hypothesis with further analysis below.

Next, we examine cross-sectional differences in total time online on the home device, and whether and how this relationship may have changed between 2008 and 2013. The existing literature studying Internet technology has found that adoption of most Internet technology frontiers is predicted by more income and more education, and (up to a point) younger ages and larger families. Most standard models of the adoption of new products presume that the same set of factors predicts both adoption and the extent of *use* of new technology. However, we observe the Internet many years after most households first used it. The Internet also holds the potential to respond to a different set of forces because it generally consumes leisure time and not money.

We present the results of a simple OLS regression of time online per week on demographics, and show the results in Table 5. These results show that total time is sensitive to income. For example, in 2008, looking at the income endpoints, those with incomes greater than $100,000 spend 835 minutes of time online per week while those with incomes less than $15,000 spend 979 minutes of time online. A similar monotonicity appears in 2013. This confirms H1, namely, total time online declines with increased income. In Figure 1, we show how this relationship compares across our two years of 2008 and 2013. Although we get a statistical rejection of H3, it is clear that there is no important qualitative change in the relationship between time online and income over this period. Hence, in spite of different access technologies and extensive changes in patterns of use in later years, the role of income as a determinant of total time online for the home device is consistent with what has been previously identified in the literature.[5]

[Table 5 about here]

[Figure 1 about here]

While our data do not provide information on non-adopters, information about household adoption of the Internet is readily available for this time period from other sources.[6] Household use of the Internet increases monotonically with income (using slightly different aggregate income categories). That is, adoption of broadband is a normal good. In 2008 the percentage of adults who use the Internet is 54%, 78%, 88%, and 95% for income levels of, respectively, less than $30,000, $30,000 to $50,000, $50,000 to $75,000, and greater than $75,000. In 2013 these rates are, respectively, 72%, 86%, 93%, and 97%. At this aggregate level, and qualitatively similar to what Goldfarb-Prince observed, we also observe "attention inferiority," decline in the amount of online attention with higher income. The contrasts with prior observation are worthwhile to highlight. Prior observations were based on use of the dial-up Internet, while the recent observation reflects adoption and use of broadband, a great deal more activity and time devoted to online activity, and two distinct years of data. This finding reinforces the conclusion that this effect is a stable relationship; hence, we find yet more evidence of "persistent attention inferiority."

Other demographic determinants of time online are generally weak and inconsistent over the two years. We see a positive relationship between more education and total time in 2013, but the relationship is not monotonic in 2008. The finding in 2008 likely resulted from the poor measurement of education in 2008. However, other inconsistencies are more difficult to explain. Large households also spend more time online, but the relationship is only strong in 2008. In

[5] Note that the *levels* of hours are higher than reported in prior surveys, such as Brynjolfsson and Oh (2015), which uses Forrester surveys of households' self-reported number of hours online. In their 2008 data, the average per household is 8.79 hours per week. In contrast to this prior work, here we report active weeks online. That imparts a slight upward bias, but not enough to account for the difference. On average, households are active approximately 83% of the weeks. Moreover, we observe only one device, rather than all devices in a household, which should impart a downward bias in our estimate. We suspect the main issue is measurement. Our data is measured with a number of defaults and not self-reported.

[6] http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/#internet-usage-by-household-income

2013 only the presence of children captures this effect. Total time also declines with the age of head of household in 2008, but no such monotonicity appears in the coefficients for 2013.

These findings support an interesting puzzle. Consider the stability of the relationship between income and total time and the inconsistency in the relationship between age of head of household and total time. They contrast with findings – from surveys done by the Pew Charitable Trusts – about smartphone and tablet ownership, which show (unconditional) monotonicity in adoption of these devices in income (increasing) and age of households (decreasing).[7] If tablets and smartphone ownership *caused* users to substitute time online on devices for the time online on PCs, then we would expect younger and higher income households to adjust their total time more, and that is not observed in these data. (While we do see the expected monotonicity in the education of households in 2013, the baseline in 2008 is poorly measured, and provides no useful information.) Overall, therefore, we observe a puzzle – namely, a decline in time online on the PC, as expected if outside devices caused it, but not the demographic associations that would be indicative of such a cause. Hence, the adoption of tablets and smartphones may have had little to do with even the small decline we observe in time online for the home device.

A back-of-the-envelope calculation can illustrate the implications of the estimates for the predominant role of income. Though ComScore does not seek to compile a representative sample of US households, they do select from a wide range of income, regional, and demographic backgrounds. As long as the error in measurement *from within each income category* is random, then the conditional estimates for each category can be projected to the US household population. So we ask, if the US household population behaves like our sample, what does this imply for the scale of total time online for users from different income groups? Conditional on this assumption, we make such a calculation.

As noted above, other sources (Pew) provide estimates for the adoption of the Internet for four mutually exclusive income categories. We combine the Pew survey with standard US Census estimates of the fraction of households from each income group. Aggregating the total estimates for time use (from our study) into these four income groups,[8] we calculate total time

[7] See, for example, http://www.pewinternet.org/2015/10/29/the-demographics-of-device-ownership/.
[8] The estimate for under $30,000 take a weighted average from three groups, under $15k, $15k-$25k, and half of group making $25k to $35k. The estimate for $30k-$50k take a weight average from the $35k to $50k and half the

online from all US household PCs for all users as 24.4 million hours per week in 2008, and 25.5 million hours in 2013.[9] That makes for a total size of 1.26 billion hours in 2008 and 1.32 billion hours in 2013. Despite the decline in online time per household, the online total time from PCs went up between 2008 and 2013 due to increasing adoption of the Internet, especially among lower income groups.

These estimates about income also imply what type of customer suppliers compete for. The majority of total time online *seen by online suppliers* reflects higher incomes. In 2008 the estimate of total time divides into four categories: 20.6% of the time comes from the lowest income group (under $35,000), 18.2% come from the next group ($35,000-$50,000), and 19% from the next ($50,000-$75,000), while 42.2% come from the highest income group (above $75,000). The percentages for 2013 are, respectively, 25.0%, 17.9%, 17.8%, and 39.3%. In both cases, we note a similar qualitative pattern, with more time coming from high income participants. The fraction of time from the highest income category declined mildly between the two years due to higher adoption rates by low income households.[10]

How is that time divided among different interests, and how does its breadth and depth appear to suppliers? That depends on the distribution among the population of users. We turn to this next.

5.2. Online Attention Allocation Patterns

In this subsection, we present findings concerning our measures of fundamental browsing behavior (how), in terms of breadth and depth. Figure 2 presents the unconditional joint density of our measures of breadth and depth for 2008 and 2013, using all the observed machine weeks

---

$25-$35k. The estimate for over $75,000 takes a weighted average for the estimates for the two income levels above $75,000 (i.e., for $75,000-$100,000, and for $100,000 and up). In all cases, the weights come from number of households, and the estimates for total number of households comes from 2010 US Census.

[9] The number in the text is the sum across all groups (indexed by i) in the total time online. Total time online for group *i* = (Total US households in income group *i*)*(adoption rate for group *i*)*(Point estimate for total time for group *i*).

[10] The cross-sectional differences in adoption rates alter the composition of user attention for which a supplier of content potentially competes in 2008 and 2013. Though higher income households each have lower total PC time per household, more such households use the Internet in high percentages. In 2013 the composition changes due to increasing adoption, again, especially, among lower income groups.

of data. Here, we see a very well-behaved joint distribution that strongly resembles a joint normal. However, it is the comparison of the graphs over time that generates a particularly striking finding – the distribution of these measures of online attention allocation is essentially unchanged during this five year time period! The summary statistics in Section 4 showed that the means of each measure were very similar and the features of the demographics in each sample also resembled one another, but Figure 2 clearly indicates that the similarity goes well beyond just the means – the entire distributions are nearly identical, a property we call "persistent attention distribution."

[Figure 2 about here]

Despite the striking visual similarity, we can reject the null hypothesis that the breadth and depth are statistically indistinguishable, likely because our combined sample size is over three million. Tables 6a and 6b present statistical tests of the means of our measures of breadth and depth across years and a Kolmogorov-Smirnov test for the equality of distribution functions across years, respectively. While not statistically identical, these differences are economically insignificant. The mean of household breadth is 3.5% greater in 2013 and household depth is greater by 1%.

[Table 6a, 6b about here]

A possible concern about our finding of persistent attention distribution is that the measures of online attention allocation may be strongly driven by a household's total time online on the home device. For example, we may worry that households spending the most time online would have greater depth and perhaps more breadth. In short, we are concerned that a household's location within the distribution presented in Figure 2 arises merely from income's influence on total time. To address this concern, we break total time online on the home device into quartiles, and recreate our joint distribution for each quartile. The results are in the Appendix (Figure A.1). Here we see that, while not statistically identical, the joint distribution of our measures of a household's browsing behavior is strikingly consistent across the quartiles.

Further, we see that within quartile, this joint distribution is again highly stable between 2008 and 2013.

Another concern with this finding centers on multitasking. Specifically, when browsing online, households can simultaneously visit multiple sites, e.g., via multiple tabs on a browser or even multiple open windows on the PC. We worry that changes in multitasking behavior mask/negate changes in breadth and depth over time. We address this concern in two ways. First, we note that time spent multitasking only changed negligibly between 2008 and 2013 in our data. The percentage of time spent multitasking is 19.3% in 2008 and 19.7% in 2013. Second, we break time multitasking on the home device into quartiles, and recreate our joint distribution for each quartile. The results are also in the Appendix (Figure A.2). It is not surprising that within a given year the level of multitasking correlates with distinct distributions of breadth and (especially) depth. For this exercise we are more interested in the comparison across years. Here we see again that, while not statistically identical, the joint distribution of our measures of a household's browsing behavior is strikingly consistent across the quartiles. Further, we see that within a quartile, this joint distribution is again highly stable between 2008 and 2013.

As shown in our summary statistics in Section 4, there are some differences in the demographic profiles between our sample in 2008 and 2013. Another concern could be that online attention allocation behavior, conditional on demographics, did change over this time period, but, by some lucky coincidence, the changes in behavior offset the demographic changes in the samples from the two years. To address this possibility, we assess if and how our measures of online attention relate to our demographics, namely: income, age, education, household size, and presence of children. This analysis not only addresses a potential concern with our vertical finding, but also directly shows if and how our measures of breadth and depth vary horizontally, i.e., with respect to demographics.

Table 7 presents a set of seemingly-unrelated-regressions (SURs) for our measures of breadth and depth. We do not observe monotonic estimates with respect to income. Indeed, both depth and breadth are virtually independent of income levels after controlling for total time online. The demographics that meaningfully correlate with breadth are education (in 2013), age of head of household, and household size. In particular, more educated households, larger households, and younger households visit a larger variety of sites in both 2008 and 2013. In contrast, depth is largely independent of demographics.

[Table 7 about here]

These estimates do not help much in explaining why the two figures from 2008 and 2013 look so similar. Broadly speaking, demographics explain less than 15% (for depth) and less than 3% (for breadth) of the variation in our household classifications. Households that are larger and have more education and income are less likely to be classified as dwellers (i.e., someone who stays at a web site for a long period), but the economic significance of these effects is modest. Households with older heads of household and more education are less likely to be classified as focused, but the economic significance of these effects is also modest, therefore insufficient to shape the ultimate outcome by much.

Demographics have a modest role in causing changes in breadth and depth, and even the total amount of time online. While income's impact is monotone and consistent for time online, demographics as a whole explain a very limited amount of variation in all three of these dependent variables. As far as income's relationship to breadth and depth is concerned, a household's income mildly shapes its total time online and mostly through this channel, its breadth and depth. Since breadth and depth do not appear to have the same measured determinants as total time online, we note an interesting implication that builds on our earlier observations about adoption rates across income groups. Because households' depth of use is orthogonal to income and the changing rates of adoption across income had no consequence for depth, suppliers have observed little change in the depth of their users over our sample period.

5.3. Online Attention Category Shares

As noted above, the period spanning 2008 to 2013 saw large changes in the supply of web sites, particularly with regard to online video. Consequently, we may see notable changes in *where* households allocate their time, despite remaining stable in *how* they allocate their time.

We classified the top 1000 sites from both 2008 and 2013 by categories established by Webby and measured the share of attention garnered by each category for both years. We present these shares in Figure 3. Here we see that, in 2008, Chat is by far the largest category, attracting over 25% of households' attention; however, this category saw a dramatic shift by 2013, dropping to less than 2% in 2013. Other sites imitated similar functionality, especially in social media. Attention allocated to News sites also decreased, from roughly 10% down to 5%, again, because Social Media tended to offer News. We observe that Social Media and Video have the largest increases of attention, to 26% and 16%, respectively. Interestingly, three-quarters of the drop in share for Chat and News is reflected in the increased shares of Social Media and Video. Some of this is not surprising, since popular social media sites, such as Facebook, attempted to offer functionality that replicated what chat and news services used to provide as stand-alone services.

[Figure 3 about here]

Table 8 contains the top 20 sites of 2008 and 2013. A quick glance at these rankings and the change between 2008 and 2013 further confirms what we see in Figure 3. Particularly noteworthy is the mass exodus of Chat users and the rise in Video use.

[Table 8 about here]

While Figure 3 and Table 8 suggest a large vertical change in where households went online, at least on the home device, a possible concern is that little changed at the very top sites, which garner a large share of time online. In particular, it could be that category shares changed, but top sites did not, and it is top sites that largely drive our measures of time allocation. Revisiting Figure 3, it appears this is not the case. In that figure, the dark bars represent the share of top 5 sites from 2008 within each category. In 2008, the top 5 sites accounted for 42.5% of all browsing, but in 2013 these same sites account for only 20%. Here, we see significant turnover, as only web services show any persistence in this share between the two years. Hence, it appears high usage and the persistence of the very top sites do not drive our findings.

Our analysis of where households go online more closely follows that of total time online, in that demographics do appear to be predictive in both years. Table 9 contains multinomial logit results that show higher income consistently predicts some preference for online services in both years. Higher income households prefer services that examine credit history, offer educational services, support (online and web-based) games, provide news, support online banking, offer online shopping, provide online sports services, and supply online video services. Nonetheless, demand for several online services does not appear to be sensitive to income, including chat, credit history, social media, and pornography.

[Table 9 about here]

This is further evidence that online user behavior contains a mix of fixed determinants, while also varying in the cross section. The latter variance follows measureable differences in demographic determinants. We also observe changes in website choice over time, again following measureable differences in demographics. In the face of such predictable variance in behavior, the stability of breadth and depth seems all the more surprising.

5.4. Evaluating the Predictions of the Standard Model: Is a Behavioral Component Missing?

Between 2008 and 2013 we see a remarkable lack of change in both the breadth and depth of households' browsing habits. These results are difficult to rationalize in the context of the standard model with negligible setup costs. Such a model predicts an increase in the breadth of household browsing when supply increases, and makes no prediction about the depth of household browsing. The standard model with setup costs more easily rationalizes the results: the supply of new sites increases breadth on the home device, while alternative devices decreases breadth on the home device, resulting in an ambiguous net effect. With respect to the depth of household browsing, the standard model without setup costs is agnostic about depth, while the standard model with positive setup costs predicts no change in the depth of household browsing.

While the standard model with setup costs can rationalize the lack of change in the depth of household browsing between 2008 and 2013, that model also predicts that all sessions be the same length. Consequently, all sessions either fall above or below a given threshold. Our data do not exhibit anything close to constant session length; rather, we see a wide mix of sessions of different lengths, *where the proportion has not changed across years*.

An asymmetric utility function that captures how a household values each site differently can explain the mix of sessions of different lengths that we observe. However, we believe that our empirical findings point toward a static theory of household browsing behavior. In part, this is because the demand side did not react to a massive change in the environment of supply. For example, the vast change in the menu of supply from 2008 to 2013 did not change the breadth or depth of household browsing. The vast change in supply altered only where households allocated their online time.

To summarize, there is a discrete nature to households' site choices and a limit to the number of sites that can be visited. The continuous utility function of the standard model without setup costs ignores these features and is at odds with our empirical results. We do not observe households splitting time into more numerous and shorter site visits, as predicted by the standard model without setup costs. The standard model augmented with positive setup costs performs better: it predicts that even with increasing supply, there will be a finite number of sites visited.

The standard model augmented with setup costs does fall short, however, due to the lack of a clear prediction: it offers no guidance as to whether household breadth of browsing will increase or decrease. The unchanging breadth *and* depth of household browsing patterns invites an alternative theory of household browsing behavior, one that can explain this constancy.

What would such a theory contain? Our results do not settle the question. Theory might be behavioral or hierarchical, where a household receives exogenous "slots" of time that are allocated to brief leisure activities such as watching television, reading, or browsing online. If, for example, online behavior is largely driven by a constant and exogenous nature of offline activities, then that would explain the remarkable stability of household browsing behavior over time, despite vast changes in the amount and type of supply over the same period.

Field work conducted by anthropologists and researchers on user-machine design has shown behavior consistent with exogenous offline activities determining the time spent in online activities. Such researchers have documented the periodic – or "bursty" – use of many online sources (Lindley, Meek, Sellen, Harper, 2012, Kawsaw and Brush, 2013). This work also documents the "plasticity" of online attention, as an activity that arises in the midst of other household activities as a "filler" activity (Rattenbury, Nafus and Anderson, 2008, Adar, Teevan, Dumais, 2009). This type of field work provides an explanation for the consistency of breadth and depth patterns within a household in spite of large changes in the available options. It explains the consistency as a result of unchanging household habits, which shape availability of time, and shape the availability of slots of time. These theories would hypothesize that the slots do not change much, because they cannot change much, even as the supply of online web sites does change.

6. Implications for Online Competition

In this paper, we complete the first characterization of the aggregate supply of online attention. We believe our findings speak to a number of important questions across several fields of research. For example, how do these findings change the approach for valuing online consumer surplus in a labor-leisure framework? How does this approach reconcile household aggregation of time and activity into demand for Internet access with competition between online activities? How does heterogeneity in the breadth and depth of activities translate into demand for more speed? In this section, we summarize our findings in Section 5 and discuss their main implications.

We summarize our findings in Table 10, and state the results as follows. First, total time online at the primary home device has only modestly declined, and the decline is generally consistent across income groups. This decline is not sufficient to change the predominant role played by higher income users, which arises because adoption rates for the Internet as of 2013 were still notably higher for this group, compared to lower income groups. Second, the way in which households allocate their online attention, as measured by the concentration of sites visited (breadth) and time spent in "long" sessions (depth), has remained remarkably stable. In addition, neither of these measures is well-predicted by total time online or major demographics. Lastly, the period between 2008 and 2013 saw major changes in online category shares, with social media and video experiencing significant increases while chat and news experienced significant declines. These choices of application are predicted by exogenous factors, such as income.

[Table 10 about here]

Our findings suggest that new points of contact – in the form of additional computers, tablets and smartphones – are substituting time from the primary home device, but only modestly. Consequently, as total time across all devices strongly increased during this time (e.g., Allen 2015), it appears this increase manifested as time online at additional devices, largely in addition to the relatively stable use of the home device. Hence, any new value stemming from additional time online appears to be largely coming from time on new, alternative devices.

This adds up to a surprising characterization of the supply of attention by households and the demand for online activities. The amount of time spent online varies with income and largely not the menu of supply of online websites.

We find the stability of online attention patterns over this time period to be especially striking, given the explosion of online video content and the growth of secondary devices during this time. In this context, we highlight three key takeaways. First, this finding shows that any changes in value households received from these developments did *not* arise from a change in the *way* households allocated their online attention. Therefore, even if many households shifted their attention to more sites with video offerings, which tend to demand more time, it appears these shifts are at the expense of attention at other sites, at which the household was already spending significant time. Second, this result suggests that household online attention via secondary devices has not altered the basic pattern of online attention for the primary home device. This implies that households are not systematically distributing their attention across devices in a way that, e.g., shifts "touristy" or focused sessions to secondary devices. Lastly, this result implies that, despite a large influx of new sites and content offerings, households are not increasing the spread of their attention in response, at least at the device level, and not altering the time focused on online consumption.

Taken together, our results add up to a striking model of competition for attention. As expected, there were large changes in the supply of web sites for households to visit, and many households responded to that expanded choice. Presumably existing websites also improved, as they fought to keep the attention of current users. However, the competition for users was constrained by a virtually unmovable feature of demand – the breadth of sites households are willing to visit, and the length of slots households were willing to sustain at a given site.

What model of demand is consistent with these properties? We speculate that hierarchical models of sequences of decision making – where households first choose devices for access and then choose how to allocate it, facing setup costs and other non-price frictions – will be a fruitful avenue for further developments. While there certainly will be debate about the proper, specific modeling features, our empirics are the first to identify these fundamental patterns of online attention, which future models should be capable of producing.

We also speculate that these findings will place an important constraint on models of household substitution between competing online activities. Competition for attention imposes a costly bottleneck on supplier behavior, and suppliers are competing for a finite supply of user time. Household substitution determines this competitive situation. In other words, the size of the market for online competition should be thought of as segmented along length of time, and so too should entry and related entrepreneurial behavior. This segmentation arises *in addition* to more familiar notions of differentiated competition, such as the function and application of the online site.

We also believe research should test between models that impose static specifications on labor/leisure substitution and models consistent with the microdata of household decision making. Lastly, these findings place important constraints on models of online/offline substitution, and raise several open questions. What is the appropriate model of heterogeneity in consumer surplus for demand for online access, when online activity involves heterogeneity in frictions? What type of model of labor/leisure online/offline substitution is consistent with *persistent attention distribution* and *persistent attention inferiority*? How much does of the desire to reduce frictions – e.g., by improving the operation of search tools – motivate household demand for more broadband speed, in addition to the performance improvements in specific applications, such as gaming? We leave these questions for future work.

7. Conclusions

This study uses extensive microdata on user online activity to characterize the links between user allocation of attention and online competition in the absence of prices. We characterize household heterogeneity in allocation of attention at any point in time, how households substitute between sources of supply over time, and develop implications for aggregate demand changes in the face of an increasing supply of options.

Our findings suggest that aggregate demand mixes fixed properties with flexible choices. First, income plays an important role in determining the allocation of time to the Internet. We find further evidence for *persistent attention inferiority*, namely, that higher income households spend less total time online per week. Relatedly, we also find that the level of time on the home device only mildly responds to the menu of available web sites and other devices – it slightly declines between 2008 and 2013, despite large increases in online activity via smartphones and tablets over this time. Most surprising, breadth and depth of online use have remained remarkably *stable* over the five years, namely, *persistent attention distribution*.

These findings can serve as an important guide for future modeling of demand for online attention as competition for that attention. We observe little user substitution of time or concatenation of time. That pattern is consistent with unchanging household habits, which shape availability of time and shape the availability of slots of time.

Our results also raise questions about the competition for user attention. If slots of time do not change much, then firms have strong incentives to respond to this feature of the market with products and services tailored to this feature of demand. But that raises an open question about the costs of delivering services tailored to that feature. Many observers have noticed that the costs of web site development have fallen dramatically, and technical improvement have enabled new services based on video delivery over the Internet. However, changes in supply conditions will not, nor ever has, created its own demand, and we would expect markets to equilibrate accordingly. Competition for attention imposes a costly bottleneck on supplier behavior, and, to put it simply, suppliers are competing for a finite supply of user time, but generally lack the ability to use price discounts as an instrument for attracting user attention. Models of online competition should accommodate these features and investigate how they shape online supplier behavior.

Finally, given our discovery of remarkable stability in how households allocate their scarce attention, we hypothesize that such stability in behavior may also exist as changes occur in other markets for attention, such as in television and radio. For example, increases in the supply of television content and devices through which to consume that content will likely cause households to switch to that new content (a change in “where?”), may cause a modest decline in the amount of attention allocated to the original device used for consumption (a change in “how much?"), but may not change *how* households fundamentally choose to disperse attention across content and *how* households choose their intensity of attention to content.

References

Adar, Eytan, Jaime Teevan, Susan T. Duma, 2009, "Resonance on the Web: Web Dynamics and Revisitation Patterns," *CHI 2009.*

Allen, R. 2015. Mobile Internet Trends Mary Meeker 2015 1. Smart Insights. http://www.smartinsights.com/internet-marketing-statistics/insights-from-kpcb-us-and-global-internet-trends-2015-report/attachment/mobile-internet-trends-mary-meeker-2015-1/

Anderson, Simon, 2012. Advertising on the Internet. In *the Oxford Handbook of the Digital Economy*, eds. Martin Peitz and Joel Waldfogel, Oxford University Press.

Athey, S., Calvano, E. and Gans, J. 2013. The Impact of the Internet on Advertising Markets for News Media, Stanford working paper.

Athey, S. and Mobius, M. 2012. The Impact of News Aggregators on Internet News Consumption, Stanford working paper.

Baye, M., De los Santos, B., and Wildenbeest, M. 2016. What's in a Name? Measuring Prominence and Its Impact on Organic Traffic from Search Engines. *Information Economics and Policy* 34: 44-57.

Becker, G. 1965. A Theory of the Allocation of Time. *The Economic Journal* 75: 439-517.

Brynjolfsson, Erik, and JooHee Oh, 2012, The Attention Economy: Measuring the Value of Free Digital Services on the Internet. In *33rd International Conference on Information Systems (ICIS 2012)*. Orlando, Florida: December.

Chiou, Lesley, and Catherine Tucker, 2015, Digital Content Aggregation Platforms: The Case of the News Media. Working paper.

Gabaix, Xavier. 2014. A Sparsity-Based Model of Bounded Rationality. *Quarterly Journal of Economics* 129: 1661-1710.

Goldfarb, A. and Prince J. 2008. Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide. *Information Economics and Policy* 20: 2-15.

Goolsbee, Austan, and Peter Klenow, 2006, Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet. American Economic Review, Papers and Proceeedings. May. pp 108-113.

Greenstein, Shane. 2015. How the Internet Became Commercial. Privatization, Innovation, and the Birth of a New Network. Princeton Press; Princeton, NJ

Hauser, J., Urban, G., and Weinberg, B. 1993. How Consumers Allocate Their Time When Searching for Information. Journal of Marketing Research 30: 452-466.

Hitt, L. and Tambe, P. 2007. Broadband Adoption and Content Consumption. *Information Economics and Policy* 19: 362-378.

Kawsar, Fahim, and A. J. Bernheim Brush, 2013, "Home Computing Unplugged: Why, Where and When People Use Different Connected Devices at Home," UbiComp, p 627-636

Lanham, R. 2006. The Economics of Attention: Style and Substance in the Age of Information. University of Chicago Press.

Lieber, E., and C. Syverson, 2012, "Online versus Offline Competition," In *the Oxford Handbook of the Digital Economy*, eds. Martin Peitz and Joel Waldfogel, Oxford University Press.

Lindley, Siân, Sam Meek, Abigail Sellen, and Richard Harper, 2012, "It's Simply Integral to What I do': Enquiries into how the Web is Weaved into Everyday Life," International World Wide Web Conference, Committee (IW3C2). *WWW 2012*, April 16–20, 2012, Lyon, France

Nevo, Aviv, John L, Turner, and Jonathan W. Williams, 2015, Usage Based Pricing and Demand for Residential Broadband, Econometrica, Forthcoming.

Peitz, M., and J. Waldfogel, 2012, *The Oxford Handbook on the Digital Economy*, Oxford University Press.

Ratchford, B., Lee, M.S., and Talukdar, D. 2003. The Impact of the Internet on Information Search for Automobiles. *Journal of Marketing Research* 40: 193-209.

Tye Rattenbury, Dawn Nafus, and Ken Anderson, 2008, "Plastic: A Metaphor for Integrated Technologies, *UbiComp'08. Pp232-241*

Rosston, Greg, Scott Savage, and Don Waldman, 2010. "Household Demand for Broadband Internet in 2010," The B.E. Journal of Economic and Policy Analysis (Advances), 10(1), Article 79.

Savage, Scott, and Don Waldman, 2009. "Ability, Location and Household Demand for Internet Bandwidth," International Journal of Industrial Organization 27 (2) pp. 166-174.

Simon, H. 1971. Designing Organizations for an Information-Rich World. Computers, Communication, and the Public Interest, ed. Martin Greenberger. The Johns Hopkins Press.

Wallsten, S. 2013. What are We Not Doing When We're Online? NBER Working paper.

Webster, James, 2014. The Marketplace of attention: How Audiences Take Shape in a Digital Age, MIT Press.

Figures

Figure 1
Total Time Online by Income (2008, 2013)

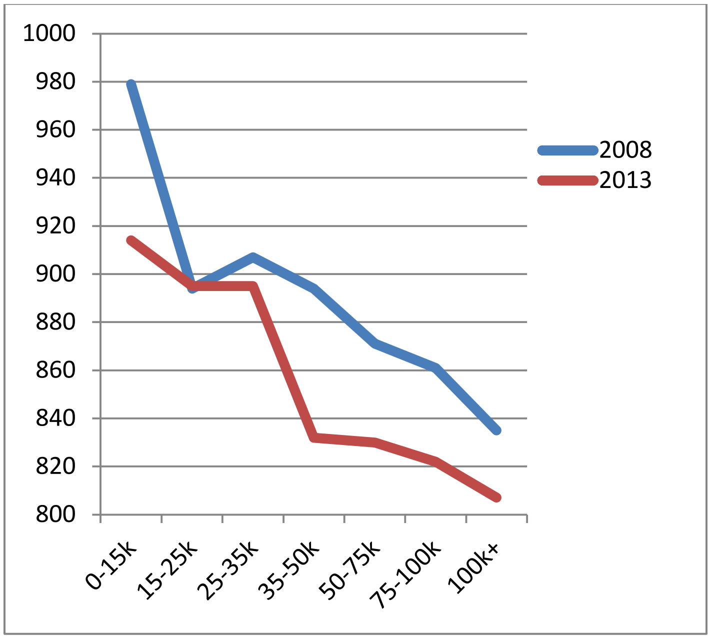


Figure 2
Unconditional Distribution of Online Attention (2008 vs. 2013)

2008	2013

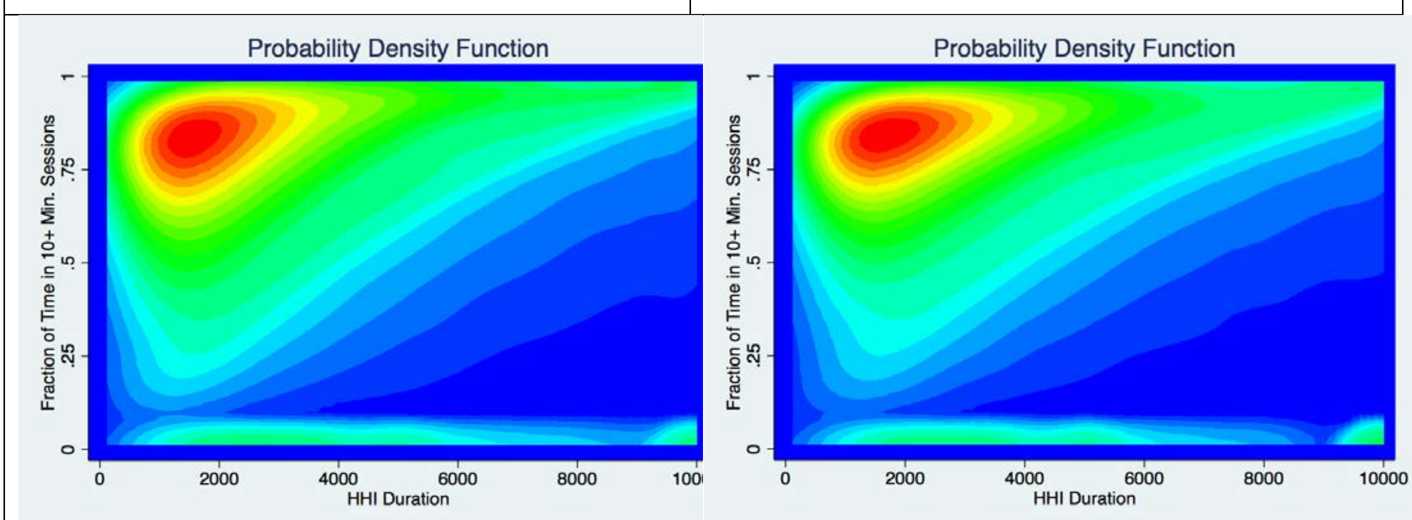
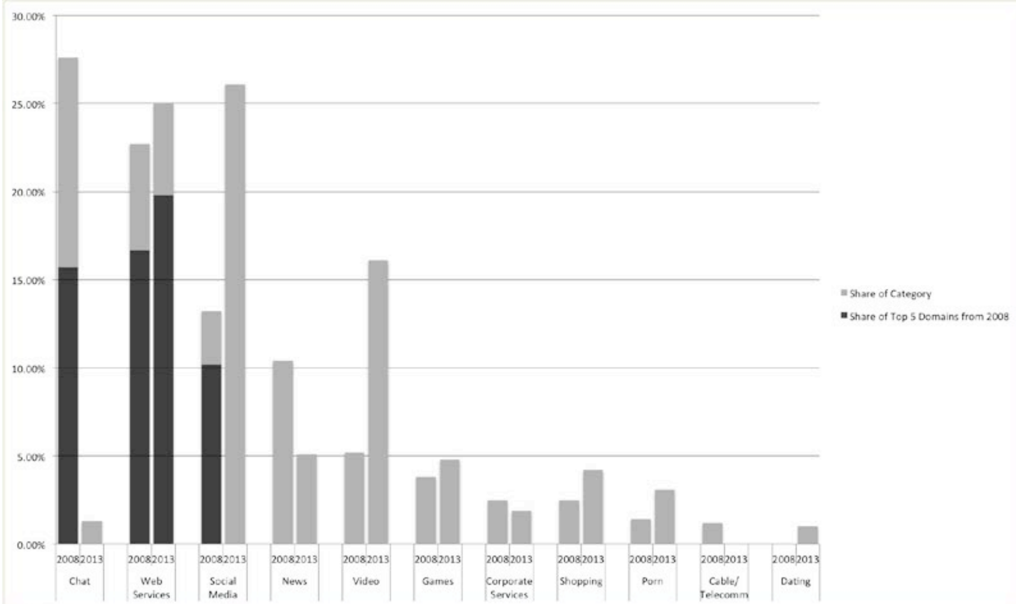


Figure 3
Changes in Attention Allocation across the Top 1000 Sites by Category (2008 - 2013)



Tables

Table 1
Simplified Household Types for Allocation of Online Attention

	High C	**Low C**
High L	Focused Dweller	Unfocused Dweller
Low L	Focused Tourist	Unfocused Tourist

Table 2
Summary of Standard Model's Predictions in Response to Two Shocks

	Small setup costs	Large setup costs ($F >> 0$)
Shock 1: New Sites	$\Delta TO_i \geq 0$ $\Delta C_i < 0$ ΔL_i (No prediction)	$\Delta TO_i \geq 0$ $\Delta C_i \leq 0$ $\Delta L_i = 0$
Shock 2: New Device	$\Delta TO_i \leq 0$ $\Delta C_i = 0$ ΔL_i (No prediction)	$\Delta TO_i \leq 0$ $\Delta C_i \geq 0$ $\Delta L_i = 0$

Table 3
Household Summary Statistics

Variable	2008 N = 40,590		2013 N =32,750	
	Mean	Std. Dev.	Mean	Std. Dev.
Income < $15k	0.14	0.34	0.12	0.33
Income $15k-$25k	0.08	0.27	0.10	0.30
Income $25k-$35k	0.09	0.29	0.11	0.31
Income $35-$50k	0.11	0.31	0.15	0.35
Income $50-$75k	0.23	0.42	0.21	0.40
Income $75-$100k	0.16	0.36	0.13	0.34
Income $100k+	0.20	0.40	0.19	0.39
Age of Head of Household 18-20	0.00	0.07	0.05	0.21
Age of Head of Household 21-24	0.02	0.14	0.07	0.26
Age of Head of Household 25-29	0.05	0.22	0.08	0.27
Age of Head of Household 30-34	0.07	0.26	0.10	0.30
Age of Head of Household 35-39	0.11	0.31	0.08	0.28
Age of Head of Household 40-44	0.15	0.35	0.10	0.31
Age of Head of Household 45-49	0.17	0.38	0.12	0.33
Age of Head of Household 50-54	0.15	0.35	0.12	0.33
Age of Head of Household 55-59	0.10	0.30	0.09	0.29
Age of Head of Household 60-64	0.07	0.25	0.07	0.25
Age of Head of Household 65+	0.10	0.30	0.12	0.32
HH size = 1	0.07	0.25	0.12	0.32
HH size = 2	0.34	0.47	0.25	0.43
HH size = 3	0.25	0.43	0.21	0.40
HH size = 4	0.18	0.39	0.19	0.39
HH size = 5	0.11	0.31	0.16	0.37
HH size = 6+	0.05	0.22	0.07	0.27
Education < High School	0.00	0.01	0	0
Education High School	0.00	0.06	0.03	0.17
Education Some College	0.00	0.06	0.19	0.40
Education Associate Degree	0.00	0.02	0.16	0.37
Education Bachelor’s Degree	0.00	0.06	0.11	0.32
Education Graduate Degree	0.00	0.04	0.01	0.08
Education Unknown	.99	0.11	0.49	.50
Children Dummy	.68	.47	.73	.44

Table 4

Summary Statistics of Browsing Behavior

	Year = 2008 *N =1,721,820*			
Variable	**Mean**	**S.D.**	**Min**	**Max**
Minutes online per week	884	1281	1	10080
Unique sites visited per week	41	44	1	3936
Focus (HHI across sites)	2868	2026	33	10000
Propensity to Dwell (Fraction of sessions > 10 minutes)	0.75	0.23	0	1
	Year = 2013 *N = 1,360,683*			
Minutes online per week	849	1091	1	10078
Unique sites visited per week	41	47	1	7525
Focus (HHI across sites)	2968	2061	1.51	10000
Propensity to Dwell (Fraction of sessions > 10 minutes)	.76	.22	0	1

Table 5
<u>Linear Regression - Time Per Week on Demographics</u>

	2008	**2013**
Covariate	**Minutes per Week**	**Minutes per Week**
Income $15k-$25k	-80*** (-3.83)	-19 (-0.95)
Income $25-$35k	-73*** (-3.57)	-19 (-0.96)
Income $35k-$50k	-91*** (-4.73)	-79*** (-4.49)
Income $50k-$75k	-118*** (-7.16)	-85*** (-5.08)
Income $75k-$100k	-131*** (-7.46)	-95*** (-5.25)
Income $100k+	-166*** (-9.90)	-124*** (-7.14)
Education High School	262 (1.84)	-
Education Some College	289 (1.97)	18 (0.64)
Education Associate Degree	189 (1.12)	13 (0.46)
Education Bachelor's Degree	348 (2.34)	80** (2.72)
Education Graduate Degree	248 (1.63)	131 (1.91)
HH Size = 2	-8 (-0.38)	-35* (-2.03)
HH Size = 3	10 (0.44)	-35 (-1.86)
HH Size = 4	27 (1.14)	-10 (-0.48)
HH Size = 5	75** (2.86)	1 (0.05)
HH Size = 6	114*** (3.69)	-21 (-0.87)
Age of Head of Household 21-24	-387*** (-4.20)	9 (0.34)
Age of Head of Household 25-29	-434*** (-4.88)	-16 (-0.62)
Age of Head of Household 30-34	-478*** (-5.42)	-36 (-1.47)
Age of Head of Household 35-39	-402*** (-4.58)	-21 (-0.84)
Age of Head of Household 40-44	-361*** (-4.11)	-18 (-0.71)
Age of Head of Household 45-49	-382*** (-4.36)	41 (1.69)

Age of Head of Household 50-54	-408^{***} (-4.66)	53^{*} (2.12)
Age of Head of Household 55-59	-502^{***} (-5.71)	14 (0.54)
Age of Head of Household 60-64	-531^{***} (-6.01)	11 (0.40)
Age of Head of Household 65+	-551^{***} (-6.28)	15 (0.59)
Children	3 (0.25)	132^{***} (10.46)
Constant	959^{***} (6.12)	800^{***} (21.53)
R-Squared	0.01	0.01
N	1,710,147	1,359,331

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Std errors clustered at the machine level.

Table 6a
Test of Equality of Means Across Years

	Dependent variable	**Dependent variable**
	Breadth (HHI of time across sites)	**Depth (Fraction of Sessions > 10 minutes)**
2013	139*** (12.27)	0.01*** (12.67)
Demographic Controls	Y	Y
Control for Time Online	Y	Y
N	3,069,478	3,069,478

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6b

Two-Sample Kolmogorov-Smirnov Test for Equality of Distribution Functions

	Variable	**Variable**
	Breadth (HHI of time across sites)	**Depth (Fraction of Sessions > 10 minutes)**
p-value	0.00	0.00
N	3,069,478	3,069,478

Table 7
SUR – Fraction of Sessions > Ten Minutes and Time HHI Across Sites

	2008	2008	2013	2013
Covariate	**HHI**	**Fraction > 10**	**HHI**	**Fraction > 10**
Income \$15k-\$25k	10 (1.37)	-0.00*** (-3.84)	22** (2.98)	0.00* (2.45)
Income \$25-\$35k	7 (0.99)	-0.01*** (-11.54)	1 (0.10)	-0.00 (-0.04)
Income \$35k-\$50k	-8. (-1.32)	-0.01*** (-14.78)	11 (1.57)	-0.00*** (-4.20)
Income \$50k-\$75k	-30*** (-5.52)	-0.01*** (-19.44)	16* (2.51)	-0.00*** (-4.10)
Income \$75k-\$100k	-1 (-0.26)	-0.01*** (-23.77)	-28*** (-3.94)	-0.00 (-1.76)
Income \$100k+	-43*** (-7.61)	-0.02*** (-28.05)	-14* (-2.12)	-0.00*** (-6.68)
Education High School	624*** (4.30)	0.09*** (6.17)	-	-
Education Some College	530*** (3.65)	0.07*** (5.01)	-12 (-1.08)	-0.01*** (-10.18)
Education Associate Degree	403* (2.49)	0.10*** (6.05)	-65*** (-5.85)	-0.01*** (-11.85)
Education Bachelor's Degree	299* (2.05)	0.09*** (5.95)	-99*** (-8.60)	-0.01*** (-9.63)
Education Graduate Degree	309* (2.10)	0.10*** (6.33)	-126*** (-5.32)	-0.02*** (-6.70)
HH Size = 2	-44*** (-6.54)	-0.00 (-0.59)	-20** (-2.84)	-0.00 (-0.29)
HH Size = 3	-58*** (-7.21)	-0.00 (-0.30)	-18* (-2.34)	-0.00 (-0.71)
HH Size = 4	-71*** (-8.68)	0.00 (0.53)	-18* (-2.19)	0.00 (1.34)
HH Size = 5	-103*** (-11.75)	0.00** (2.94)	-36*** (-4.31)	-0.00 (-0.52)
HH Size = 6	-235** (-22.92)	0.00*** (4.31)	-50*** (-5.16)	-0.00 (-1.59)
Age of Head of Household 21-24	87** (3.25)	-0.00* (-2.57)	-20 (-1.85)	-0.00*** (-3.62)
Age of Head of Household 25-29	50* (2.00)	-0.01* (-2.41)	-33** (-3.15)	-0.01*** (-7.44)
Age of Head of	100*** (4.03)	-0.00 (-1.06)	-0 (-0.02)	-0.00 (-0.78)

Household 30-34				
Age of Head of Household 35-39	105*** (4.27)	0.00 (0.90)	-8 (-0.77)	-0.00* (-2.54)
Age of Head of Household 40-44	185*** (7.51)	0.00 (1.52)	51*** (5.12)	-0.00*** (-4.26)
Age of Head of Household 45-49	232*** (9.43)	0.00 (0.92)	-0 (-0.04)	-0.00*** (-4.38)
Age of Head of Household 50-54	233*** (9.47)	-0.00 (-0.81)	-48*** (-4.87)	-0.01*** (-6.22)
Age of Head of Household 55-59	199*** (8.04)	-0.01*** (-3.47)	20* (1.98)	-0.01*** (-6.16)
Age of Head of Household 60-64	304*** (12.18)	-0.01* (-2.49)	16 (1.52)	-0.01*** (-4.81)
Age of Head of Household 65+	360*** (14.56)	-0.01** (-2.78)	53*** (5.41)	-0.01*** (-7.30)
Children	-59*** (-12.78)	-0.00 (-1.66)	-142*** (-27.01)	-0.00 (-1.05)
Minutes per Week	-0 (-0.37)	0.00*** (531.17)	-0*** (-181.12)	0.00*** (438.72)
Constant	2652*** (18.26)	0.62*** (41.25)	3346*** (228.47)	0.71*** (473.26)
N	1,710,147	1,710,147	1,359,331	1,359,331
R-Squared	0.00	0.14	0.03	0.13

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Note that across years the education dummies are relative to no high school in 2008 and relative to high school in 2013. Std errors not clustered.

Table 8
The Top 20 Sites of 2008 and 2013 (by Total Time Allocated)

2008 Top 20 Sites	Category	2013 Top 20 Sites	Category
myspace.com	Social Media	facebook.com	Social Media
yahoo.com	News	youtube.com	Video
yahoomessenger.exe	Chat	google.com	Web Services
aim6.exe	Chat	yahoo.com	News
google.com	Web Services	tumblr.com	Personal Blog
msnmsgr.exe	Chat	msn.com	News
youtube.com	Video	aol.com	News
msn.com	News	craigslist.org	Shopping
aol.com	News	bing.com	Web Services
aim.exe	Chat	ebay.com	Shopping
facebook.com	Social Media	amazon.com	Shopping
live.com	News	twitter.com	Social Media
msn.com-prop	Chat	yahoomessenger.exe	Chat
myspaceim.exe	Chat	go.com	Sports
ebay.com	Shopping	wikipedia.org	Web Services
waol.exe	Chat	live.com	News
starware.com	Corporate Services	skype.exe	Chat
pogo.com	Games	reddit.com	Social Media
craigslist.org	Shopping	outlook.com	Web Services
go.com	Sports	netflix.com	Video

Table 9
Multinomial Logit: Category of Site Visited by Income Levels

	Dependent Variable	Dependent Variable	Dependent Variable	Dependent Variable	Dependent Variable	Dependent Variable
Covariates	**2008**	**2013**	**2008**	**2013**	**2008**	**2013**
	Chat	***Chat***	***Credit History***	***Credit History***	***Education***	***Education***
Income $15k-$25k	-0.06** (-2.62)	-0.21* (-2.28)	0.11 (0.82)	0.05 (1.70)	-0.06 (-1.62)	0.02 (0.27)
Income $25k-$35k	-0.01 (-0.24)	-0.08 (-0.95)	0.24* (2.01)	0.03 (0.90)	0.03 (0.99)	0.21** (3.26)
Income $35k-$50k	-0.01 (-0.30)	-0.17* (-2.01)	0.34** (3.03)	-0.00 (-0.12)	0.03 (0.83)	0.09 (1.51)
Income $50k-$75k	0.03 (1.73)	-0.05 (-0.73)	0.28** (2.84)	0.05 (1.79)	0.06* (2.42)	0.22*** (3.89)
Income $75k-$100k	0.05 (2.85)	-0.11 (-1.29)	0.26* (2.46)	0.09** (3.08)	0.07* (2.43)	0.38*** (6.34)
Income $100k+	0.12*** (6.36)	-0.11 (-1.47)	0.012 (0.11)	0.26*** (9.87)	0.16*** (6.09)	0.43*** (7.79)
Constant	-1.92*** (-136.18)	-4.46*** (-76.96)	-5.53*** (-68.94)	-2.40*** (-111.76)	-2.76*** (-133.59)	-3.99*** (-86.94)
	2008	**2013**	**2008**	**2013**	**2008**	**2013**
	Financial Services	***Financial Services***	***Games***	***Games***	***News***	***News***
Income $15k-$25k	0.06 (1.85)	0.01 (0.54)	-0.08*** (-3.97)	0.02 (0.91)	0.02 (1.54)	0.06 (1.53)
Income $25k-$35k	0.09** (2.81)	-0.04* (-2.20)	-0.08*** (-4.03)	0.02 (1.13)	0.06*** (4.05)	0.14*** (3.62)
Income $35k-$50k	0.13*** (4.10)	-0.03 (-1.74)	-0.08*** (-4.22)	0.03 (1.51)	0.06*** (4.10)	0.07* (2.04)
Income $50k-$75k	0.09*** (3.59)	-0.11*** (-6.70)	-0.06*** (-3.66)	0.07*** (4.11)	0.10*** (8.23)	0.22*** (6.70)
Income $75k-$100k	0.07* (2.46)	-0.15*** (-8.39)	-0.11*** (-6.54)	0.12*** (6.29)	0.11*** (7.95)	0.26*** (7.16)
Income $100k+	0.05 (1.77)	-0.26*** (-15.37)	-0.08*** (-4.96)	0.15*** (9.07)	0.17*** (13.64)	0.34*** (10.09)
Constant	-2.79*** (-133.08)	-1.09*** (-88.27)	-1.55*** (-128.68)	-1.30*** (-97.25)	-1.00*** (-102.92)	-2.92*** (-106.69)
	2008	**2013**	**2008**	**2013**	**2008**	**2013**
	Online Banking	***Online Banking***	***Porn***	***Porn***	***Shopping***	***Shopping***
Income $15k-$25k	0.027 (0.89)	-0.01 (-0.23)	-0.08*** (-3.36)	0.01 (0.73)	0.03 (1.89)	-0.02 (-1.19)
Income $25k-$35k	0.12*** (4.20)	-0.04 (-0.75)	-0.04* (-2.13)	-0.06*** (-3.55)	0.07*** (4.50)	0.00 (0.00)
Income $35k-$50k	0.11*** (3.84)	-0.07 (-1.59)	-0.14*** (-6.08)	-0.09*** (-5.97)	0.10*** (6.57)	-0.05** (-3.07)
Income $50k-$75k	0.21*** (9.16)	0.03 (0.81)	0.02 (0.86)	-0.09*** (-6.28)	0.15*** (11.74)	-0.00 (-0.10)
Income $75k-$100k	0.23*** (9.36)	0.06 (1.38)	0.02 (1.17)	-0.09*** (-5.44)	0.12*** (9.09)	0.04** (2.65)
Income $100k+	0.16***	0.20***	0.01	-0.14***	0.16***	0.03*

	(6.65)	(4.92)	(0.33)	(-9.18)	(12.29)	(2.12)
Constant	-2.52*** (-136.40)	-3.24*** (-101.58)	-1.99*** (-136.85)	-0.84*** (-74.23)	-1.08*** (-107.70)	-0.85*** (-75.26)
	2008	**2013**	**2008**	**2013**	**2008**	**2013**
	Social Media	***Social Media***	***Sports***	***Sports***	***Video***	***Video***
Income $15k-$25k	-0.03 (-1.52)	-0.01 (-0.61)	0.11** (2.86)	-0.04 (-1.13)	0.02 (0.75)	-0.01 (-0.34)
Income $25k-$35k	-0.04* (-1.98)	-0.02 (-1.24)	0.16*** (4.45)	-0.04 (-1.08)	0.07*** (3.38)	-0.01 (-0.32)
Income $35k-$50k	-0.07*** (-3.45)	-0.05*** (-3.41)	0.17*** (4.89)	-0.03 (-0.81)	0.06** (3.10)	0.00 (0.01)
Income $50k-$75k	-0.05** (-2.93)	-0.03 (-1.80)	0.25*** (8.75)	-0.00 (-0.05)	0.09*** (6.02)	0.00 (0.28)
Income $75k-$100k	-0.06*** (-3.39)	-0.03 (-1.55)	0.26*** (8.54)	0.08* (2.47)	0.11*** (6.76)	0.02 (1.31)
Income $100k+	-0.07*** (-3.89)	-0.03 (-1.73)	0.33*** (11.53)	0.11*** (3.61)	0.15*** (9.22)	0.03* (2.08)
Constant	-1.75*** (-133.57)	-0.95*** (-81.23)	-3.03*** (-129.04)	-2.7*** (-109.16)	-1.62*** (-130.65)	-1.24*** (-95.00)
N	819,753	711,593	819,753	711,593	819,753	711,593
	Dependent Variable	Dependent Variable	Dependent Variable	Dependent Variable	Dependent Variable	Dependent Variable
Covariates	**2008**	**2013**	**2008**	**2013**	**2008**	**2013**
	Chat	***Chat***	***Credit History***	***Credit History***	***Education***	***Education***
Income $15k-$25k	-0.06** (-2.62)	-0.21* (-2.28)	0.11 (0.82)	0.05 (1.70)	-0.06 (-1.62)	0.02 (0.27)
Income $25k-$35k	-0.01 (-0.24)	-0.08 (-0.95)	0.24* (2.01)	0.03 (0.90)	0.03 (0.99)	0.21** (3.26)
Income $35k-$50k	-0.01 (-0.30)	-0.17* (-2.01)	0.34** (3.03)	-0.00 (-0.12)	0.03 (0.83)	0.09 (1.51)
Income $50k-$75k	0.03 (1.73)	-0.05 (-0.73)	0.28** (2.84)	0.05 (1.79)	0.06* (2.42)	0.22*** (3.89)
Income $75k-$100k	0.05 (2.85)	-0.11 (-1.29)	0.26* (2.46)	0.09** (3.08)	0.07* (2.43)	0.38*** (6.34)
Income $100k+	0.12*** (6.36)	-0.11 (-1.47)	0.012 (0.11)	0.26*** (9.87)	0.16*** (6.09)	0.43*** (7.79)
Constant	-1.92*** (-136.18)	-4.46*** (-76.96)	-5.53*** (-68.94)	-2.40*** (-111.76)	-2.76*** (-133.59)	-3.99*** (-86.94)
	2008	**2013**	**2008**	**2013**	**2008**	**2013**
	Financial Services	***Financial Services***	***Games***	***Games***	***News***	***News***
Income $15k-$25k	0.06 (1.85)	0.01 (0.54)	-0.08*** (-3.97)	0.02 (0.91)	0.02 (1.54)	0.06 (1.53)
Income $25k-$35k	0.09** (2.81)	-0.04* (-2.20)	-0.08*** (-4.03)	0.02 (1.13)	0.06*** (4.05)	0.14*** (3.62)
Income $35k-$50k	0.13*** (4.10)	-0.03 (-1.74)	-0.08*** (-4.22)	0.03 (1.51)	0.06*** (4.10)	0.07* (2.04)

Income \$50k-\$75k	0.09*** (3.59)	-0.11*** (-6.70)	-0.06*** (-3.66)	0.07*** (4.11)	0.10*** (8.23)	0.22*** (6.70)
Income \$75k-\$100k	0.07* (2.46)	-0.15*** (-8.39)	-0.11*** (-6.54)	0.12*** (6.29)	0.11*** (7.95)	0.26*** (7.16)
Income \$100k+	0.05 (1.77)	-0.26*** (-15.37)	-0.08*** (-4.96)	0.15*** (9.07)	0.17*** (13.64)	0.34*** (10.09)
Constant	-2.79*** (-133.08)	-1.09*** (-88.27)	-1.55*** (-128.68)	-1.30*** (-97.25)	-1.00*** (-102.92)	-2.92*** (-106.69)
	2008	**2013**	**2008**	**2013**	**2008**	**2013**
	Online Banking	***Online Banking***	***Porn***	***Porn***	***Shopping***	***Shopping***
Income \$15k-\$25k	0.027 (0.89)	-0.01 (-0.23)	-0.08*** (-3.36)	0.01 (0.73)	0.03 (1.89)	-0.02 (-1.19)
Income \$25k-\$35k	0.12*** (4.20)	-0.04 (-0.75)	-0.04* (-2.13)	-0.06*** (-3.55)	0.07*** (4.50)	0.00 (0.00)
Income \$35k-\$50k	0.11*** (3.84)	-0.07 (-1.59)	-0.14*** (-6.08)	-0.09*** (-5.97)	0.10*** (6.57)	-0.05** (-3.07)
Income \$50k-\$75k	0.21*** (9.16)	0.03 (0.81)	0.02 (0.86)	-0.09*** (-6.28)	0.15*** (11.74)	-0.00 (-0.10)
Income \$75k-\$100k	0.23*** (9.36)	0.06 (1.38)	0.02 (1.17)	-0.09*** (-5.44)	0.12*** (9.09)	0.04** (2.65)
Income \$100k+	0.16*** (6.65)	0.20*** (4.92)	0.01 (0.33)	-0.14*** (-9.18)	0.16*** (12.29)	0.03* (2.12)
Constant	-2.52*** (-136.40)	-3.24*** (-101.58)	-1.99*** (-136.85)	-0.84*** (-74.23)	-1.08*** (-107.70)	-0.85*** (-75.26)
	2008	**2013**	**2008**	**2013**	**2008**	**2013**
	Social Media	***Social Media***	***Sports***	***Sports***	***Video***	***Video***
Income \$15k-\$25k	-0.03 (-1.52)	-0.01 (-0.61)	0.11** (2.86)	-0.04 (-1.13)	0.02 (0.75)	-0.01 (-0.34)
Income \$25k-\$35k	-0.04* (-1.98)	-0.02 (-1.24)	0.16*** (4.45)	-0.04 (-1.08)	0.07*** (3.38)	-0.01 (-0.32)
Income \$35k-\$50k	-0.07*** (-3.45)	-0.05*** (-3.41)	0.17*** (4.89)	-0.03 (-0.81)	0.06** (3.10)	0.00 (0.01)
Income \$50k-\$75k	-0.05** (-2.93)	-0.03 (-1.80)	0.25*** (8.75)	-0.00 (-0.05)	0.09*** (6.02)	0.00 (0.28)
Income \$75k-\$100k	-0.06*** (-3.39)	-0.03 (-1.55)	0.26*** (8.54)	0.08* (2.47)	0.11*** (6.76)	0.02 (1.31)
Income \$100k+	-0.07*** (-3.89)	-0.03 (-1.73)	0.33*** (11.53)	0.11*** (3.61)	0.15*** (9.22)	0.03* (2.08)
Constant	-1.75*** (-133.57)	-0.95*** (-81.23)	-3.03*** (-129.04)	-2.7*** (-109.16)	-1.62*** (-130.65)	-1.24*** (-95.00)
N	819,753	711,593	819,753	711,593	819,753	711,593

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Numeraire category: Web Services
Std errors clustered at machine-week level.

Table 10
Hypotheses and Findings

Hypothesis	Description	Finding	Source
H1. $TO_x(S_t, X_{it}) < 0$.	Total time declines with income	Confirmed.	Table 4 Figure 1
H2. $TO(S_t, X_{it}) - TO(S_{t-1}, X_{it-1}) = 0$.	Total time does not change over time with new supply.	Total time slightly declines.	Table 3
H3. $TO_x(S_t, X_{it}) – TO_x(S_{t-1}, X_{it-1}) = 0$.	The relationship between income and total time does not change with new supply.	Very little change in relationship.	Figure 1
H4. $C_x(S_t, X_{it}) = 0$ and $L_x(S_t, X_{it}) = 0$,	Breadth/depth does not vary/decline with income.	Breadth/depth do not vary with income.	Table 5
H5. $C(S_t, X_{it}) - C(S_{t-1}, X_{it-1}) = 0$, and $L(S_t, X_{it}) - L(S_{t-1}, X_{it-1}) = 0$.	Breadth/depth does not change with new supply.	Breadth/depth does not vary meaningfully with new supply.	Figure 2