The Empirical Economics of Online Attention

Andre Boik, Shane Greenstein, and Jeffrey Prince*

December 2015

Preliminary. Not for quotation without Permission.

Abstract

We model and measure how households allocate online attention, and assess if and how online attention changed between 2008 and 2013, a time of large increases in online offerings, e.g., video and access points. We calculate our measures using click-stream data for thousands of U.S. households. We find that general measures of breadth and depth of online attention are remarkably stable over this period, while shares of domain categories markedly change – with video and social media expanding, and chat and news contracting. This frames a surprising contrast: The level of total time online varies with income at any point in time, but this relationship is impervious to changes in the menu of available web sites. A key implication is that increasingly valuable offerings change where households go online, but not their general (i.e., breadth/depth) online attention patterns.

* University of California, Davis, Department of Economics; Harvard Business School, Department of Technology and Operations Management; and Indiana University, Department of Business Economics and Public Policy, Kelley School of Business. We thank Kelly School of Business and the Harvard Business School for funding. We thank seminar audiences at Georgetown, Harvard and Northwestern for comments and suggestions. Philip Marx for excellent research assistance. We are responsible for all errors.

1. Introduction

"…in 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).

When originally crafted, the above quote was directed toward firms' information systems. Despite predating it by over two decades, this insight can easily be applied to the commercial Internet as well – the Internet contains vast amounts of information and users' attention is the scarce resource that must be allocated. However, unlike a typical firm's in-house information system, the information on the Internet is produced by many different firms, and allocation of consumers' attention is the result of firm competition for that attention.

The means of competition for attention on the Internet differs from most other markets, as users pay for online access and unless they face a binding cap on usage, no price shapes any other marginal decision. Present evidence suggests only a small fraction of users face such constraints across the majority of their surfing (Nevo, Turner, Williams, 2015), which means the vast majority of Internet web sites typically do not compete on price. In fact, only one of the top twenty domains (Netflix) is a subscription service, i.e., where the price of allocating time is explicit. There is now plenty of evidence demonstrating the value of speed and user demand for broadband access, which, presumably, users spread over a vast array of content (Rosston, Savage, and Waldman, 2010). As of this writing, however, economists generally have not focused on priceless online competition except in the theoretical literature on competition for attention (Athey, Calvano and Gans, 2013). There has been almost no empirical work on that competition except in the context of conflicts between news aggregators and news sites (Chiou and Tucker, 2015). Hence, little is known about the "what, where and how" behind the core economics of competition for online attention. In particular, what do users do with that bandwidth, where do they allocate their time, and how does user behavior shape online competition for attention?

Despite the absence of prices at users' marginal decisions, competition among Internet domains has much in common with other competitive settings. Users make choices about where to allocate their time, and the time has value. In some cases (e.g., electronic commerce), the firms try to convert that attention into sales. In other cases (e.g., most media), firms try to convert that attention into advertising sales. Firms compete for users by investing in web page quality and other aspects of their business related to the services displayed on the pages. Over time, new firms enter with new offerings, and users can respond by making new choices. That motivates our research: At the heart of this competitive environment is the allocation of attention. This is the aspect that has not received much empirical attention, and where economists need to focus analysis. In this paper we focus our analysis on the consumer side of the empirical foundations of the market for online attention.

Our specific research question is: How did a large increase in the availability of online offerings alter the volume and allocation of online attention? While the question examines a specific time period – US households across the years 2008 and 2013 – the question is quite general. Internet has experienced increases in online offerings throughout its existence, so the analysis potentially has very general implications. We focus on measuring and analyzing one other rather general factor, households' online attention, and how it changed. The specific setting is rich with factors that should change household allocations. This time period saw a massive expansion in online video offerings, social media, and points of contact (e.g., tablets, smartphones), among other changes.

To conduct our analysis, we start by fixing ideas with a simple economic model of online time allocation for a home computer, based on the seminal work of Becker, 1965. Using this model, we highlight theoretical ambiguities as to predicted changes in online attention with increased online offerings. We then characterize three basic types of online attention measurements, designed to answer the questions: How much? How is it allocated? and Where is it allocated? In doing so, we create novel measures of online attention allocation designed to capture the breadth and depth of a household's online attention. We then take these measurement constructs to our data.

For our empirics, we utilize ComScore data on the web browsing behavior of thousands of households in the United States. These data track households over an entire year, recording all of their web browsing behavior on the household's home computer, as well as some key demographics. Using these data, we measure changes between 2008 and 2013 in total time online, online attention patterns, and the types of domains visited. To measure online attention patterns, we calculate the weekly market concentration (in terms of time) for domains (our measure of breadth, or "focus") and the weekly fraction of domain visits that lasted at least 10 minutes (our measure of depth, or "dwelling").

This frames a surprising contrast: The level of total time online varies with income at any point in time, but this relationship is impervious to changes in the menu of available web sites. We see this in several findings. First, total time online on the household's home computer has declined, but only modestly, and the decline is generally consistent across income groups. Second, the way in which households allocate their online attention, as measured by the concentration of domains visited (focus) and time spent in "long" sessions (dwelling), 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.

A key implication is that increasingly valuable offerings change where households go online, but not their general (i.e., breadth/depth) online attention patterns. With regard to total time, it appears that new points of contact only modestly substitute time away from the primary home device. Hence, although we are limited by only observing the home computer, our results suggest that any new value stemming from additional total time online (across all devices) appears to be largely coming from time on new, alternative devices.

These insights suggest that our attention metrics (focus, dwelling) represent a stable, underlying feature of households' optimal allocation of attention, given their apparent immunity to major changes in online offerings. In addition, the fact that these measures are poorly predicted by changes in demographics suggests that these online attention patterns are indicative of fundamental household behavioral characteristics outside what demographics can capture. Lastly, and more specifically to this time period, our findings concerning online category shares are consistent with social media, and possibly video, becoming a substitute for both chat and news.

We are able to say much more with regard to competition for the reallocation of online attention. Our results imply that reallocation does not take the form of changes in concentration of domain visits or proportion of long/intense visits. Instead, reallocation of online attention came almost entirely in the form of changes in how that concentration/intensity portfolio is filled. This implies that at any point in time there are a fixed set of "slots" of attention to allocate. Firm entry and exit does not alter the total number of slots open for competition. Rather, firms compete for given slots of time from users.

We note that our notion of value contrasts sharply with prior work on the value of household time online. Prior work has characterized the value of online attention in terms of its consumer surplus or the opportunity cost of work time (Goolsbee and Klenow, 2006, Brynjolfsson and Oh, 2012). It has also considered how users trade-off between online and offline leisure time, recognizing the user pays an opportunity cost of online time by withdrawing from other leisure activity (Wallsten, 2015). In contrast, we focus on the value generated by users' allocation of attention to the suppliers of online web sites, and focus on competition for that value. That focus motivates our focus on the allocation of user time, and leads to a very different analysis of the core economics than prior research.

We also contrast with 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 domains. Much of the prior work pertaining to online advertising has focused on #2, namely, the principals of targeting ads. This is largely driven by firms tapping into "big data" and extensive information about users' private lives. The marketing literature on targeting tends not to focus on behavioral changes by consumers as supply changes. In contrast, our analysis focuses on the core economic 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 very different conceptualization about the nature of competition for attention.

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 increases.

The start of our time period is near the end of the first diffusion of broadband networks. By 2007 close to 62 million US households had adopted broadband access for their household Internet needs, while by 2013 the numbers were 73 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 of 2007, and it is widely credited with catalyzing entry of Android-based phones the following year, and 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 of 2006. Social media was also quite young. For example, Twitter entered in March, 2006, while Facebook starts in February, 2004, and starts allowing widespread use in September, 2006. By 2013 social media had become a mainstream online application, and, as our data will show, was widely used.

3. A Model of Online Attention

In this section, we present a simple model of online attention, followed by summary measures one can use to characterize how it is allocated. We use the insights from the model along with the summary measures we construct in the empirical analyses that follow.

3.1. The Model

Our model of online attention follows the basic structure of the seminal work by Becker (1965). However, given household visits to online domains generally do not carry a price, our focus will be entirely on time decisions across available domains for a given device. In particular, household *i* chooses the amount of time to spend at each Internet domain (t_{ij}) on its "home device" to maximize its utility. Hence, each household solves:

$$(1)\quad max_{t_{i1},\ldots,t_{iJ}} U\left(t_{i1},\ldots,t_{iJ},T_i-\left(t_{i1}+\cdots+t_{iJ}\right);\vec{W}\right)$$

s.t. $t_{i1}\geq 0,\ldots,t_{iJ}\geq 0, T_i\geq (t_{i1}+\cdots+t_{iJ})$

In equation (1), $\overrightarrow{W}$ represents all relevant features (i.e., content, subscription fee – if any, etc.) for the available web domains. Further, T_i represents all time available to household i, 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.

The necessary conditions in solving (1) are:

(2) $U_1 - U_{J+1} \leq 0, t_{i1} \geq 0, (U_1 - U_{J+1}) * \ t_{i1} = 0$

…

$U_J - U_{J+1} \leq 0, t_{iJ} \geq 0, (U_J - U_{J+1}) * \ t_{iJ} = 0$

Notably, there are two key departures in our model from Becker (1965). First, ours is a partial analysis, focusing on time spent online using a home device. Second, we consider shocks to the choice set and their potential effects, discussed in the next subsection.

3.2. Effects of Two Model Shocks

Over the time period of our data, two important shocks to our model of online attention occurred. First, a wave of new domains entered the worldwide web, and many of these new domains 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 there certainly were domain exits during the time we analyze, the net change in domains was certainly positive, with a notable increase in online video available. Within our model, this influx of domains manifests as an increase in J to J* and a change in the full list of domains – and their characteristics – comprising the J* total domains.

The second shock to our model was due to the release of a new batch of connected devices – in particular, tablets and smartphones. Given our model is for the home device, this shock essentially altered the composition of the composite good within the model. Now, time spent not online on the home device could include time spent online on a tablet or smartphone.

We adapt the previously-specified utility model to account for an increase in the available domains to J* and the possibility of time spent at some domain *j* through alternative devices, denoted t_{ij}^{dev}. Each household now solves:

$$(3)\quad max_{t_{i1},\ldots,t_{iJ^*},t_{i1}^{dev},\ldots,t_{iJ^*}^{dev}} U\left(t_{i1},\ldots,t_{iJ^*},t_{i1}^{dev},\ldots,t_{iJ^*}^{dev},T_i-\left(t_{i1}+\cdots+t_{iJ^*}^{dev}\right);\vec{W}\right)$$

subject to the usual time constraints. The necessary conditions in solving equation (3) are:

$$(4)\quad U_1-U_{J^*+1}\leq 0, t_{i1}\geq 0, (U_1-U_{J^*+1}) *\ t_{i1}=0$$

…

$$U_{J^*}-U_{J^*+1}\leq 0, t_{iJ^*}^{dev}\geq 0, (U_{J^*}-U_{J^*+1}) * t_{iJ^*}^{dev}\ =0$$

These necessary conditions implicitly define a household-level "attention allocation function" t_{ij}^{*}. From the perspective of households, the value created by additional domains and ability to browse from additional devices can manifest itself in several possible re-allocations of time. The first is a re-allocation between online and offline activities, and the second is a re-allocation of time across domains conditional on some total amount of time online. Assuming no change in either the underlying utility function across time or in the utility from the outside option, one feasible and natural outcome is that the new domains act as imperfect substitutes to existing domains, leading to an increase in total time online but a decrease in time allocated to "old" domains. If "new" domains are instead perfect substitutes to old domains, total time online remains unchanged but there is a complete shift of attention from old content to new content. Note that in terms of the utility model, allocation of time to old or new domains via secondary devices can essentially be treated as new domains, and so the intuition still applies: we may see either no change or a decrease in attention allocated to old domains.

To summarize, in our empirical work, we expect changes in online activity, with superior (and younger) web sites replacing older choices. The net effect of new websites and new devices could increase or decrease the amount of time users spend online on older devices, such as a PC.

3.3 Characterizing Online Attention

The vectors of times at each domain within our model are a conceptual construct; solving for these vectors and matching them to the data would be a highly unwieldy task, with limited promise for pointed insights. Consequently, in this subsection we lay out three basic types of online attention measurements, and detail the components of each within the context of our model. These measurement lend themselves to hypothesis testing. Specifically, these allow us to measure online attention in terms of three questions: How much? How is it allocated? and Where is it allocated?

We begin with "How much?" Using our framework, measuring how much time a household spends online for the home device over a given period (e.g., a week) is straightforward. The model produces the following identity for time online for household i (TO_i) when there are J domains:

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

Our next type of online attention measurements seek to answer the question "How is it allocated?" Answering this question is less straightforward – one could conceivably construct many different measures of online attention allocation that are informative. For example, we could measure the total number of unique domains visited, or the average time spent at each domain. We contend that there are two natural dimensions to consider when measuring how attention is allocated online – breadth and depth. That is, how is attention allocated across domains, and how intensely is it allocated within a domain?

Our measure of breadth stems from the classic literature in industrial organization. Specifically, we measure breadth using a Herfindahl-Herschman index for time spent at domains visited by household i, denoted C_i. We define C_i as:

$$(6)\ \ C_i = 10{,}000 * \sum_j \left(\frac{t_{ij}^*}{TO_i}\right)^2$$

Defined this way, our measure of breadth captures the level of concentration (in terms of time at domains) household i exhibits in its domain visits. A high value for C_i indicates a breadth of visits that is highly concentrated at a small number of domains, whereas a low value for C_i indicates a breadth of visits that is unconcentrated, i.e., spread out across relatively many domains.

Our measure of depth takes inspiration from an early constraint on YouTube, specifically the cap on video length of 10 minutes, which lasted until mid-2010. We measure depth as the fraction of domain visits by household i that lasted at least 10 minutes, denoted L_i. To calculate L_i, we must decompose the optimal time spent at each domain during the given time period (e.g., a week). To see this, suppose $t_{i1}^* = 30$. Hence, time spent at domain #1 during the observed week was 30 minutes. However, this measurement does not distinguish between the 30 minutes being comprised of 6 separate visits lasting 5 minutes each and one visit lasting 30 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 at domain j for household i. Hence, the length of $\overrightarrow{S_{ij}}$ is the number of separate visits made by household i to domain j. Next, let t_{ijk}^* be the optimal time spent by household i at domain 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:

$$(7)\ \ L_i = \frac{\sum_j \sum_k 1(t_{ijk}^* > 10)}{\sum_j \sum_k 1(t_{ijk}^* > 0)}$$

As defined, L_i is the proportion of total domain visits that lasted more than ten minutes for household i.

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 domains, 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 households propensity to dwell at the domains it visits – households with a high value for L_i tend to dwell at domains 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.

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

Now that we have detailed our measures of online attention in terms of "how much?" and "how is it allocated?," we turn to our last measures for "where is it allocated?" For this measure, we calculate shares of total time online on the home device for different domain 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 domains in category c. Formally, we have:

$$(8)\ TS_c = \frac{\sum_i \sum_{j \in c} t_{ij}^*}{\sum_i TO_i}$$

This measure also suggests one approach to measuring the extent of competition. We expect new entry to lead to turnover when users direct their attention to new categories of web sites. One measure of competition is the fraction of attention that moves to these new categories.

In what follows, we take 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 5 through 8 – to the data. By doing so, we can observe if and how they have changed during the span of time our data cover.

3.4. Hypothesis development

Hypotheses need to 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: $TO_{it} = \mathrm{TO}(S_t, X_{it})$

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

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

What are the properties of this allocation? Goldfarb and Prince (2008) have shown that the internet violates the standard intuition when it comes to income and adoption and use. In Goldfarb and Prince (2008), those with high income are more likely to adopt but they do not use the Internet as intensively due to the outside option value of their leisure time. In this setting, if X_{it} is income, the Goldfarb-Prince effect would appear as:

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

We seek to learn whether this income effect holds in our measures of the attention economy, and on a very different data set than previously used. A further question is whether the relationship in H1 has changed over time. That is, has the improvement in devices attracted user attention away from the improving web sites on PCs, or vice versa? 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 could be either higher or lower. If we reject H2, then an interesting question focuses on whether the income effect has changed over time. That is, despite changes 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. $TO_x(S_t, X_{it}) – TO_x(S_{t-1}, X_{it-1}) = 0$.

H1, H2, and H3 test hypotheses about the extensive margin. It is possible to ask a similar question about the determinants of the intensive margin. In particular, does greater online time lead to greater breadth and depth? If so, then – once again, assuming X is income – we would expect larger X to lead to 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 importance of supply conditions. 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? Once again, 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.

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 home computer. The information collected includes the domains visited on the machine, how much time was spent at each domain, and the number of pages visited within the domain. 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. Importantly, we delete households that have fewer than 6 months of at least 5 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.

Summary statistics of our demographic measures are presented in Table 2. These demographics include household income thresholds, 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 be targeted more towards higher income households, but 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 younger. In addition, Comscore provides no information on the speed of the broadband connection.

[Table 2 about here]

Summary statistics of our key variables representing browsing types such as the concentration of time across domains and the fraction of sessions that exceed 10 minutes are presented in Table 3. On average a household spends roughly 15 hours online per week in 2008 and 14 hours online in 2013. Perhaps surprisingly, our measures of browsing behavior are virtually identical across years, with 75% of sessions lasting over 10 minutes and households' allocation of time across domains being quite concentrated with an HHI of approximately 2,900. We discuss these similarities in greater detail in the next two sections.

[Table 3 about here]

5. Empirical Analysis

We take our utility framework and measures to characterize online attention to the 2008 and 2013 data. Households optimally allocate time across online domains and offline activities. This allocation maps to our data in terms of a total amount of time online, and a joint distribution of how that time is distributed across: number of sessions, unique domain visits, and time per session. As discussed in Section 2, to capture heterogeneity in online time allocation across households conditional on their time online, we generate intuitive measures of fundamental browsing behavior conditional on an amount of time online: focus (a measure of time concentration over domains) and propensity to dwell (a measure of time spent at a given domain).

In this section, we present three types of results that shed light on three corresponding basic questions pertaining to online attention: How much? How? 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 make comparisons across 2008 and 2013, and discuss key insights from these comparisons in Section 6.

5.1 Total Time Online

Our data do allow us to conduct measurements and analyses that are informative about households' total time online and how it has changed over the tumultuous period between 2008 and 2013. Since our data are at the home device level, we are limited in our ability to draw conclusions about the total time spent online by a household (across all devices). We only observe time spent on the PC.

First, our summary statistics show that the average household spends approximately 2 hours per day on the Internet. Our theory predicted that time on the PC could go up or down over time. We see, in fact, that total time online on the primary home device declined by approximately 5% between 2008 and 2013 (rejecting 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. As later data will show a large increase in the menu of choices on the intensive margin, we stress that users devoted attention to other devices *in spite of improvement in the quality* of sites online from the PC.

Our data also allow us to examine how total time online on the home device relates to demographics, 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. However, the Internet seems to be different because it generally consumes leisure time and not money. Most standard models of adoption predict that the extent of *use* of Internet technology is increasing in the same factors that predict adoption.

In this data we see a Goldfarb-Prince effect in any given year. This is shown in Table 4 and Figure 1, and, confirming H1, total time online declines with income. Hence, we find that the determinants of total time online for the home device, particularly income, are consistent with those previously identified in the literature. We present the results of a simple regression of time online per week on demographics (Table 4). We illustrate this relationship further in Figure 1, which illustrates that we cannot reject the null of H3. 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. Other demographic determinants of time online also line up with the existing literature, as follows.

[Table 4 about here]

[Figure 1 about here]

Our findings and data relate the Goldfarb-Prince effect to its underlying determinants. We see that the relationship between total time and income remains stable across time. On first blush, this suggests the effect is a function of user allocation of attention. More to the point, for a given amount of time a household spends online, its allocation of that time (i.e. concentration and length) does not depend on income. However, at the same time, total time is declining in income and in a similar way across time periods despite a change in supply. Hence, the Goldfarb-Prince effect appears to be a stable relationship of the extensive margin of a time-allocation problem at the household level.

5.2. Online Attention Allocation Patterns

In this subsection, we present analyses of our basic measures of households' allocation of online attention: the Herfindahl-Herschman index for time spent at domains visited (our measure of focus) and the proportion of total sessions that lasted more than ten minutes (our measure of propensity to dwell). Figure 2 presents the unconditional joint density of our measures of focus and propensity to dwell for 2008 and 2013. Here, we see a very well-behaved joint distribution that strongly resembles a joint normal. However, it is the comparison of the graphs that generates a particularly striking finding – the distribution of these measures of online attention allocation is essentially unchanged during this time period! The summary statistics in Section 3 showed that the means of each measure were very similar, but Figure 2 clearly indicates that the similarity goes well beyond just the means – the entire distributions are nearly identical. We cannot reject the null for H5.

[Figure 2 about here]

A possible concern with the measures of online attention allocation we have created is that they 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 be more likely to dwell and perhaps be less likely to be focused. In short, it could be the case that total time online strongly predicts a household's location within the distribution presented in Figure 1, limiting the informativeness of the measures in the figure. To address this possibility, we break total time online on the home device into quartiles, and recreate our joint distribution for each quartile. The results are in Figure 3. Here we see that, while not 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. This suggests we cannot reject the null for H4, though more investigation is needed.

[Figure 3 about here]

As shown in our summary statistics in Section 3, there are some differences in the demographic profiles between our sample in 2008 and 2013. Consequently, it could be that online attention allocation patterns, conditional on demographics, did change over this time period, but the changes are offset by the demographic changes in our samples. 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.

Table 5 presents a set of seemingly-unrelated-regressions (SURs) for our measures of focus and propensity to dwell. Here we see that both measures are virtually independent of income levels after controlling for total time online. The only demographics that meaningfully correlate with focus are lower levels of education and older heads of households. In addition, we see that households' propensity to dwell is remarkably independent of demographics.

[Table 5 about here]

Broadly speaking, the percentage of variation in our household classifications explained by demographics is less than 20% (for dwellers) and less than 3% (for whether the household is focused). Households that are larger, have more education and income are less likely to be classified as dwellers, but the economic significance of these effects are 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 again modest.

These are quite striking findings about the role of demographics in breadth and depth in light of our earlier results about total time. Demographics help shape total time online far more than its composition. From the previous subsection, it appears that little has changed with regard to *how* households allocate their online attention, at least on their primary home device.

5.3. Online Attention Category Shares

As noted above, the period spanning 2008 to 2013 saw large changes in the supply of website domains, 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 domains from both 2008 to 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 4. 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. Attention allocated to News domains also sees a decrease, from roughly 10% down to 5%. We observe the largest increases of attention being allocated towards Social Media and Video, 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.

[Figure 4 about here]

Table 6 contains the top 20 domains of 2008 and 2013. A quick glance at these rankings and the change between 2008 and 2013 further confirms what we see in Figure 4. Particularly noteworthy is the mass exodus of chat and the rise in video.

[Table 6 about here]

6. Discussion

In this section, we summarize our findings in Section 5 and discuss their main implications, focusing on notable changes between 2008 and 2013 – a time of substantial change in the Internet ecosystem. We summarize our findings in Table 7, and state the results as follows. First, total time online at the primary home device has declined, but only modestly, and the decline is generally consistent across income groups. Second, the way in which households allocate their online attention, as measured by the concentration of domains visited (focus) and time spent in "long" sessions (dwelling), 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.

Altogether, this adds up to a surprising characterization of the attention economy. User allocation of a given amount of time online varies with supply conditions and not income, while the amount of time spent online varies with income and not the menu of supply.

Taken together, our findings have several important implications concerning how households adjusted their online attention to gain value from changes in the Internet ecosystem between 2008 and 2013. First, our findings concerning total time online for the primary home device suggest that new points of contact – in the form of additional computers, tablets and smartphones – are substituting time away 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 coming on top of a relatively stable home device. Hence, any new value stemming from additional time online appears to be largely coming from time on new, alternative devices.

Our findings regarding how households allocate their time online (i.e., focus and dwelling) suggest that these attention metrics represent a stable, underlying feature of households' optimal allocation of attention. In addition, the fact that these measures are poorly predicted by changes in demographics suggests that these online attention patterns are indicative of fundamental household behavioral characteristics outside what demographics can capture. For example, these results indicate that the distribution of online attention patterns is almost identical for the wealthiest and poorest households.

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 achieved resulting 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 domains with video offerings, which tend to demand more dwelling, it appears these shifts are at the expense of attention at other domains at which the household was already dwelling. Second, this result suggests that households' online attention via secondary devices has not been such that it alters 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 domains and content offerings, households are not increasing the spread of their attention in response, at least at the device level.

Our last set of findings, regarding where households allocate their time, indicate that households likely achieved additional value between 2008 and 2013 – beyond simple increases in total time online – by reallocating their attention across domains. That is, households changed where they allocated their time online in terms of the types of domains they visited. The changes in category shares are consistent with social media, and possibly video, becoming a substitute for both chat and news. It is important to note, however, that these figures correspond to attention allocation through the household's home computer. The time period of 2008 to 2013 also saw a dramatic increase in the use of handheld devices capable of browsing the Internet; some of the changes in attention allocation presented in Figure 4 may also represent substitution to handheld devices. The category of chat, for example, has moved away from instant messenger software on the home computer towards text messaging software on devices.

7. Conclusions

In sum, any changes in value households achieved in response to changes in the Internet ecosystem between 2008 and 2013 appear to have come via changes in "how much" and "where" but not "how."

References

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/

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

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.

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.

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

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

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.

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.

Figures

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

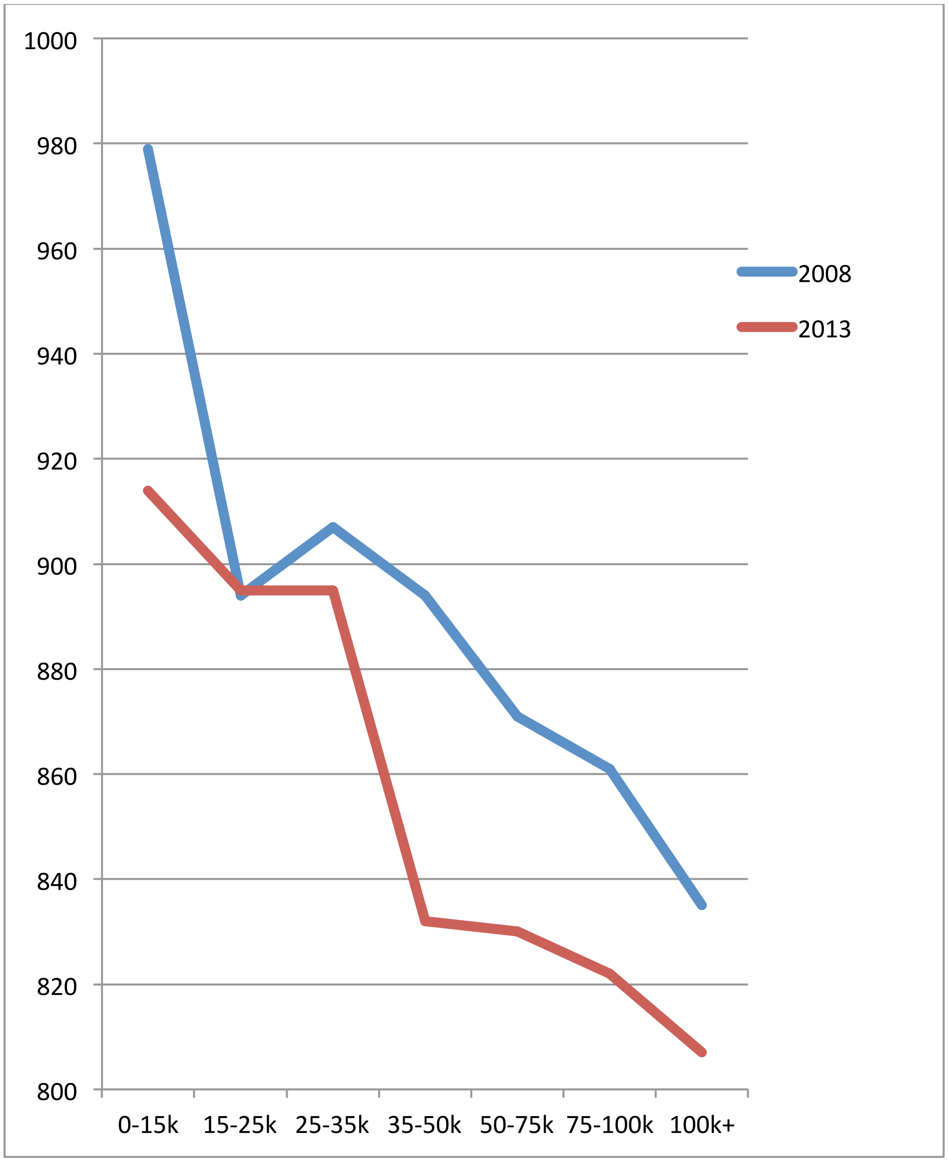


Figure 2

Unconditional Distribution of Online Attention (2008 vs. 2013)

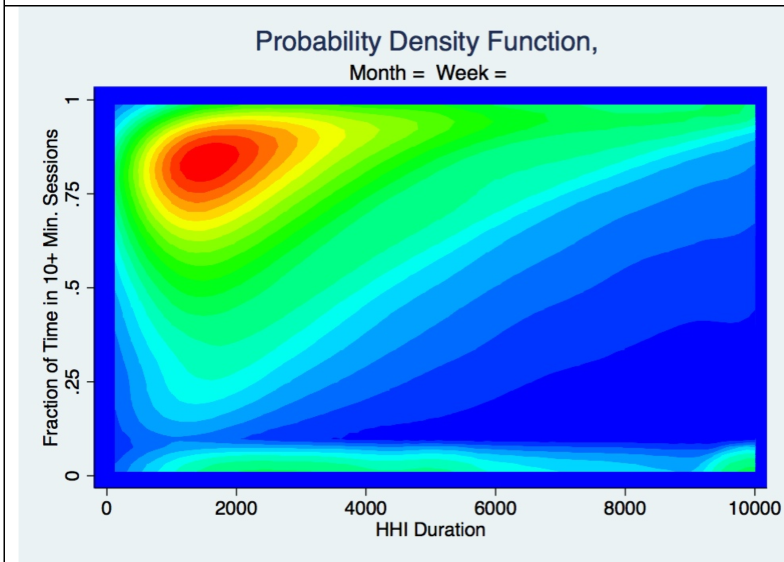
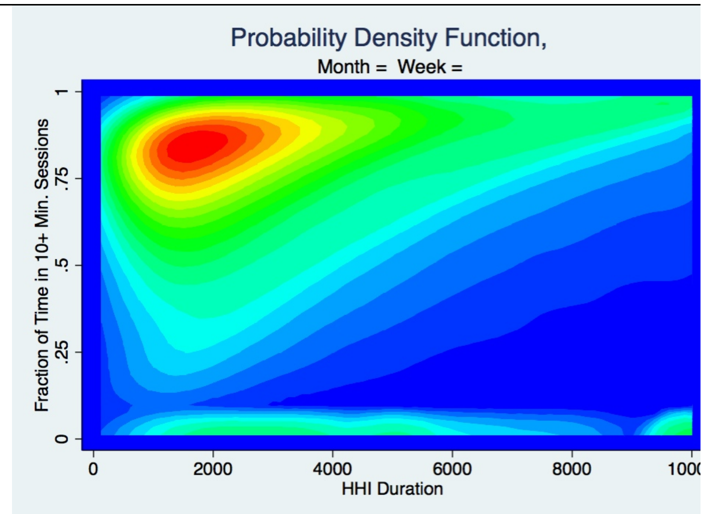
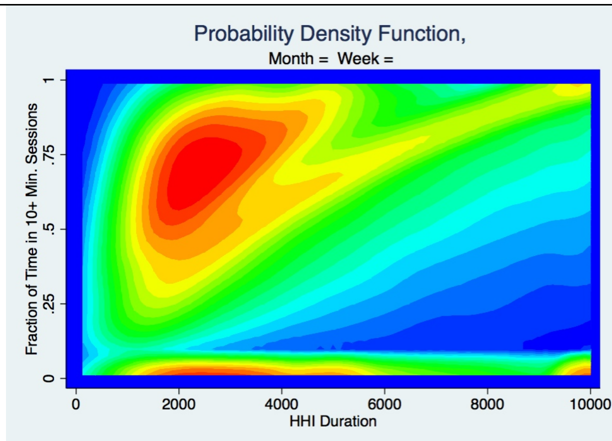
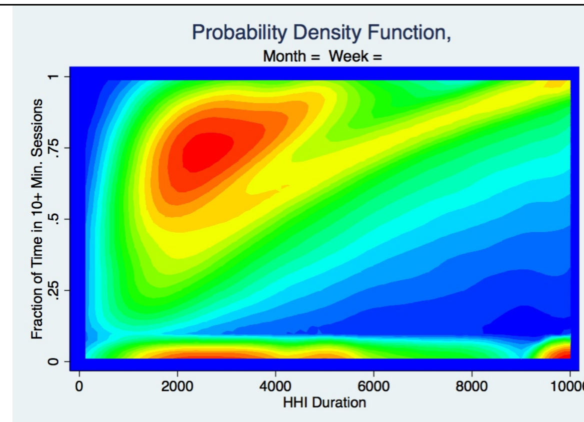
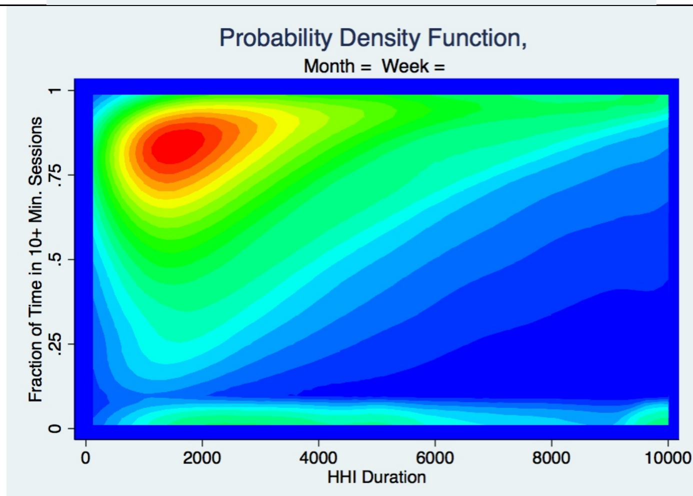
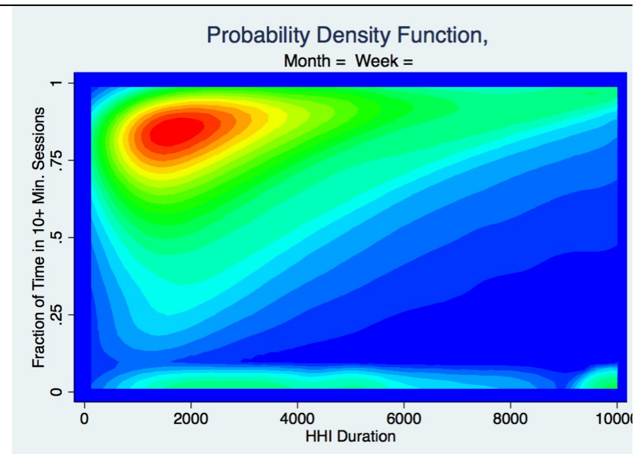
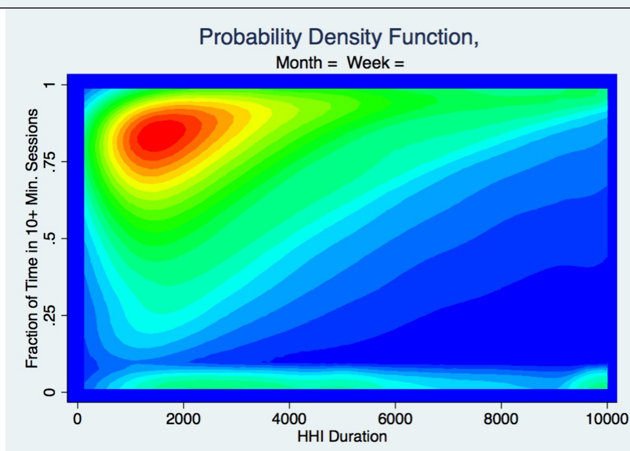
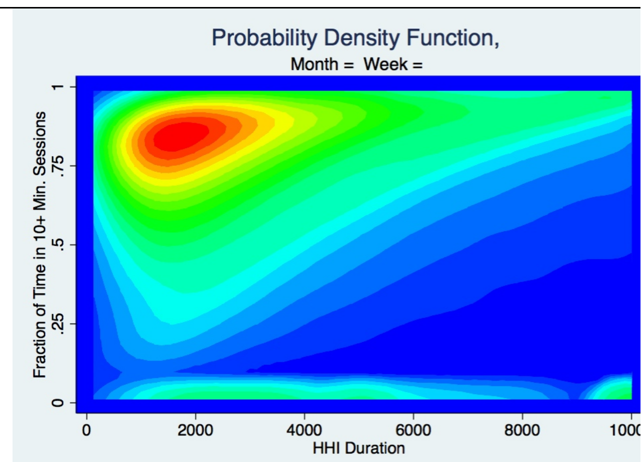
2008	2013
	

Figure 3

Distribution of Online Attention for Households by Quartiles of Total Minutes Online

	2008	2013
1st Quartile of Total Time Online		
2nd Quartile of Total Time Online		
3rd Quartile of Total Time Online		

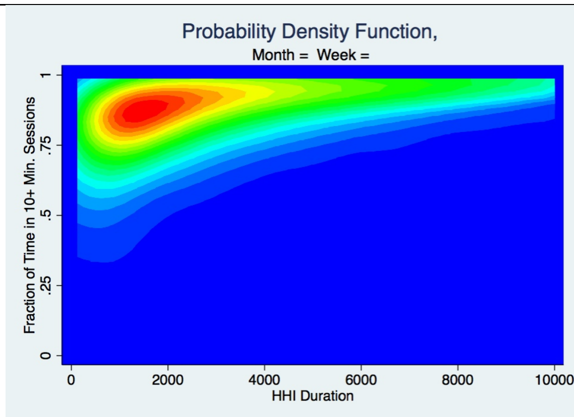
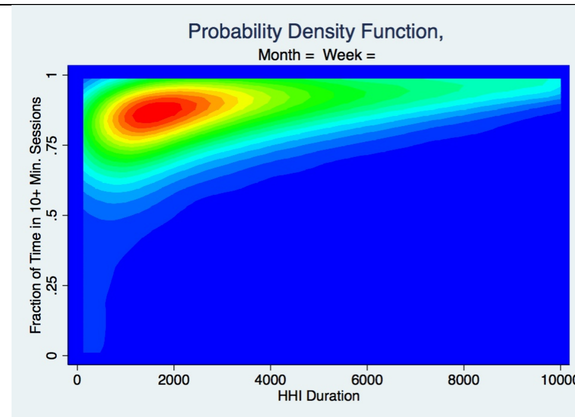
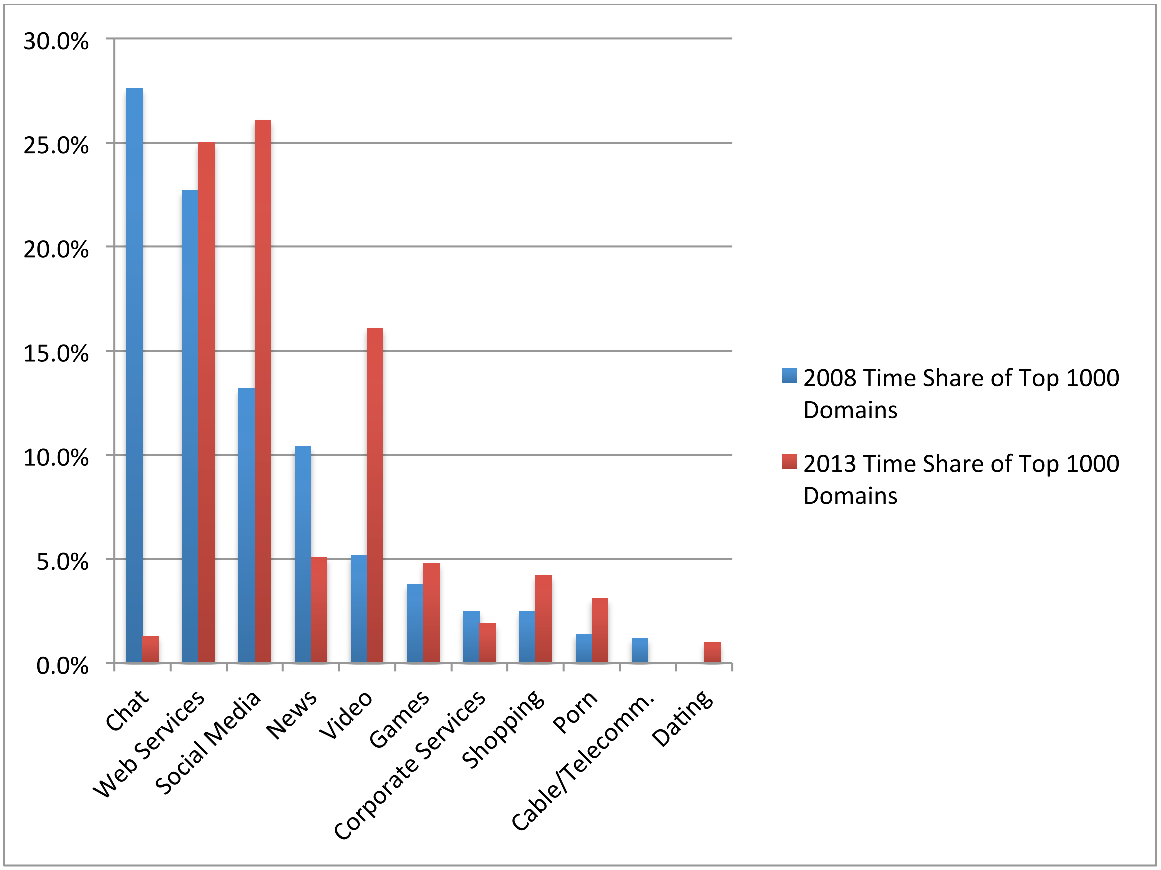
4th Quartile of Total Time Online		

Figure 4
Changes in Attention Allocation across the Top 1000 Domains by Category (2008 - 2013)



Tables

Table 2

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	0.07	0.25	0.07	0.25

60-64				
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		

Table 3

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
Focus (HHI of time across domains)	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
Focus (HHI of time across domains)	2968	2061	1.51	10000
Propensity to Dwell (Fraction of sessions > 10 minutes)	.76	.22	0	1

Table 4
Linear Regression - Time per week on demographics

	2008	2013
Covariate	Minutes per Week	Minutes per Week
Education High School	262.3 (1.84)	-
Education Some College	288.6* (1.97)	17.69 (0.64)
Education Associate Degree	188.7 (1.12)	12.84 (0.46)
Education Bachelor's Degree	348.1* (2.34)	79.60** (2.72)
Education Graduate Degree	248.3 (1.63)	131.3 (1.91)
HH Size = 2	-7.566 (-0.38)	-35.22* (-2.03)
HH Size = 3	10.38 (0.44)	-35.28 (-1.86)
HH Size = 4	27.27 (1.14)	-9.752 (-0.48)
HH Size = 5	74.72** (2.86)	1.002 (0.05)
HH Size = 6	113.6*** (3.69)	-21.04 (-0.87)
Age of Head of Household 21-24	-387.1*** (-4.20)	9.291 (0.34)

Age of Head of Household 25-29	-434.1*** (-4.88)	-15.89 (-0.62)
Age of Head of Household 30-34	-477.5*** (-5.42)	-36.37 (-1.47)
Age of Head of Household 35-39	-402.4*** (-4.58)	-21.14 (-0.84)
Age of Head of Household 40-44	-360.7*** (-4.11)	-17.66 (-0.71)
Age of Head of Household 45-49	-381.5*** (-4.36)	41.42 (1.69)
Age of Head of Household 50-54	-408.1*** (-4.66)	52.50* (2.12)
Age of Head of Household 55-59	-501.6*** (-5.71)	13.65 (0.54)
Age of Head of Household 60-64	-531.0*** (-6.01)	10.62 (0.40)
Age of Head of Household 65+	-550.6*** (-6.28)	14.60 (0.59)
Income $15k-$25k	-80.25*** (-3.83)	-18.85 (-0.95)
Income $25-$35k	-73.01*** (-3.57)	-18.67 (-0.96)
Income $35k-$50k	-91.39*** (-4.73)	-79.30*** (-4.49)
Income $50k-$75k	-117.7*** (-7.16)	-84.90*** (-5.08)

Income \$75k-\$100k	-131.3*** (-7.46)	-94.81*** (-5.25)
Income \$100k+	-165.5*** (-9.90)	-124.1*** (-7.14)
Children	3.388 (0.25)	132.3*** (10.46)
Constant	958.6*** (6.12)	799.9*** (21.53)
R-Squared	0.01	0.01
N	1,710,147	1,359,331

Table 5
SUR – Fraction of Sessions > 10 Minutes and Time HHI Across Domains

	2008	2008	2013	2013
Covariate	**HHI**	**Fraction > 10**	**HHI**	**Fraction > 10**
Education High School	624.2*** (4.30)	0.0922*** (6.17)	-	-
Education Some College	530.3*** (3.65)	0.0749*** (5.01)	-11.73 (-1.08)	-0.0114*** (-10.18)
Education Associate Degree	402.9* (2.49)	0.101*** (6.05)	-64.78*** (-5.85)	-0.0135*** (-11.85)
Education Bachelor’s Degree	299.2* (2.05)	0.0892*** (5.95)	-99.05*** (-8.60)	-0.0114*** (-9.63)
Education Graduate Degree	308.6* (2.10)	0.0960*** (6.33)	-125.7*** (-5.32)	-0.0163*** (-6.70)
HH Size = 2	-44.25*** (-6.54)	-0.000408 (-0.59)	-20.15** (-2.84)	-0.000213 (-0.29)
HH Size = 3	-57.55*** (-7.21)	-0.000247 (-0.30)	-18.03* (-2.34)	-0.000567 (-0.71)
HH Size = 4	-70.93*** (-8.68)	0.000446 (0.53)	-17.64* (-2.19)	0.00111 (1.34)
HH Size = 5	-102.7*** (-11.75)	0.00264** (2.94)	-35.72*** (-4.31)	-0.000443 (-0.52)
HH Size = 6	-235.4*** (-22.92)	0.00455*** (4.31)	-49.57*** (-5.16)	-0.00157 (-1.59)

Age of Head of Household 21-24	86.58^{**} (3.25)	-0.00704^{*} (-2.57)	-19.72 (-1.85)	-0.00398^{***} (-3.62)
Age of Head of Household 25-29	50.39^{*} (2.00)	-0.00624^{*} (-2.41)	-32.97^{**} (-3.15)	-0.00800^{***} (-7.44)
Age of Head of Household 30-34	100.4^{***} (4.03)	-0.00273 (-1.06)	-0.159 (-0.02)	-0.000806 (-0.78)
Age of Head of Household 35-39	105.4^{***} (4.27)	0.00228 (0.90)	-7.925 (-0.77)	-0.00270^{*} (-2.54)
Age of Head of Household 40-44	184.7^{***} (7.51)	0.00384 (1.52)	51.09^{***} (5.12)	-0.00437^{***} (-4.26)
Age of Head of Household 45-49	231.6^{***} (9.43)	0.00232 (0.92)	-0.367 (-0.04)	-0.00440^{***} (-4.38)
Age of Head of Household 50-54	232.9^{***} (9.47)	-0.00205 (-0.81)	-47.54^{***} (-4.87)	-0.00625^{***} (-6.22)
Age of Head of Household 55-59	199.0^{***} (8.04)	-0.00883^{***} (-3.47)	20.14^{*} (1.98)	-0.00644^{***} (-6.16)
Age of Head of Household 60-64	304.2^{***} (12.18)	-0.00640^{*} (-2.49)	16.32 (1.52)	-0.00531^{***} (-4.81)
Age of Head of	360.0^{***} (14.56)	-0.00707^{**} (-2.78)	53.28^{***} (5.41)	-0.00740^{***} (-7.30)

Household 65+				
Income \$15k-\$25k	9.556 (1.37)	-0.00276*** (-3.84)	22.29** (2.98)	0.00189* (2.45)
Income \$25-\$35k	6.577 (0.99)	-0.00787*** (-11.54)	0.721 (0.10)	-0.0000278 (-0.04)
Income \$35k-\$50k	-8.455 (-1.32)	-0.00975*** (-14.78)	10.70 (1.57)	-0.00295*** (-4.20)
Income \$50k-\$75k	-29.68*** (-5.52)	-0.0108*** (-19.44)	16.06* (2.51)	-0.00270*** (-4.10)
Income \$75k-\$100k	-1.538 (-0.26)	-0.0142*** (-23.77)	-27.86*** (-3.94)	-0.00128 (-1.76)
Income \$100k+	-42.53*** (-7.61)	-0.0161*** (-28.05)	-14.25* (-2.12)	-0.00461*** (-6.68)
Children	-58.62*** (-12.78)	-0.000783 (-1.66)	-142.4*** (-27.01)	-0.000568 (-1.05)
Minutes per Week	-0.000443 (-0.37)	0.0000662*** (531.17)	-0.290*** (-181.12)	0.0000724*** (438.72)
Constant	2652.0*** (18.26)	0.617*** (41.25)	3346.2*** (228.47)	0.713*** (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 6

The Top 20 Domains of 2008 and 2013 (by Total Time Allocated)

2008 Top 20 Domains	Category	2013 Top 20 Domains	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 7

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 changes over time with new supply.	Total time 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 changes with new supply.	No 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/declines with income.	Breadth/depth do not vary with income.	Figure 3
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 with new supply.	Figure 2