The Empirical Economics of Online Attention

Andre Boik, Shane Greenstein, and Jeffrey Prince*

April 2016

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. We illustrate how this finding is difficult to reconcile with standard models of optimal time allocation, and suggest alternatives that may be more suitable. The fact that increasingly valuable offerings change where households go online, but not their general (i.e., breadth/depth) online attention patterns, has important implications for competition and welfare.

* 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 Kelley School of Business and the Harvard Business School for funding. We thank seminar audiences at Georgetown, Harvard and Northwestern, and conference participants at the American Economic Association Annual Meetings and the International Industrial Organization Conference 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).

First articulated about information systems, Herb Simon brought attention to a broad economic principle that applies to any situation with abundant information. The principle remains relevant today, even more so for the supply of information by the commercial Internet. Scarce users' attention must be allocated across the Internet's vast supply of web sites. It is not an exaggeration to say that 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 there is only a finite amount of such time to allocate. 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, potentially substituting one source of supply for another.

First impressions mislead. This situation lacks one of the standard hallmarks of competitive situations. Relative prices largely do not determine user choice among options, nor do prices determine competitive outcomes. Most households pay for monthly service, then allocate among endless options without further expenditure. Unless a household faces 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). In fact, as we will show below, only one of the top twenty domains (Netflix) is a subscription service, i.e., where the price of a web site plays an explicit role in decision making.

How should economic analysis characterize the links between user allocation of attention and online competition in the absence of prices? An empirically grounded theory of competition would have to characterize market demand. That depends on three interrelated aspects of users: how users allocate the amount and division of attention across multiple sources of information; how users adjust the allocation to a change in supply; and how this change in the allocation shapes the competition for their attention. From those three building blocks, it should be possible to characterize: household heterogeneity in allocation of attention *at any point in time*, how households substitute between sources of supply *over time*, and, by extension, how aggregate demand changes in the face of increasing supply of options. Finally, such results could inform a theory about how the allocation of online attention shapes competitive behavior, such as entry of new sites or building of new features to attract attention.

The goal of this paper is to make such a characterization of demand, based on empirically grounded observation. We examine a specific context, the adjustment of attention at US households to the enormous changes in supply of online options between the years 2008 and 2013. We choose these starting and ending periods because over 70% of US households were on broadband connections by 2008, and in the intervening years US households experienced a massive expansion in online video offerings, social media, and points of contact (e.g., tablets, smartphones), among other changes. These years allow us to examine household reaction to large changes in supply of content.

Specifically, we examine a dataset of more than 30 thousand primary home computers at US households in 2008 and 2013. These data come from ComScore, a firm that tracks households over an entire year, recording all of the web domains visited, as well as some key demographics. Our unit of observation is choices made by households over the course of a week. We calculate the weekly market for online attention (total time), its 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"). In addition, we measure shares of attention for different domain categories (e.g., social media). Using these measures of online attention, we examine how they vary over the time period, as supply changes.

Our findings suggest that aggregate demand has very specific properties. 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 (Goldfarb and Prince, 2008), but does so on more extensive detail and later years. We find that higher income households spend less total time online per week. Our results suggest that a household making $25-35K a year spends 92 *more* minutes a week online than a household making $100K or more a year in income, and differences vary monotonically over intermediate income levels. 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. At the same time, the monotonic negative relationship between income and total time remains *stable*, exhibiting the same slope of sensitivity to income. The change is generally similar across income groups, and is consistent with a simple hypothesis about the changing allocation of time across devices. That is, any new value stemming from additional total time online (across all devices) appears to be largely coming from time on new, alternative devices.

Despite the evidence of some economic determinants of total online time, we see evidence that its allocation is sensitive to different factors. Breadth and depth are not well-predicted by income, but there are roles for major demographics, such as family education, household size, age of head of household, and presence of children. More remarkable, both depth and breadth do not meaningfully change in spite of massive changes in supply. 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 expectation was that depth would increase, and more tentatively, that breadth would increase as well. . Our findings do little to confirm what we expected. Rather, focus and dwelling has remained remarkably *stable* over the five years. While there is a statistical difference in the joint distribution of focus and dwelling, it is just that – statistical, and driven by our large sample. The size of the difference is remarkably small, with little implied economic consequence. Also remarkable is that these measures are so stable despite households changing the web sites they visit a great deal. Between 2008 and 2013 online categories, such as social media, and possibly video, become a substitute for both chat and news, and this substitution is readily apparent in our dataset. In summary, new offerings *did* alter where households went online, but 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 particularly which sites to visit are distinct from the determinants of breadth and depth.

These findings have important implications for competition to reallocate 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. Because the demographics of household heterogeneity did not dramatically change between 2008 and 2013, aggregate demand only mildly changed, as total time online change. Altogether, as we illustrate in our theoretical development, these findings suggest that at any point in time there are a fixed set of "slots" of attention to allocate, and very limited substitution by households between different "slots" of different lengths. Stated starkly, firm entry and exit compete for given slots of time from users.

Our results merit attention for numerous reasons. First of all, the commercial Internet is a big market, 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 advertising revenue, and trillions in transaction revenue in online sales. Yet, despite the shared features with other competitive US markets, user choice among many web sites remains largely uninformed by prices, and analysis has not built on this simple fact. This leaves a gap in analysis about how commercial firms compete for user attention.

As of this writing, economists generally have not focused on priceless online competition except for a few theoretical pieces on competition for attention (Athey, Calvano and Gans, 2013).There has been almost no formal statistical work on the competition for attention except 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). No work has characterized the entire allocation of a household – the "what, where and how" behind the core economics of competition for online attention, nor built a model of aggregate demand from such data. We address this gap.

Our study also relates to the extensive literature on the economic allocation of time. We ask whether user patterns of behavior are consistent with the predictions of a basic theoretical model of the allocation of time. In this study we 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 much of its roots in Becker (1965). Using this model, we highlight theoretical ambiguities as to predicted changes in online attention with increased online offerings. We then 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. We use these to characterize three basic types of online attention measurements – How much? How is it allocated? and Where is it allocated?

These findings and this outlook build on prior work on the value of household time online, and contrasts with it. Several studies provide evidence demonstrating the demand for, and market value of, speed in broadband access, which users spread over a vast array of content (Rosston, Savage, and Waldman, 2010, Hitt and Tambe, 2007). Prior work also 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). In addition, prior work has 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 (Webster, 2014, 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 attention. That focus leads to a very different analysis of the core economics.

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 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 very different conceptualization about competition for attention.

Though we depart from some of the existing economics literature, our findings are not much of a departure from field work conducted by anthropologists and researchers on user-machine design. That line of research has documented the periodic – or "bursty" – use of many online sources, consistent with our findings about the breadth of session times (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 also view our work as a bridge between economic analysis and conversations taking place within other domains 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 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. Netflix and Hulu both began offering streaming services in 2008. 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. 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

We propose a standard model of online attention following the basic structure of the seminal work by Becker (1965) on the allocation of time, and which has been adapted by others in various ways to examine household demand for broadband (e.g. Savage and Waldman 2009). Critical to our model is that visits to online domains do not carry a price; rather, the cost of a domain 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 domain. The setup cost can be interpreted as either a necessary minimum time cost to absorb the information at a domain, a cognitive cost of switching domains, a time cost of waiting for a new domain to load, or so on. The point is that the existence of any such cost will generate continuous visits to domains that end only when the time slot has expired or the marginal utility from additional time spent at the domain falls below the marginal utility of visiting some other domain net of the switching cost.

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

$$(1)\ max_{t_{i1},\dots,t_{iJ}} U\big(t_{i1},\dots,t_{iJ}, T_i - (t_{i1}+\dots+t_{iJ}); \vec{W}\big) - \textstyle\sum_j^J 1(t_{ij}>0)F$$

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

where *F* is the setup cost of visiting a domain. In equation (1), $\vec{W}$ 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* in 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 no structure is placed 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 the 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.[1] The model produces the following identity for time online for household i (TO_i) when there are J domains:

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

Next, we consider measures for breadth and depth of online time allocation. 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:

[1] 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, 2015), 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.

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

Defined this way, our measure of breadth captures the level of concentration (in terms of time at domains) household i exhibits in its domain 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 domains, whereas a low value for C_i indicates a breadth of visits that is unconcentrated, i.e., spread out across relatively many domains. It also can measure changes over time: C_i gets larger as a household substitutes a larger fraction of its time into fewer web 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. If the setup cost is strictly positive, the standard model suggests households spend all of their time at each domain 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 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:

$$(3)\ L_i = \frac{\sum_j \sum_k 1(t_{ijk}^* > 10)}{\sum_j \sum_k 1(t_{ijk}^* > 0)}$$

As defined, L_i is the proportion of total domain visits that lasted more than ten minutes for household i. 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 at a household, with an increase in L indicating that a household has substituted a large fraction of its time into longer sessions.

An illustration can help build intuition for how these 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 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 of heterogeneity across households.

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 consider one last measure: "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:

$$(1)\ TS_c = \frac{\sum_i \sum_{j \in c} t_{ij}^*}{\sum_i TO_i}$$

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. 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 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 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. 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.

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

$$(2)\ max_{t_{i1},\ldots,t_{iJ^*},t_{i1}^{dev},\ldots,t_{iJ^*}^{dev}} U\left(t_{i1},\ldots,t_{iJ^*},t_{i1}^{dev},\ldots,t_{iJ^*}^{dev},T_i-\left(t_{i1}+\cdots+t_{iJ^*}^{dev}\right);\vec{W}\right) - \sum_j^J 1(t_{ij}>0)F$$

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

The household faces more domain choices and the option to consume them on an alternative device. We assume setup costs affect the alternative device as they do for the home device, which implies the solution closely mirrors that without additional domains or an additional device. We ask how these two changes impact three key outcomes within our model: total time, breadth, and depth. That is, in terms of time online, we ask how these changes impact how much, 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 domains, quasilinear in an unchanging offline outside option, and the setup costs are small, then an increase in the number of domains weakly increases the total amount of time online, and decreases the concentration of time spent across domains on the home device.[2] 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 domain 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* domains before the shock and will continue to visit the same number of domains after the shock, so that the concentration of time across domains is unchanged. If the household was not constrained before

[2] The details of the microeconomics behind this prediction and those that follow can be found in the Appendix.

the shock, then concentration of time across domains will fall. Additionally, the marginal effect of the introduction of an alternative device is to weakly increase concentration: any domain visits substituted towards the alternative device increase the time share of the domains viewed on the home device.

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

	Small setup costs	Large setup costs ($F >> 0$)
Shock 1: New Domains	$\Delta TO_i \geq 0$ $\Delta C_i < 0$ 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 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 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 domains, our model is silent with respect to how households will reallocate attention across different types of domain 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 suggests a predicted increase in L_i.

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 be providing 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 5 through 8. 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

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 households with high income are more likely to adopt, but they do not use the Internet as intensively. They hypothesize that this is 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. $\mathrm{TO}_x(S_t, X_{it}) < 0$.

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 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 vice versa? The null hypothesis specifies no change in total time:

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

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. $\mathrm{TO}_x(S_t, X_{it}) – \mathrm{TO}_x(S_{t-1}, X_{it-1}) = 0$.

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 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 reaction of households to growth in 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? We expect new sources of supply to increase depth and breadth, and 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 rejecting the null hypothesis becomes. That is because the null defines household stability in the allocation of breadth and depth in spite of changes in options available to households, and presumably the growth in options becomes much larger with the passage of longer time. Substitution can arise from a vast array of endless possibilities, either splitting up a large moment of time into many smaller units of time or it can arise from taking many small units and putting them together into one long unit. After fives 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 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 3. 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 except to indicate that virtually of them are not dial-up.

[Table 1 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 4. 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 2 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, 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.

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.

We present the results of a simple regression of time online per week on demographics, and show the results in Table 3. In these data we see a Goldfarb-Prince effect in any given year, and the evidence is much stronger due to our access to a much large set of data over more households, and over multiple years.

We confirm H1, namely, 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. 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 qualitative change in the relationship between time online and income over this period. For 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.

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. Large households also spend more time online, but the relationship is only strong in 2008. In 2013 only the presence of children captures this effect. Total time is also declining in the age of head of household in 2008, but no such monotonicity arises in 2013.

[Table 3 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 largely stable across timeHence, the Goldfarb-Prince effect appears to be a stable relationship for total time at the household level.

5.2. Online Attention Allocation Patterns

In this subsection, we present analyses of focus and dwelling. 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 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 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.

Despite this, we can reject the null hypothesis that they are statistically indistinguishable, likely because our combined sample size is over three million. Tables 4a and 4b present statistical tests of the means of our measures of Focus and Propensity to Dwell 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 Focus is only 3.5% greater in 2013 and household Propensity to Dwell greater by only 1%.

[Figure 2 about here]

We are concerned 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 be more likely to dwell and perhaps be less likely to be focused. In short, we are concerned that a household's location within the distribution presented in Figure 1 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 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.

[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. 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. 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 focus are lower levels of education, older heads of households, and household size. In contrast, households' propensity to dwell is largely independent of demographics. In particular, more educated households, larger households, and younger households visit a larger variety of sites in both 2008 and 2013.

[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 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.

These are quite striking findings about the role of demographics in breadth and depth in light of our earlier results about total time. Income of households helps 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]

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 results are more easily rationalized by the standard model with setup costs: the supply of new domains 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.

From 2008 to 2013, we observe no change in the depth of household browsing. Again, the standard model with setup costs rationalizes the data better than the model without setup costs. However, the standard model with setup costs predicts all sessions be of the same length so that all sessions either fall above or below a given threshold. We do not see similar length. We see a mix of sessions of different lengths, *where the proportion has not changed across years*.

This mix of sessions of different lengths could be explained by an asymmetric utility function that captures how a household values each domain differently. However, we believe that our empirical findings point towards 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 households' browsing.

To summarize, there is a discrete nature to households' domain choices and a bound on the number of domains 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 of time into more numerous and shorter domain visits, as predicted by the standard model without setup costs. The standard model augmented with a positive setup costs performs better: it predicts that even with increasing supply, there will be a finite number of domains 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 households' breadth of browsing will increase or decrease. The unchanging breadth *and* depth to households' browsing patterns invites an alternative theory of household browsing behavior, one that can explain this constancy.

What would such a theory contain? It might be behavioral or one where a household receives exogenous “slots” of time which 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 have observed behavior consistent with the 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). It also documents the “plasticity” of online attention, as an activity that arises within the midst of 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 it 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 supply of online web sites does change.

6. Implications for online competition

What are the implications for online competition? To begin, we summarize our findings in Section 5 and discuss their main implications, focusing on no changes between 2008 and 2013 – a time of substantial change in the Internet ecosystem. We summarize our findings in Table 8, 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. 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.

Our findings also 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.

Altogether, this adds up to a surprising characterization of the supply of attention by households and the demand for online activities. 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.

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.

This last set of findings suggest that households likely achieved additional value between 2008 and 2013 by reallocating their attention across domains with only mild changes in the total time being allocated. 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.

This adds up to a striking model of competition for attention. As expected, there were large changes in the supply of web sites for households to choose, and many households responded to that expanded choice. Presumably existing web sites also improved, as they fought to keep the attention of existing users from households. However, the competition for users was constrained by virtually unmoveable feature of demand – the length of slots households were willing to sustain at a given site, and the ultimate desire of households to experience a wide variety of content. Competition for household attention *did* alter the total time on the home PC, shifting some of it to other devices, but competition for households *did not* lengthen or shorten the actual time online, nor did it change the effective consumption of variety.

7. Conclusions

To be written…

8. Appendix

This appendix provides more details behind the predictions of the standard model under small or large setup costs summarized in Table 2. We assume utility is symmetric among domains, has decreasing marginal utility across domains, and is linear in the normalized and constant outside option of offline activities so that $U\left(t_{i1},\dots,t_{iJ*},T_i-\left(t_{i1}+\cdots+t_{iJ*}\right);\overrightarrow{W}\right)$ can be written $U\left(t_{i1},\dots,t_{iJ*};\overrightarrow{W}\right)$ + $T_i -$ $\left(t_{i1}+\cdots+t_{iJ*}\right)$ and symmetry giving $U\left(t_{ik},\dots,t_{i-k};\overrightarrow{W}\right)=U\left(t_{i-k},\dots,t_{ik};\overrightarrow{W}\right)$ and $U_k\left(t_{ik},\dots,t_{i-k};\overrightarrow{W}\right)=U_k\left(t_{i-k},\dots,t_{ik};\overrightarrow{W}\right)$ $\forall$ i,t_{ik},t_{i-k}.

8.1 Total time online, $\boldsymbol{TO_i}$

When setup costs are small ($F = 0$), the household visits all available domains and allocates equal amounts of time to all of them. When the number of domains increases from J to J*, the total amount of time online TO_i increases by a factor of $\frac{J*}{J}$, and when an alternative device becomes available, TO_i is unchanged because it represents time on the home device. The standard model with small setup costs predicts total time online on the home device weakly increases in response to the two shocks when setup costs are small.

When setup costs are large ($F > 0$), the household may be constrained to visit a subset of domains even before the number of domains increases from J to J*. When new domains arrive, households will either allocate the same amount of time to online activities or increase their time spent online. When an alternative device becomes available, a household that was previously constrained to visit a subset of domains may substitute some of those visits towards the alternative device, decreasing time online on the home device. If the household was not previously constrained, then additional time spent on the alternative device does not affect time on the home device. Therefore the effect of the two shocks on total time online in the presence of large setup costs is ambiguous.

8.2 Breadth of browsing, C_i

When setup costs are small ($F = 0$), the household visits all available domains and allocates equal amounts of time to all of them. When domains increase from J to J*, the household visits as much as twice as many domains and spends no more time per domain as the household did before. The household's breadth of browsing weakly declines. When an alternative device becomes available, the household continues to visit all domains on the home device in equal amounts so that the breadth of browsing does not change. The standard model predicts the concentration of time across domains will fall in response to the two shocks when setup costs are small.

When setup costs are large ($F > 0$), the household may be constrained to visit a subset of domains even before the number of domains increases from J to J*. When new domains arrive, households will either visit the same number of domains and spend equal amounts of time per domain or will increase the number of domains visited and spend equal amounts of time per domain; concentration of time across domains weakly declines. The alternative device either causes no change in the allocation of time on the home device or results in the household substituting domain visits to the alternative device, in which case concentration of time across domains on the home device weakly increases. The net effect of the two shocks on the household's breadth of browsing is ambiguous when setup costs are large.

8.3 Depth of browsing, L_i

When setup costs are small ($F = 0$), the household can visit the same domain any number of different times to reach the same amount of total time spent at that domain: for instance, a single session of 10 minutes or ten sessions of 1 minute each. The standard model with small setup costs does not make a prediction about the depth of browsing.

When setup costs are large ($F > 0$), time spent across all visited domains is identical and expended in a single, continuous session so as to incur the setup cost only once. The household may be constrained to visit a subset of domains even before the number of domains increases from J to J*. When new domains arrive, households will either visit the same number of domains and spend equal amounts of time per domain as before or will increase the number of domains visited and spend equal amounts of time per domain as before; in either case, the depth of browsing remains unchanged. The alternative device either causes no change in the allocation of time on the home device or results in the household substituting domain visits to the alternative device, but leaving the depth of browsing of the remaining visits on the home device unchanged. The two shocks are not predicted to have any effect on the depth of browsing when setup costs are large.

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/

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.

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.

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.

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.

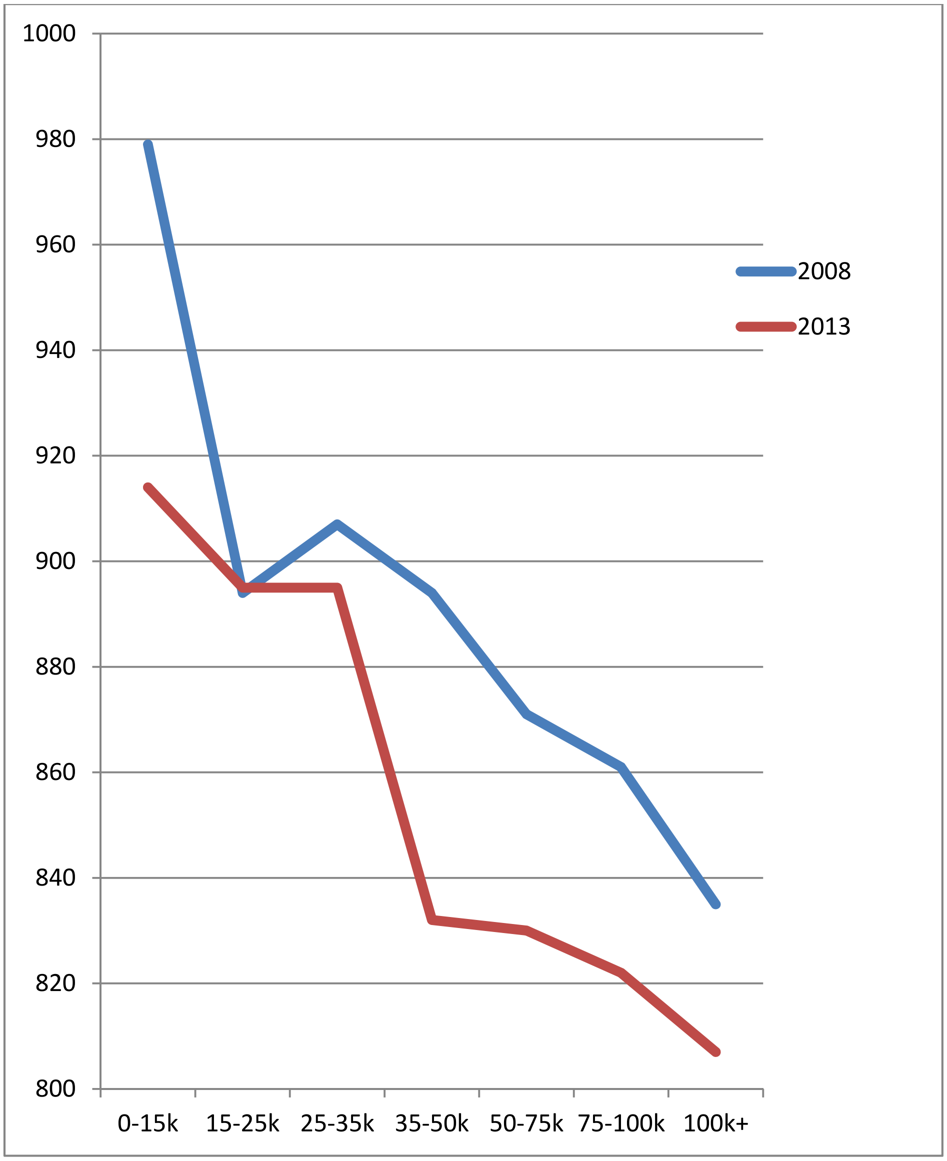
Wallsten, S. 2013. What are We Not Doing When We're Online? NBER Working paper.

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)



1000
980
960
940
920
900
880
860
840
820
800
0-15k
15-25k
25-35k
35-50k
50-75k
75-100k
100k+
2008
2013

Figure 2

Unconditional Distribution of Online Attention (2008 vs. 2013)

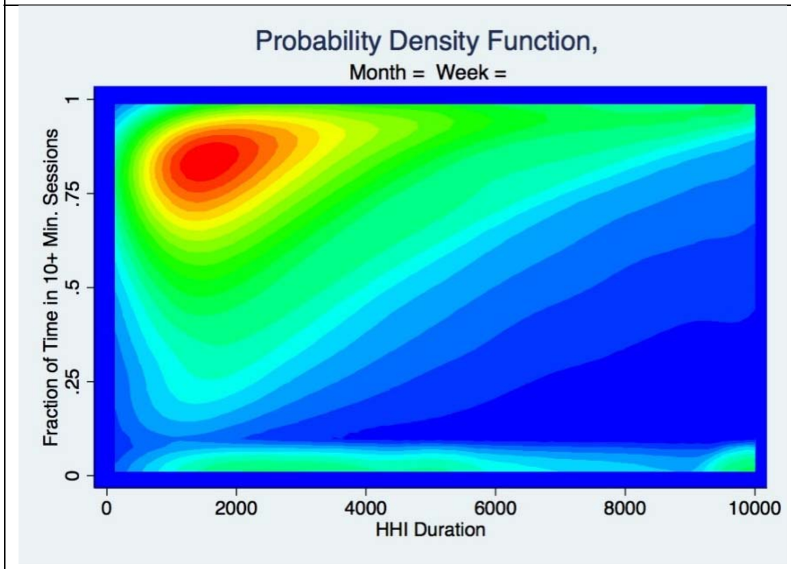
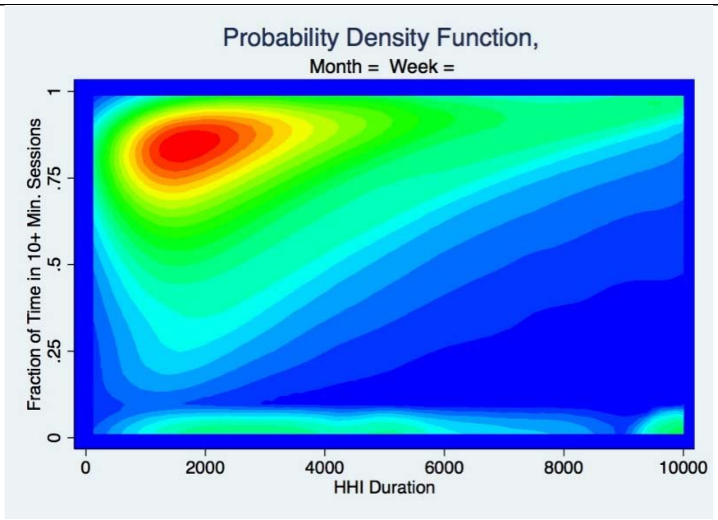
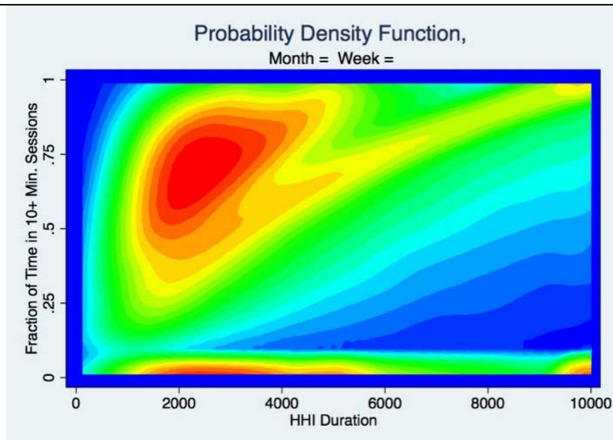
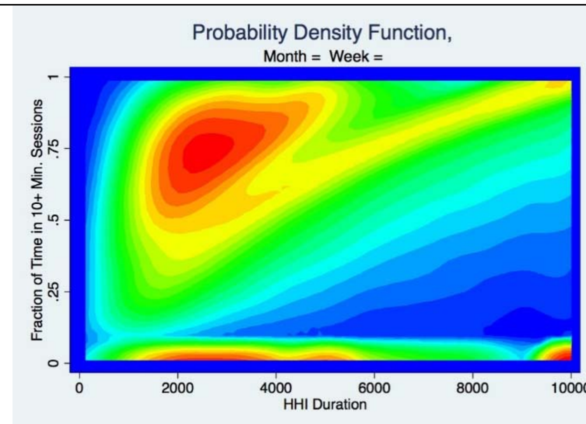
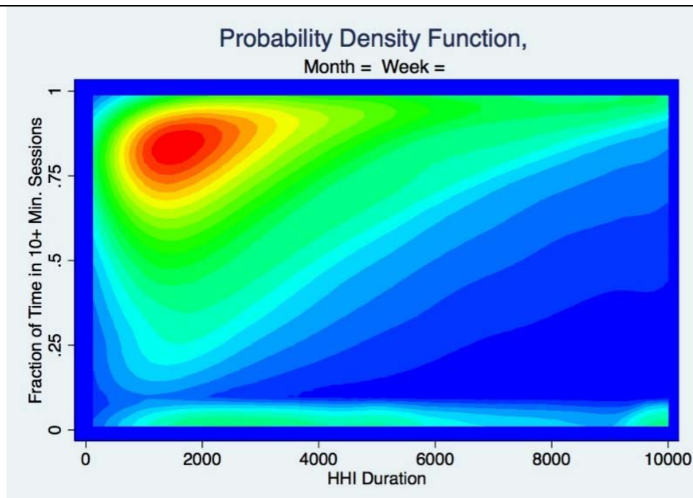
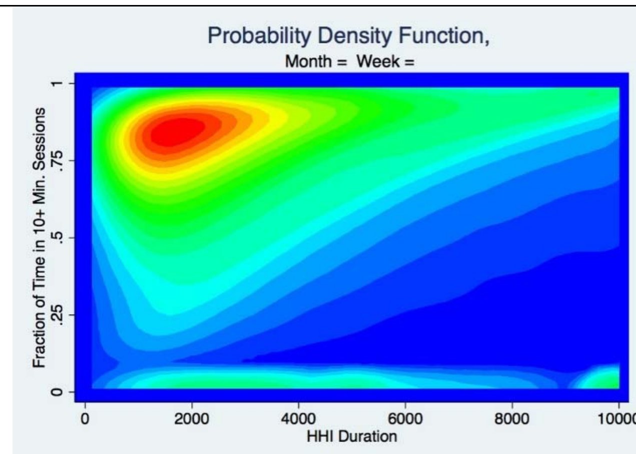
2008	2013
	

Figure 3

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

<table>
<tr><th></th><th>2008</th><th>2013</th></tr>
<tr><td>1st Quartile of Total Time Online</td><td>
</td><td>
</td></tr>
<tr><td>2nd Quartile of Total Time Online</td><td>
</td><td>
</td></tr>
</table>

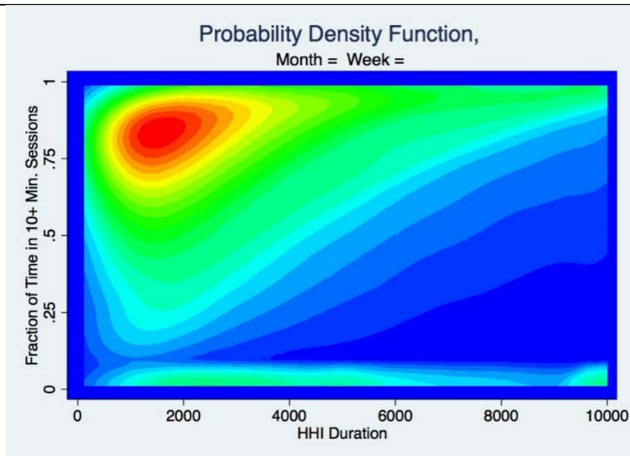
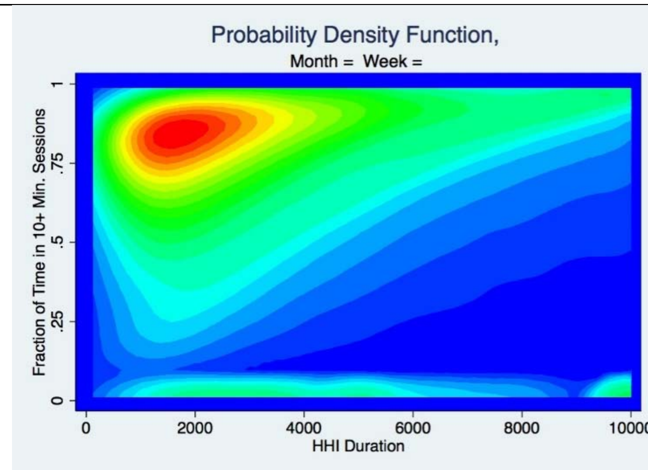
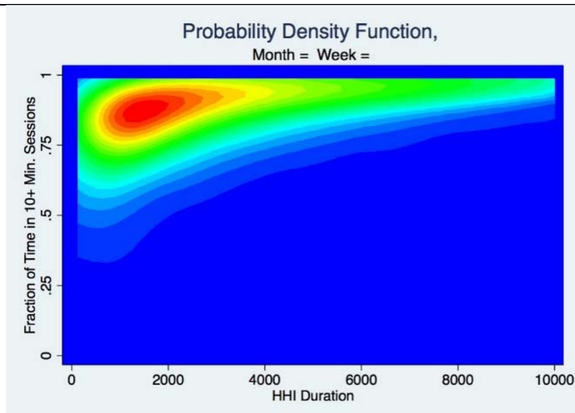
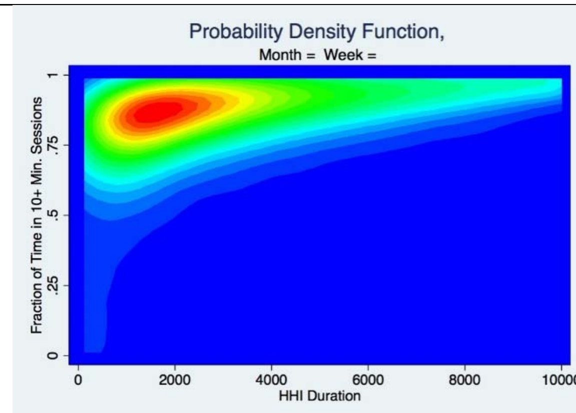
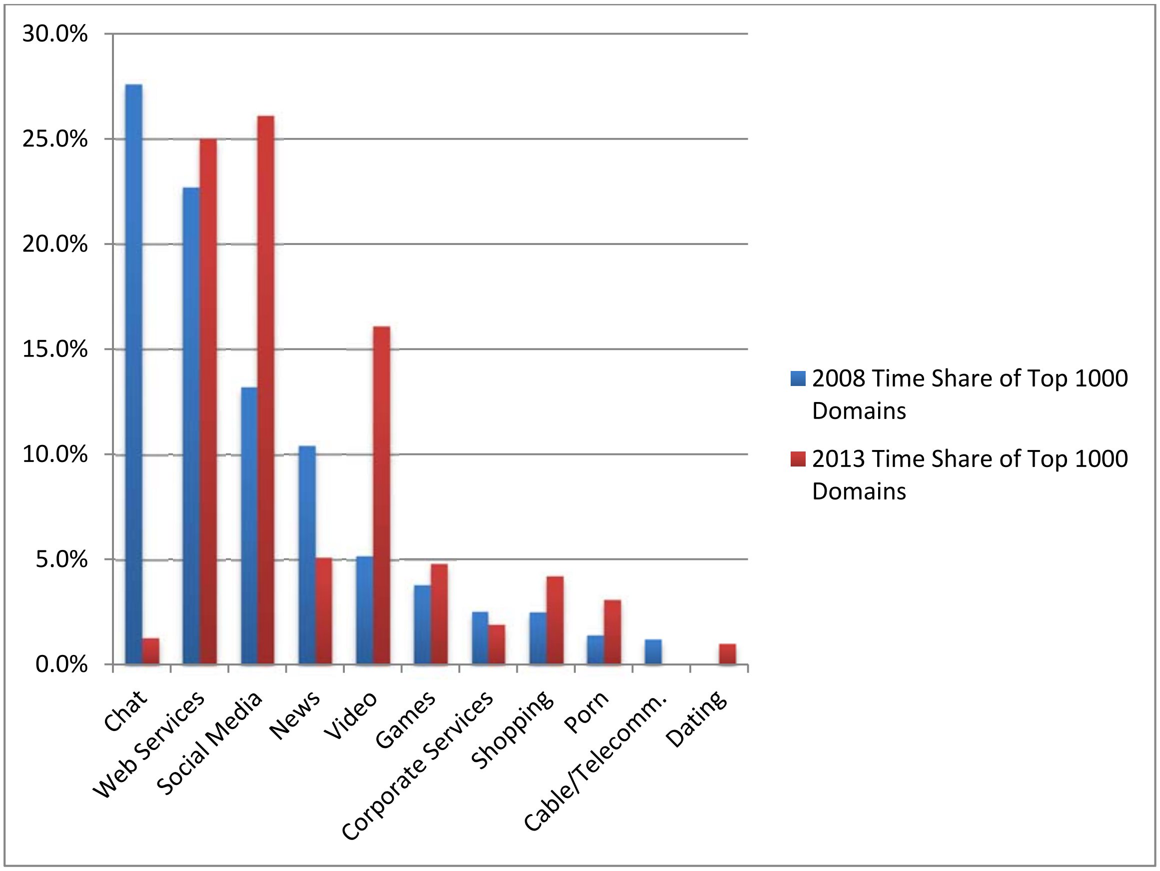
3rd Quartile of Total Time Online		
4th Quartile of Total Time Online		

Figure 4

<u>Changes in Attention Allocation across the Top 1000 Domains by Category (2008 - 2013)</u>



Tables

Table 1

<u>Household Summary Statistics</u>

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		

Table 2

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 domains visited per week	41	44	1	3936
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
Unique domains visited per week	41	47	1	7525
Focus (HHI of time across domains)	2968	2061	1.51	10000
Propensity to Dwell (Fraction of sessions > 10 minutes)	.76	.22	0	1

Table 3

Linear Regression - Time per week on demographics

	2008	2013
Covariate	**Minutes per Week**	**Minutes per Week**
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)
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)
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

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4a

Test of equality of means across years

	Dependent variable	**Dependent variable**
	Focus (HHI of time across domains)	**Propensity to Dwell (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 4b

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

	Variable	**Variable**
	Focus (HHI of time across domains)	**Propensity to Dwell (Fraction of Sessions > 10 minutes)**
p-value	0.00	0.00
N	3,069,478	3,069,478

Table 5

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

	2008	**2008**	**2013**	**2013**
Covariate	**HHI**	**Fraction > 10**	**HHI**	**Fraction > 10**
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)
Education High	624.2*** (4.30)	0.0922*** (6.17)	-	-

School				
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 Household 65+	360.0*** (14.56)	-0.00707** (-2.78)	53.28*** (5.41)	-0.00740*** (-7.30)

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

Linear Regression - Unique domains visited per week

	2008	2013
Covariate	Unique domains visited per week	Unique domains visited per week
Time per week	0.0181*** (89.75)	0.0252*** (91.81)
Income \$15k-\$25k	0.748 (1.31)	-0.969 (-1.57)
Income \$25-\$35k	0.906 (1.56)	-0.183 (-0.28)
Income \$35k-\$50k	1.092* (2.06)	-0.895 (-1.63)
Income \$50k-\$75k	1.661*** (3.75)	-0.840 (-1.58)
Income \$75k-\$100k	1.199* (2.54)	-1.086 (-1.87)
Income \$100k+	1.977*** (4.37)	-0.204 (-0.36)
Education High School	-6.908 (-0.81)	-

Education Some College	-13.02□(-1.65)	0.590□(0.62)
Education Associate Degree	-9.897□(-1.22)	1.325□(1.40)
Education Bachelor's Degree	-5.593□(-0.71)	2.968**□(3.01)
Education Graduate Degree	-9.223□(-1.15)	7.603**□(2.84)
HH Size = 2	-0.195□(-0.35)	-0.0245□(-0.04)
HH Size = 3	0.718□(1.07)	0.0539□(0.09)
HH Size = 4	0.299□(0.45)	-0.520□(-0.82)
HH Size = 5	1.000□(1.39)	0.273□(0.41)

HH Size = 6	3.780*** (4.25)	0.213 (0.28)
Age of Head of Household 21-24	-3.730 (-1.51)	1.093 (1.31)
Age of Head of Household 25-29	-4.907* (-2.10)	2.007* (2.52)
Age of Head of Household 30-34	-5.497* (-2.39)	-0.155 (-0.21)
Age of Head of Household 35-39	-5.465* (-2.38)	0.650 (0.83)
Age of Head of Household 40-44	-6.853** (-3.01)	0.131 (0.17)
Age of Head of Household 45-49	-7.135** (-3.13)	1.336 (1.71)
Age of Head of Household 50-54	-7.055** (-3.09)	1.582* (2.09)
Age of Head of Household 55-59	-7.829*** (-3.42)	0.968 (1.24)

Age of Head of Household 60-64	-9.177***□(-4.00)	1.706*□(2.05)
Age of Head of Household 65+	-8.979***□(-3.93)	0.596□(0.79)
Children	0.932*□(2.43)	2.384***□(6.01)
Constant	34.29***□(4.27)	16.38***□(12.74)
R-Squared	0.28	0.34
N	1,710,147	1,359,331

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8

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