The Value of Online Scarcity Signals *

Pascal Courty† Sinan Ozel‡

April 9, 2016

Abstract

Online retailers use scarcity cues to increase sales. Many fear that these pressure tactics are meant to manipulate behavioral biases by creating a sense of urgency. At the same time, scarcity cues could also convey valuable information. We measure the value of scarcity messages in the context of air travel. Airlines and metasearch engines respectively post messages on availability (number of seats left at a given price) and recommendations (buy/wait). We measure the economic value of scarcity messages as the increase in expected utility derived by optimally timing the purchase of a ticket. Scarcity messages increase the expected utility of sophisticated and unsophisticated consumers by about 4 percent. We document how this figure is influenced by retailer incentives and competition.

Keywords: Scarcity, Persuasion, Online Recommendations, Price Discrimination, Airline Ticket.

JEL Classification: L1.

WORK IN PROGRESS — DO NOT CIRCULATE OR QUOTE

*We thank

†University of Victoria and CEPR; pcourty@uvic.ca.

‡University of Victoria; sozel@uvic.ca.

1 Introduction

Scarcity cues and pressure tactics are widely used by online retailers to increase sales (Aggarwal et al., 2011; Aguirre-Rodriguez, 2013). Scarcity is conveyed in essentially two ways (Gierl and Huettl, 2010): limited-quantity offers (remaining stock left, flash sales, auction...) and limited time offers (expiry sales, Black Friday, closing time, countdown timer...). According to marketers and social scientists, scarcity creates a sense of urgency, it increases desirability and gives a perceived benefit of acting quickly (Worchel et al., 1975; Lynn, 1991; Verhallen and Robben, 1994; Mullainathan and Shafir, 2013). Many fear that sellers manipulate the psychology of consumers. At the same time, marketers warn us that such tactics are effective only if the sender is trustworthy and her messages are credible. To the extent that scarcity messages deliver information that is not available otherwise, a Bayesian consumer could benefit from messages that are meant to manipulate behavioral consumers subject to decision biases.

This paper measures the informational value of scarcity messages in the context of air travel. Airfares vary dramatically from day to day. Travelers have to choose when to book an airline ticket without knowing what will happen to prices and whether it would be wise to postpone their decisions in the hope of better fares. Airlines try to influence travelers by presenting scarcity signals next to airfares. For example, an American Airline flight displayed on Expedia can mention that there is a limited number of seats available at the posted price. These signals may be pressure tactics with no informational content. Alternatively, they may contain information that could help travellers time their booking decisions. Metasearch engines also post information on what is likely to happen to the overall level of prices on a given itinerary (given route and departure date). For example, Kayak gives recommendation on whether to buy or wait.

We present a simple framework to evaluate the economic value of signals. The model shows that a message may be informative in a statistical sense (it helps predict future prices) although this information has no economic value. In our context, information has economic value only if it influences some traveler's decision to purchase a ticket. For example, the American Airline signal mentioned above may increase consumers' posterior that the flight will not be available the following week. This information, however, has no economics impact if it does not influence travelers' booking decisions.

To formalize this notion of economic value, we consider a rational consumer who can postpone her decision to purchase a ticket by one week. If the consumer expects prices to increase on average, which is often the case for average airfares, her decision depends on how much she values traveling: she waits when prices are above a threshold value and purchase otherwise. We define the economic value of a signal as the increase in utility to the Bayesian traveler who benefit the most from the message. This is a relevant benchmark if consumers respond to messages as expected utility maximizers. Although we cannot test this assumption (because we do not observe consumer bookings), our measure establishes a relevant benchmark because it delivers an upper bound for the value of signals.

We show that the consumer who value the signal the most is the marginal consumer, that is, the consumer who is indifferent (in the absence of signal) between buying today and postponing her decision by a week. Information on prices alone is sufficient to compute the economic value of the signal. The signal is not valuable if and only if the marginal consumer is the same under the two signal realizations. This will be the case, for example, when the two posterior distributions are identical for low price change, although they could differ for high price change, and the signal could predict, for example, the chance that the flight will be available in the future.

We find that scarcity messages are valuable. The average Expedia message increases the expected utility of an unsophisticated traveler, who does not condition her decision on any publicly observable information, by 4.34 percent. For a consumer who conditions her decision on the number of days remaining till departure, the maximum increase in expected utility is 4.12 percent. We also compare the value of information as a function of seller incentives. Airlines on Expedia may use the messages to optimize revenue. Kayak, however, only benefits when a traveler purchases a ticket. Finally, we investigate how the value of messages depend on competition, measured by the number of airlines serving a route, and find that the value of the signal increases with competition.

This work is related to five main strands of literature: (a) Price discrimination and Bayesian persuasion (Lewis and Sappington, 1994; Gentzkow and Kamenica, 2011; Dana, 1998; Deneckere and Peck, 2012; Milgrom and Roberts, 1986); (b) Exploitation of behavioral biases (DellaVigna, 2009; DellaVigna and Gentzkow, 2009); (c) The informational content of Buy/Sell recommendations by financial analysts (Stickel, 1995; Busse et al., 2012; Malmendier and Shanthikumar, 2007); (d) Airline revenue management (McAfee and Te Velde, 2006; Bilotkach and Rupp, 2011; Escobari, 2012; Escobari and Jindapon, 2014; Gallego and Van Ryzin, 1994); (d) Scarcity theory in psychology and marketing (Aggarwal et al., 2011; Aguirre-Rodriguez, 2013; Brock, 1968; Lynn, 1989; Lynn and Bogert, 1996; Lynn, 1991; van Herpen et al., 2014).

The rest of this paper is organized as follows. The next section presents a model of consumer decision making under price uncertainty and derives a measure of the value of information. The following section presents our case study, the data and descriptive statistics. Section 4 presents our main results and the last section concludes.

2 Informational Value of Scarcity Signals

An airline sells a ticket to a traveller who has a fixed value $v \geq 0$. The current price is p_0 and the price next period changes according to $p_1 = (1+r)p_0$ where the growth rate in price, $r \in [-1, \infty)$, is a random variable distributed with c.d.f. $F^u()$. The firm sends a scarcity signal $s \in \{g, b\}$ such that $Pr(s = b) = \tau_b$. The posteriors about the growth rate conditional on the signal realization s, are $F^s()$ such that the three distributions are ordered by first order stochastic dominance (FOSD), $F^b(r) \leq F^u(r) \leq F^g(r)$ for any $r \geq -1$, and

$$F^u(r) = (1 - \tau_b)F^g(r) + \tau_b F^b(r).$$

We denote $H(r) = \int_{-1}^{r} \left(F^g(x) - F^b(x)\right) dx$. We address the following issues: (a) For which value of v does the consumer buy early and when does she wait? (b) How does this decision change with the signal? (c) Derive a measure of the value of the signal. In order to establish a benchmark, we consider a risk neutral Bayesian consumer.

The consumer's date zero utility is $U^0(v) = v - p_0$. In period one, the consumer purchases when $v \geq p_1$, that is, for growth rate realization $r \leq \frac{v}{p_0} - 1$. For ease of notation, we denote $\mathbf{r}(v) = \frac{v}{p_0} - 1$ and $\mathbf{v}(r) = p_0(1+r)$. The expected utility from waiting given information $s \in \{u, b, g\}$ is $U^s(v) = vF^s(\mathbf{r}(v)) - p_0 \int_{-1}^{\mathbf{r}(v)} (1+r)dF^s(r)$. Integration by part gives $\int_{-1}^{r}(1+x)dF^s(x) = (1+r)F^s(r) - \int_{-1}^{r} F^s(x)dx$ and after replacing this expression in $U^s(v)$, we obtain

$$U^s(v) = p_0 \int_{-1}^{\mathbf{r}(v)} F^s(r)dr.$$

If $Er^s < 0$, the price decreases in expectation, and we have $U^s(v) > v - E(p_1|s) > v - p_0 = U^0(v)$ for all v. The consumer waits independently of her valuation. Otherwise, there exists a solution to $U^s(v) = U^0(v)$ which corresponds to the marginal consumer indifferent between buying and waiting. We denote $\rho^s = \infty$ if $Er^s < 0$ and otherwise ρ^s is the solution to

$$\rho^s = \int_{-1}^{\rho^s} F^s(r)dr. \tag{1}$$

Lemma 1. *There exist a unique triplet $\rho^g \geq \rho^u \geq \rho^b \geq 0$. When consumer v has belief $F^s()$, she waits if $v \in [0, \mathbf{v}(\rho^s)]$ and buys early if $v \in (\mathbf{v}(\rho^s), \infty)$.*

The solution to equation (1) is independent of p_0. Holding the distribution of growth rate $F^s()$ constant, the threhold growth rate, ρ^s, does not depend on the level of prices. This implies that in the empirical application we will only need to estimate the function $F^s()$ in order to derive ρ^s. The next Proposition describes how the consumer changes her decision as a function of the signal realization and her valuation.

Proposition 1. *(a) Consumer $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^u)]$ waits without signals. With scarcity signals, she switches to buy when the signal is bad. (b) Consumer $v \in [\mathbf{v}(\rho^u), \mathbf{v}(\rho^g)]$ buys early without a signal. With scarcity signals, she switches to wait when the signal is good. (c) Consumer $v \notin [\mathbf{v}(\rho^b), \mathbf{v}(\rho^g)]$ does the same with and without a signal.*

The introduction of scarcity signals changes both the decision to wait (timing of purchase) and the decision to purchase (a consumer who waits may not buy late). We say that a signal has no economic value if it cannot improve the consumer's decision independently of her valuation v. That is, a consumer is not willing to pay anything for a signal that has no economic value.

Corollary 1. *The signal has no economic value if and only if $H(\rho^u) = 0$.*

The condition in the Corollary is equivalent to $\rho^g = \rho^u = \rho^b$ and to $F^g(r) = F^b(r)$ for $r \leq \rho^u$. The signal has no economic value when the marginal consumer is the same independently of the signal realization. Note that the signal could still be statistically informative. In particular, it could predict availability which is the case when $\lim_\infty F^g(r) > \lim_\infty F^b(r)$. This information, however, is not useful to the consumer if the condition in Corollary 1 holds. Somewhat surprisingly, the signal can be valuable even when prices decrease on average ($Er^u < 0$) and the consumer wait independently of her valuation without signals. A special case where the signal has no value to any consumer happens when $\rho^g = \rho^b = \infty$ which is equivalent to $Er^b \leq 0$ (this implies the condition in the Corollary).

Next, we derive the consumer's benefit from the signal. Denote the utility gain from scarcity signals by $\Delta U(v)$. Using Proposition 1, we obtain

$$\Delta U(v) = \begin{cases} \tau_b(v - p_0) + (1-\tau_b)p_0 \int_{-1}^{\mathbf{r}(v)} F^g(r)dr - p_0 \int_{-1}^{\mathbf{r}(v)} F^u(r)dr, & \text{if } v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^u)] \\ (1-\tau_b)\left(p_0 \int_{-1}^{\mathbf{r}(v)} F^g(r)dr - (v - p_0)\right), & \text{if } v \in [\mathbf{v}(\rho^u), \mathbf{v}(\rho^g)]. \end{cases} \tag{2}$$

The function $\Delta U(v)$ reaches a maximum at $\mathbf{v}(\rho^u)$.[1]

Corollary 2. *Consumer with value $\mathbf{v}(\rho^u)$ receives the highest utility gain from the signal.*

[1] Note that $\Delta U(v)'' = \frac{1-\tau_b}{p_0} f^g(\mathbf{r}(v)) > 0$ for $v \in [\mathbf{v}(\rho^u), \mathbf{v}(\rho^g)]$. For $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^u)]$, $\Delta U(v)'' = \frac{1}{p_0}((1-\tau_b)f^g(\mathbf{r}(v)) - f^u(\mathbf{r}(v)))$ cannot be signed in general. In the case where $F^u()$ and $F^g()$ are uniform with the same support, we have $\Delta U(v)'' < 0$.

Using identity (1), we obtain $\Delta U(\mathbf{v}(\rho^u)) = p_0(1-\tau_b)\int_0^{\rho^u}(F^g(r) - F^u(r))dr$. We define the value of the signal, I, as the utility gain to the marginal consumer measured in relative term

$$I = \frac{\Delta U(\mathbf{v}(\rho^u))}{U^0}.$$

As explained above, only consumers with a valuation $v \in (\mathbf{v}(\rho^b), \mathbf{v}(\rho^g))$ benefit (in expectation) from the signal. Expression I is the value of information for the marginal consumer and it is an upper bound on the value of information across all consumers. After replacement, we have

$$I = \frac{\tau_b(1-\tau_b)}{\rho^u} H(\rho^u). \tag{3}$$

As expected, we have $I = 0$ when the condition in Lemma 1 holds ($H(\rho^u) = 0$). The value of information has the following properties: It is independent of p_0. And ceteris paribus, it increases as there is more uncertainty about the signal realization ($\tau_b(1-\tau_b)$ large), as the consumer has lower threshold (ρ^u small) and as the signal shifts the posterior further apart ($H()$ large). We illustrate with two examples.

Example 1: Change in the Mass Probability at Zero. The distribution of price growth rate displays a large mass probability at $r = 0$ and this mass is smaller under the bad signal than under the good one. Assume the density of growth rates are such that $f^u(r) = f^b(r) = f^g(r)$ for $r \neq 0$ and $Pr(\rho^g = 0) - Pr(\rho^b = 0) = \Delta$. We have $H(r) = r\Delta$ for $r > 0$ and the value of information simplifies to

$$\frac{\Delta U(\mathbf{v}(\rho^u))}{U^0} = (1-\tau_b)\tau_b\Delta.$$

The value of information is independent of the prior $F^u()$ and threshold ρ^u. It increases as the distribution in the good state shifts by a larger amount (Δ large).

Example 2: First Order Stochastic Dominance with Constant Shift. Assume that the conditional posterior are horizontal shifts of the prior: $F^u(r) = F^g(r-\delta_g) = F^b(r+\delta_b)$.[2] The expected discount associated with the good signal, δ_g, is equal to the expected in growth rate under the good signal. We have $H(r) = \frac{1}{\tau_b}\int_r^{r+\delta_g} F(r)dr \approx \frac{\delta_g}{\tau_b}F^u(r)$ where the approximation holds for δ_g small. The value of information simplifies to

$$\frac{\Delta U(\mathbf{v}(\rho^u))}{U^0} \approx (1-\tau_b)\delta_g\frac{F(\rho^u)}{\rho^u}.$$

[2]The assumption is required only for $r \in [-1, \rho^u]$. A special case is when r^u and r^g are uniform $[-r_l, r_h]$ and $[-r_l - \delta_g, r_h - \delta_g]$ respectively. When $s = u$, for example, ρ^u is the highest solution to the quadratic equation $\frac{(x-r_l)^2}{2(r_h - r_l)} = x$, which exist as long as $\frac{r_h - r_l}{2} \geq 0$.

The value of information increases as the good state has a large impact on the distribution of growth rate (δ_g large). The value of the signal is proportional to $(1-\tau_b)\delta_g$. The consumer $\mathbf{v}(\rho^u)$ cares only about the product of the probability that the good realization be drawn and the impact of the good realization on the posterior distribution.[3] Note that this holds only for consumer $\mathbf{v}(\rho^u)$. Holding constant $(1-\tau_b)\delta_g$, the consumers with valuation below $\mathbf{v}(\rho^u)$ prefers a signal with low τ_b. The opposite holds for consumer $v \in [\mathbf{v}(\rho^u), \mathbf{v}(\rho^g)]$.[4]

3 Case Study, Data, and Descriptive Statistics

A traveler can search for a plane ticket directly at the airline's branded site, at an Online Travel Agent (OTA) such as Expedia, Priceline, Travelocity and Orbitz which control 95% of the U.S. OTA market, or visit a meta-search engine such as Kayak, Hopper or Chipmunk that rely on big data analytics to show a variety of price comparisons. Travelers can find a wide variety of advices on the Internet to save on airfare from 'experts', forums, meta-search sites, and blog posts. It is widely accepted that finding the cheapest fare for a given itinerary has a lot to do with timing. Delaying purchasing a ticket can be profitable, especially before 3 weeks prior to departure, because drops in fares, due to slow sales (Escobari, 2012), temporary promotions, or in response to competitive pressure, are not uncommon.

Casual observation suggests that many consumers actively search for low fares. For example, they compare prices across sellers, sign up for fare alerts, and make searches on multiple days.[5] Hopper.com reports that most customers purchase a ticket within two weeks of their initial search. Li et al. (2014) report that about 19 percent of consumers are strategic in that expectations about future prices influence their decision to buy or wait. Beyond that, we are not aware of systematic empirical research on how consumers search for fare.[6]

Sellers offer travelers a variety of information to influence the timing of purchase. A *scarcity*

[3] This is relevant as descriptive statistics indicate that the Kayak signal would correspond to the former case and the Expedia signal to the later one (need to also compare the product $(1-\tau_b)\delta_g$ for the two signals). Despite these differences, the marginal consumer may value the two signals the same.

[4] Take the case of consumer $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^u)]$. Rewrite $\Delta U(v) = \tau_b \left(v - p_0 - p_0 \int_{-1}^{\mathbf{r}(v)} F^u(r)dr\right) + (1-\tau_b)p_0 \int_{-1}^{\mathbf{r}(v)} \left(F^g(r) - F^u(r)\right) dr$. The second term is approximated by $(1-\tau_b)p_0\delta_g F(\mathbf{r}(v))$ which is proportional to the product $(1-\tau_b)\delta_g$. The first term, however, decreases with τ_b since $v - p_0 - p_0 \int_{-1}^{\mathbf{r}(v)} F^u(r)dr < 0$ for $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^u)]$.

[5] *Fare alerts* are email notices sent to subscribers when ticket prices plunges or when it is a good time to purchase a ticket. OTAs, meta-search engines, as well as specialists such as Airfarewatchdog.com, offer fare alert options for specific routes or for other options such as departure city only.

[6] There is much theoretical research on the real option of delaying purchase with fluctuating prices Ho et al. (1998).

signal is defined as information on the number of seats left at a given price. It is used by OTAs and airlines.[7] Kayak instead posts BUY or WAIT *recommendations*, supported by a price forecast, along with a confidence level.[8] The signal realization set is different for scarcity signals (airline/OTA) and recommendations (Kayak). For the former, the realization is a number of seats available at the current price or no signal in which case it is not possible to distinguish whether the seller does not use signals or whether the signal did not report any scarcity. For the later, a signal realization is always provided, at least for departure dates that are not too far in the future. Although the Expedia signals are framed in term of availability, we have shown in the model section that this information is important to a Bayesian consumer only if it changes the marginal consumer, and this cannot happen unless the signal influences the distribution of price changes for values below ρ^u.

3.1 Data Collection

We use a web-scrapping script to collect data on airfares and signals. Many sellers send on a daily basis scarcity signals for a large number of travel itineraries. We select a small subset of sellers, routes, and travel dates. As with past research the sampling is constrained by the time horizon and restrictions on query processing (Edelman, 2012). We end up running daily queries for travels plans that take place at most 100 days in the future. Our dataset is similar to past studies using Internet airfares (Bilotkach and Rupp, 2011; Escobari, 2012; McAfee and Te Velde, 2006) with the shared caveat that what will be learned is sample-specific. We collect signals and prices from two sellers. Following Escobari (2010) and Bilotkach et al. (2010), we use Expedia which is one of the largest OTA worldwide. We also collect BUY/WAIT recommendations from the Kayak metasearch engine. For each seller, we conduct a number of specific searches, or *travel queries*, which comprises a route and departure date (along with a return date in the case of a return trip). For Expedia we consider one-way trips while for Kayak we consider return-trips because Kayak does not display recommendations for one-way trips.

The travel queries span 10 routes (city pairs) and 30 departure and return dates. See Appendix (Section 9) for the list of travel queries. We selected routes with significant gains from purchase timing (high average fares and high variability in fares) and such that sellers are likely to have some information about future fare changes (we include monopoly routes where

[7] Expedia reports "According to the data that we receive from the airline, there are very few tickets currently available at this price. While limited availability can be an indicator that he price for this flight may increase, this is not always the case."

[8] For example, "Price may rise within 7 days. 80% Confidence: Our model has been 80% accurate on forecasting whether these fares will rise or stay within $20 of the current price over the next 7 days. The forecast is based on analysis of historical price changes and is not a guarantee of future results."

competitive dynamics is not a major source of price variations). Many of our routes are the same as the sample of monopoly routes used by (Bilotkach and Rupp, 2011) and rank low on the FAA measures of competition.[9]

For Expedia, a travel query may return a large number of *fight* options. We collect the prices and signals for each option displayed. If we find different flight options for a given travel query on different query dates, we construct an *availability* variable that is equal to zero for all the query dates for which that flight option was not available and one otherwise. For Kayak, we collect for each travel query the lowest price and the BUY/WAIT signal as well as the reported confidence level (see Footnote 8). We conduct travel queries for both sellers each day starting on July 1st 2015 over 30 days.[10] Denote *day-in-advance* (DiA) the number of days between the query day and the departure day. The sample of query and departure dates is such DiA runs from 1 to 100 days.

Figure 1 and 2 present the basic nature of the data. Both figures plot the price and signal realization as a function of DiA. Figure 1 does so for Kayak. It plots the value of the minimum price for a given route (the panels correspond to different routes). Figure 2, corresponding to Expedia, plots the price of a selected set of flights corresponding to a given query (the panels correspond to different flights). Although the signal rarely varies from day to day for a given query, there is much variation across queries in Kayak. For Expedia, we also see much variation in the signal value across flight options for the same query.

3.2 Descriptive Statistics

Table 1 presents summary statistics on the main variables. For Expedia, the price increases on average by 6 percent of the next 7 days. For Kayak, the price increases by 2 percent. The lower mean price growth in Kayak is due to two reasons. Kayak covers observations with lower DiA than Expedia. Moreover, Kayak reports only the minimum price over all flights available in a query. This excludes Expedia flights that have large price increases. As expected, the average price growth increases with the length of the window used to compute changes (1 versus 7 or 14 days).

For Kayak, there is a large chance that the bad signal is sent ($\tau_b = .86$) and the good signal shifts the posterior by a large amount: the price decreases by 1 percent on average over the following week (3 points difference relative to prior). The bad signal instead has a small impact

[9] We also looked at Hopper.com's list of U.S. hubs with highest expected saving.

[10] Fare sales typically last for a few days. Daily price collection minimizes the probability of 'missing' such a fare sale which are available to travelers who check fares on a daily basis.

on the posterior (1 point difference relative to prior). The pattern is reversed for Expedia. There is a small chance of scarcity signal ($\tau_b = .33$) and the signal shifts the posteriors from 6 to 12 percent under the bad signal and to 4 percent under the good signal.

Figure 3 plots the distribution of price change over the next 7 days for Kayak and Expedia. As expected, the Kayak distribution is shifted to the left relative to that of Expedia. Figures 4 and 5 plot the distribution of price change conditional on the signal. The signal shifts the posterior distribution by significant amounts in both cases and the two figures share some similarities: The good signal CDF FOSD the bad signal CDF. Moreover, the probability that price stay the same (jump at $r_7 = 0$) is higher with the good signal. There are two main differences between the two figures. For Kayak, the good signal has a much greater impact on price decreases. For Expedia, we observe a jump at $r_7 = \infty$ and this is because Expedia flights can be non-available (which is coded as an infinite price increase). The probability of non-availability is higher with the bad signal.

The two distributions presented in Figures 4 and 5 are averages over all DiA. One would like to make sure that the patterns observed on these figures remain for subsamples of DiA where airfares are stable. The concern is that low DiA could be associated with more frequent bad signals and higher price growth. Figure 6 reproduces Figure 4 but only for DiA greater than 56 days (more than 8 weeks prior to departure). The main patterns found in Figure 4 remain altough slightly attenuated. We conclude that the signal contains information that is not solely about the changes that take place in the last few weeks before departure.

Tables 2 and 3 report key quantiles of the distributions of price returns for Expedia and Kayak broken down by week. Recall that the signal has no economic value when $Er^b < 0$. This is never the case for Expedia and not the case for most weeks in Kayak with important exceptions in week 5 and 6. In these weeks, $Er^b < 0$ and $I = 0$. Note also, that $Er^g \leq 0$ occurs in both Kayak and Expedia. All consumers (independently of v) should wait under the good signal when this is the case. Finally, the FOSD assumption holds in general with some exception in week 5, 6 and 7 for Expedia and in week 2 for Kayak. Thus, Figures 4 and 5 conceal heterogeneity that could be important when we compute the value of information. This is relevant because a violation of the FOSD assumption can imply negative value of information (which means that the consumer should do the opposite from what the signal advices).

Recall that Example 1 presented the case where the signal shifts the CDF of price returns only at zero. This appears not a bad approximation for Expedia (see Figures 4 and 6). Using the formula for I specific to Example 2, we plug the values $\tau_b = .33$ from Table 1 and approximation $\Delta = .2$ from Figure 4, and obtain the value $I = .042$. This rough approximation suggests that the signal increases the consumer utility by 4.2 percent.

4 Results

PRELIMINARY

4.1 Baseline Value of Information

We compute $(\rho^u, \rho^b, \rho^g, I)$ for an unsophisticated consumer who does not condition her decision to buy/wait on any public information. In order to eyeball ρ^s, Figure 7 plots the functions $f_0(r) = r$ and $f_1(r) = \int_{-1}^{r} F^s(x)dx$. According to equation (1), ρ^s is found where these two functions intercept. Repeating this for the two posteriors (conditional on the signal realization), Table 4 reports the values of (ρ^u, ρ^b, ρ^g).

A consumer with a valuation in the interval $[\mathbf{v}(\rho^g), \mathbf{v}(\rho^b)]$ values the signal because the signal influences her decision to buy or wait. Consumer who value traveling between 6.2 and 11 percent more than the value of the ticket would wait without a signal but prefer to buy early when the signal is bad. Instead, consumers in the range 11 to 16.1 percent would buy early without a signal and change their decision to wait when the signal is good.

We use equation (3) to compute $I = \frac{\tau_b(1-\tau_b)}{\rho^u} H(\rho^u)$. Table 4 reports the value of I. The signal increases the utility of the consumer with valuation $\mathbf{v}(\rho^u)$ by about 4.34 percent. This is not a negligible amount that confirms the approximation presented in the previous Section. Interestingly, the marginal consumer gains the signal not because the signal predicts events when prices decrease. (The two posteriors are very close to one another for $r \leq 0$ on Figure 4.) Instead, the signal helps predict the event that prices remain constant. The consumer gains form waiting when the signal is good because she ends up paying a lower price when the growth rate is positive but lower than her threshold $\rho^u = .11$.

4.2 Consumer Sophistication, Supplier Incentives and Competition

We report the value of I after controlling for a set of conditioning variables to document the influence of traveler sophistication, retailer incentives and competition.

Consumer Sophistication: A concern is that the signal is correlated with public information that is also correlated with price changes. For example, the probability of the Expedia scarcity signal increases with DiA and prices also increase close to the departure date. A sophisticated consumer, who conditions her decision on DiA, may not benefit from the signal if DiA is a sufficient statistics for the signal. Similarly, a sophisticated consumer may condition her posterior on route, airline, or other publicly observable variable. We investigate whether a sophisticated consumer still benefits from the information in the signal. Table 4 reports the value of I for three subsets of DiA. The value of the signal is still significant, although the magnitude of the numbers decreases slightly, when the traveler conditions her purchase decision on DiA. Interestingly, the value of information is negative for $DiA < 28$. This is because the FOSD assumption is violated for low DiA. Since the signal realizations (g, b) play a symmetric role in equation 3, the traveler should do the opposite of what the signal suggests in order to earn the absolute value of the value of I reported.

To put the reported values of I into perspective, we compute how much a sophisticated traveler, who understands that the distribution of price returns depend on DiA, gains relative to an unsophisticated one, who does not base her decision on DiA. The marginal unsophisticated traveler is located at the $\rho^u = .11$ computed using the entire sample. Now, say that the traveler is told that the distribution of returns is different for $DiA \leq 28$ and for $DiA > 28$. The utility gain from becoming sophisticated (learning about DiA) is $I = \frac{\tau_b(1-\tau_b)}{\rho^u} H(\rho^u)$ where τ_b is the fraction of observations in the sample with $DiA \leq 28$ (29 percent), $\rho^u = .11$, $F^u()$ is the distribution of price return in the entire sample, and $F^g()$ is the distribution of price return for $DiA > 28$. Table 4 reports the value of I in this counterfactual thought experiment to be 5.87 percent. The consumer gains more from conditioning her purchase decision on DiA than on the Expedia signal. Still, the gains from the Expedia signal are significant even when compared to the value of sophistication.

Supplier Incentives: The middle panel in Table 4 reports the value of information for the top five airlines in the sample. When a query reports a flight that involves multiple legs offered by different carriers we report it as 'multiple airlines'. The value of information varies a little across airlines.

The second to last panel reports the value of information in the Kayak sample. Kayak's incentives differ from Expedia. While airlines are competing on a route, Kayak benefits if the consumer buys any flight independently of the choice made. Recall from Table 3 that the value of information for Kayak is zero for week 5 and 6. This is because the consumer always wait independently of the signal realization. Over all observations, we find $I = 1.90$. This is a relatively small value of I.

The Kayak signal could also inform a consumer who is purchasing a ticket on Expedia. We can compute the value of the Kayak signal on the matching queries in Expedia. We find that that the Kayak signal increases the consumer utility by a very small margin ($I = 0.56$). This confirms the hypothesis that signals are more informative on the decisions they are supposed to influence.

Competition: We measure the level of competition in the fundamental market at the route level as the number of airlines active on the route. Figure 8 plots the values of I by route against the number of airlines competing on that route. Although the two variable are positively correlated (.22), it is difficult to draw conclusions given the small number of routes in our sample.

In a separate analysis (not reported here) we find that the number of airlines on a route predict the probability that price will not change over the next seven days and the effect is both economically and statistically significant even after controlling for various fixed effects (DiA, route, airline and even flight fixed effect). This offers additional evidence that the information contained in the signal increases with competition.

5 Conclusion

This paper computes the value of scarcity signals to a Bayesian risk neutral consumer in the context of the travel industry. We find that scarcity signals can be valuable to both an unsophisticated traveler (who does not condition her decision on publicly available information) and to a sophisticated one.

These results demonstrate that scarcity signals can have economic value. It is also consistent with the view that scarcity cues cannot be effective if the sender is not trustworthy and her signal not informative. An important issue for future work is to study how consumers respond to scarcity signals: as risk neutral Bayesian as assumed in this work or are they subject to behavioral bias as psychologists and marketing scholars argue?

This work has not addressed the supply side of scarcity signals. What information do airlines and search engines signal (relative to what they know) and how do they try to influence consumers? What is the impact of competition on the information available to consumers?

References

Aggarwal, P., S. Y. Jun, and J. H. Huh. 2011. Scarcity messages. *Journal of Advertising* 40 (3): 19–30.

Aguirre-Rodriguez, A. 2013. The effect of consumer persuasion knowledge on scarcity appeal persuasiveness. *Journal of Advertising* 42 (4): 371–379.

Bilotkach, V., and N. G. Rupp. 2011. A guide to booking airline tickets online. *Available at SSRN 1966729*.

Brock, T. C. 1968. Implications of commodity theory for value change. *Psychological foundations of attitudes* 1:243–275.

Busse, J. A., T. Clifton Green, and N. Jegadeesh. 2012, May. Buy-side trades and sell-side recommendations: Interactions and information content. *Journal of Financial Markets* 15 (2): 207–232.

Dana, Jr, J. D. 1998. Advance-purchase discounts and price discrimination in competitive markets. *Journal of Political Economy* 106 (2): 395–422.

DellaVigna, S. 2009. Psychology and economics: Evidence from the field. *Journal of Economic Literature* 47 (2): 315–372.

DellaVigna, S., and M. Gentzkow. 2009. Persuasion: empirical evidence. Technical report, National Bureau of Economic Research.

Deneckere, R., and J. Peck. 2012. Dynamic competition with random demand and costless search: A theory of price posting. *Econometrica* 80 (3): 1185–1247.

Edelman, B. 2012. Using internet data for economic research. *The Journal of Economic Perspectives*:189–206.

Escobari, D. 2012. Dynamic pricing, advance sales and aggregate demand learning in airlines. *The Journal of Industrial Economics* 60 (4): 697–724.

Escobari, D., and P. Jindapon. 2014. Price discrimination through refund contracts in airlines. *International Journal of Industrial Organization* 34:1–8.

Gallego, G., and G. Van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management science* 40 (8): 999–1020.

Gentzkow, M., and E. Kamenica. 2011. Bayesian persuasion. *American Economic Review* 101 (6): 2590–2615.

Gierl, H., and V. Huettl. 2010. Are scarce products always more attractive? the interaction of different types of scarcity signals with products' suitability for conspicuous consumption. *International Journal of Research in Marketing* 27 (3): 225–235.

Ho, T.-H., C. S. Tang, and D. R. Bell. 1998. Rational shopping behavior and the option value of variable pricing. *Management Science* 44 (12-part-2): S145–S160.

Lewis, T. R., and D. E. Sappington. 1994. Supplying information to facilitate price discrimination. *International Economic Review*:309–327.

Li, J., N. Granados, and S. Netessine. 2014. Are consumers strategic? structural estimation from the air-travel industry. *Management Science* 60 (9): 2114–2137.

Lynn, M. 1989, June. Scarcity effects on desirability: Mediated by assumed expensiveness? *Journal of Economic Psychology* 10 (2): 257–274. 00116.

Lynn, M. 1991. Scarcity effects on value: A quantitative review of the commodity theory literature. *Psychology & Marketing* 8 (1): 43–57.

Lynn, M., and P. Bogert. 1996, November. The Effect of Scarcity on Anticipated Price Appreciation. *Journal of Applied Social Psychology* 26 (22): 1978–1984. 00022.

Malmendier, U., and D. Shanthikumar. 2007, August. Are small investors naive about incentives? *Journal of Financial Economics* 85 (2): 457–489.

McAfee, R. P., and V. Te Velde. 2006. Dynamic pricing in the airline industry. *forthcoming in Handbook on Economics and Information Systems, Ed: TJ Hendershott, Elsevier*.

Milgrom, P., and J. Roberts. 1986. Relying on the information of interested parties. *The RAND Journal of Economics*:18–32.

Mullainathan, S., and E. Shafir. 2013. *Scarcity: Why having too little means so much*. Macmillan.

Stickel, S. E. 1995. The anatomy of the performance of buy and sell recommendations. *Financial Analysts Journal* 51 (5): 25–39.

van Herpen, E., R. Pieters, and M. Zeelenberg. 2014, September. When less sells more or less: The scarcity principle in wine choice. *Food Quality and Preference* 36:153–160. 00001.

Verhallen, T. M., and H. S. Robben. 1994. Scarcity and preference: An experiment on unavailability and product evaluation. *Journal of economic psychology* 15 (2): 315–331.

Worchel, S., J. Lee, and A. Adewole. 1975. Effects of supply and demand on ratings of object value. *Journal of Personality and Social Psychology* 32 (5): 906.

6 Figures

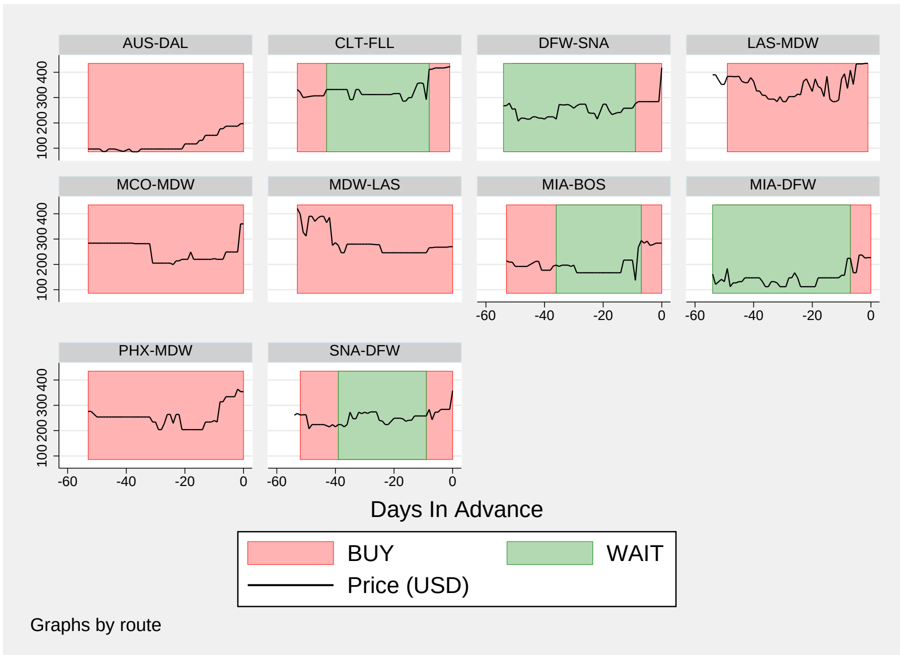

Figure 1: Each panel shows for a different query the Kayak signal (green/red shade) and price (black line) as a function of DiA (horizontal axis).

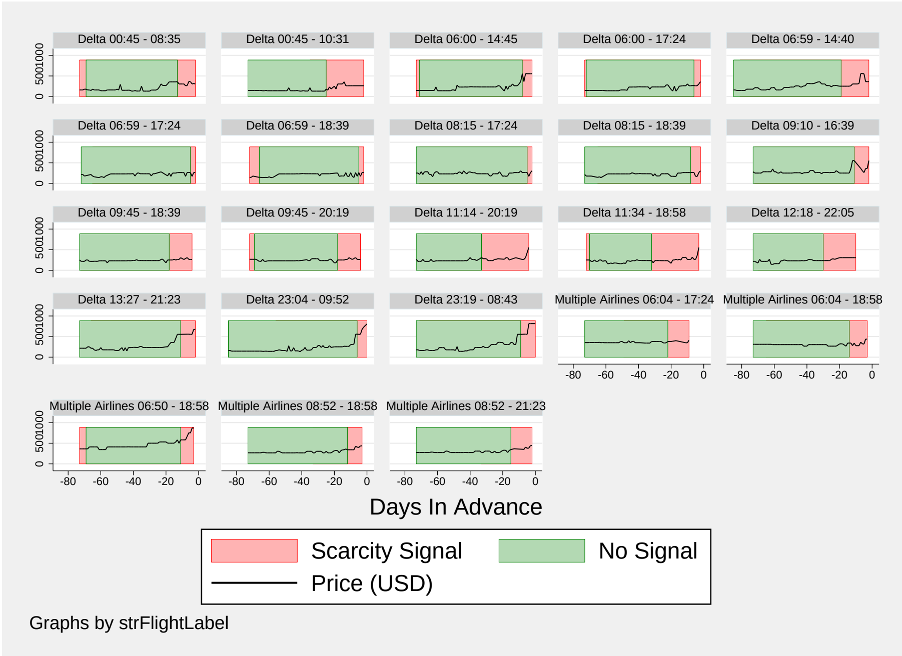

Figure 2: Each panel shows for a given Expedia querie, some of the flights available (panels), and for each flight, non-availability (no value), the value of the signal (green/red shade) and price (black line) as a function of DiA (horizontal axis).

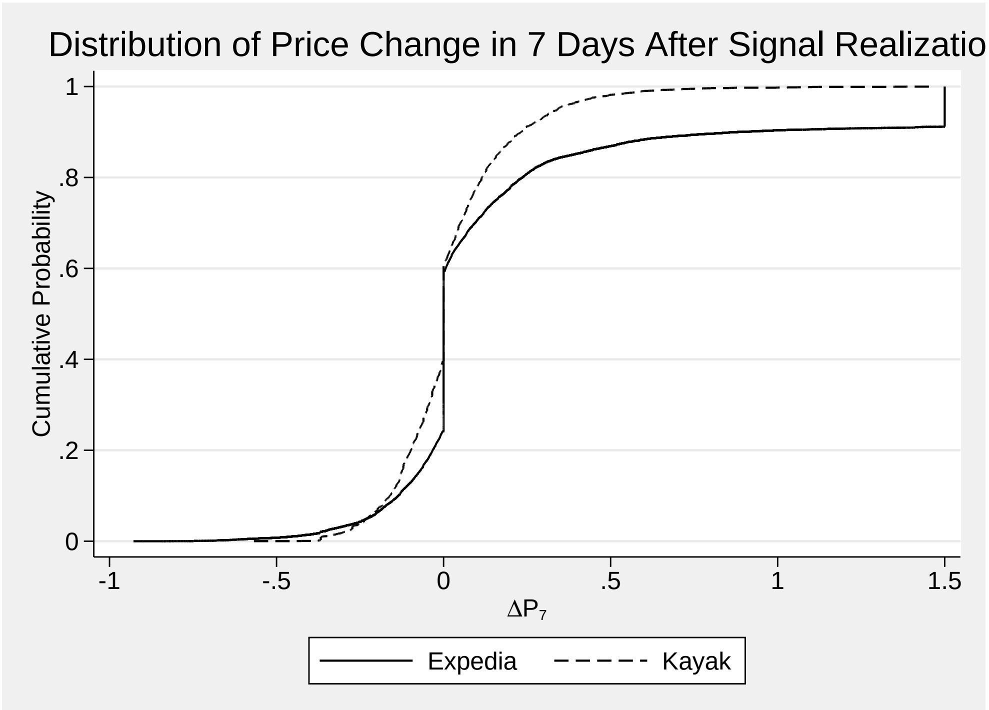

Figure 3: Comparison of Kayak and Expedia distributions of price change over 7 day period.

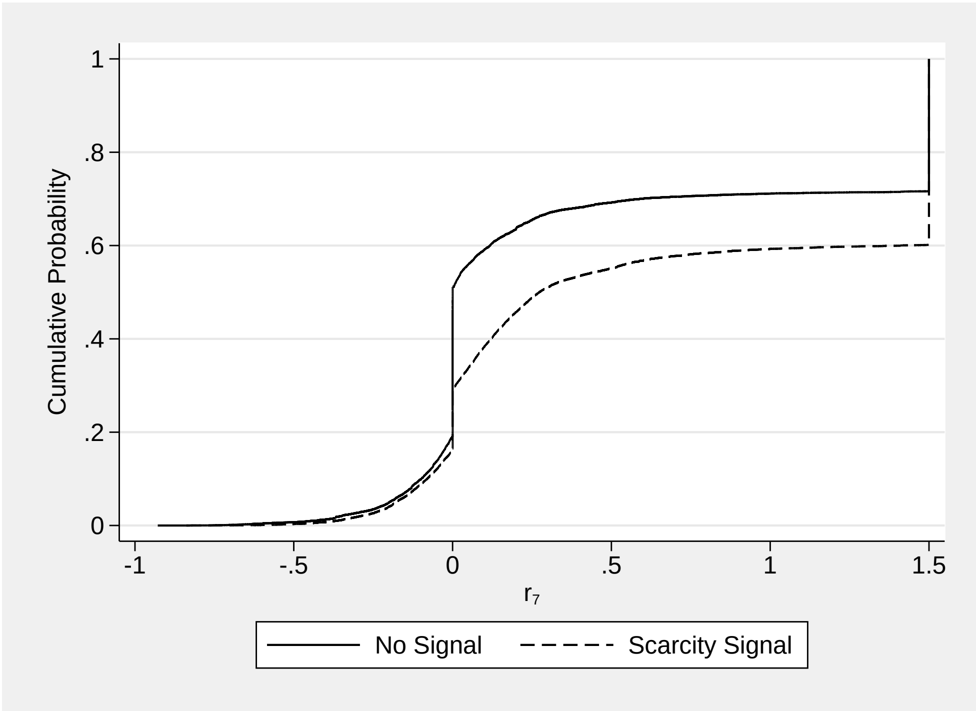

Figure 4: Distributions of Expedia percentage price change over 7 day period as a function of Expedia signal realization.

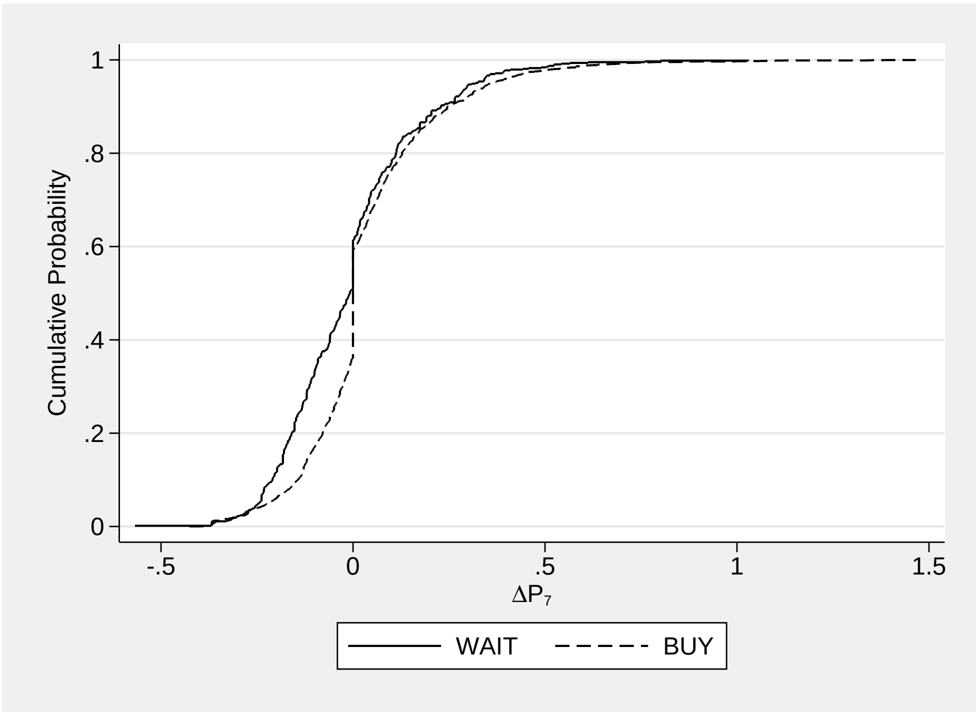

Figure 5: Distributions of Kayak percentage price change over 7 day period as a function of Kayak signal realization.

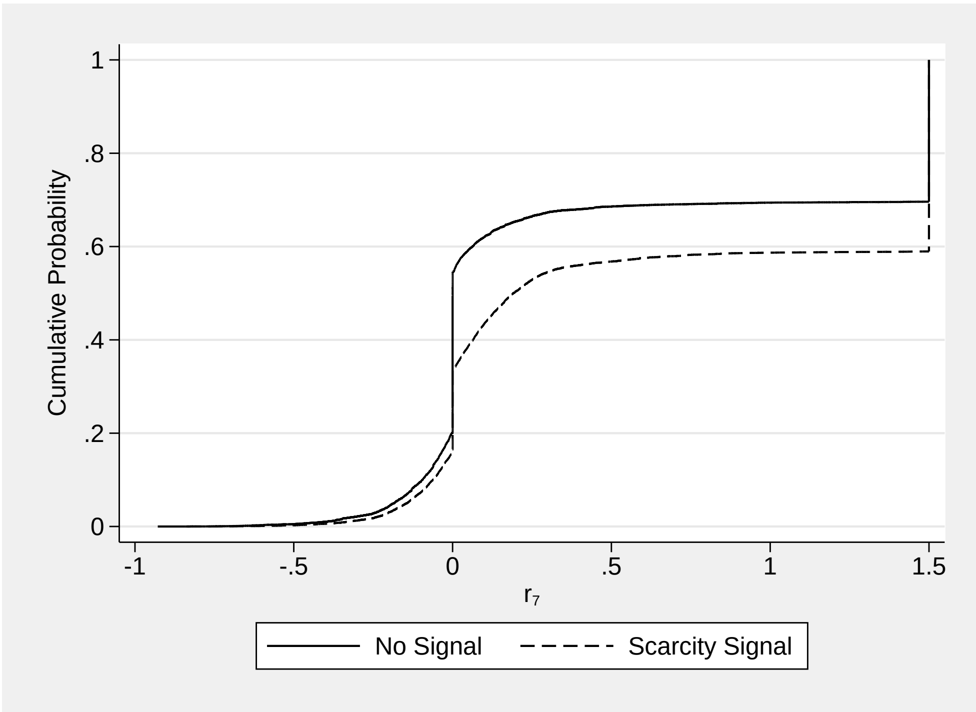

Figure 6: Distributions of Expedia percentage price change for $DiA \geq 56$ over 7 day period as a function of Expedia signal realization.

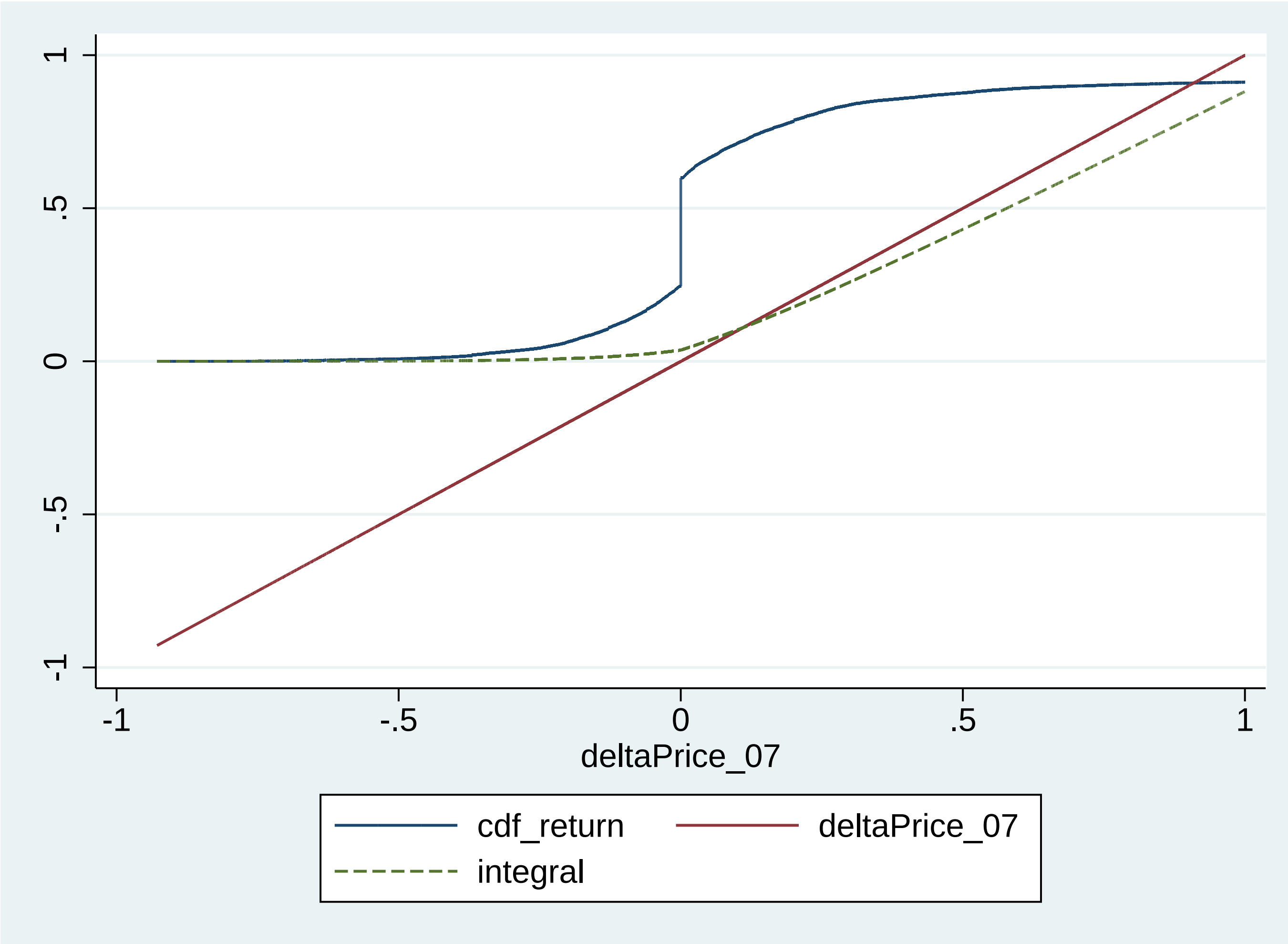

Figure 7: Functions $F^u(r)$, $f_1(r) = \int_{-1}^{r} F^u(x)dx$, and $f_0(r) = r$ computed using entire sample. ρ^u is such that $f_1(\rho) = f_0(\rho)$.

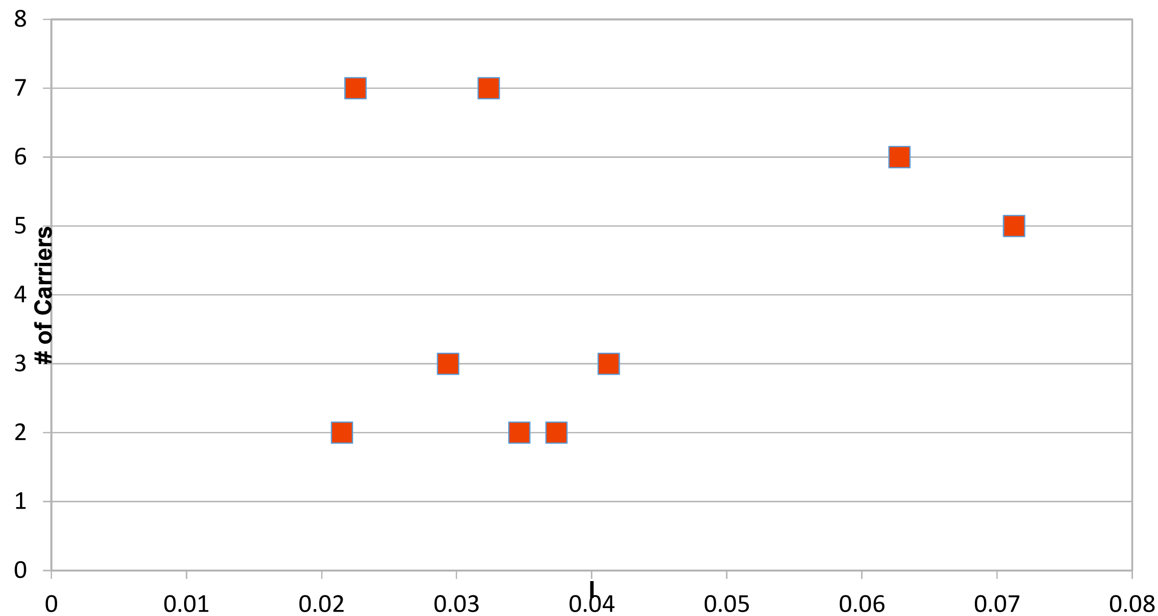

Figure 8: Information value and number of airlines by routes

7 Tables

Table 1: Expedia: Signals, Prices and Availability

Variable	Obs	Mean	Std. Dev.	P25	P50	P75
Expedia Sample						
Signal (1: BAD, 0: GOOD)	566041	.33	.47	0	0	1
Price	566041	299.43	160.41	165.1	274.2	409.1
Price | BAD	189429	295.38	151.07	179.6	267.2	368.6
Price | GOOD	376612	301.47	164.88	161.6	279.1	409.1
r_1	474138	.02	.18	0	0	0
r_1| BAD	146754	.04	.22	0	0	.03
r_1| GOOD	327384	0	.16	0	0	0
r_7	357894	.06	.31	-.01	0	.08
r_7| BAD	101142	.12	.36	-.01	.01	.2
r_7| GOOD	256752	.04	.29	-.01	0	.02
r_{14}	273990	.11	.42	-.04	0	.16
r_{14}| BAD	72920	.18	.48	-.03	.06	.27
r_{14}| GOOD	201070	.08	.39	-.04	0	.1
Kayak Sample						
Signal Sent (BUY/WAIT combined)	5384	.86	.35	1	1	1
Price	6244	260.06	96.45	204	259	312
Price | BUY (BAD)	4617	261.47	100.98	204	259	314
Price | WAIT (GOOD)	767	247.35	72.2	192	258	284
r_1	5876	.01	.11	0	0	0
r_1| BAD	4334	.01	.12	0	0	0
r_1| GOOD	733	-.03	.11	-.05	0	0
r_7	4984	.02	.19	-.07	0	.08
r_7| BAD	3599	.03	.2	-.05	0	.09
r_7| GOOD	628	-.01	.18	-.13	-.01	.07
r_{14}	4006	.03	.24	-.11	0	.12
r_{14}| BAD	2838	.06	.26	-.08	0	.14
r_{14}| GOOD	483	-.01	.22	-.16	-.06	.07

Table 2: Expedia r_{t+7} by Signal

Variable	Obs	Mean	Std. Dev.	P10	P25	P50	P75	P90
Entire Sample								
r_{t+7}	357894	.06	.31	-.15	-.01	0	.08	.29
$r_{t+7}\|B$	101152	.12	.36	-.15	-.01	.01	.2	.47
$r_{t+7}\|G$	256742	.04	.29	-.15	-.01	0	.02	.23
Week 1								
r_{t+7}	11474	.3	.56	-.13	0	.2	.49	.84
$r_{t+7}\|B$	4642	.36	.66	-.21	0	.21	.56	1.03
$r_{t+7}\|G$	6832	.25	.47	-.04	0	.19	.36	.68
Week 2								
r_{t+7}	27221	.24	.49	-.13	0	.14	.36	.68
$r_{t+7}\|B$	11286	.26	.49	-.17	0	.16	.43	.77
$r_{t+7}\|G$	15935	.23	.49	-.05	0	.13	.32	.61
Week 3								
r_{t+7}	29653	.15	.41	-.15	0	.02	.23	.52
$r_{t+7}\|B$	11397	.18	.39	-.17	0	.1	.29	.57
$r_{t+7}\|G$	18256	.13	.43	-.13	0	0	.18	.44
Week 4								
r_{t+7}	30376	.08	.35	-.18	-.02	0	.12	.35
$r_{t+7}\|B$	10601	.12	.35	-.18	-.03	.03	.22	.47
$r_{t+7}\|G$	19775	.06	.35	-.17	-.01	0	.04	.27
Week 5								
r_{t+7}	30084	.03	.29	-.19	-.04	0	.05	.24
$r_{t+7}\|B$	9393	.08	.32	-.17	-.05	0	.16	.35
$r_{t+7}\|G$	20691	.01	.27	-.2	-.04	0	0	.17
Week 6								
r_{t+7}	29740	.03	.25	-.15	-.02	0	.04	.22
$r_{t+7}\|B$	8956	.07	.26	-.15	-.03	0	.14	.31
$r_{t+7}\|G$	20784	.01	.24	-.15	-.02	0	0	.15
Week 7								
r_{t+7}	30133	.02	.22	-.14	-.02	0	.03	.2
$r_{t+7}\|B$	8044	.06	.24	-.15	-.02	0	.13	.28
$r_{t+7}\|G$	22089	.01	.21	-.14	-.02	0	0	.14
Week 8								
r_{t+7}	30384	.02	.22	-.14	-.01	0	.03	.19
$r_{t+7}\|B$	7368	.06	.23	-.12	-.01	0	.12	.26
$r_{t+7}\|G$	23016	.01	.22	-.14	-.02	0	0	.14
Week 9								
r_{t+7}	29557	.02	.23	-.14	-.03	0	.01	.18
$r_{t+7}\|B$	6813	.06	.27	-.13	-.02	0	.11	.26
$r_{t+7}\|G$	22744	0	.21	-.15	-.03	0	0	.13
Week 10								
r_{t+7}	26336	.03	.24	-.14	-.02	0	.01	.17
$r_{t+7}\|B$	5555	.07	.29	-.12	0	0	.12	.28
$r_{t+7}\|G$	20781	.01	.23	-.14	-.02	0	0	.13

Table 3: Kayak r_{t+7} by Signal

Variable	Obs	Mean	Std. Dev.	P10	P25	P50	P75	P90
				Entire Sample				
r_{t+7}	4984	.02	.19	-.16	-.07	0	.08	.23
$r_{t+7}\|B$	3599	.03	.2	-.14	-.05	0	.09	.25
$r_{t+7}\|G$	628	-.01	.18	-.21	-.13	-.01	.07	.23
				Week 1				
r_{t+7}	216	.27	.4	0	.07	.14	.34	.69
$r_{t+7}\|B$	143	.32	.45	.02	.1	.19	.39	.75
$r_{t+7}\|G$	64	.14	.23	-.07	.01	.1	.22	.39
				Week 2				
r_{t+7}	361	.16	.23	-.1	0	.14	.3	.45
$r_{t+7}\|B$	269	.18	.23	-.09	0	.15	.32	.46
$r_{t+7}\|G$	78	.13	.21	-.17	.01	.13	.29	.36
				Week 3				
r_{t+7}	375	.08	.21	-.12	-.03	.03	.19	.34
$r_{t+7}\|B$	310	.09	.22	-.12	-.03	.03	.2	.35
$r_{t+7}\|G$	45	.03	.15	-.16	-.06	0	.11	.27
				Week 4				
r_{t+7}	384	.02	.17	-.15	-.07	0	.09	.22
$r_{t+7}\|B$	329	.03	.16	-.13	-.04	0	.1	.22
$r_{t+7}\|G$	47	-.08	.15	-.27	-.2	-.11	.01	.06
				Week 5				
r_{t+7}	374	-.02	.15	-.19	-.11	0	.02	.18
$r_{t+7}\|B$	303	0	.16	-.18	-.09	0	.04	.18
$r_{t+7}\|G$	44	-.09	.11	-.23	-.16	-.13	-.02	.04
				Week 6				
r_{t+7}	373	-.01	.14	-.18	-.1	0	.04	.16
$r_{t+7}\|B$	305	-.01	.13	-.16	-.07	0	.04	.16
$r_{t+7}\|G$	54	-.04	.17	-.24	-.16	-.04	.04	.11
				Week 7				
r_{t+7}	377	.01	.15	-.17	-.07	0	.06	.19
$r_{t+7}\|B$	316	.01	.16	-.15	-.06	0	.06	.19
$r_{t+7}\|G$	38	-.05	.14	-.23	-.18	-.05	.06	.17

Table 4: Values of (ρ^u, ρ^b, ρ^g), τ_b, I

	ρ^g	ρ^u	ρ^b	τ_b	I $(*100)$
		Expedia: by DiA			
All DiA	.161	.110	.062	.33	4.34
$DiA < 28$	.040	.039	.037	.47	-2.99
$28 \leq DiA < 54$	.734	.240	.107	.33	4.12
$54 \leq DiA < 86$	.688	.304	.251	.25	2.07
		Expedia: $DiA \lesseqgtr 0$ Counterfactual			
All DiA	.287	.110	.039	.29	5.87
		Expedia: By Airlines			
Delta	.221	.129	.073	.36	5.07
American Airline	.194	.143	.082	.37	4.15
Multiple Airlines	.092	.065	.039	.34	2.76
US Airways	.104	.089	.064	.32	2.74
United Airline	.253	.151	.077	.25	4.91
		Kayak: by DiA			
All DiA	.807	.173	.142	.86	1.90
$21 \leq DiA < 28$	.31	.080	.061	.87	6.44
		Kayak Signal On Expedia Flights			
All DiA	.172	.123	.114	.84	0.56

8 Appendix: Proofs

Proof: The function $G^s(x) = x - \int_{-1}^{x} F^s(r)dr$ is increasing in x. We have $G^s(-1) < 0$ and $\lim_\infty G^s(x) = Er^s$. If $Er^s > 0$, equation (1) has a unique solution r^s for $s \in \{u, b, g\}$ and this solution is positive since $G^s(0) < 0$. Moreover, $F^b(r) \leq F^u(r) \leq F^g(r)$ implies $\rho^g \geq \rho^u \geq \rho^b$. Finally, $U^0(v) - U^u(v) = p_0 G^s(\mathbf{r}(v))$. Thus, consumer v stricly prefers to buy early when her belief is $F^s()$ if and only if $G^s(\mathbf{r}(v)) > 0$, or $v \in (0, \mathbf{v}(\rho^s))$. $\square$

Proof: (a) For $v \in [0, \mathbf{v}(\rho^b))$ we have $U^u(v) > U^0(v)$ and $U^s(v) > U^0(v)$. Thus, the consumer prefers to wait with or without signal. (b) For $v \in (\mathbf{v}(\rho^b), \mathbf{v}(\rho^u))$ we have $U^u(v) > U^0(v)$, $U^g(v) > U^0(v)$ and $U^b(v) < U^0(v)$. Thus, the consumer buys early only if the signal is bad. (c) For $v \in (\mathbf{v}(\rho^u), \mathbf{v}(\rho^g))$ we have $U^0(v) > U^u(v)$, $U^g(v) > U^0(v)$ and $U^b(v) < U^0(v)$. Thus, the consumer buys early if the signal is bad or without a signal. (d) For $v > \mathbf{v}(\rho^g)$ we have $U^0(v) > U^u(v)$ and $U^s(v) < U^0(v)$. Thus, the consumer prefers to buy early with or without signal. $\square$

Proof: $\Delta U(v)$ increases from $\mathbf{v}(\rho^b)$ to $\mathbf{v}(\rho^u)$, and decreases from $\mathbf{v}(\rho^u)$ to $\mathbf{v}(\rho^g)$. $\frac{\partial}{\partial v}\Delta U(v) = \tau_b + (1-\tau_b)F^g(\mathbf{r}(v)) - F^u(\mathbf{r}(v)) \geq 0$ for $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^u)]$ and $\frac{\partial}{\partial v}\Delta U(v) = (1-\tau_b)(F^g(\mathbf{r}(v)) - 1) \leq 0$ for $v \in [\mathbf{v}(\rho^u), \mathbf{v}(\rho^g)]$. $\square$

9 Appendix: Dataset

We use the following terminology: A *route* is an origin/destination pair defined by the three letter international airport codes; A *querie* consist of a route and departure date; Each query is submitted on different *booking dates* and return a number of flight options; Each *flight* is identified by a querie, a departure time and a carrier. Therefore, the nesting goes from route to query to flight. We construct the variable days-in-advance DiA as the difference between departure date and booking date.