The Value of Online Scarcity Signals *

Pascal Courty[†] Sinan Ozel[‡]

October 25, 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 economic value of scarcity messages posted by Expedia. A signal reveals information on the number of seats available at the posted price. Consumers can use this information to optimally time when they purchase a ticket. The maximum gain in expected utility for a naive consumer, who does not use publicly available information, is 8 percent. For a sophisticated consumer the increase is between 4-7 percent. Scarcity signals have a negligible impact on seller revenue and consumption.

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 even if the messages 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 the purchase 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 sometimes mentions 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 travelers time their booking decisions.

We develop a Bayesian rational framework to evaluate the economic value of signals. The rationality assumption may not apply to all travelers. But one can derive powerful predictions that can be tested with publicly available information. An important contribution of this paper is to show data on prices and signal realization alone is sufficient to compute the economic value of the signal to a rational consumer. We implement the theory using online information that was collected using a web scraping script (Edelman, 2012). Our approach offers a relevant benchmark for the consumers who respond to messages as expected utility maximizers. Although we cannot test this assumption (because we do not observe consumer bookings), our measure delivers an upper bound for consumer benefits.

To formalize the 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 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 information. The model shows that a signal may be informative in a statistical sense (it helps predict future prices) although this information has no economic value. In the model, information has economic value only if it influences the 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, may not influence travelers' booking decisions. 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. The signal is not valuable if and only if the marginal consumer is the same under the two signal realizations.

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.

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, derives a measure of the value of information and computes the impact of the signal on seller revenues and consumption. The following section presents our case study, the data and descriptive statistics. Section 4 presents our main results and the last section concludes.

2 A Model of Scarcity Signals

An airline sells a ticket to a traveler. The traveler knows her willingness to pay v at the time of purchase. Let $G(v)$ denote the cumulative distribution of v. 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. $F^n()$. The firm sends a scarcity signal that takes realization bad (b) and good (g) such that the bad realization occurs with probability τ_b. The posteriors about the growth rate conditional on the signal realization are $F^b(r)$ and $F^g(r)$. Bayes rule requires that

$$F^n(r) = (1-\tau_b)F^g(r) + \tau_b F^b(r). \tag{1}$$

For example, the good state could imply a better distribution of price in the sense of first order stochastic dominance, $F^g(r) \geq F^b(r)$. Note, however, that this assumption is not required in the analysis. We denote the mean of the three expected growth rates $\overline{r}^s$, for $s \in \{n, g, b\}$ and $H(r) = \int_{-1}^{r} \left(F^g(y) - F^b(y)\right) dy$. The consumer is risk neutral consumer and Bayesian. The model addresses the following issues: (a) For which value of v does the consumer buy early? (b) How does the signal realization influence this decision? (c) What is the utility increase associated with the signal? (d) What is the change in supplier revenues and units sold?

We illustrate some of the results with two examples that are used for a didactic end but also to show that some results hold only for specific functional forms. In the first example, the signal delivers information on the probability that the price remains constant. The distribution of price growth rate has a mass probability at zero and this mass is smaller under the bad signal than under the good one. The distributions of growth rates are such that $F_1^n(r) = F_1^b(r)$ for $r < 0$ and $F_1^b(r) = F_1^n(r) - (1-\tau_b)x$ for $r \geq 0$. This implies $H_1(r) = xr$ for $r > 0$. In the second example, the signal shifts the cumulative distribution function by a constant. The conditional posterior are horizontal shifts of the prior: $F_2^b(r) = F_2^n(r-x)$. We have $H_2(r) = \frac{1}{1-\tau_b}\int_{r-x}^{r} F^n(y)dy$. The expected growth rate under the bad signal is equal to the expected in growth rate under the prior plus x.[1]

2.1 Informational Value of Scarcity Signals

The consumer's expected utility from buying in current period is $v - p_0$. Next period, the consumer purchases when $v \geq p_1$, that is, for price returns $r \leq \frac{v}{p_0} - 1$. We denote $\mathbf{r}(v) = \frac{v}{p_0} - 1$ and $\mathbf{v}(r) = p_0(1+r)$ to facilitate going back and forth from valuation to equivalent return. The

[1] Using equation (1), we obtain $F_1^g(r) = F_1^n(r) + \tau_b r$ for $r \geq 0$ and $F_2^g(r) = \frac{F_2^n(r) - \tau_b F_2^n(r-x)}{1-\tau_b}$.

expected utility from waiting given belief F^s is $E(Max(v-p_1,0)|s) = p_0 \int_{-1}^{\mathbf{r}(v)} F^s(r)dr$.[2] If $\overline{r}^s < 0$, the price decreases in expectation, and we have $E(Max(v-p_1,0)|s) > v - E(p_1|s) > v - p_0$ for all v. The consumer waits independently of her valuation. Otherwise, there exists a solution v^s to the indifference condition

$$E(Max(v^s - p_1, 0)|s) = v^s - p_0.$$

This is the valuation of the consumer indifferent between buying and waiting. We also call that consumer the marginal consumer. The indifference condition is rewritten $\int_{-1}^{\mathbf{r}(v^s)} F^s(r)dr = \mathbf{r}(v^s)$. We define $\rho^s = \infty$ if $\overline{r}^s < 0$. Otherwise, ρ^s is the solution to

$$\rho^s = \int_{-1}^{\rho^s} F^s(r)dr. \tag{2}$$

The solution to equation (2) is independent of the initial price p_0.

Lemma 1. *There exist a unique a triplet (ρ^b, ρ^n, ρ^g) such that $Min(\rho^g, \rho^b) \leq \rho^n \leq Max(\rho^g, \rho^b)$. When consumer v has belief $F^s()$, she waits if $v \in [0, \mathbf{v}(\rho^s))$ and buys early if $v \in (\mathbf{v}(\rho^s), \infty)$. Her expected utility is:*

$$U^s(v) = \begin{cases} p_0 \int_{-1}^{\mathbf{r}(v)} F^s(r)dr, & \text{if } v \in [0, \mathbf{v}(\rho^s)] \\ v - p_0, & \text{if } v \in [\mathbf{v}(\rho^s), \infty]. \end{cases} \tag{3}$$

Lemma 1 says that (ρ^b, ρ^g) lie on each side of ρ^n. For example, a sufficient condition for $\rho^b \leq \rho^g$ is $F^g(r) \geq F^b(r)$ for all r. In the rest of the analysis, we label the two states such that $\rho^b \leq \rho^g$, which is a matter of convention. Under this usage, we obtain the intuitive outcome that the bad signal triggers some consumers to change their decision to wait and the good signal triggers some consumers to change their decision to buy, in a sense that is formally defined in the following Proposition.

Proposition 1. *(a) Consumer $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$ waits without signals. With scarcity signals, she switches to buy when the signal is bad. (b) Consumer $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$ 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 in period one). Consumer v's

[2] We have $E(Max(v-p_1,0)|s) = vF^s(\mathbf{r}(v)) - p_0 \int_{-1}^{\mathbf{r}(v)} (1+r)dF^s(r) = p_0 \int_{-1}^{\mathbf{r}(v)} F^s(r)dr$.

gain from the signal is $\Delta U(v) = U^n(v) - \left(\tau_b U^b(v) + (1-\tau_b)U^g(v)\right)$. We show in the Appendix that

$$\Delta U(v) = \begin{cases} \tau_b \left(v - p_0 - p_0 \int_{-1}^{\mathbf{r}(v)} F^b(r)dr\right), & \text{if } v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)] \\ (1-\tau_b)\left(p_0 \int_{-1}^{\mathbf{r}(v)} F^g(r)dr - (v - p_0)\right), & \text{if } v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]. \end{cases} \tag{4}$$

We say that a signal has no economic value if it cannot improve the consumer's decision independently of her valuation v. A consumer is not willing to pay anything for a signal that has no economic value.

Corollary 1. *A signal has no economic value if and only if* $\rho^g = \rho^n = \rho^b$.

The signal has no economic value when the marginal consumer is the same independently of the signal realization. A random signal, for example, has no economic value: We have $F^g() = F^b()$ which implies the condition in the Corollary. Another special case where the signal has no value to any consumer happens when $\rho^g = \rho^b = \infty$ which is equivalent to $\overline{r}^b \leq 0$ (this implies the condition in the Corollary). The consumer waits independently of her valuation or of the signal realization. The signal is worthless but could still be statistically informative (for example, if $\overline{r}^b \neq \overline{r}^g$). Interestingly, a signal that helps predict availability ($\lim_\infty F^g(r) > \lim_\infty F^b(r)$) is not useful to the consumer if the condition in Corollary 1 holds. On the contrary, and somewhat counter to intuition, the signal can be valuable even when prices decrease on average ($\overline{r}^n < 0$).

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

Using identity (2), we obtain $\Delta U(\mathbf{v}(\rho^n)) = p_0(1-\tau_b)\int_0^{\rho^n}(F^g(r) - F^n(r))dr$. We define the value of the signal, I, as the relative utility change to consumer $\mathbf{v}(\rho^n)$

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

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

$$I = \frac{\tau_b(1-\tau_b)}{\rho^n} H(\rho^n). \tag{5}$$

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

Take example 1. Since $H_1(r) = rx$ for $r > 0$, the value of information simplifies to $I_1 = (1-\tau_b)\tau_b x$. The value of information is independent of the prior $F_1^n()$ and of threshold ρ^n. It increases as the signal shifts the probability of no price change by a larger amount (x large). Take example 2. We have $H_2(r) \approx \frac{x}{1-\tau_b} F_2^n(r)$, where the approximation holds for x small, and $I \approx \tau_b x \frac{F_2(\rho^n)}{\rho^n}$. The value of information increases as the signal shifts the distributions of growth rate further appart (x large). It is proportional to $\tau_b x$. The consumer $\mathbf{v}(\rho^n)$ cares only about the product of the probability that the bad realization be drawn and the impact of the bad realization on the posterior distribution. Note that this holds only for consumer $\mathbf{v}(\rho^n)$. Holding constant $\tau_b x$, the consumers with valuation below $\mathbf{v}(\rho^n)$ prefers a signal with low τ_b. The opposite holds for consumer $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$.[3]

2.2 Supplier Revenues and Consumption

The model has only two periods and a single consumer. Within this restricted framework, it only makes sense to discuss the impact of the signal on expected consumer revenue and the probability that the consumer buys. The bigger picture is profits and welfare. But investigating the impact of scarcity signals on profits would require modeling dynamic trade-offs that are beyond the scope of this model.[4] Denote by $\Delta R(v)$ the difference in revenue, received from the consumer with valuation v, with and without a signal. $\Delta R(v) = 0$ for $v \notin [\mathbf{v}(\rho^b), \mathbf{v}(\rho^g)]$ and

$$\Delta R(v) = \begin{cases} p_0 \tau_b \left(1 - \int_{-1}^{\mathbf{r}(v)} (1+r) dF^b(r)\right), & \text{if } v \in (\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)) \\ -p_0(1-\tau_b)\left(1 - \int_{-1}^{\mathbf{r}(v)} (1+r) dF^g(r)\right), & \text{if } v \in (\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)). \end{cases} \tag{6}$$

Take the top line in the above equation. Traveler $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$ waits without the signal. The supplier earns $p_0 \int_{-1}^{\mathbf{r}(v)} (1+r) dF^n(r)$. With a signal the traveler buys early when the realization is bad and waits otherwise. The supplier earns p_0 in the former case and $p_0 \int_{-1}^{\mathbf{r}(v)} (1+r) dF^g(r)$ in the latter one. The expected supplier revenues with a signal are $\tau_b p_0 + (1-\tau_b) p_0 \int_{-1}^{\mathbf{r}(v)} (1+r) dF^g(r)$.

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

[4] Under the assumption of zero cost of capacity, change in welfare is equal to change in consumer surplus $v \Delta C(v)$. However, there is an opportunity cost of capacity under dynamic revenue management. The supplier sells capacity to consumers who arrive continuously till the departure date. Computing welfare is beyond the scope of the current work.

Taking the difference between the two revenues, and recalling that $dF^n(r) = \tau_b dF^b(r) + (1 - \tau_b)dF^g(r)$, gives the top expression in equation (6). The revenues for $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$ are computed similarly.

Lemma 2 in the Appendix shows that the function $\Delta R(v)$ is equal to zero up to $\mathbf{v}(\rho^b)$, at which point it jumps to a positive value, decreases up to $\mathbf{v}(\rho^n)$ where it drops to a negative value, then increases up to $\mathbf{v}(\rho^g)$ where it is still negative and where it finally jumps back to zero. The supplier loses from travelers with valuation $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$. She gains from travelers with valuation $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$ under a condition that holds in our application. Figure 9 in the Appendix plots $\Delta\overline{R}(v)$ for a distribution corresponding to example 1.

The signal also changes the expected probability of purchase. This is important to the seller because inventory is central to revenue management.

$$\Delta C(v) = \begin{cases} \tau_b(1 - F^b(\mathbf{r}(v)), & \text{if } v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)] \\ -(1-\tau_b)(1 - F^g(\mathbf{r}(v)), & \text{if } v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]. \end{cases} \tag{7}$$

The signal changes the composition of consumers who end up travelling: A consumer with valuation $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$ is more likely to consume and the opposite holds for a consumer with valuation above $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$. Stated differently, the signal has a positive impact on the set of consumers who wait in the sense that the average consumer in that pool has a higher valuation.

Averaging across all consumer valuations, the expected change in supplier revenues is $\Delta\overline{R} = \int_{\mathbf{v}(\rho^b)}^{\mathbf{v}(\rho^g)} \Delta R(v) dG(v)$. The expected change in consumption (quantity sold) is $\Delta\overline{C} = \int_{\mathbf{v}(\rho^b)}^{\mathbf{v}(\rho^g)} C(v) dG(v)$. The sign and magnitude of $\Delta\overline{R}$ and $\Delta\overline{C}$ depend on the four primitives $(\tau_b, F^n(), F^b(), G())$. It is not possible to evaluate these expressions in the absence of information about $G()$. One can make progress, however, by looking at 'small' signals. Takes the family of binary signals that are generated through a linear combination of the prior F^n and a perturbation S, as in, $F(r, y) = F^n(r) + yS(r)$. The prior's cumulative distribution is $F^n(r) = F(r, 0)$. The bad signal occurs with probability τ_b and the associated posterior cumulative distribution is $F^b(r|x) = F(r, -x(1 - \tau_b))$. Equation (1) says that the good signal's posterior is $F^g(r|x) = F(r, x\tau_b)$. We assume that the posteriors are well defined cumulative distribution function.[5] A signal with $x = 0$ does not convey any information. As the value of x increases so does the weight put on S. We denote $\Delta\overline{R}(x) = R(x) - R(0)$ the changes in expected revenues

[5] Take Example 1 as an illustration. We have $S_1(r) = 1$ if $r \geq 0$ and we obtain $F_1^b(r|x) = F^n(r) - (1 - \tau_b)x$ if $r \geq 0$ and $F^b(r|x) = F^n(r)$ otherwise. The parameter x measures the change in the probability that the price remains constant.

associated to signal x. $\Delta\overline{C}(x)$ is similarly defined.

Proposition 2. *For small x, the signal has no first order impact on revenue and consumption ($\Delta\overline{R}'(0) = \Delta\overline{C}'(0) = 0$). Revenue increases ($\Delta\overline{R}''(0) > 0$) if and only if $\frac{f(\rho^n)}{1-F(\rho^n)} - \frac{g'(v^n)}{g(v^n)}p_0 > \frac{2}{1+\rho^n}$. Consumption increases ($\Delta\overline{C}''(0) > 0$) if and only if $\frac{f(\rho^n)}{1-F(\rho^n)} - \frac{g'(v^n)}{g(v^n)}p_0 > 0$.*

Corollary 2 is related to the Bayesian persuasion literature. Gentzkow and Kamenica (2011) consider a sender who can inform a single receiver. They characterize the optimal multi-dimensional signal that satisfies Bayes plausibility (condition in equation 1). Our application has multiple receivers. Each receiver's private type is her valuation v. The optimal signal depends on the type distribution $G(v)$ which is unknown. To get around this problem, we consider only 'small' changes in posteriors that influence only the receivers in a small neighborhood of the marginal consumer (type v^n). We also restrict to binary signals. This is not restrictive when signals are coarse which holds with our application. A contribution of Proposition 2 is to relate the primitive of the problem to empirically testable conditions. When $g'(v^n) = 0$, the signal increases consumption. It also increases revenue if and only if $\frac{f(\rho^n)}{1-F(\rho^n)} > \frac{2}{1+\rho^n}$. In general, the signal is more likely to increase revenue and consumption when there are more consumers to the left of the marginal consumer than to the right ($g' < 0$).[6] This is simply because the seller always looses from the consumers on the right of the marginal consumer and benefits (at least for some) of the consumers on the left. Note that the characteristic of the signal ($\tau_b, S()$) do not influence whether the signal is profitable or not. It influences the scale of the impact.[7] A signal is more profitable if $\tau_b = .5$ and if $S()$ puts more weight on low growth rate realizations.

3 Case Study, Data, and Descriptive Statistics

3.1 Online Travel Booking and Exedia Scarcity Signals

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' blogpost, forums or meta-search sites. It is widely accepted that finding the cheapest fare for a given itinerary has a lot to do

[6]The seller's revenue in the absence of signal, $R(v) = p_0(1 - G(p_0))$, is concave when $-\frac{g'(v^n)}{g(v^n)}p_0 < 2$, which does not exclude the possibility that $g'(v^n) < 0$.

[7]In the Appendix, we show that $\Delta\overline{R}''(0)$ and $\Delta\overline{C}''(0)$ are proportional to $\tau_b(1 - \tau_b)\left(\int_{-1}^{\rho^n} S(r)dr\right)^2$.

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.[8] 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.[9]

A *scarcity* signal is defined as information on the number of seats left at a given price. It is used by OTAs and airlines.[10] When no signal is posted, it is not possible to distinguish whether the seller does not use signals or whether the signal did not report any scarcity. The Expedia signals are framed in term of availability of seat at a given price. This may be because airlines revenue management systems make only a number of a tickets available for each fare class. Airlines and OTAs may warn travelers when the number of tickets remaining at a given price is low. It is also possible that either or both strategically manipulate the information they send consumers. These strategic and vertical (between the airline and the OTA) issues are not covered here. The model takes the signal, and posterior distributions, as given.

3.2 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. Following Escobari (2010) and Bilotkach et al. (2010), we use Expedia which is one of the largest OTA

[8] *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.

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

[10] 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.”

worldwide. We conduct a number of specific searches, or *travel queries*, for one-way trips. A query comprises a route and departure date. The methodology used to collect the data is described in the appendix (Section 10).

The travel queries span 10 routes (city pairs) and 22 departure dates. We selected routes with a single non-stop carrier and significant gains from purchase timing according to hopper.com. The selected routes rank low on the FAA measures of competition. For these routes, we expect sellers to have more information about future fare changes because price randomness associated with competitive dynamics is less important. Many of our routes are the same as the sample of monopoly routes used by (Bilotkach and Rupp, 2011).

A travel query may return a large number of *fight* options. We collect the prices and signals for each option displayed. The signal is a dummy variable that is equal to one if a scarcity message is posted. If we find different flight options for a given travel query on different booking dates, we construct an *availability* variable that is equal to zero for all the booking dates for which that flight option was not available and one otherwise. We conduct travel queries each day starting between the 19th of July and the 26th of October, 2015.[11] We started with a given set of departure date and added new ones as existing ones expire. Denote *day-in-advance* (DiA) the number of days between the booking day and the departure day. Due to the sampling methodology, DiA is about evenly distributed between 1 to 100 days.

Figure 1 presents the basic nature of the data. The figure plots the price and signal realization as a function of DiA for 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, we see much variation in the signal value across flight options for the same query.

3.3 Descriptive Statistics

Table 1 presents summary statistics on the main variables. The price increases on average by 6 percent over the next 7 days. As expected, the average price growth increases with the length of the window used to compute changes (1 versus 7 or 14 days). 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.

The top panel on Figure 2 plots the distribution of price change conditional on the signal. The signal shifts the posterior distribution by significant amounts: The good signal CDF first-

[11] 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.

order stochastically dominates (FOSD) the bad signal CDF. Moreover, the probability that the price stays the same (jump at $r_7 = 0$) is higher with the good signal. There is also 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 distribution presented on the top panel on Figures 2 are averages over all DiA. One would like to make sure that the patterns observed on this figure 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. The bottom panel (Figure 2) reproduces the top panel but only for DiA greater than 56 days (more than 8 weeks prior to departure). The main patterns found on the top panel remain although 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.

Table 2 reports key quantiles of the distributions of price returns broken down by week. Recall that the signal has no economic value when $\overline{r}^b < 0$. This is never the case. The two distributions are ordered by FOSD with some exception (in weeks 2-7, the first decile is weakly smaller in the bad state). Thus, Figure 2 conceals heterogeneity that could be important when we compute the value of information. This is relevant because a violation of FOSD can imply negative value of information (which means that the consumer should do the opposite from what the signal advices).

Figure 2 reveals that the two PDFs differ mostly at $r = 0$: Under the good signal there is a greater probability of no price change. Thus, Example 1 may not be a bad approximation of what the signal does to the two posterior CDFs. Using the formula for I specific to Example 1, we plug the values $\tau_b = .33$ from Table 1 and approximation $x = .2$ from Figure 2, and obtain the value $I = .042$. This rough approximation suggests that the signal increases the consumer utility by 4.2 percent.

4 Results

Given the close connection between the model and data, the empirical analysis uses simple statistics: The empirical prior and the two posteriors $(F^n(), F^b(), F^g())$ are computed to estimate non-parametrically the values of ρ^g, ρ^n, ρ^g and I. We estimate these values using the empirical distributions and report equal-tail confidence intervals that are computed using 'case resampling' bootstrapping.[12] In the computations presented in the Tables, we normalize all expressions assuming $p_0 = 1$ which implies $\mathbf{v}(r) = 1 + r$. The value of v is expressed in units of p_0. Since the average ticket price in the sample is close to \$300, we sometimes multiply the results by that amounts to present monetary figures.

4.1 Baseline Value of Information (Table 6)

We compute $(\rho^n, \rho^b, \rho^g, I)$ for an unsophisticated consumer who does not condition her decision to buy/wait on public information. In order to eyeball ρ^s, Figure 3 plots the function $f_1(r) = \int_{-1}^{r} F^s(x)dx$ and the forty-five degree line. According to equation (2), ρ^s is found where these two intercept. Repeating this for the two posteriors (conditional on the signal realization), the top line of Table 6 reports the values of (ρ^n, ρ^b, ρ^g) for the entire sample.

Consumer who value traveling between 2.9 and 4.5 percent more than the price of the ticket would wait without a signal but prefer to buy early when the signal is bad. Instead, consumers in the range 4.5 to 5.7 percent would buy early without a signal and change their decision to wait when the signal is good. Stated in dollar amounts, a consumer with a valuation in the interval $[\mathbf{v}(\rho^g), \mathbf{v}(\rho^b)] = [\$308.7, \$317.1]$ respond the signal by changing their decision depending on the realization.

We use equation (5) to compute $I = \frac{\tau_b(1-\tau_b)}{\rho^n} H(\rho^n)$. Table 6 column 6 reports the value of I. The signal increases the utility of the consumer with valuation $\mathbf{v}(\rho^n)$ by about 8.3 percent. This is not a negligible amount that is higher than the approximation presented in the previous Section because that approximation took into account only the change in probability that the price remain constant. This higher figure says that the signal does not only help the traveller predicts events when prices are constant ($r = 0$) but also when prices are more likely to decrease (on Figure 2, the good posterior lays above the bad one for $r < 0$).

Figure 4 plots the percentage utility change for the consumers who respond to the signal. The consumer who benefits the most correspond to the marginal consumer used to compute the value of I. The value of information is positive, peaks at ρ^n, and has a tent shape. The average gain amongst the consumers who respond to the signal, $\frac{\Delta \overline{U}}{\rho^g - \rho^b}$, is reported in Table 6 as 0.00186 corresponding to a dollar value of \$.56. The average percentage utility increase amongst the consumers who respond to the signal is $\frac{\Delta \overline{U}}{\overline{U}} = 4\%$. The percentage utility increase is large relative to the absolute utility increase because the consumers who respond to the signal receive

[12] We re-compute all statistics reported for each bootstrap sample. The bootstrap samples depend on the original data subsample used to compute the statistic. The data subsamples are: entire sample, decomposition by DiA, DiA28-56XCarrier, DiA28-56XRoute. For each data subsample, we draw different bootstrap samples.

a small surplus in the abscence of signal (their willingness to pay is close to the initial price).

Although we find that the signal has value in our sample, one may argue that any signal is valuable. As a robustness check, we make sure that this is not the case. Corollary 1 says that a random signal should have no value. To conduct a placebo test, we draw a thousand replication of a random signal vector that has the same τ_b as the one from the Expedia sample. Since these placebo signals contain no information, we would expect that the value of I should be close to zero in most cases. We compute the value of I for each signal and the 95 percent confidence interval. We check that the value of I in the Expedia sample falls outside this interval in order to confirm that the information value of the Expedia signal is not due to chance.

4.2 Consumer Sophistication and Synthetic Signals (Table 6)

The signal is correlated with public information that is also correlated with price changes. For example, the probability of the Expedia scarcity signal decreases with DiA and prices 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. In order to investigate whether a sophisticated consumer still benefits from the information in the signal, we report the value of I after controlling for a set of conditioning variables to document the influence of traveler sophistication on the value of information.

Table 6 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 (I lies within the range $3.7 - 7.4$), when the traveler conditions her purchase decision on DiA. The traveller benefits most from signals sent 5 to 8 weeks in advance. The information value of the signal also remains positive and significant after conditioning on airline or routes. The value of I varies across airlines in the range $6.3 - 11.6$. The variation across routes is in the range $4.4 - 8.5$.

To put the reported values of I into perspective, we conduct the following thought experiment. Take a world without signal and consider two travellers: one use information on DiA to condition purchasing decision and the other doesn't. We compute how much the 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. This is like creating a 'synthetic' signal for public information and we can do so for DiA, Route, Carrier. We report the values of ρ^g, ρ^n, ρ^g, I for these synthetic signals.

To put these values for the Expedia signal into perspetive, we compute the value of a synthetic signal that uses public information. Assume a traveler is told that the distribution of returns is different for $DiA \leq 28$ and for $DiA > 28$. The utility gain from becoming sophisticated (learn about the distribution of returns conditional on DiA) is $I = \frac{\tau_b(1-\tau_b)}{\rho^n}H(\rho^n)$ where ρ^n is the marginal unsophisticated traveler (.11), τ_b is the fraction of observations in the sample with $DiA \leq 28$ (29 percent), $F^n()$ is the distribution of price return in the entire sample, and $H()$ is computed using for $F^g()$ the distribution of price return for $DiA > 28$. In this counterfactual thought experiment, we have $I = 10.4\%$ which is a little more than what the consumer gains more from conditioning her purchase decision on the Expedia signal.

4.3 Revenue and Consumption (Table 7 and 8)

The impact of the signal on revenue and consumption depends on the consumer's valuation. To illustrate, Figure 5 plots for the increase revenues in $\Delta R(v)$ (equation 6) computed over the entire sample. Based on this graph, we can make the following points: (a) revenues change is positive up to ρ^n, negative after and drops at ρ^n. (b) Although the signal has a large impact on individual revenue, this impact largely cancel out once averaged across all consumers. To better understand the impact of the signal, we take arithmetic average of revenue and consumption across all consumers (this assumes uniform distribution of valuation $g(v) = 1$).[13] The main findings for the change in revenue are:

1. The change in revenue for individual travellers is large and significant. For example, the positive increase in revenue $\Delta R(\mathbf{v}(\rho^b)) = \tau_b(1 + \rho^b)(1 - F^b(\rho^b))$ in absolute terms is .245 for the entire sample. The largest decrease in revenue, $\Delta R(\mathbf{v}(\rho_u^+))$, is .33.

2. Column 5 reports the average change in revenue (measured as a fraction of p_0) is $\frac{\Delta \overline{R}}{\rho^g - \rho^b}$ while column 6 reports the percentage increase in revenues for the consumers who respond is $\frac{\Delta \overline{R}}{R}$. The numbers are small and non-significant in most cases. In four cases, the numbers are negative and significant.

3. Column 4 reports the value of $\frac{f(\rho^n)}{1-F(\rho^n)} - \frac{2}{1+\rho^n}$.[14] The value for this expression suggests that profits decrease with the signal (in all but two cases). Interestingly, the change in

[13] The formula for this expression is derived in the appendix.

[14] This expression has to be greater than $\frac{g'(v^n)}{g(v^n)}p_0$ for the use of signal to be profitable. In order to interpret this expression, assume $g()$ is uniform on the left and right of ρ_u with $g(\rho_b)$ and $g(\rho_g)$ denoting the values of the density function on each side of ρ^n. The threshold ratio of consumer density, $\gamma = \frac{g(\rho_b)}{g(\rho_g)}$ such that $\int_{\mathbf{v}(\rho^b)}^{\mathbf{v}(\rho^g)} \Delta R(v)dv = 0$ is $\gamma = \frac{-\int_{\mathbf{v}(\rho_u^+)}^{\mathbf{v}(\rho^g)} \Delta \overline{R}(v)dv}{\int_{\mathbf{v}(\rho_b)}^{\mathbf{v}(\rho_u^-)} \Delta \overline{R}(v)dv}$.

revenue in column 5 and 6 that are statistically significant are entirely consistent with the condition in column 4.

Turning to consumption, the impact of the signal on the the percentage increase in consumption $\frac{\Delta C(v)}{C(v)}$ (normalized equation 7) looks very similar as $\frac{\Delta R(v)}{R(v)}$. The impact of the signal on consumption has similar properties:

1. The total change in consumption $\Delta\overline{C}$ is always small and insignificant in all but one case.
2. The average change in consumption is $\frac{\Delta\overline{C}}{\rho^g-\rho^b}$. The percentage change in consumption is $\frac{\Delta\overline{C}}{\overline{C}}$. Both numbers are small and typically insignificant.

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. We also find that scarcity signals have no impact on average revenue and average consumption.

These results demonstrate that scarcity signals can have economic value. It is also consistent with the view that scarcity cues cannot be effective if signals are not informative. The results reveal that scarcity signals have a very small impact on consumer welfare, of the order of a few dollars for a ticket that costs on average $300. But signals can have a significant impact on expected utility because signals influence the decision of consumers who do not receive much surplus in the absence of signal.

Important issues are left for future work. One is to study how consumers respond to scarcity signals: is their behavior consistent with 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 communicate (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 M. Pejcinovska. 2007. Distribution of airline tickets: a tale of two market structures. *Available at SSRN 1031747*.

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.

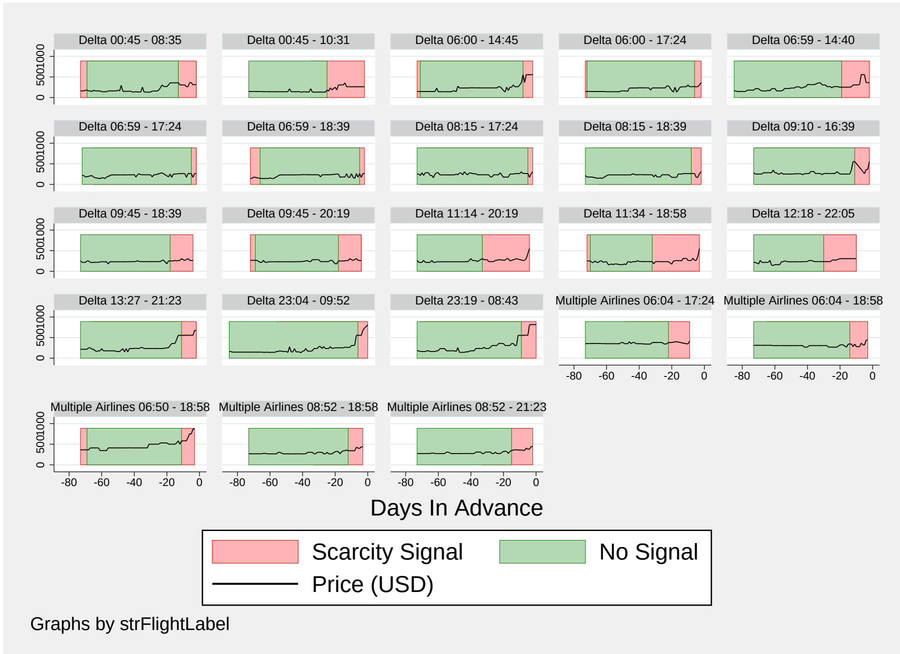
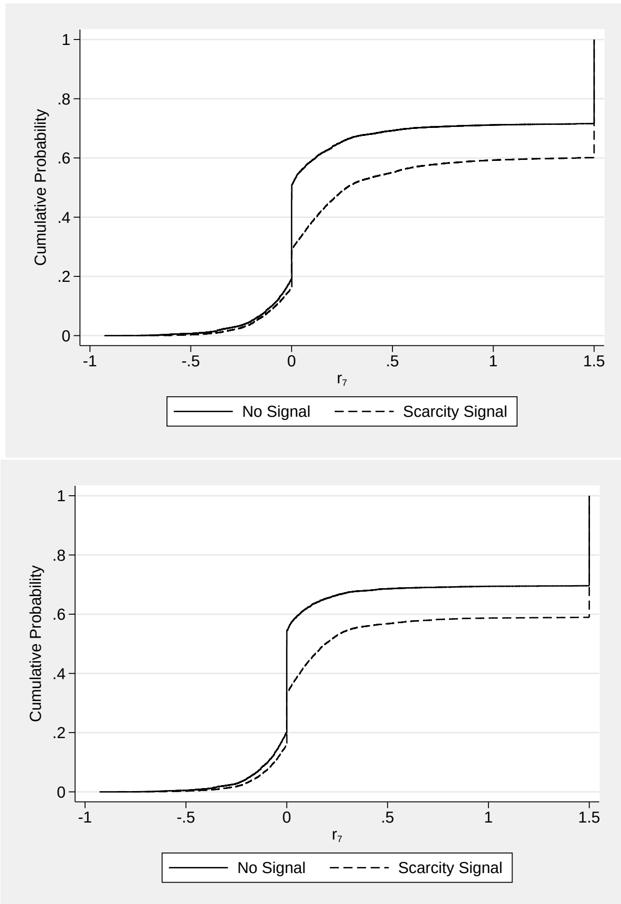


Figure 1: Each panel shows for a given Expedia querie, some of the flights available (panels), and for each flight, non-availability (blank), the value of the signal (green/red shade) and price (black line) as a function of DiA (horizontal axis).

Figure 2: Distributions of Expedia percentage price change over 7 day period as a function of Expedia signal realization, F^n and F^b (top figure is all DiA; bottom one is $DIA \geq 56$).



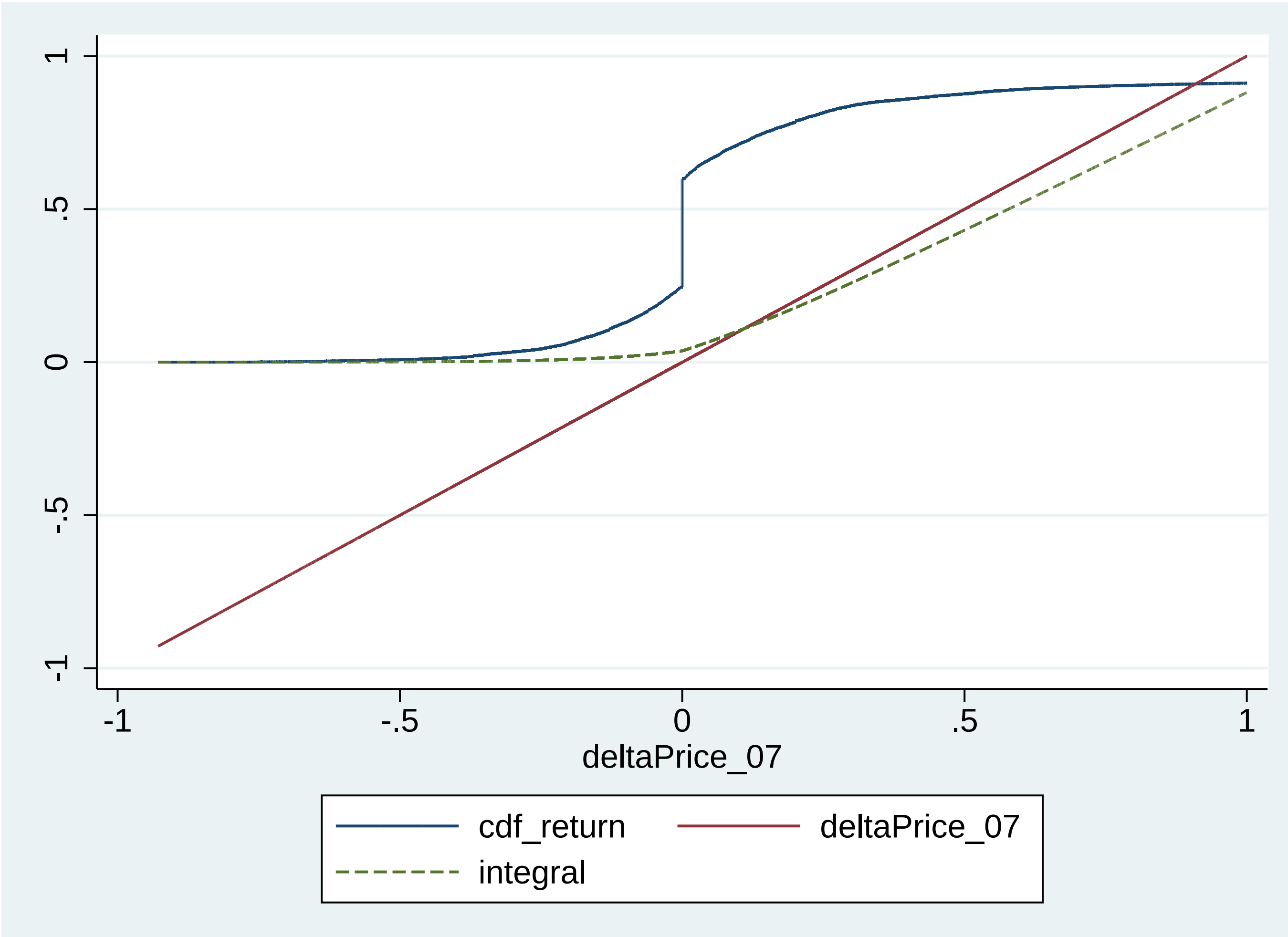


Figure 3: Computation of ρ^n: Functions $f_1(r) = \int_{-1}^{r} F^n(x)dx$ and $f_0(r) = r$ are computed using entire sample. ρ^n is such that $f_1(\rho) = f_0(\rho)$.

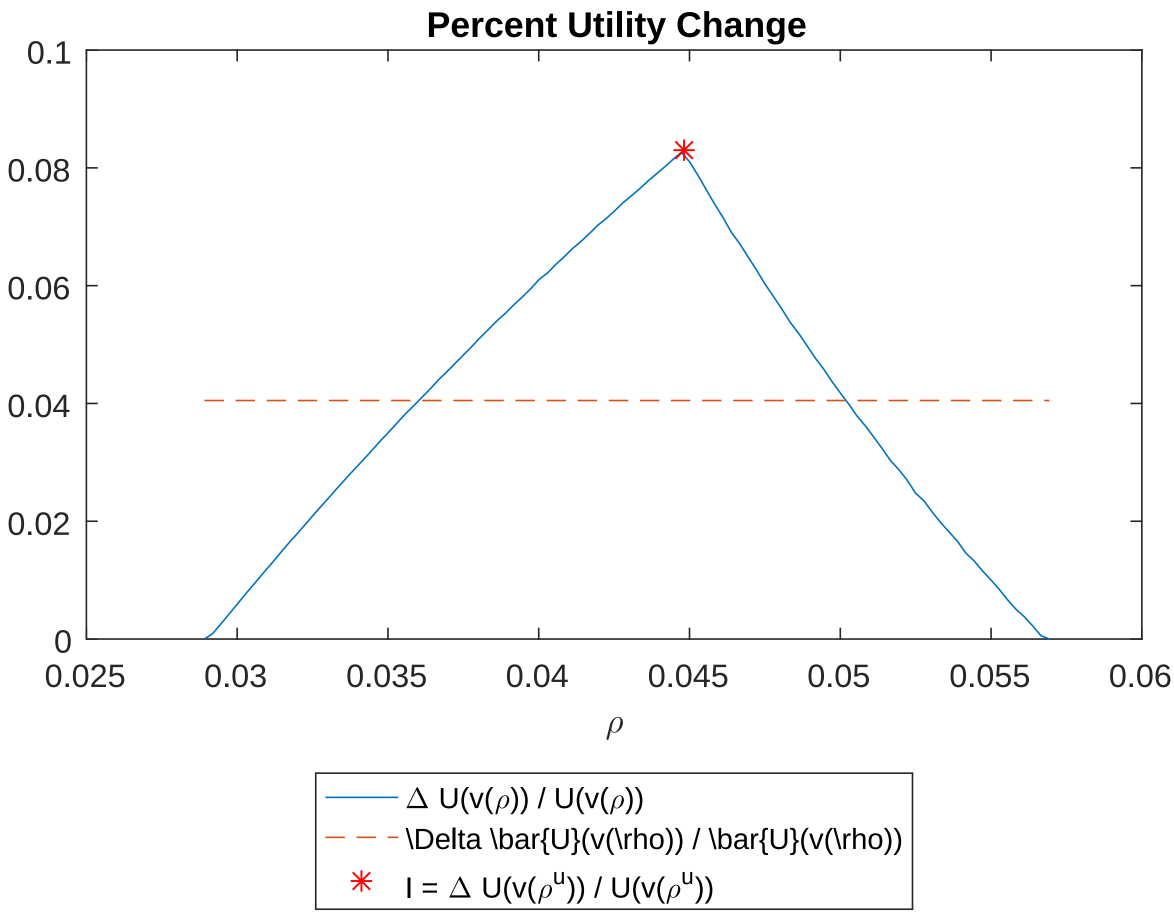


Figure 4: Percentage utility increase $\frac{\Delta U(\mathbf{v}(r))}{U(\mathbf{v}(r))}$. The red dashed line plots the average across all $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^g)]$. The red star presents the maximum percentage utility increase.

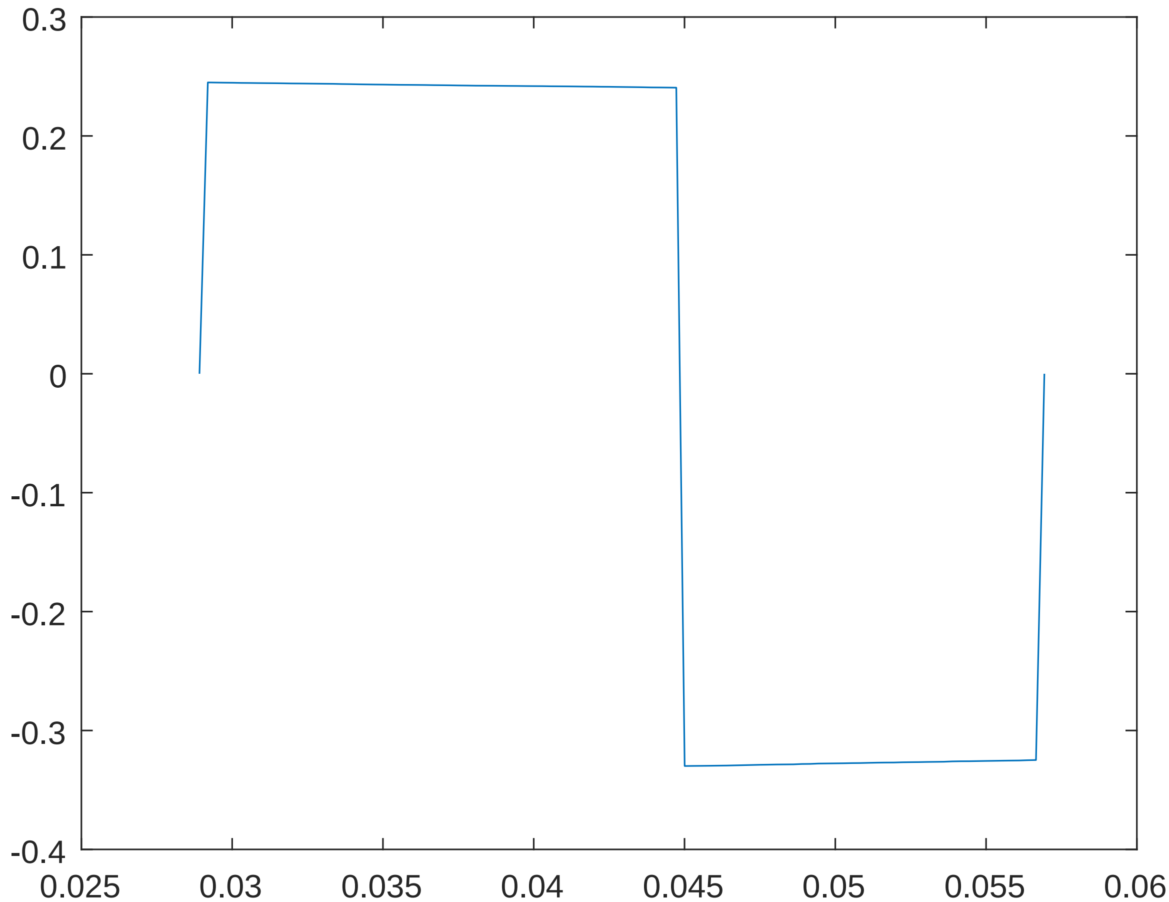


Figure 5: Change in Revenue $R(v)$ (measured as a fraction of p_0).

6 Tables

Table 1: Signals, Prices and Availability

Variable	Obs	Mean	Std. Dev.	P25	P50	P75
Signal (1: BAD, 0: GOOD)	539506	.33	.47	0	0	1
Days in Advance	539506	47.66	26.91	24	47	71
Price	539506	299.91	160.07	165.1	274.6	409.1
Price | BAD	180071	293.58	148.93	179.2	266.6	366.2
Price | GOOD	359435	303.08	165.28	162.1	281.6	409.1
r_1	455910	.02	.19	0	0	0
r_1| BAD	141828	.04	.22	0	0	.03
r_1| GOOD	314082	0	.16	0	0	0
r_7	345129	.06	.32	-.01	0	.08
r_7| BAD	98109	.12	.36	-.01	.01	.2
r_7| GOOD	247020	.04	.29	-.01	0	.02
r_{14}	266187	.11	.42	-.04	0	.16
r_{14}| BAD	71179	.18	.48	-.04	.06	.27
r_{14}| GOOD	195008	.08	.39	-.04	0	.1

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

Expedia Signal

Expedia-50	ρ^u	ρ^b	ρ^g	τ_b	I	$\frac{\Delta \bar{U}}{\rho^g - \rho^b}$	$\frac{\Delta \bar{U}}{\bar{U}}$	N
all	0.0448	0.0291	0.0568	0.334	0.0829	0.00186	0.0405	539506
	(0.044 - 0.046)	(0.0282 - 0.0301)	(0.0554 - 0.0586)		(0.0786 - 0.0873)	(0.00174 - 0.00196)	(0.0384 - 0.0426)	
lt28	0.0234	0.0212	0.0257	0.469	0.0367	0.000436	0.0181	153842
	(0.0224 - 0.0243)	(0.0201 - 0.022)	(0.024 - 0.0272)		(0.0218 - 0.0478)	(0.000253 - 0.000582)	(0.0115 - 0.0238)	
28to56	0.0672	0.0418	0.0875	0.327	0.0743	0.00249	0.0364	165325
	(0.065 - 0.0694)	(0.0398 - 0.0435)	(0.0842 - 0.0913)		(0.069 - 0.0808)	(0.00231 - 0.00275)	(0.0338 - 0.0395)	
56to84	0.0569	0.0316	0.0702	0.25	0.0675	0.00192	0.0348	155047
	(0.0551 - 0.059)	(0.0289 - 0.0336)	(0.0673 - 0.0734)		(0.0625 - 0.0747)	(0.00177 - 0.00212)	(0.0321 - 0.0386)	
Delta	0.0607	0.04	0.0793	0.362	0.0824	0.00249	0.0395	157598
	(0.0587 - 0.062)	(0.0382 - 0.0418)	(0.0756 - 0.0815)		(0.0746 - 0.0889)	(0.00224 - 0.00269)	(0.0361 - 0.0426)	
American Airlines	0.0544	0.0306	0.0773	0.371	0.116	0.00317	0.0548	112925
	(0.0516 - 0.0566)	(0.0285 - 0.0326)	(0.0723 - 0.0809)		(0.107 - 0.124)	(0.00288 - 0.00342)	(0.0512 - 0.0587)	
Multiple Airlines	0.0208	0.0156	0.0243	0.339	0.0632	0.000658	0.0311	149417
	(0.02 - 0.0215)	(0.0147 - 0.0163)	(0.0231 - 0.0252)		(0.0529 - 0.0704)	(0.000532 - 0.000725)	(0.026 - 0.0345)	
US Airways	0.0509	0.0357	0.0613	0.318	0.0681	0.00171	0.0332	48562
	(0.0471 - 0.0545)	(0.0317 - 0.0397)	(0.0557 - 0.0657)		(0.058 - 0.0825)	(0.00143 - 0.00205)	(0.0283 - 0.04)	
United	0.0639	0.0253	0.0853	0.258	0.107	0.00342	0.055	39220
	(0.0592 - 0.0675)	(0.0222 - 0.0279)	(0.0779 - 0.0902)		(0.0975 - 0.115)	(0.00304 - 0.00373)	(0.0508 - 0.0594)	
28to56_MIA-BOS	0.0663	0.0475	0.0873	0.434	0.0778	0.00251	0.0357	21792
	(0.0604 - 0.0702)	(0.0414 - 0.0531)	(0.0754 - 0.0952)		(0.0564 - 0.0954)	(0.00167 - 0.00313)	(0.0264 - 0.0429)	
28to56_MIA-DFW	0.104	0.0608	0.137	0.345	0.0842	0.00428	0.0406	20522
	(0.0938 - 0.112)	(0.0545 - 0.0692)	(0.12 - 0.149)		(0.0665 - 0.0933)	(0.00332 - 0.00497)	(0.0323 - 0.0447)	
28to56_SNA-DFW	0.0813	0.0289	0.111	0.199	0.0854	0.00331	0.0432	15595
	(0.0725 - 0.0914)	(0.023 - 0.0342)	(0.0952 - 0.128)		(0.0729 - 0.102)	(0.00268 - 0.00424)	(0.0363 - 0.0526)	
28to56_PHX-MDW	0.0506	0.0403	0.0585	0.362	0.0436	0.00108	0.021	13451
	(0.0455 - 0.0552)	(0.0329 - 0.045)	(0.0516 - 0.0648)		(0.0306 - 0.0653)	(0.000729 - 0.00162)	(0.0143 - 0.0315)	

1

Figure 6: Impact of Scarcity Signal on Consumer Utility

Expedia-50

Expedia-50 Revenue Estimates	$\Delta R(v(\rho^b))$	$\Delta R(v(\rho^{u+}))$	Revenue Drop	$\frac{f(\rho^u)}{1-F(\rho^u)} - \frac{2}{1+\rho^u}$	$\frac{\Delta \overline{R}}{\rho^g - \rho^b}$	$\frac{\Delta \overline{R}}{\overline{R}}$	N
all	0.245	-0.33	0.575	-0.686	-0.0024	-0.0035	539506
	(0.243 - 0.247)	(-0.332 - -0.326)	(0.57 - 0.578)	(-0.727 - -0.639)	(-0.0036 - 0.0047)	(-0.0051 - 0.0067)	
lt28	0.403	-0.393	0.796	-1.67	-0.0016	-0.0026	153842
	(0.401 - 0.405)	(-0.399 - -0.39)	(0.793 - 0.802)	(-1.68 - -1.65)	(-0.0095 - 0.0071)	(-0.015 - 0.011)	
28to56	0.209	-0.272	0.481	0.385	-0.0075	-0.0097	165325
	(0.207 - 0.212)	(-0.278 - -0.268)	(0.475 - 0.489)	(0.328 - 0.467)	(-0.014 - -0.0061)	(-0.018 - -0.0079)	
56to84	0.16	-0.311	0.471	0.388	-0.0062	-0.0088	155047
	(0.157 - 0.163)	(-0.315 - -0.307)	(0.465 - 0.477)	(0.348 - 0.465)	(-0.012 - -0.0047)	(-0.017 - -0.0066)	
Delta	0.255	-0.295	0.55	-0.524	-0.0028	-0.0038	157598
	(0.252 - 0.257)	(-0.3 - -0.291)	(0.545 - 0.556)	(-0.586 - -0.458)	(-0.0042 - 0.0057)	(-0.0056 - 0.0073)	
American Airlines	0.278	-0.303	0.582	-0.544	-0.0068	-0.0093	112925
	(0.275 - 0.281)	(-0.309 - -0.295)	(0.574 - 0.587)	(-0.621 - -0.442)	(-0.014 - 0.0013)	(-0.019 - 0.0017)	
Multiple Airlines	0.258	-0.389	0.647	-1.32	-0.00073	-0.0012	149417
	(0.255 - 0.26)	(-0.394 - -0.385)	(0.642 - 0.653)	(-1.36 - -1.3)	(-0.0043 - 0.0063)	(-0.007 - 0.01)	
US Airways	0.238	-0.358	0.596	-0.491	-0.0069	-0.01	48562
	(0.233 - 0.244)	(-0.368 - -0.348)	(0.584 - 0.608)	(-0.608 - -0.362)	(-0.02 - 0.00027)	(-0.03 - 0.0002)	
United	0.186	-0.355	0.541	-0.322	-0.0091	-0.013	39220
	(0.181 - 0.191)	(-0.364 - -0.347)	(0.53 - 0.551)	(-0.461 - -0.197)	(-0.02 - -0.0007)	(-0.03 - -0.0016)	
28to56_MIA-BOS	0.29	-0.271	0.561	-0.528	-0.0074	-0.0096	21792
	(0.284 - 0.299)	(-0.281 - -0.26)	(0.549 - 0.575)	(-0.676 - -0.392)	(-0.018 - 0.0055)	(-0.023 - 0.0067)	
28to56_MIA-DFW	0.222	-0.311	0.533	-0.07	-0.012	-0.016	20522
	(0.214 - 0.23)	(-0.32 - -0.301)	(0.517 - 0.546)	(-0.296 - 0.15)	(-0.022 - -0.0062)	(-0.03 - -0.0084)	
28to56_SNA-DFW	0.139	-0.261	0.4	0.905	-0.0064	-0.0083	15595
	(0.133 - 0.145)	(-0.276 - -0.253)	(0.392 - 0.415)	(0.734 - 1.01)	(-0.018 - 0.0053)	(-0.023 - 0.0066)	
28to56_PHX-MDW	0.225	-0.305	0.53	-0.616	-0.0035	-0.0048	13451
	(0.215 - 0.234)	(-0.317 - -0.288)	(0.511 - 0.543)	(-0.781 - -0.424)	(-0.018 - 0.012)	(-0.025 - 0.016)	

1

Figure 7: Impact of Scarcity Signal on Firm Revenue

Expedia-50

Expedia-50 Consumption Estimates	$\Delta\overline{C}$	$\frac{\Delta\overline{C}}{\rho^g-\rho^b}$	$\overline{C}$	$\frac{\Delta\overline{C}}{\overline{C}}$	N
all	-2.6e-06	0.019	-9.5e-05	-0.00014	539506
	(-8.4e-05 - 0.00011)	(0.018 - 0.02)	(-0.0031 - 0.0038)	(-0.0045 - 0.0056)	
lt28	2.8e-05	0.0028	0.0063	0.01	153842
	(1.4e-05 - 0.00013)	(0.0016 - 0.0037)	(0.0027 - 0.028)	(0.0042 - 0.044)	
28to56	-5.4e-06	0.034	-0.00012	-0.00016	165325
	(-0.00015 - 0.00016)	(0.032 - 0.038)	(-0.0033 - 0.0034)	(-0.0044 - 0.0045)	
56to84	-3.3e-05	0.027	-0.00086	-0.0012	155047
	(-0.0002 - 0.0001)	(0.025 - 0.03)	(-0.0051 - 0.0026)	(-0.0072 - 0.0037)	
Delta	2e-05	0.028	0.00051	0.0007	157598
	(-7.7e-05 - 0.00028)	(0.025 - 0.031)	(-0.0018 - 0.0073)	(-0.0024 - 0.01)	
American Airlines	5.8e-05	0.033	0.0012	0.0017	112925
	(-0.00033 - 0.00063)	(0.03 - 0.036)	(-0.0073 - 0.014)	(-0.01 - 0.019)	
Multiple Airlines	-2.3e-07	0.0054	-2.6e-05	-4.2e-05	149417
	(-5.1e-05 - 4.5e-05)	(0.0043 - 0.0059)	(-0.0059 - 0.0052)	(-0.0096 - 0.0083)	
US Airways	0.00012	0.017	0.0045	0.0069	48562
	(-0.00015 - 0.00088)	(0.014 - 0.02)	(-0.0072 - 0.035)	(-0.011 - 0.053)	
United	0.0003	0.04	0.0049	0.0074	39220
	(2.5e-05 - 0.0011)	(0.035 - 0.044)	(0.00054 - 0.018)	(0.00085 - 0.027)	
28to56_MIA-BOS	-0.00025	0.03	-0.0063	-0.0083	21792
	(-0.0012 - 0.0005)	(0.019 - 0.038)	(-0.031 - 0.013)	(-0.041 - 0.017)	
28to56_MIA-DFW	0.00026	0.055	0.0034	0.0047	20522
	(-0.00043 - 0.0016)	(0.042 - 0.065)	(-0.0049 - 0.022)	(-0.0068 - 0.03)	
28to56_SNA-DFW	-0.00043	0.062	-0.0052	-0.0069	15595
	(-0.0025 - 0.0014)	(0.049 - 0.078)	(-0.029 - 0.016)	(-0.039 - 0.021)	
28to56_PHX-MDW	-4e-05	0.013	-0.0022	-0.0031	13451
	(-0.0006 - 0.00037)	(0.0086 - 0.019)	(-0.036 - 0.019)	(-0.052 - 0.027)	

1

Figure 8: Impact of Scarcity Signal on Consumption

7 Notations

Table 3: Notations

$g(v), G(v)$	Consumer valuation density and cdf
$s \in \{n, b, g\}$	State (no signal, bad, good)
$p_1 = (1+r^s)p_0$	Prices (period 0 and 1)
$\overline{r}^s$	Expected growth rate
$F^s()$	Distributions of growth rate (eq. 1)
$H(r)$	$\int_{-1}^{r}(F^g(y) - F^b(y))dy$
$\mathbf{v}(r)$, $\mathbf{r}(v)$	$\mathbf{v}(r) = p_0(1+r)$, $\mathbf{r}(v) = \frac{v}{p_0} - 1$
ρ^s	Marginal consumer (eq. 2)
$U^s(v)$	Expected utility of consumer v with belief F^s (eq. 3)
I	Maximum value of information (eq. 5)
$R(v)$, $C(v)$	Consumer v revenue and consumption under prior ($s = n$)
$\overline{U}$, $\overline{R}$, $\overline{C}$	$\overline{U} = \int_{\mathbf{v}(\rho^b)}^{\mathbf{v}(\rho^g)} U^n(v)dv$ and same for $\overline{R}$, $\overline{C}$
$\Delta U(v)$, $\Delta R(v)$, $\Delta C(v)$	Impact of signal on individual outcomes (see eq. 4, 6 and 7)
$\Delta\overline{U}$, $\Delta\overline{R}$, $\Delta\overline{C}$	Average across all v (e.g. $\Delta\overline{U} = \int_{\mathbf{v}(\rho^b)}^{\mathbf{v}(\rho^g)} \Delta U(v)dv$)
$F(r,y)$, $S(r)$, $F^s(r\|x)$	Small signal (Proposition 2)
$\Delta\overline{U}(x)$, $\Delta\overline{R}(x)$, $\Delta\overline{C}(x)$	Utility, Revenue Consumption with small signal

In the result section, we normalize all expression assuming $p_0 = 1$, $\mathbf{v}(r) = 1 + r$ and $\mathbf{r}(v) = v - 1$. The value of v is thus expressed in units of p_0. Averages are computed over the individuals who respond to the signal, $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^g)]$, and using arithmetic average ($g(v) = 1$). We obtain:[15]

Table 4: Utility, Revenue, Consumption

	$v \leq \mathbf{v}(\rho^n)$	$v \geq \mathbf{v}(\rho^n)$	Average ($\overline{U}$, $\overline{R}$, $\overline{C}$)
$U^n(v)$	$\int_{-1}^{v-1} F^n(x)dx$	$v-1$	$\int_{\rho^b}^{\rho^n}\left(\int_{-1}^{r} F^n(x)dx\right)dr + \int_{\rho^n}^{\rho^g} rdr$
$R(v)$	$\int_{-1}^{v-1}(1+r)dF^n(r)$	1	$\int_{\rho^b}^{\rho^n}\left((1+r)F^n(r) - \int_{-1}^{r} F^n(x)dx\right)dr + \rho^g - \rho^n$
$C(v)$	$F(v-1)$	1	$\int_{\rho^b}^{\rho^n} F^n(r)dr + \rho^g - \rho^n$

Using equations 4, 6 and 7, we obtain $\Delta\overline{U}$, $\Delta\overline{R}$, $\Delta\overline{C}$. For example, $\Delta\overline{U} = \tau_b \int_{\rho^b}^{\rho^n}\left(r - \int_{-1}^{r} F^b(x)dx\right)dr + (1-\tau_b)\int_{\rho^n}^{\rho^g}\left(\int_{-1}^{r} F^g(x)dx - r\right)dr$.

[15] Note that $\int_{-1}^{r}(1+x)dF(x) = (1+r)F(r) - \int_{-1}^{r} F(x)dx$.

8 Appendix: Proofs

Proof of Lemma 1 The function $G^s(x) = x - \int_{-1}^{x} F^s(r)dr$ is increasing in x. We have $G^s(-1) < 0$ and $\lim_\infty G^s(x) = \lim_\infty \left(x - \int_{-1}^{x} F^s(r)dr\right) = \lim_\infty \left(x + \int_{-1}^{x} rdF^s(r) - xF^s(x)\right) = \overline{r}^s$. If $\overline{r}^s > 0$, equation (2) has a unique solution ρ^s for $s \in \{n, b, g\}$ and this solution is positive since $G^s(0) < 0$.

Next, we prove by contradiction that $Min(\rho^g, \rho^b) \leq \rho^n \leq Max(\rho^g, \rho^b)$. Assume, for example, that $Min(\rho^g, \rho^b) > \rho^n$. The following three obsevations (a) $\rho^n < \rho^g$, $G^g(x)$ strictly increasing, and (c) $G^g(\rho^g) = 0$ imply that $G^g(\rho^n) < 0$. The same reasoning implies that $G^b(\rho^n) < 0$. Thus $\tau_b G^b(\rho^n) + (1 - \tau^b)G^g(\rho^n) < 0$ or $\rho^n < \int_{-1}^{\rho^n} F^n(r)dr$. A contradiction. The same logic proves the other inequality.

Finally, $v - p_0 - E(Max(v - p_1, 0)|s) = p_0 G^s(\mathbf{r}(v))$. 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 $v \in (0, \mathbf{v}(\rho^s))$. □

Proof of Proposition 1: (a) For $v \in [0, \mathbf{v}(\rho^b))$ we have $E(Max(v - p_1, 0)|s) > v - p_0$. Thus, the consumer prefers to wait with or without signal. (b) For $v \in (\mathbf{v}(\rho^b), \mathbf{v}(\rho^n))$ we have $E(Max(v - p_1, 0)|n) > v - p_0$, $E(Max(v - p_1, 0)|g) > v - p_0$ and $E(Max(v - p_1, 0)|b) < v - p_0$. Thus, the consumer buys early only if the signal is bad. (c) For $v \in (\mathbf{v}(\rho^n), \mathbf{v}(\rho^g))$ we have $v - p_0 > E(Max(v - p_1, 0)|n)$, $v - p_0 > E(Max(v - p_1, 0)|b)$ and $E(Max(v - p_1, 0)|g) > v - p_0$. Thus, the consumer waits only if the signal is good. (d) For $v > \mathbf{v}(\rho^g)$ we have $E(Max(v - p_1, 0)|s) < v - p_0$. Thus, the consumer prefers to buy early with or without signal. □

Derivation of equation (4) for $\Delta U(v)$: Consumer $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$ waits without a signal. Her expected utility is $p_0 \int_{-1}^{\mathbf{r}(v)} F^n(r)dr$. With the signal, she wait when the realization is good. Her expected utility is $p_0 \int_{-1}^{\mathbf{r}(v)} F^g(r)dr$. She buys early when the realization is bad and receive utility $v - p_0$. Taking expectation, we obtain:

$$\Delta U(v) = \tau_b(v - p_0) + (1 - \tau_b)p_0 \int_{-1}^{\mathbf{r}(v)} F^g(r)dr - p_0 \int_{-1}^{\mathbf{r}(v)} F^n(r)dr$$

for $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$. Using identity 1, we obtain the top part in equation (4). The same reasoning applies to the bottom part of equation (4) corresponding to $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$. Note that $\Delta U(v)'' = \frac{1-\tau_b}{p_0} f^g(\mathbf{r}(v)) > 0$ for $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$. For $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$, $\Delta U(v)'' = \frac{1}{p_0}\left((1 - \tau_b) f^g(\mathbf{r}(v)) - f^n(\mathbf{r}(v))\right)$ cannot be signed in general. In the case where $F^n()$ and $F^g()$

are uniform with the same support, we have $\Delta U(v)'' < 0$.

Proof of Corollary 1: If $\rho^g = \rho^n = \rho^b$, no consumer changes her decision and the signal has no economic value. To prove the reverse assume $\rho^g = \rho^n = \rho^b$ does not hold. Lemma 1 implies $\rho^g > \rho^b$ or $H(\rho^n) > 0$. Equation 4 is rewritten as $\Delta U(\mathbf{v}(\rho^n)) = p_0(1-\tau_b)H(\rho^n) > 0$. A contradiction. □

Proof of Corollary 2: $\frac{\partial}{\partial v}\Delta U(v) = \tau_b - \tau_b F^b(\mathbf{r}(v)) \geq 0$ for $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$ and $\frac{\partial}{\partial v}\Delta U(v) = (1-\tau_b)(F^g(\mathbf{r}(v)) - 1) \leq 0$ for $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$. Thus, $\Delta U(v)$ reaches a maximum at $\mathbf{v}(\rho^n)$. □

Properties of the Revenue Function: Using the identity $\int_{-1}^{\rho}(1+r)dF(r) = (1+\rho)F(\rho) - \int_{-1}^{\rho} F(r)dr$, we can rewrite equation (6) as

$$\Delta R(v) = \begin{cases} p_0\tau_b \left(1 - (1+\mathbf{r}(v))F^b(\mathbf{r}(v)) + \int_{-1}^{\mathbf{r}(v)} F^b(r)dr\right), & \text{if } v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)] \\ -p_0(1-\tau_b)\left(1 - (1+\mathbf{r}(v))F^g(\mathbf{r}(v)) + \int_{-1}^{\mathbf{r}(v)} F^g(r)dr\right), & \text{if } v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]. \end{cases} \tag{8}$$

The supplier's profit function has the following properties.

Lemma 2. *(a)* $\Delta R'(v) < 0$ *for* $v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)]$.

$\Delta R'(v) > 0$ *for* $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$.

(b) $\Delta R(\mathbf{v}(\rho^b)) = p_0\tau_b(1+\rho^b)(1-F^b(\rho^b)) > 0$,

$\Delta R(\mathbf{v}(\rho^g)) = -p_0(1-\tau_b)(1+\rho^g)(1-F^g(\rho^g)) < 0$.

$\lim_{\rho\to\rho_u^+} \Delta R(\mathbf{v}((\rho)) < 0$.

$\lim_{\rho\to\rho_u^+} \Delta R(\mathbf{v}((\rho)) - \lim_{\rho\to\rho_u^-} \Delta R(\mathbf{v}((\rho)) = -p_0(1+\rho^n)(1-F^n(\rho^n)) < 0$.

$\lim_{\rho\to\rho_u^-} \Delta R(\mathbf{v}((\rho)) > 0$ *if and only if* $(1+\rho^b)(1-F^b(\rho^n)) > \int_{\rho^b}^{\rho^n}(F^b(\rho^n) - F^b(x))dx$.

(c) $\Delta R(v) < 0$ *for* $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$.

Note that the Lemma implies that $\Delta R(\mathbf{v}(\rho^b)) = Max_v \Delta R(v)$ and $\lim_{\rho\to\rho_u^+} \Delta R(\mathbf{v}((\rho)) = Min_v \Delta R(v)$.

The jump at $\mathbf{v}(\rho^b)$ is equal to $\Delta R(\mathbf{v}(\rho^b)) = p_0\tau_b(1+\rho^b)(1-F^b(\rho^b))$. This is because traveler $v = \mathbf{v}(\rho^b)$ waits in the absence of signal and switches to buy when the signal realization is bad.

The signal enables the supplier to extract that traveler's marginal valuation, $p_0(1+\rho_b)$, when the bad state occurs, which happens with probability τ_b, and when the price is above the traveler's valuation, which happens with probability $1 - F^b(\rho^b)$.

Proof of Lemma 2 (a) Take derivatives with respect to v in equation (4)

$$\Delta R(v) = \begin{cases} p_0\tau_b \left(1 - \int_{-1}^{\mathbf{r}(v)} (1+r)dF^b(r)\right), & \text{if } v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)] \\ -p_0(1-\tau_b)\left(1 - \int_{-1}^{\mathbf{r}(v)} (1+r)dF^g(r)\right), & \text{if } v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]. \end{cases}$$

$$\Delta R'(v) = \begin{cases} -\tau_b(1+\mathbf{r}(v))f^b(\mathbf{r}(v)) < 0, & \text{if } v \in [\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)] \\ (1-\tau_b)(1+\mathbf{r}(v))f^g(\mathbf{r}(v)) > 0, & \text{if } v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]. \end{cases}$$

(b) From equation (8), we have $\Delta R(\mathbf{v}(\rho^b)) = p_0\tau_b\left(1 - (1+\rho^b)F^b(\rho^b) + \int_{-1}^{\rho^b} F^b(r)dr\right)$ and since $\int_{-1}^{\rho^b} F^b(r)dr = \rho^b$, we obtain, $\Delta R(\mathbf{v}(\rho^b)) = p_0\tau_b(1+\rho^b)(1-F^b(\rho^b))$. Applying the same logic one obtains $\Delta R(\mathbf{v}(\rho^g))$.

Evaluate equation (8) at ρ_u^+, add ρ^n and substract $\int_{-1}^{\rho^n} F^n(r)dr$ (recall $\rho^n = \int_{-1}^{\rho^n} F^n(r)dr$), to obtain, $\lim_{\rho\to\rho_u^+} \Delta R(\mathbf{v}((\rho)) = -p_0(1-\tau_b)\left((1+\rho^n)(1-F^g(\rho^n)) + \int_{-1}^{\rho^n}(F^g(r) - F^n(r))dr\right) < 0$ where the inequality follows from $F^g(r) > F^n(r)$.

Evaluate equation (8) at ρ_u^-, $\lim_{\rho\to\rho_u^-} \Delta R(\mathbf{v}((\rho)) = p_0\tau_b\left(1 - (1+\rho^n)F^b(\rho^n) + \int_{-1}^{\rho^n} F^b(r))dr\right)$. But since $\rho^b = \int_{-1}^{\rho^b} F^b(r))dr$, we obtain $\lim_{\rho\to\rho_u^-} \Delta R(\mathbf{v}((\rho)) = p_0\tau_b\left(1 + \rho^b - (1+\rho^n)F^b(\rho^n) + \int_{\rho^b}^{\rho^n} F^b(r))dr\right)$ or $\lim_{\rho\to\rho_u^-} \Delta R(\mathbf{v}((\rho)) = p_0\tau_b\left((1+\rho^b)(1-F^b(\rho^n)) - \int_{\rho^b}^{\rho^n}(F^b(\rho^n) - F^b(r))dr\right)$. And $\lim_{\rho\to\rho_u^-} \Delta R(\mathbf{v}((\rho)) > 0$ is equivalent to $(1+\rho^b)(1-F^b(\rho^n)) > \int_{\rho^b}^{\rho^n}(F^b(\rho^n) - F^b(r))dr$. Although one can construct examples such that this condition is violated, this does not happen in our application because $\rho^n - \rho^b$ is small relative to $1 - F^b(\rho^n)$.

(c) $\Delta R(v) < 0$ for $v \in [\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$ follows from $\Delta R(v)$ monotone in $[\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)]$, $\Delta R(\mathbf{v}(\rho^g)) < 0$ and $\lim_{\rho\to\rho_u^+} \Delta R(\mathbf{v}((\rho)) < 0$. □

Proof of Proposition 2: Let $\rho(y)$ denote the indifferent consumer when the cumulative distribution of price returns is $F(r,y)$. Specifically, $\rho(y)$ is defined by equation (2) after replacing $F^s(r)$ with $F(r,y)$. For ease of notation, we denote $\rho^n = \rho(0)$ and $v^n = \mathbf{v}(\rho(0))$. When the cumulative distribution of price returns is $F(r,y)$, the supplier revenues are

$$R(y) = p_0 \int_0^{\mathbf{v}(\rho(y))} \left(\int_{-1}^{\mathbf{r}(v)} (1+r) dF(r,y) \right) dG(v) + p_0 \left(1 - G(\mathbf{v}(\rho(y)))\right).$$

This is because a traveler with a valuation below $\mathbf{v}(\rho(y))$ waits, and subsequently purchases if the price is below her valuation. A traveler with a valuation above $\mathbf{v}(\rho(y))$ buys early. Expected revenues with signal x are $\tau_b R(-x(1-\tau_b)) + (1-\tau_b)R(x\tau_b)$ and

$$\Delta\overline{R}(x) = \tau_b R(-x(1-\tau_b)) + (1-\tau_b)R(x\tau_b) - R(0).$$

We have $\Delta\overline{R}'(x) = \tau_b(1-\tau_b)(-R'(-x(1-\tau_b)) + R'(x\tau_b))$ and $\Delta\overline{R}'(0) = 0$. There is no first-order impact of the signal on profits.

We have $\Delta\overline{R}''(x) = \tau_b(1-\tau_b)((1-\tau_b)R''(-x(1-\tau_b)) + \tau_b R''(x\tau_b))$ and $\Delta\overline{R}''(0) = \tau_b(1-\tau_b)R''(0)$. Denote $R(y) = p_0 K(\rho(y), y)$. Differentiating twice with respect to y gives $R''(y) = p_0 \frac{d^2K}{dy^2}(y)$ with

$$\frac{dK}{dy} = K_1\rho' + K_2$$
$$\frac{d^2K}{dy^2} = K_1\rho'' + K_{11}\left(\rho'\right)^2 + K_{22} + 2K_{12}\rho'$$

where K_1, for example, denotes the derivative of K with respect to its first argument. We have

$$K_1(\rho, y) = p_0 g(\mathbf{v}(\rho) \left(\int_{-1}^{\rho} (1+r) d(F(r,y)) - 1 \right)$$
$$K_2(\rho, y) = \int_0^{\mathbf{v}(\rho)} \int_{-1}^{\mathbf{r}(v)} (1+r) dS(r) dG(v)$$
$$K_{11}(\rho, y) = p_0 g(\mathbf{v}(\rho))(1+\rho) \left(f(\rho, y) - \frac{g'(\mathbf{v}(\rho))}{g(\mathbf{v}(\rho))} p_0 (1 - F(\rho, y) \right)$$
$$K_{22}(\rho, y) = 0$$
$$K_{12}(\rho, y) = p_0 g(\mathbf{v}(\rho)) \left((1+\rho)S(\rho) - \int_{-1}^{\rho} S(r) dr \right)$$

where the last expression uses the identity $\int_{-1}^{\rho}(1+r)dS(r) = (1+\rho)S(\rho) - \int_{-1}^{\rho} S(r)dr$ is obtained using integration by part. Next, take $\int_{-1}^{\rho}(1+r)d(F(r,y)) - 1$ in K_1. Apply the same integration by part identity, and evaluate at $\rho = \rho(y)$ to obtain $\int_{-1}^{\rho(y)}(1+r)dF(r,y) - 1 = -(1+\rho(y))(1-$

$F(\rho(y), y))$ and consequently

$$K_1(\rho(y), y) = -p_0 g(\mathbf{v}(\rho(y))(1+\rho(y))(1-F(r,y)).$$

Using equation (2), we evaluate:

$$\begin{aligned} \rho'(0) &= \frac{\int_{-1}^{\rho(y)} S(r)dr}{1-F(\rho,0)} \\ \rho''(0) &= \frac{2S(r)\int_{-1}^{\rho(y)} S(r)dr}{(1-F(\rho,0))^2} \end{aligned} \tag{9}$$

Replacing the expressions for K_1, K_{11}, K_{22}, K_{12}, ρ' and ρ'' gives:

$$\left.\frac{d^2K}{dy^2}\right|_{y=0} = p_0(1+\rho^n)g(v^n)\frac{\left(\int_{-1}^{\rho^n} S(r)dr\right)^2}{1-F^n(\rho^n)}\left(\frac{f(\rho^n)}{1-F^n(\rho^n)} - \frac{g'(v^n)}{g(v^n)}p_0 - \frac{2}{1+\rho^n}\right)$$

from which we conclude

$$\Delta\overline{R}''(0) = p_0^2(1+\rho^n)g(v^n)\tau_b(1-\tau_b)\frac{\left(\int_{-1}^{\rho^n} S(r)dr\right)^2}{1-F^n(\rho^n)}\left(\frac{f(\rho^n)}{1-F^n(\rho^n)} - \frac{g'(v^n)}{g(v^n)}p_0 - \frac{2}{1+\rho^n}\right)$$

A similar argument applies for consumption where $R(y)$ is replaced with $C(y) = \int_0^{\mathbf{v}(\rho(y))} F(\mathbf{r}(v))dG(v) + 1 - G(\mathbf{v}(\rho(y)))$. We obtain $\Delta\overline{C}''(0) = \tau_b(1-\tau_b)C''(0)$.

$$\Delta\overline{C}''(0) = p_0 g(v^n)\tau_b(1-\tau_b)\frac{\left(\int_{-1}^{\rho^n} S(r)dr\right)^2}{1-F^n(\rho^n)}\left(\frac{f(\rho^n)}{1-F^n(\rho^n)} - \frac{g'(v^n)}{g(v^n)}p_0\right)$$

□

9 Appendix: Example

We have $\alpha > 0$, $\beta > 0$, $\delta > 0$ and

$$F^n(r) = \begin{cases} 0, & \text{if } r \notin [-\alpha, \frac{\alpha}{\beta}(1-\beta-\delta)] \\ \beta + \frac{\beta}{\alpha} r, & \text{if } r \in [-\alpha, 0] \\ \beta + \delta + \frac{\beta}{\alpha} r, & \text{if } r \in [0, \frac{\alpha}{\beta}(1-\beta-\delta)] \end{cases}$$

and $S(r) = xI(r > 0)$. Thus $F^b(r) = \beta + \delta - x(1-\tau_b) + \frac{\beta}{\alpha} r$ for $r \in [0, \frac{\alpha}{\beta}(1-\beta-\delta+x(1-\tau_b))]$ and $F^g(r) = \beta + \delta + x\tau_b + \frac{\beta}{\alpha} r$ for $r \in [0, \frac{\alpha}{\beta}(1-\beta-\delta-x\tau_b)]$. We assume that $2\beta + \delta + \tau_b x < 1$ so that $\overline{r}^b > 0$ which implies that (ρ^b, ρ^n, ρ^g) are finite.

Using equation (2), we obtain $\rho^n = \frac{\alpha}{\beta}\left(1-\beta-\delta-\sqrt{(1-2\beta-\delta)(1-\delta)}\right)$,

$$\rho^b = \frac{\alpha}{\beta}\left(1-\beta-\delta+x(1-\tau_b)-\sqrt{(1-2\beta-\delta+x(1-\tau_b))(1-\delta+x(1-\tau_b))}\right)$$

$$\rho^g = \frac{\alpha}{\beta}\left(1-\beta-\delta-x\tau_b-\sqrt{(1-2\beta-\delta-x\tau_b)(1-\delta-x\tau_b)}\right).$$

Using Equation (6), we obtain

$$\Delta\overline{R}(v) = \begin{cases} p_0\tau_b\left(1-\beta-\delta+(1-\tau_b)x+\frac{\alpha\beta}{2}+\frac{\beta}{2\alpha}-\frac{\beta}{2\alpha}v^2\right) & \text{if } v \in (\mathbf{v}(\rho^b), \mathbf{v}(\rho^n)) \\ -p_0(1-\tau_b)\left(1-\beta-\delta-\tau_b x+\frac{\alpha\beta}{2}+\frac{\beta}{2\alpha}-\frac{\beta}{2\alpha}v^2\right), & \text{if } v \in (\mathbf{v}(\rho^n), \mathbf{v}(\rho^g)). \end{cases}$$

Figure 9 sets $\alpha = .3$, $\beta = .3$, $\delta = .3$, $\tau = .5$ and $x = .1$. The distribution of consumer valuation is uniform, $g'(v) = 0$, for $v \in [\mathbf{v}(\rho^b)), \mathbf{v}(\rho^g))]$. The condition in Proposition 2, $\frac{f(\rho^n)}{1-F^n(\rho^n)} - \frac{g'(v^n)}{g(v^n)}p_0 - \frac{2}{1+\rho^n} > 0$, holds. The firm benefits from the signal. $\square$

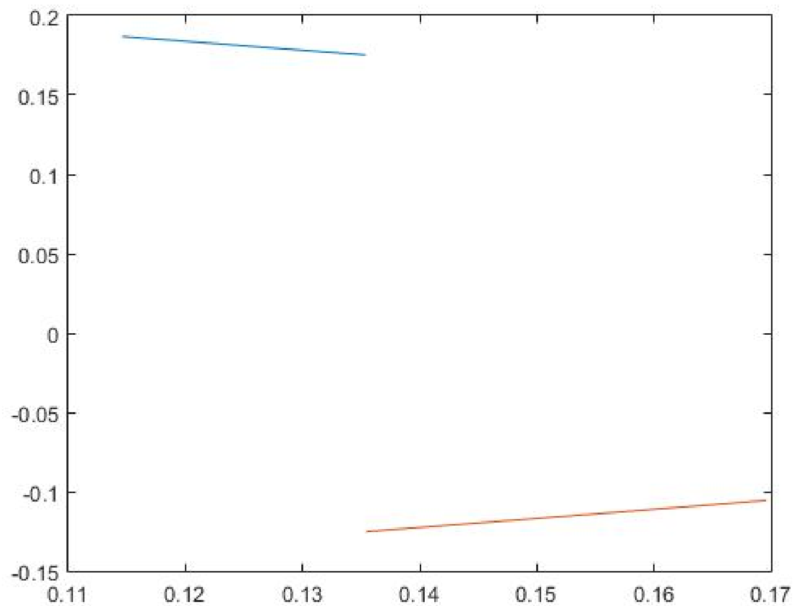


Figure 9: $\Delta R(v)$ for Example 1 with uniform $F^n(r)$. (See Appendix for derivations.)

10 Appendix: Expedia Dataset

We collect the price and scarcity signal for one-way travel from the Expedia website. It may be that Expedia employs cookies to track visitors, and set an individual price for someone who returns to the site. Although our scraper accepts cookies from Expedia, it also delete the cookies after each query. Therefore, Expedia cannot identify the scraper through cookies. The price we collect is the price that Expedia would return a fresh query.

A *route* is a pair composed of an origin and a destination airport. We use the standard three-letter designations for airports. A *query* is a pair composed of a *route* and a *departure date*. Each query is submitted on different *booking dates* and return a number of flight options. Each *flight* is identified by a query, a departure and arrival time, the number of layovers and the carrier(s). Therefore, the nesting goes from route to query to flight. For each flight, we collect the price and signal. We construct the variable days-in-advance DiA as the number of days between the departure and booking dates.

Route Selection: We selected routes on the basis of four criteria: (a) Routes with a dominant carrier. (b) Busy routes in term of passenger. (c) Routes previously selected in airline literature. (d) Routes with large potential price savings. We implemented these selection criteria as follows. To find routes operated by a dominant carrier, we use the T-100 data bank from the Bureau of Transportation Statistics web site.[16] The T-100 data bank includes the number of airlines per route, as well as total number of passengers. All of our ten routes are operated by a single airline. Although there is a single carrier offering a direct flight, it is possible to combine two or more flights to make the city-pair travel. Within this set of routes, we selected the top 25 subset in terms of number of passengers transported. A similar approach was previously used by Bilotkach and Rupp (2011). Indeed, many of our routes are the same as Bilotkach and Rupp (2011) and by Bilotkach and Pejcinovska (2007). Finally, 7 of our 10 routes have destinations listed as domestic destinations with largest potential savings by Hopper.com.[17] For these routes, signals are likely to have a greater economic value for customers.

Sampling: We run queries daily for about 100 days, starting on July 19th and ending on October 26th, 2015. The first departure date is the 10th of August, which is approximately three weeks from our first booking date (19th of July). The following 11 departures dates are 8 days apart (18th of August, 26th of August, and so on... till November 5th). A flight expires when its

[16] www.transtats.bts.gov

[17] Hopper.com is an aggregator website that collects price data and reports the best days in advance to purchase tickets.

Table 5: Observation Count

Route	N
AUS-DAL	12635
CLT-FLL	80099
DFW-SNA	56752
LAS-MDW	48524
MCO-MDW	54519
MDW-LAS	57053
MIA-BOS	72854
MIA-DFW	62791
PHX-MDW	43616
SNA-DFW	52539
Total	541382

Source: expedia_v1.0.dta

Airline	N
Alaska Airlines	23534
American Airline	112925
Delta	157598
Frontier Airline	1734
JetBlue Airways	4158
Multiple Airline	149417
Sun Country Airl	142
US Airways	48562
United	39220
Virgin America	4092
Total	541382

Source: expedia_v1.0.dta

departure date has past. When this happens, a new query is automatically added 8 days after the last querry added. Using this rolling window sampling method, we end up with 22 departure dates that cover fairly evenly the seven days of the week. Combined with 10 routes, this adds to 220 distinct queries. For some queries (those with a departure date between Octobeer 26th and November 5th), we obtain time series (price, signal) for each flight that cover 100 days. The time series are shorter for the rest of the queries.

Table 5 reports the observation count broken down by route and airline. The more busy routes offer more flight options and end up with more observations. There are 9 main airlines serving the 10 routes. The flights observations have between one and 14 weeks in advance, and because of the rolling departure date sampling methodology, the number observations is roughly evenly distributed with about 40K observations per week in advance. The number of observations is about 9K for the first departure date, peaks to about 47K on the 10th departure date and then decreases to about 2K on the last one (22nd departure date). This is because the early and late departure dates are queried less frequently.

Table 6 shows that there are on average 179 flights displayed for a given query made on a given booking date. This average varies a little by route. Table 7 shows that the signal can take five positive values: 1, 2,.., 5 seats left at a given price. The frequency of sending signals with high values is slightly lower. In most of the empirical analysis we assume that the signal takes a binary value: 0 if Expedia does not report a number of seats left at the posted price) and 1 otherwise.

Table 6: Number of flights per queryXbooking date (mean/min/max)

route	**mean**	**min**	**max**
AUS-DAL	24.0979	9	42
CLT-FLL	187.0916	105	288
DFW-SNA	173.3658	56	304
LAS-MDW	175.6805	25	289
MCO-MDW	179.5969	45	289
MDW-LAS	179.6123	35	282
MIA-BOS	210.2532	76	372
MIA-DFW	203.59	99	349
PHX-MDW	148.1207	26	237
SNA-DFW	163.4948	66	278
Total	178.9122	9	372

Table 7: Observation count for each signal realization (number of seat left at posted price)

Signal Realization	**N**
0	359435
1	45611
2	39744
3	35083
4	31978
5	27655
Total	539506