The Nature of Sales in Online Markets: Asymmetric Consumer Information or Benefits to Bulk Shopping?

Jihui Chen*
University of North Dakota

Jeffrey A. Livingston†
Bentley College

Patrick A. Scholten‡
Bentley College

July, 2006

Abstract

Price dispersion - firms charging different prices for the same product - is widely observed in both online and traditional offline markets. While most price dispersion is explained by stylized clearinghouse models such as Varian (1980), these models do not explain why prices in offline markets are lower on weekends than during the work week, and before Christmas than after Christmas. We argue that price dispersion online is fully explained by clearinghouse models. First, because search and travel costs are lower online, these anomalous pricing patterns disappear. Second, prices charged by firms, price dispersion, the number of firms posting prices, and the minimum price in the online markets for several products vary in ways that are all consistent with the predictions of clearinghouse models.

JEL Codes: L100, L130 and L150.

Keywords: Price Dispersion, Price Competition, Internet, and Information

*Department of Economics, College of Business and Public Administration, P.O. Box 8369, University of North Dakota, Grand Forks, ND 58201, susan.chen@und.nodak.edu.

†Department of Economics, 175 Forest Street, Bentley College, Waltham, MA 02452-4705, jlivingston@bentley.edu.

‡Corresponding author: Department of Economics, 175 Forest Street, Bentley College, Waltham, MA 02452-4705, pscholten@bentley.edu, fax: (781) 891-2896.

1 Introduction

> *"Internet search engines have already reshaped the way people shop for airline tickets, electronics and other products. In November, almost 40 million Americans visited comparison-shopping sites, an increase of 13% from the same period last year, according to comScore Media Metrix, an Internet-research firm. The search engines can zip across the Internet locating the cheapest prices among online retailers." (Wall Street Journal, "Getting Your Own Book Deal; New Breed of Search Engine Focuses Strictly on Finding The Cheapest Available Copy," D1, 12/21/2004)*

As the above quote suggests, the Internet has reshaped the way in which consumers shop for a wide array of products and services. For example, Internet price comparison sites, which charge firms a fee to list prices for products, allow consumers to simultaneously acquire price information for homogeneous products from multiple sellers. Traditional restrictions on when consumers can shop in offline markets is no longer relevant in online markets: consumers can shop anytime of the day provided they have Internet access. In addition, a search for a particular product at an Internet price comparison site results in price information from multiple sellers; thus, consumers who access these sites can easily identify and purchase from the low-price seller.

One might conjecture that better informed consumers would lead to markets that closely resemble competitive markets. Price dispersion, however, is a well-documented phenomenon in both offline and online retails markets (for an overview of this literature see Baye, Morgan and Scholten (2005)). One explanation for price dispersion in these retail markets is that sellers independently offer periodic markdowns on prices; i.e., place items on "sale."

"Clearinghouse" models offer conditions under which equilibrium price dispersion will persist (c.f., Varian 1980; Rosenthal 1980; Narasimhan 1988; Baye and Morgan 2001; and Baye, Morgan and Scholten 2004b). These models postulate that some, but not necessarily all, consumers access an information clearinghouse to observe the price distribution of all sellers. These "informed" consumers are assumed to purchase from the seller offering the lowest price. The remaining "uninformed" consumers do not know the identity of the low-price seller, so they randomly select a firm from which to buy (and pay the average price). Accordingly, sellers have no incentive to compete prices down to marginal cost; they are able to attract a portion of the uninformed consumers even when offering a price in the upper portion of the support.

Warner and Barsky (1995) find empirical evidence of price dispersion in traditional shopping malls for eight consumer products. They find that sales frequently occur during weekends as well as during the shopping period prior to the Christmas holiday; even though demand is relatively high in each of these periods. They argue that such counter-cyclical pricing patterns can only be explained by clearinghouse models under stringent assumptions. Alternatively, they suggest that during these high intensity shopping periods, customers will be willing to invest more effort into learning the identity of the low-price firm because the associated search and travel costs can be shared across several purchases. The shopping mall environment, where a large number of stores are clustered together, makes doing so particularly easy. This encourages retailers to offer lower prices during weekends and in the weeks leading up to Christmas:

> Because consumers are more vigilant and better informed in the high demand states, individual retailers perceive their demand to be more elastic in such periods. The optimal markup of price over marginal cost is thus lower, and the market achieves an outcome closer to that of perfect competition (Warner and Barsky 1995).

The incentives that drive retail managers to behave in the ways that Warner and Barsky (1995) find should not exist in online markets. Unlike offline retail markets – where consumers are constrained by retailers' normal hours of operation – consumers can access price and product information anytime of the day from any computer (or handheld device) with Internet access, and the costs of doing so are extremely low. As a result the benefits of bulk shopping to save on search and travel costs by concentrating shopping on weekends are greatly reduced in online markets. Since consumers will be no better informed on any particular day, we should observe a stable day-to-day price distribution. During the holiday shopping period, however, we do have reasons to expect that online retail sellers will alter their behavior, but for reasons different from offline sellers. These reasons, we will argue, are consistent with the predictions of clearinghouse models, as we will discuss shortly.

To analyze whether this is in fact the case, we conduct a study that is similar to that of Warner and Barsky, but in online markets. Our dataset is collected from CNet's Shopper.com – an online price comparison site. We collected data on the daily prices set by sellers of fifteen different products from November 28, 2003 through January 31, 2004 to examine how observed price distributions fluctuate by day of the week and between the pre- and post-holiday shopping period. As predicted, we find that average prices do not systematically vary by day of the week. Also, average prices are statistically higher, not lower, during the pre-Christmas compared to post-Christmas shopping season. The new shopping environment appears to have altered the Warner and Barsky (1995) result.

Study of the pre-Christmas period allows us to further examine how well the clearinghouse models explain firm behavior in online markets. While it is a well-established fact that demand increases during this period, we conjecture that consumers' opportunity cost of time is also high during this period. Therefore, consumers will spend less time shopping to secure the lowest market price, so there will be more consumers who are uninformed about the identity of the firm offering this price. Varian (1980) has very clear predictions about how an increase in the number of uninformed consumers impacts not only the price distribution, but several other market characteristics, including the standard deviation of prices (our measure of price dispersion), the number of firms that advertise a price, and the minimum price offered. We empirically examine each of these market characteristics and find that there are no aspects of firm behavior that are inconsistent with the predictions of the Varian model. Also, none of these other market characteristics vary systematically with the day of the week, further confirming the hypothesis that bulk shopping incentives play no role in online markets.[1]

The remainder of the paper proceeds as follows. Section 2 presents an

[1]Bulk shopping is also unlikely to explain the dispersion in our sample since 1) on average these products are relatively expensive (over $500) and 2) there are no obvious product complementarities.

overview of the Varian model and provides comparative static predictions. Section 3 provides a description of the empirical methodology and data, and Section 4 presents and discusses the empirical results. Section 5 concludes.

2 Theory: Overview of the Varian Clearinghouse Model

Varian (1980) offers one of the first clearinghouse models. There are two types of consumers who purchase a homogeneous product: informed (I) and uninformed (M). Both consumer types demand one unit of the good provided the observed price p is less than the reservation price r. Each of the I consumers access an information clearinghouse, such as CNet's Shopper.com. Hence, type I consumers buy the product and pay $p_{\min}$. In contrast, each of the M consumers are unaware of or do not access an information clearinghouse to buy the product. Therefore, M consumers buy from a randomly selected firm and pay the expected market price, $\overline{p}$.

It is assumed that n identical firms compete to for these $I+M$ consumers. The I informed consumers all buy from the firm that offers the lowest price, but the M uniformed consumers are equally distributed among all firms. Therefore, firms face a trade-off: they can charge a high price and earn a larger profit from each person they sell to, but sell to only $\frac{M}{n}$ consumers. Alternatively, a firm can set a price that is lower than any other firm offers, earning a lower profit for each sale made, but sell to $I + \frac{M}{n}$ consumers.

Varian (1980) shows that there is no symmetric pure strategy Nash equilibrium of this game. Instead, each firm i plays mixed strategies by selecting p_i from a common distribution, F,[2] that is common knowledge over support $[p^*, r]$. This randomization – which Varian (1980) interprets as firms periodically offering "sale" prices – causes the identity of low-price firm to change frequently. This makes it difficult for consumers to identify the low-price firm, giving firms an incentive to continue to randomize their price rather than descending into Bertrand competition.

If market conditions are stable in the Varian model, the price distribution from which firms are drawing will not change. Changing marketing conditions, however, will impact the price distribution in predictable ways. Table 1 reports the key comparative static results in Varian (1980) and (1981)[3] for several key parameters that impact the price distribution as M changes. Not surprisingly, there is a positive relationship between the number of uninformed consumers, M, and the number of firms listing prices, n. The intuition behind this result is straightforward: the market is able to sustain a larger number of firms as more uninformed consumers are present in the market. An increase in M also causes the expected price, $\overline{p}$, to increase. However, the impact on the expected minimum price, $\overline{p}_{\min}$, is ambiguous. Finally, we measure price dispersion with the standard deviation of prices. The following proposition describes how the standard deviation of prices changes with M.

Proposition 1 *When* $E(p) < \frac{\int_{p^*}^{r} \frac{\partial (pF(p))}{\partial M} dp}{\int_{p^*}^{r} \frac{\partial F(p)}{\partial M} dp}$, *under the Varian price distri-*

[2] In Varian (1980) the functional form of the distribution is $F = 1 - \left(\frac{rM}{nI} \left(\frac{1}{p} - \frac{1}{r} \right) \right)^{\frac{1}{n-1}}$, where r is the market reservation price, M is the number of uninformed consumers, I is the number of informed consumers, n is the number of firms listing prices, and p is price.

[3] Varian (1980) contained an error in the calculation of the effect of an increase in M on the minimum price, which was subsequently corrected in Varian (1981).

bution, F, the standard deviation of price is increasing in the number of uninformed consumers, M.

The proof of Proposition 1 is contained in the Appendix.

3 Empirical Strategy, Model and Results

Our primary objective is to explore whether the two pricing patterns that Warner and Barsky (1995) find in traditional offline markets that can only be explained by imposing stringent assumptions on clearinghouse models are also present in online markets. We therefore employ a similar methodology. Accordingly, we analyze daily price data immediately prior to and following the holiday shopping period for 15 popular online products – consumer electronics and computer-related products – such as MP3 players and digital cameras. We use a popular Internet price comparison site – CNet's Shopper.com – to collect the price distributions between November 28, 2003 and January 31, 2004. Thus, like Warner and Barsky (1995) we are able to study how prices charged for a particular product change over the intense holiday shopping period. Our sample consists of 22,051 firm-product-date-level price observations. Product-market-level variables, however, result in 898 observations (the number of firms is, for instance, a market level variable).[4]

After examining whether the weekend effect observed in Warner and Barsky disappears in online markets, we turn our attention to the holiday shopping period to determine whether the market variables vary in ways

[4] Not all products were listed on everyday in the sample. Thus, we have fewer than 975 product-dates.

that are consistent with the predictions of Varian's clearinghouse model as described in Section 2. Recall, the Varian model predicts that when there are more uninformed consumers in a market, the average price, number of firms and standard deviation of price should increase, and the expected minimum price for a product should change, but in an ambiguous direction. To test these predictions we formulated various regression models.

Summary statistics for these variables for all 15 products in our sample are presented in Table 2. The average weighted price of all 22,051 price observations is $510.81; making the average product in our sample expensive. The D-Link PC adapter card is the least expensive product in our sample with an average price of $42.96. The most expensive product in our sample is the Canon GL2 camcorder with an average price of $2,322.04. There is considerable price dispersion within each product in the sample: the average standard deviation of prices is greater than zero for each of the 15 products in our sample. Since the standard deviation is used to measure the variability in prices, it is inappropriate to make relative dispersion comparisons across products. This is an issue to which we will return when formulating a regression model. There is considerable firm participation: on average, more than 24 firms are listing prices daily at Shopper.com.

Using each of the four variables listed in Table 2 as dependent variables we construct four independent regressions to measure how each market characteristic varies by day of the week and before the holiday shopping period. Table 3 presents the summary statistics on the day of the week and holiday shopping period. Price observations are (fairly) uniformly distributed across days of the week: Sundays represent about 13 percent and Saturday 16 percent of the price observations in our sample. The day of the week dummy variables in the regressions will permit us to test whether the weekend effect observed in Warner and Barsky exists in online markets. To examine whether the counter-cyclical pricing pattern (lower prices during the high demand holiday shopping seasons) observed at shopping malls in Warner and Barsky exists we divided our sample into two time periods: the pre-holiday shopping period (November 28 - December 24) and the post-holiday shopping period (December 25 - January 31). About 44 percent of the price observations occur during the pre-holiday shopping season. For the purposes of our regression, we construct dummy variables to indicate the time period in which a price observation falls.

Our first task is to use OLS with Huber-White robust standard errors to control for heteroscedasticity that is likely to occur across products and firms to replicate the counter-cyclical pricing patterns observed in Warner and Barsky. Equation 1 is constructed to address whether average prices systematically vary by day of week and between the pre- and post-holiday shopping season.

$$P_{ijt} = \alpha^{(1)} + \boldsymbol{\beta}^{(1)} DOW_{ijt} + \theta^{(1)} SEA_{ijt} + \mu_i^{(1)} + \gamma_j^{(1)} + \tau_t^{(1)} + \varepsilon_{ijt}^{(1)} \qquad (1)$$

where P_{ijt} = the natural logarithm of the price charged by firm $i = 1, ..., n$, for product $j = 1, ..., m$, on date $t = 1, ..., T$; DOW_{ijt} is a vector of dummy variables indicating the day of the week, SEA_{ijt} is a dummy variable that equals one if the observation occurs during the holiday shopping period, μ_i

are firm fixed effects, γ_j are product fixed effects, τ_t are time fixed effects, and $\varepsilon_{ijt}^{(1)}$ is an error term. Since the same set of independent variables will be estimated using different dependent variables, we use parenthetical superscripts to distinguish coefficient estimates across regressions. Given that our data is daily, the presence of autocorrelation is likely. To control for autocorrelation, we re-calculate OLS using Newey-West standard errors with a four lag error structure.[5]

Table 4 presents the results of our OLS regression with the two different standard error computations for five different model specifications. Each model builds successively on the previous by conditioning on relevant variables. Model 1 is a pooled regression that conditions on exogenous dummy variables for each day of the week except for the omitted category of Tuesday.[6] Model 2 conditions on the exogenous pre-holiday shopping season. We conjecture that the pre-holiday shopping season is associated with high opportunity cost of time and more uninformed consumers. Models 3 - 5 sequentially add controls for unobservable heterogeneity stemming from several other factors. Equilibrium price dispersion observed in the Varian model assumes that firms are that homogenous. Accordingly, in Model 3 we condition on firm dummy variables to control for unobserved heterogeneity stemming

[5] Using the Newey-West procedure in this circumstance is cumbersome, because our panel has three dimensions. For the purposes of computing the Newey-West standard errors, we define the cross-section as the interaction between firms and products. An alternative would be to assume that covariances are correlated within firms, and estimate standard errors that cluster by firm. Such a procedure yields the similar results as the Newey-West procedure. Also, our panel is unbalanced because some firms did not advertise a price for every product and every date in the panel. We force the estimation to ignore such gaps.

[6] We follow Warner and Barsky (1995) by omitting Tuesday. This avoids the commonly known "dummy variable trap."

from firm differences. Table 2 also suggests that products are at various stages of their life cycle. There are six products where the difference between maximum and minimum number of firms listing daily prices across time exceeds 10. Consequently, it is important to condition on unobservable heterogeneity resulting from time variations. We condition on this additional set of time dummy variables in Model 4. Finally, Model 5 conditions on unobservable product heterogeneity, which is apparent by examining the differences in average prices across products.

Our results are quite clear: unlike in traditional offline markets, there is no difference between the average price offered by firms during a weekend and during the weekday. This finding is robustly observed in Models 1 - 5. There is no statistically significant average price difference by day of the week. Since search costs are much lower on the Internet, there is no incentive for consumers to shop in bulk, and the difference between prices on weekends and during the rest of the week disappears.

Table 4 illustrates that we have mixed support for the hypothesis that average prices will be higher during the pre-holiday shopping period. In contrast to Varian, Models 2 - 3 suggest that prices are statistically lower during the pre-holiday shopping period relative to the post-holiday period. This statistical significance, however, disappears when conditioning on product life cycle effects through the unobservable heterogeneity over time in Model 4. Surprisingly, Model 5 suggests that conditioning on all of the observable exogenous and unobservable variables, the average price is statistically higher by a magnitude of 2.2 percent during the pre-holiday shopping period. Thus, although not result against all specifications, when all unobserved heterogeneity is properly controlled for, our results are consistent with the comparative static result in Varian. Namely, that higher average prices during the pre-holiday shopping season is associated with an increase in the number of uninformed consumers.

It is clear that the empirically observed counter-cyclical pricing patterns observed by Warner and Barsky at shopping malls is not present at Shopper.com. Therefore, while the bulk shopping and increasing returns to shopping technology explain the pricing patterns at shopping malls, this explanation seems inadequate in online markets like Shopper.com.

To further explore how well the Varian model explains the behavior in prices at Shopper.com, we empirically examine other comparative static results from Table 1. Again, our strategy is to use the exogenous change in between holiday shopping periods that we assume are associated with an increase in the number of uninformed consumers and examine the impact on other market variables.

Recall, Table 2 highlights the comparative static results of the Varian model with testable implications. Under the assumption that the number of uninformed consumers increases, we should observe 1) higher price dispersion – measured by the standard deviation in prices – during the pre-holiday shopping period relative to the post-holiday period; 2) more firms listing prices at Shopper.com during the pre-holiday shopping period; and 3) an ambiguous impact on the expected minimum price. Thus, the Varian model results in clear comparative static predictions on the remaining market variables with the exception of the expected minimum price. In addition, we should not observe any statistically significant day-to-day fluctuations in these variables.

To examine whether price dispersion varies between the pre- and post-holiday shopping seasons, it is necessary to construct a measure of price dispersion that is not explained by heterogeneities due to firms, products and time. Such a measure of price dispersion can be constructed by implementing a two step approach. In step one, we use OLS to regress P_{ijt} on firm, product and daily time dummy variables as in Equation 2.

$$P_{ijt} = \alpha^{(2)} + \mu_i^{(2)} + \gamma_j^{(2)} + \tau_t^{(2)} + \varepsilon_{ijt}^{(2)} \tag{2}$$

Saving the residuals from Equation 2, $\widehat{P_{ijt}}$, and computing the standard deviation of $\widehat{P_{ijt}}$ results plus the estimate constant in measure of price dispersion that is "purged" of the effects of firm, product and times heterogeneities. The resulting adjusted standard deviation, σ_{jt}, for each product-date in the sample can then be used in step 2 to test the comparative statics in Varian. That is, we use OLS with Huber-White Robust standard errors and Newey-West standard errors with a four lag error structure to estimate the following equation.[7]

$$\sigma_{jt} = \alpha^{(3)} + \boldsymbol{\beta}^{(3)} DOW_{jt} + \theta^{(3)} SEA_{jt} + \varepsilon_{jt}^{(3)} \tag{3}$$

Table 5 reports the results for the two models represented by Equation 3.[8] Model 1, which conditions the adjusted standard deviation of price on the day-of-week dummy variables. The results indicate that price dispersion using this measure does not systematically vary by day of the week in any

[7] In this case, however, our panel has two dimensions so the cross-section can be defined properly.

[8] Recall, given the two step approach, it is unnecessary to include firm, product and daily dummy variables in these specification since the dependent variable, σ_{jt}, is purged of these effects.

statistically meaningful way. Model 2 illustrates that the lack of the day of week effect is robust to further conditioning on the pre-holiday period dummy variable. Notice, the pre-holiday period dummy variable in Model 2 is positive, as predicted by Proposition 1; however, this estimate is not statistically significant. Lack of statistical significance on this variable may not be surprising given that daily time dummy variables were conditioned upon in constructing a measure of price dispersion that is free from unobservable heterogeneity due to time – like product life cycle effects. We interpret this result to provide some evidence that price dispersion varies in the manner that is consistent with the assumed increase in the number of uninformed consumer during the pre-holiday shopping period.

We examine whether the number of firms systematically varies by day of the week and between the pre- and post-holiday seasons by estimating the following OLS regression with the two different standard error computations.[9]

$$N_{jt} = \alpha^{(4)} + \boldsymbol{\beta}^{(4)} DOW_{jt} + \theta^{(4)} SEA_{jt} + \gamma_j^{(4)} + \tau_t^{(4)} + \varepsilon_{jt}^{(4)} \tag{4}$$

where N_{jt} is the number of firms that advertised a price for product j on date t, DOW_{jt} is the usual day-of-the-week dummy variable for product j on date t, SEA_{jt} is a dummy variable indicating whether the date is in the pre-holiday season, γ_j is a product j dummy variable, and τ_t is a daily dummy variable for time t.

Table 6 presents the estimates of our number of firms regression. It is readily apparent in Models 1 - 3 that there is no statistically significant day-

[9] The Newey-West again uses a four lag error structure.

of-the-week effect. In Model 4, however, the coefficient estimate on Friday is statistically greater than the number of firms listing prices on Tuesday. On average, about three firms list prices on Friday than on Tuesday. The pre-holiday period dummy variable is positive in Models 2 - 4, and when all unobservable heterogeneity is properly controlled for in Model 4, it is statistically significant using both the Huber-White robust standard errors and Newey-West standard errors with a four lag error structure. Nonetheless, taken together Table 6 provides evidence in favor of a Varian-style clearinghouse model when the number of uninformed consumers is relatively high during the pre-holiday shopping period.

As a final step, we examine whether the expected minimum price varies systematically by day of the week and between the pre- and post-holiday periods. The Varian model yields no clear prediction on how the expected minimum price varies with the number of uninformed consumers. The expected sign on the $\theta^{(5)}$ is ambiguous. In the same way that our measure of price dispersion – the standard deviation of prices – needed adjusting by conditioning on firm, product and time dummy variables in a step one regression, we follow the same procedure for the expected minimum price. That is, we estimate Equation 2 to obtained $\widehat{P_{ijt}}$ for each observation. We construct our dependent variable for the expected minimum price as follows: $P_{jt}^{\min} = \min(\widehat{P_{ijt}})$. The expected minimum price, therefore, consists of a single observation for each product j and time t. In step 2, we use OLS to estimate the parameters of the following equation:

$$P_{jt}^{\min} = \alpha^{(5)} + \boldsymbol{\beta}^{(5)} DOW_{jt} + \theta^{(5)} SEA_{jt} + \varepsilon_{jt}^{(5)} \tag{5}$$

where the right-hand side variables are all as previously defined and the Newey-West standard errors are computed using a four lag error structure.

Table 7 presents the coefficient estimates of Equation 5. Models 1 - 2 suggest that no day of the week effect is present. There is no statistically significant difference in the minimum price by day of the week. This is consistent with the prediction of a stable day-to-day price distribution in Varian. Interestingly, Model 2 provides mixed results on the statistical significance of the expected minimum price. Using the Huber-White robust standard errors, the pre-holiday coefficient estimate of -0.011 is statistically significant at the 5 percent level, but the Newey-West standard errors with a four lag error structure renders the coefficient estimate statistically insignificant. Without further data and assumptions on the Varian model, we are unable to make a clear prediction.

In sum, the Varian model predicts that when there are more uniformed consumers in the market, firms should, on average, charge higher prices. In addition, the price dispersion should be greater when the standard deviation of prices is used to measure dispersion and more firms should list prices at Internet price comparison sites. We find evidence supporting these hypotheses, suggesting that the Varian model does an remarkable job of describing why price dispersion arises in online markets. Also, none of our market characteristics appear to vary with the day of the week. With no exception, the day of the week does not impact any of our average price or price dispersion variables. Only the number of firms is higher on Friday compared to Tuesday; by an order of magnitude of around 3 firms. Overall, however, the market is quite static from one part of the week to the next, confirming the hypothesis that no weekend effect is prevalent in online markets.

4 Conclusion

Shopping online, particularly during the holiday season, is becoming increasingly popular. According to the latest data, the 2003-04 holiday shopping season yielded a total of $8.8 billion in online shopping purchases, a 24 percent increase from the same period the year before. As shopping shifts from traditional offline markets to online markets, it is important to understand how these two market settings differ. The important study of offline markets conducted by Warner and Barsky (1995) argued that most aspects of firm behavior in such markets could be explained by "clearinghouse" models such as the model presented by Varian (1980), but they found two aspects of pricing behavior that were inconsistent with this model: prices were systematically higher on weekends than during the work week, and also before Christmas than after Christmas. They argue this occurs because consumers tend to concentrate their shopping for multiple products during these times in order to save on search costs and transaction costs.

Since these costs are greatly reduced when shopping online, we expect that these patterns will not be present in online markets, and in fact we find that they are not. By empirically examining daily price adjustments for fifteen retail consumer electronics and computer products spanning the pre- and post-holiday shopping season from November 28, 2003 until January 31, 2004, we find that prices were statistically the same over the weekend and weekdays, and that prices were typically higher, not lower, before Christmas than after Christmas. Also, we show that observed variation before and after Christmas in the standard deviation of prices (a measure of price dispersion) and the number of firms advertising prices are consistent with the predictions of Varian (1980). While online markets may not be perfectly described by the neoclassical model, clearinghouse models appear to adequately describe the reason why prices are dispersed in these markets: the presence of consumers who are unaware of the identity of the firm offering the lowest available price.

5 Appendix

Proof. The proof of Proposition 1 – that the standard deviation of prices is increasing in the number of uninformed consumers, M, in Varian's (1980) clearinghouse model – follows. The variance is defined as

$$\sigma^2 = E(p^2) - (E\,(p))^2\,.$$

Rewriting the first term and integrating the second term of the variance by parts yields the expression

$$\int_{p^*}^{r} p^2 f\,(p)\,dp - \left(r - \int_{p^*}^{r} F(p)dp\right)^2.$$

Integrating the first term by parts and manipulating the resulting expression results in

$$= 2r\int_{p^*}^{r} F(p)dp - 2\int_{p^*}^{r} pF(p)dp - \left(\int_{p^*}^{r} F(p)dp\right)^2.$$

Applying Leibniz rule it is straight forward to show that

$$\begin{aligned}
\frac{\partial\sigma^2}{\partial M} &= 2r\left(\int_{p^*}^{r}\frac{\partial F(p)}{\partial M}dp\right) - 2\int_{p^*}^{r}\frac{\partial\left(pF(p)\right)}{\partial M}dp - 2\int_{p^*}^{r}F(p)dp\int_{p^*}^{r}\frac{\partial F(p)}{\partial M}dp \\
&= 2\left(r - \int_{p^*}^{r}F(p)dp\right)\left(\int_{p^*}^{r}\frac{\partial F(p)}{\partial M}dp\right) - 2\int_{p^*}^{r}\frac{\partial\left(pF(p)\right)}{\partial M}dp \\
&= 2E(p)\left(\int_{p^*}^{r}\frac{\partial F(p)}{\partial M}dp\right) - 2\int_{p^*}^{r}\frac{\partial\left(pF(p)\right)}{\partial M}dp.
\end{aligned}$$

Notice, since $\frac{\partial F(p)}{\partial M} < 0$, $\frac{\partial\sigma^2}{\partial M} > 0$ when

$$E\left(p\right) < \frac{\int_{p^*}^{r}\frac{\partial(pF(p))}{\partial M}dp}{\int_{p^*}^{r}\frac{\partial F(p)}{\partial M}dp}.$$

■

References

[1] Baye, M.R. and J. Morgan. 2001. "Information Gatekeepers and the Competitiveness of Homogeneous Product Markets." *American Economic Review*, 91(3), 454-474.

[2] Baye, M.R., J. Morgan, and P. Scholten. 2006. "Persistent Price Dispersion in Online Markets." forthcoming in Dennis Jansen (ed.) *The New Economy and Beyond: Past, Present and Future*, Edward Edgar.

[3] Baye, M.R., J. Morgan, and P. Scholten. 2005. "Information, Search and Price Dispersion." forthcoming in Terry Hendershott (ed.) in Handbook of Economics and Information Systems, Elsevier.

[4] Baye, M.R., J. Morgan, and P. Scholten, 2004a. "Temporal Price Dispersion: Evidence from an Online Consumer Electronic Market." *Journal of Interactive Marketing*, 18(4), 101-115.

[5] Baye, M. R., J. Morgan, and P. Scholten. 2004b. "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site." *Journal of Industrial Economics*, 52(4), 463-496.

[6] Brynjolfsson, E. and M. D. Smith. 2000. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers." *Management Science*, 46(4), 563-585.

[7] Narasimhan, C. 1988. "Competitive Promotional Strategies." *Journal of Business*, 61(4), 427-449.

[8] Newey, W.K. and K.D. West. 1987. "A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55, 703-708.

[9] Rosenthal, R.W. 1980. "A Model in which An Increase in the Number of Sellers Leads to a Higher Price." *Econometrica*, 48(6), 1575-1580.

[10] Smith, M. D. and E. Brynjolfsson. 2001. "Consumer Decision-making at an Internet Shopbot: Brand Still Matters." *Journal of Industrial Economics*, 49(4), 541-558.

[11] Sorensen, A.T. 2000. "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs." *The Journal of Political Economy*, 108(4), 833-850.

[12] Varian, H. 1980. "A Model of Sales." *American Economic Review*, 70(4), 651-659.

[13] Varian, H. 1981. "Errata, A Model of Sales." *American Economic Review*, 71, 517.

[14] Warner, E., and R. B. Barsky. 1995. "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays." *Quarterly Journal of Economics*, 110(2), 321-352.

Table 1: Comparative Static Results from Varian (1980)*

	M
n	+
$\bar{p}$	+
σ **	+
F	-
$\bar{p}_{\min}$	?

* Source: Varian (1980) and authors' calculations

** This result relies on the condition in Proposition 1.

Table 2: Summary Statistics for Market Characteristics

Product	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Price over Firms, Products and Dates					
Entire Sample	22051	$510.81	$714.03	$13.00	$2,799.99
Price by Product over Firms and Dates					
Canon GL2 Camcorder	2284	2,322.04	348.95	1619.99	2799.99
Creative Nomad Jukebox Zen NX (30GB) MP3 Player	282	270.90	13.66	233.59	299.99
D-Link Air DWL 650 Wireless PC Adapter Card	1514	42.96	8.04	24.85	84.99
Linksys WRT54G Wireless-G Broadband Router	1897	92.65	15.23	64.98	191.95
Microsoft Windows XP - Professional Operating System	2714	261.84	42.22	135.00	308.50
NETGEAR WG511 54 Mbps Wireless PC Adapter Card	789	62.82	6.63	51.00	79.99
Palm Zire 71 PDA	2182	260.77	33.23	198.00	302.49
Panasonic SV-AV30 e-wear Camcorder	1576	333.46	40.68	259.00	399.99
Quicken 2004 Premier Home & Business	2031	78.85	11.10	54.00	99.95
Roxio Easy CD & DVD Creator 6.0	1631	72.88	10.62	13.00	99.95
Sonicblue Rio S10 MP3 Player	355	73.04	8.39	55.00	114.16
Sonicblue Rio S35S MP3 Player	855	117.54	12.47	100.00	149.00
Sony CLIE' PEG-UX40 PDA	1881	511.53	47.30	348.99	601.90
Sony DCR-PC330 Camcorder	1615	1,405.03	200.49	1195.00	1699.99
Toshiba Pocket PC e740	445	340.62	80.63	184.00	449.00
Standard Deviation of Price over Products and Dates					
Entire Sample	898	60.43	94.80	0.00	371.58
Standard Deviation of Price by Product over Dates					
Canon GL2 Camcorder	62	352.57	9.77	322.92	371.58
Creative Nomad Jukebox Zen NX (30GB) MP3 Player	60	7.23	6.08	0.00	21.21
D-Link Air DWL 650 Wireless PC Adapter Card	52	7.93	1.61	6.36	10.77
Linksys WRT54G Wireless-G Broadband Router	60	15.33	1.55	10.85	22.73
Microsoft Windows XP - Professional Operating System	61	42.60	1.84	40.43	47.58
NETGEAR WG511 54 Mbps Wireless PC Adapter Card	61	6.76	0.91	5.32	8.16
Palm Zire 71 PDA	61	31.63	2.16	27.54	36.58
Panasonic SV-AV30 e-wear Camcorder	61	41.41	2.29	38.30	47.15
Quicken 2004 Premier Home & Business	62	11.07	0.85	9.36	12.23
Roxio Easy CD & DVD Creator 6.0	62	9.67	2.45	7.27	15.29
Sonicblue Rio S10 MP3 Player	57	7.96	3.29	3.35	15.40
Sonicblue Rio S35S MP3 Player	59	12.85	1.30	9.48	17.40
Sony CLIE' PEG-UX40 PDA	58	45.02	9.40	34.97	62.13
Sony DCR-PC330 Camcorder	61	203.79	3.77	194.17	212.83
Toshiba Pocket PC e740	61	92.36	44.77	71.02	330.24
Number of Firms over Products and Dates					
Entire Sample	898	24.6	12.3	1	48
Number of Firms by Product over Dates					
Canon GL2 Camcorder	62	36.8	2.3	33	42
Creative Nomad Jukebox Zen NX (30GB) MP3 Player	60	4.7	4.6	1	18
D-Link Air DWL 650 Wireless PC Adapter Card	52	29.1	2.0	26	33
Linksys WRT54G Wireless-G Broadband Router	60	31.6	2.8	25	35
Microsoft Windows XP - Professional Operating System	61	44.5	2.3	40	48
NETGEAR WG511 54 Mbps Wireless PC Adapter Card	61	12.9	1.0	12	14
Palm Zire 71 PDA	61	35.8	3.8	28	41
Panasonic SV-AV30 e-wear Camcorder	61	25.8	2.4	21	29
Quicken 2004 Premier Home & Business	62	32.8	2.2	29	37
Roxio Easy CD & DVD Creator 6.0	62	26.3	4.6	17	32
Sonicblue Rio S10 MP3 Player	57	6.2	2.8	2	12
Sonicblue Rio S35S MP3 Player	59	14.5	3.3	5	18
Sony CLIE' PEG-UX40 PDA	58	32.4	1.9	27	35
Sony DCR-PC330 Camcorder	61	26.5	2.1	23	29
Toshiba Pocket PC e740	61	7.3	3.1	4	14
Minimum Price over Products and Dates					
Entire Sample	898	$340.50	$495.23	$13.00	$1,916.95
Minimum Price by Product over Dates					
Canon GL2 Camcorder	62	1827.07	78.2	1619.99	1916.95
Creative Nomad Jukebox Zen NX (30GB) MP3 Player	60	262.67	13.5	233.59	279.00
D-Link Air DWL 650 Wireless PC Adapter Card	52	26.64	2.3	24.85	30.35
Linksys WRT54G Wireless-G Broadband Router	60	78.05	2.8	64.98	79.99
Microsoft Windows XP - Professional Operating System	61	170.56	18.6	135.00	189.00
NETGEAR WG511 54 Mbps Wireless PC Adapter Card	61	51.00	0.0	51.00	51.00
Palm Zire 71 PDA	61	207.83	4.8	198.00	221.00
Panasonic SV-AV30 e-wear Camcorder	61	271.77	8.7	259.00	288.00
Quicken 2004 Premier Home & Business	62	58.47	0.9	54.00	59.95
Roxio Easy CD & DVD Creator 6.0	62	54.59	13.8	13.00	64.00
Sonicblue Rio S10 MP3 Player	57	61.44	6.2	55.00	70.00
Sonicblue Rio S35S MP3 Player	59	103.43	2.7	100.00	109.99
Sony CLIE' PEG-UX40 PDA	58	411.99	41.6	348.99	469.00
Sony DCR-PC330 Camcorder	61	1205.79	4.7	1195.00	1209.00
Toshiba Pocket PC e740	61	226.69	37.2	184.00	259.00

Table 3: Summary Statistics for Day of the Week and Pre- and Post-Holiday Shopping Period

Variable	Number of Observations	Mean Proportion	Sample Standard Deviation
Time Periods			
November 28 - December 24 (*Holiday Shopping Period*)	22,051	0.44	0.496
December 25 - January 31 (*Post-Holiday Shopping Period*)	22,051	0.56	0.496
Day of the Week			
Sunday	22051	0.13	0.331
Monday	22051	0.14	0.350
Tuesday	22,051	0.13	0.337
Wednesday	22,051	0.13	0.335
Thursday	22,051	0.15	0.353
Friday	22,051	0.16	0.370
Saturday	22,051	0.16	0.370

Table 4: Impact of the Holiday Shopping Season and Day of the Week on Average Price

Dependent Variable: Log of Price. The sample is drawn from CNet's Shopper.com for the period November 28, 2003 to January 31, 2004. Models 1 - 5 estimate an OLS regression of the dependent variable on exogenous variables controlling for seasonal and day of the week variables. Huber-White robust and Newey-West standard errors are reported in parenthesis to the right of each coefficient estimate. *** indicates statistical significance at the 1 percent level, ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level

	Model 1					Model 2					Model 3					Model 4					Model 5				
Dummy variables for:	Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors		Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors		Coefficients	Huber-White Robust Standard Errors		Newey-West Robust Standard Errors		Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors		Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors	
Day of the Week																									
Sunday	0.016	(0.0323)		(0.0227)		0.016	(0.0323)		(0.0227)		0.013	(0.0252)		(0.0177)		0.081	(0.0701)		(0.0641)		-0.005	(0.0071)		(0.0071)	
Monday	-0.014	(0.0312)		(0.0162)		-0.011	(0.0313)		(0.0160)		-0.003	(0.0244)		(0.0124)		-0.014	(0.0711)		(0.0945)		0.003	(0.0074)		(0.0091)	
Wednesday	-0.003	(0.0320)		(0.0179)		-0.003	(0.0320)		(0.0179)		0.003	(0.0248)		(0.0140)		0.002	(0.0713)		(0.0899)		-0.001	(0.0073)		(0.0037)	
Thursday	0.028	(0.0309)		(0.0209)		0.031	(0.0310)		(0.0210)		0.025	(0.0241)		(0.0164)		0.097	(0.1143)		(0.0745)		0.006	(0.0112)		(0.0070)	
Friday	0.031	(0.0301)		(0.0212)		0.032	(0.0301)		(0.0211)		0.029	(0.0235)		(0.0165)	*	0.113	(0.0691)		(0.0933)		0.005	(0.0071)		(0.0071)	
Saturday	0.025	(0.0301)		(0.0211)		0.026	(0.0301)		(0.0210)		0.025	(0.0235)		(0.0164)		0.098	(0.0700)		(0.0744)		-0.005	(0.0071)		(0.0120)	
Holiday Shopping Season																									
Nov. 28 – Dec. 24						-0.031	(0.0164)	*	(0.0339)		-0.047	(0.0131)	***	(0.0271)	*	-0.101	(0.0888)		(0.0890)		0.022	(0.0084)	***	(0.0083)	***
Intercept	5.45186	(0.0226)	***	(0.0226)	***	5.46411	(0.0235)	***	(0.0274)	***	7.72134	(0.0208)	***	(0.0298)	***	7.69949	(0.1054)	***	(0.0678)	***	7.66581	(0.0105)	***	(0.0112)	***
Firm Fixed Effects	No					No					Yes					Yes					Yes				
Date Fixed Effects	No					No					No					Yes					Yes				
Product Fixed Effects	No					No					No					No					Yes				
No. of observations	22051					22051					22051					22051					22051				

Table 5: Impact of the Holiday Shopping Season and Day of the Week on the Adjusted Standard Deviation of Price

Dependent Variable: Adjusted Standard Deviation of Price. The sample is drawn from CNet's Shopper.com for the period November 28, 2003 to January 31, 2004. Models 1 - 2 estimate an OLS regression of the dependent variable on exogenous variables controlling for seasonal and day of the week variables. Huber-White robust and Newey-West standard errors are reported in parenthesis to the right of each coefficient estimate. *** indicates statistical significance at the 1 percent level, ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level

	Model 1					**Model 2**				
Dummy variables for:	Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors		Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors	
Day of the Week										
Sunday	1.322	(6.2097)		(4.3090)		1.237	(6.2097)		(4.3051)	
Monday	0.905	(5.9830)		(3.0541)		0.549	(5.9900)		(3.0322)	
Wednesday	-0.128	(6.1295)		(3.3173)		-0.128	(6.1243)		(3.3151)	
Thursday	0.695	(5.9054)		(3.9011)		0.364	(5.9136)		(3.9224)	
Friday	0.674	(5.8138)		(3.9846)		0.542	(5.8130)		(3.9733)	
Saturday	0.706	(5.8339)		(3.9985)		0.574	(5.8337)		(3.9900)	
Holiday Shopping Season										
Nov. 28 – Dec. 24						3.883	(3.1620)		(6.6242)	
Intercept	57.840	(4.3134)	***	(4.3134)	***	56.392	(4.4720)	***	(5.1968)	***
No. of observations		898					898			

Table 6: Impact of the Holiday Shopping Season and Day of the Week on the Number of Firms

Dependent Variable: Number of Firms. The sample is drawn from CNet's Shopper.com for the period November 28, 2003 to January 31, 2004. Models 1 - 4 estimate an OLS regression of the dependent variable on exogenous variables controlling for seasonal and day of the week variables. Huber-White robust and Newey-West standard errors are reported in parenthesis to the right of each coefficient estimate. *** indicates statistical significance at the 1 percent level, ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level

	Model 1			Model 2			Model 3			Model 4		
Dummy variables for:	Coefficients	Huber-White Robust Standard Errors	Newey-West Standard Errors	Coefficients	Huber-White Robust Standard Errors	Newey-West Standard Errors	Coefficients	Huber-White Robust Standard Errors	Newey-West Standard Errors	Coefficients	Huber-White Robust Standard Errors	Newey-West Standard Errors
Day of the Week												
Sunday	-0.230	(1.6187)	(0.8173)	-0.287	(1.6118)	(0.8068)	0.033	(4.5624)	(3.6540)	-0.541	(0.8841)	(0.8661)
Monday	0.339	(1.5776)	(0.7298)	0.102	(1.5705)	(0.7037)	0.000	(4.6615)	(1.6683)	0.000	(0.7305)	(0.3674)
Wednesday	-0.356	(1.6236)	(0.7940)	-0.356	(1.6162)	(0.7879)	0.571	(4.7092)	(6.3766)	0.571	(0.6699)	(1.3953)
Thursday	0.095	(1.5633)	(0.7163)	-0.126	(1.5571)	(0.7122)	0.033	(4.6607)	(5.8036)	-0.541	(0.8238)	(1.3115)
Friday	0.451	(1.5300)	(0.7147)	0.363	(1.5206)	(0.7005)	4.643	(4.3364)	(4.3364)	2.768	(1.1430) **	(1.1446) **
Saturday	0.416	(1.5319)	(0.7114)	0.328	(1.5223)	(0.7002)	1.608	(7.0503)	(6.0644)	1.099	(1.3947)	(1.3952)
Holiday Shopping Season												
Nov. 28 – Dec. 24				2.588	(0.8230) ***	(2.1090)	1.8	(5.2175)	(5.0120)	4.03909	(1.1335) ***	(1.2409) ***
Intercept												
Date Fixed Effects		No			No			Yes			Yes	
Product Fixed Effects		No			No			No			Yes	
No. of observations		898			898			898			898	

Table 7: Impact of the Holiday Shopping Season and Day of the Week on the Adjusted Log of Minimum Price

Dependent Variable: Adjusted Log of Minimum Price. The sample is drawn from CNet's Shopper.com for the period November 23, 2003 to January 31, 2004. Models 1 - 2 estimate an OLS regression of the dependent variable on exogenous variables controlling for seasonal and day of the week variables. Huber-White robust and Newey-West standard errors are reported in parenthesis to the right of each coefficient estimate. *** indicates statistical significance at the 1 percent level, ** indicates significance at the 5 percent level and * indicates significance at the 10 percent level

	Model 1					**Model 2**				
Dummy variables for:	Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors		Coefficients	Huber-White Robust Standard Errors		Newey-West Standard Errors	
Day of the Week										
Sunday	-0.006	(0.0100)		(0.0074)		-0.006	(0.0100)		(0.0074)	
Monday	-0.007	(0.0095)		(0.0056)		-0.006	(0.0095)		(0.0056)	
Wednesday	0.000	(0.0093)		(0.0051)		0.000	(0.0092)		(0.0051)	
Thursday	-0.001	(0.0088)		(0.0059)		0.000	(0.0088)		(0.0059)	
Friday	-0.003	(0.0090)		(0.0063)		-0.003	(0.0089)		(0.0063)	
Saturday	-0.006	(0.0092)		(0.0067)		-0.006	(0.0092)		(0.0067)	
Holiday Shopping Season										
Nov. 28 – Dec. 25						-0.011	(0.0050)	**	(0.0103)	
Intercept	7.078	(0.0064)	***	(0.0064)	***	7.082	(0.0067)	***	(0.0078)	***
No. of observations		898					898			