On Retail Gasoline Pricing Websites: Potential Sample Selection Biases and Their Implications for Empirical Research

by

Benjamin Atkinson*
Economic Policy and Enforcement Branch, Competition Bureau
50 Victoria Street
Gatineau, Quebec K1A 0C9
E-mail: Atkinson.Benjamin@cb-bc.gc.ca

May 17, 2007

Abstract

Retail gasoline price data collected from the Internet can be used to study pricing dynamics in markets where prices cycle, rising by relatively large amounts in a short period of time and then falling by several smaller increments over a longer amount of time. However, there appear to be no studies which examine the sample selection biases that might exist in these data, or their implications for empirical research. To fill this gap in the literature, two data sets have been compiled for 27 stations in Guelph, Ontario over 103 days in 2005: an unbalanced panel of 12-hourly price data collected from OntarioGasPrices.com, and a balanced panel of bi-hourly price data collected eight times per day by direct observation. It is found that the Internet data tend to accurately identify features of cycles that can be distinguished using company-operated major brand station prices (e.g., the existence of a cycle, its period, height, approximate starting days, asymmetry, etc.), while features that require individual independent station data or very high frequency data might not be well-identified (e.g., the identities of price leaders and the order in which stations change their prices).

---

* I wish to thank my PhD supervisors, Andrew Eckert and Douglas S. West for their guidance and supervision; Richard Gilbert, Stuart Landon, Robin Lindsey, and Moin Yahya for helpful comments and suggestions; Raymond and Lila Atkinson for providing food, shelter, emergency loans for gas, and the exclusive use of a vehicle during the data collection process; Michaelle Tuz-Atkinson for collecting price data from the Internet; and the City of Guelph for traffic flow data. I am responsible for any errors or omissions. The views expressed in this paper are solely my own, and are not purported to be those of the City of Guelph, the Competition Bureau or the Commissioner of Competition.

1. Introduction

Over the past several decades, a considerable amount of empirical research has been published in which retail gasoline price movements are examined. However, researchers are typically restricted by data that are limited both in terms of frequency and level of aggregation; for example, some studies use prices that are averaged across stations and collected once per week, while others use station-specific data that are collected as frequently as every 12 hours, but only for a small subset of stations in the market. While these data are useful for examining certain issues regarding price movements in retail gasoline markets, they cannot be used to examine price competition between stations across large geographic areas. This is particularly true when prices in a market follow weekly cycles, rising quickly by large amounts over one or two days, and then frequently falling by relatively small increments during the next several days. High frequency, station-specific data for an entire market are ideal in these cases, but the costs of collecting such data can be very high, both in terms of time and money.

However, retail gasoline price data for certain jurisdictions are publicly available on the Internet at little to no monetary cost, which might be feasible for answering some questions that cannot be addressed by other publicly available retail gasoline price data. For example, GasBuddy.com is a network of over 173 gasoline price information sites across Canada and the United States, which operate under location-specific domain names such as OntarioGasPrices.com and ClevelandGasPrices.com. Consumers ("price spotters") voluntarily post the brands, locations and prices of gasoline retailers, and members are identified by their nicknames; non-members are identified as "visitor". Each post is also time-stamped, but the time that the price is actually observed by the spotter is not provided. Membership is free and anonymous, and members earn 150 points per posted price (up to 750 points per day) that can be used to participate in raffles for such prizes as U.S.$250 gas cards. Further points are earned by participating in other features of the site, such as opinion polls and message forums.

There are a number of scenarios in which data collected from such websites (henceforth referred to as the "Internet data") are expected to provide reliable results for researchers. For example, one might construct a data set that not only includes prices for stations in a market, but also various product and spatial characteristics for each station, such as the number of pumps, traffic counts, service levels, and other (non-gasoline) operations; one could then examine how these characteristics might influence the general direction in which price increases and decreases tend to propagate across the market, and whether certain brands tend to price higher or lower than other brands. One could likely also examine how price uniformity and volatility are influenced by a major structural change in the market, such as the entry or exit of certain players in the market, or a merger between two key players. A researcher might even collect data for multiple markets, and use them to examine how certain features of price competition differ across the markets depending on local concentration, and the presence or absence of certain types and brands of stations.

However, Internet data might not permit answers to empirical questions that require prices for stations that are relatively less likely to be reported on these sites, or which require very high frequency data. For example, it is unlikely that the data can accurately identify either the specific leaders of price increases, which stations are more likely than others to set the minimum price in the market, or the order in which stations change their prices, because all stations are not observed in the Internet data each period; some stations might be sampled two or three times per day, while others might be spotted once every two or three weeks, on average. Thus, the first stations to change their prices are not necessarily identified in the Internet data.

While the potential benefits of Internet price data are evident, no studies have been found which examine the sample selection biases that can arise in these data, or their implications for empirical research. Thus, the first purpose of this paper is to examine the extent to which prices reported on OntarioGasPrices.com are a random sample, and then identify those factors that make a particular price or station more or less likely to be reported. These results will then be used to identify which features of price movements can and cannot be accurately identified using these data. These goals will be achieved by studying price movements in a market where prices appear to cycle, using two different data sets: one consisting of bi-hourly price data collected from 27 stations in Guelph, Ontario, eight times per day for 103 days; and the other an unbalanced panel of 12-hourly price data collected from OntarioGasPrices.com for Guelph during the same 103 days. Thus, this case study will provide guidance to researchers regarding when they can use Internet data with reasonable confidence, and when a more complete data set should be compiled instead.

To anticipate results, it is found that retail gasoline consumers in Guelph report prices of major brand stations that are price controlled by their head offices significantly more often than other stations, particularly those of small independent stations. Potential explanations for these biases relate to spatial and product differentiation. However, it does not appear that they tend to post a station's price more often as it falls relative to its rivals' prices. Consistent with these results, it is found that the Internet data tend to accurately identify features of cycles that can be distinguished using company-operated major brand station prices (e.g., the existence of a cycle, as well as its period, height, approximate starting days, asymmetry, etc.), while features that require individual independent station data or very high frequency data might not be well-identified (e.g., the identities of price leaders and the order in which stations change their prices). Also, daily mode prices tend to be more accurately measured in the Internet data than daily mean prices, in part because the small non-branded independents that tend to price above the city-wide mode price are under-represented in the Guelph data. However, daily mean prices are still well-measured by the Internet data. Finally, both the daily mode and mean prices are approximated less accurately on days when new cycles are initiated with large price increases, because it can take two days for these "restorations" to be fully reflected in the data.

This paper is organized as follows. Section 2 reviews the relevant theoretical and empirical literatures regarding retail gasoline price cycles, and which stations' prices are relatively most likely to be reported by consumers. Section 3 introduces the data, which are used in Section 4 to econometrically examine certain station characteristics that might influence which prices are reported on the Internet, such as spatial and product differentiation, and relative prices. Based on these conclusions, Section 5 uses both data sets to empirically examine the extent to which any biases that exist in the Internet data influence the pricing patterns observed in the Guelph market. Section 6 concludes.

2. Literature Review

2.1. Factors That Might Influence Which Prices Are Reported

The quality of price data collected from Internet pricing sites depends importantly on whether the same conclusions would be made with a random sample of data collected by more scientific means. However, a number of factors might affect the probability that a station's price will be sampled, which could raise issues regarding sample selection biases. These factors, which are explained below, have been divided into three categories: spatial differentiation, product differentiation, and price differentials.

2.1.1. Spatial Differentiation

First, a station's geographic location can affect the probability that a station's price will be reported by consumers. For example, Sheppard, Haining, and Plummer (1992) theoretically examine spatial pricing in interdependent oligopolistic markets, in which retail prices vary with consumer price sensitivity, the choice sets available to consumers, and consumers' awareness of prices at different stations. In an empirical examination of this and related theories, Plummer, Haining, and Sheppard (1998) find evidence that the geographic location of a station is very important to consumers in St. Cloud, Minnesota. Specifically, through consumer surveys, the authors find that consumers tend to buy gas during work and shopping trips, which suggests that stations along major commuting routes will likely be most visible to consumers, and might therefore be most likely to be reported. It also suggests that if a station is located along a major commuting route, then its price will be relatively visible on weekdays, while a station that is near a shopping center might also be relatively visible on weekends.

2.1.2. Product Differentiation

A station's particular characteristics might also influence whether its price is more or less likely to be observed than prices at other stations. As such, stations can be divided into four groups, as follows. First, Groups A and B stations operate under refiner (major) brand labels; Group A stations are price controlled by their respective brands' head offices, while Group B station managers have some independent price control. On the other hand, Groups C and D stations (independents) do not operate under a major refiner brand label; in this paper, Group C stations are distinguished from Group D stations by their relative size (higher regular-grade nozzle counts), and might also be more recognized by consumers for non-gasoline offerings, such as convenience items, groceries and other merchandise.

There are a number of reasons why consumers might be relatively more likely to observe the prices of stations in a certain group during their commutes. For example, consumers might be more inclined to report Groups A and B stations' prices in order to draw attention to their pricing practices, because they believe that the major refiner brands are "price gouging", or possibly involved in an illegal price-fixing agreement; Eckert and West (2004b) find evidence of such a perception in the comments that were posted by Vancouver consumers along with the stations' prices. On the other hand, consumers might be relatively more likely to post prices of certain "maverick" Group C stations, which tend to price closer to the wholesale price of gas, to further draw attention to differences between their prices and the prices of the major brand stations.

A consumer might also be relatively more likely to observe prices at stations that have other operations, such as convenience stores and car washes, which can attract their attention due to demand complementarities. As noted above, many Group C stations are in this category.

Finally, stations can be divided both across and within groups based on brand loyalty (see Kleit, 2003, 15).[1] Empirical evidence in this regard is provided by Plummer, Haining, and Sheppard (1998), who find that the three most important characteristics of a gas station to consumers surveyed in St. Cloud, Minnesota are quality of service, brand, and extra services, respectively.

2.1.3. Relative Ranking of a Station's Price

Perhaps the most obvious reason why a consumer might report a certain station's price is that this price is higher or lower than prices of other stations. For example, the expressed purpose of GasBuddy.com is for consumers to report the lowest prices in their markets. Consistent with this goal, Eckert and West (2004b) note that some consumers who post Internet prices for Vancouver leave messages indicating that they are attempting to point out relatively low prices. However, other consumers in their data appear to be trying to identify relatively high prices, while others appear to be attempting to demonstrate price uniformity across stations.

2.2. Edgeworth Cycle Theory

In the standard model developed by Maskin and Tirole (1988), two identical firms maximize their

---

[1] For example, Petro-Canada, Esso, and Shell have Petro-Points, Esso Extra, and Air Miles loyalty programs, respectively; Sunoco has a "Swipe and Save" program for CAA members; and Canadian Tire, 7-Eleven, and Pioneer have cash-back programs in the form of special currencies that can only be used in their stores.

present-discounted stream of profits by setting prices for a perfectly substitutable product over an infinite horizon. Marginal cost is constant, there are no fixed costs or capacity constraints, and the firm with the lowest price serves the entire market. If both firms charge the same price, then they split market demand evenly. Price competition takes place in discrete time, and prices are chosen over a finite grid. Finally, Maskin and Tirole (1988) restrict a firm's strategies to depend only on the most recent price set and current marginal cost. Using this framework, the authors prove that for a sufficiently fine price grid and a discount factor near one, many Markov perfect equilibria exist, including Edgeworth cycles.

The structure of an equilibrium cycle is described as follows. Beginning at the top of the cycle, a firm undercuts its rival's price by one unit; this strategy is played by the firm because it expects its rival to do the same in the next period, i.e., it expects to earn zero profits in the next period regardless of its strategy. These one-unit undercuts continue until one firm lowers its price to marginal cost. There is then a war of attrition of indeterminate length as each firm waits with positive probability for the other to initiate a cycle "restoration" by raising its price to the new cycle peak, which is some function of the monopoly price (i.e., each firm plays mixed strategies where the expected profits of waiting and relenting are equal). This reluctance to lead a restoration is not because it might not be followed, but rather due to the expectation that the follower will undercut this price by an incremental unit, causing the leader to make zero profits for two consecutive periods.[2] The cycle is then repeated.[3]

Eckert (2003) extends the basic model to allow firms to differ in size, which can be measured

[2] The knowledge that leading a restoration will result in two consecutive periods of zero profits also deters firms from raising their prices before they equal marginal cost.

[3] Noel (2006) extends this model to permit randomly fluctuating marginal cost, and shows computationally that if there is product or spatial differentiation, capacity constraints or three firms, then Edgeworth cycles can still exist in equilibrium. He also demonstrates that with three firms, a firm might abandon its attempt to lead a restoration if one or both of its competitors do not follow this lead quickly enough; Noel (2006) calls these failed attempts "false starts".

by the number of stations each one operates in the market; if both charge the same price, their shares of market demand are proportional to their relative sizes. He finds that the large firm tends to lead price increases, while the small firm is relatively more likely to undercut its rival's price, because the market sharing rule is biased against it when both charge the same price; given the choice of either matching or undercutting its rival's price, the undercutting strategy tends to be more profitable for the small firm, while the matching strategy tends to be more profitable for the large firm.

2.3. Empirical Studies of Cycling Markets

Most studies of retail gasoline markets use price data that are collected once every one or two weeks, and which are averaged across all sampled stations. For example, Eckert (2002) and Noel (2007a) use weekly average price data to examine retail gasoline price cycles. Such data can be useful for studying certain structural characteristics of cycles if cycle durations are longer than a week, but they cannot be used to study the price movements of particular stations.

Recognizing the deficiencies of weekly average price data, Noel (2007b) collected station-specific price data for 22 stations in Toronto, every 12 hours for 131 days in 2001. The author finds that these prices do follow cycles consistent with the Edgeworth cycle theory, that major brand stations tend to lead price increases, and that independents tend to lead price decreases. However, the fact that these 22 stations are all located within a small section of the city implies that the author is unable to examine spatial pricing patterns between stations, such as whether price decreases tend to be initiated in sections of Toronto where certain brands are relatively highly concentrated.

Since publicly available data are too limited to spatially examine intra-city retail gasoline price competition, Eckert and West (2004a-b; 2005) collected station-specific price data from Internet gasoline price sites for Vancouver and Ottawa. Eckert and West (2004a) find that the existence of cycles seems to depend on the presence of aggressive "maverick" retailers that prevent tacit collusion, such as Sunoco and Pioneer in Ottawa, and ARCO and Tempo in Vancouver. Eckert and West (2004b) find that price decreases in Vancouver appear to originate in regions where ARCO and Tempo are most highly concentrated, and restorations are usually initiated on Tuesdays and Wednesdays. Finally, Eckert and West (2005) examine whether prices in Vancouver tend to be uniform across stations, consistent with a competitive theory of gasoline pricing advanced by industry participants. While these data arguably generate reliable results for the authors' purposes, no studies have been found which empirically investigate the types of questions that can be reasonably studied using Internet gasoline prices. Such an investigation would ideally compare data collected from the Internet for a set of stations in a market to a "complete" balanced panel for the same stations during the same time period.[4]

3. The Data

3.1. Balanced Panel of Bi-Hourly, Station-Specific Price Data

Regular-grade fuel prices in cents per liter (cpl) were collected every two hours (8:00AM to 10:00PM) from August 14 to November 24, 2005 for 27 stations in Guelph, a city in southern Ontario with an approximate population of 106,000.[5] Station characteristics were also collected, including operating hours (24 hours or not), service levels (full-, self-, split-serve), capacities (regular-grade nozzle counts), and other operations (repair bay, variety store, car wash). Next, in order to quantify a station's visibility

---

[4] Other empirical studies of retail gasoline markets that use Internet data include Wang (2006) and Lewis and Marvel (2007). Wang (2006) collected prices from a website operated by the Western Australian government (which regulates the timing of price changes) to study Edgeworth cycles in Perth. Lewis and Marvel (2007) use U.S. price data collected from GasBuddy.com to examine the relationship between consumer search behavior and retail gasoline price movements.

[5] The last period during which prices were collected on November 24 began at 4:00PM.

to consumers, traffic flow data were obtained from the City of Guelph for major intersections from 2003-05, which include the estimated annual number of cars that travel in each direction of a given intersection; the most recent data for each station's nearest available intersection are used in this paper. Finally, daily rack price data for London, Ontario were obtained from MJ Ervin & Associates to be a proxy for marginal costs.

Station locations are plotted in Figure 1, and the approximate collection times for each station are also provided with this figure; for example, Station 12's price was collected at approximately 14 minutes after the hour.[6] Esso, Petro-Canada, Shell/Beaver and Sunoco are the vertically-integrated (major) brands in the city, while Canadian Tire, 7-Eleven and Pioneer are among the independents. Selected station characteristics are provided in Table 1, where major brand stations tend to be 24-hour, self-serve stations with relatively high capacities, convenience stores and car washes, and no repair bays. Stations selling gasoline under the 7-Eleven, Canadian Tire and Pioneer brands are similar to these major brand stations in terms of capacities and other characteristics, while the other independents are full-serve stations with relatively low capacities, limited hours and repair bays, but no other operations.

As argued in Section 2.1, a consumer might be more likely to report a station's price if it is price controlled by a refiner brand. Thus, an attempt was made to determine whether each station's price is set by its manager or its supplier. Representatives of the above seven brands were contacted, and six reported that each station's price is company-controlled (Station 14's manager is permitted some control); Esso refused to provide price control information. However, according to MJ Ervin & Associates Inc. (2006, Appendix A), 7-Eleven Canada controls prices at its Esso-branded stations, implying that Station 7's price is not controlled by Esso. Also, Station 17 has a Rainbow-branded car

---

[6] Esso Station 28 is excluded from the sample because it did not post its price; its price was also never observed on OntarioGasPrices.com during these 103 days.

wash and convenience store, and is described on an Esso-affiliated website as a family-operated dealer.[7] Finally, the empirical literature on station contracts suggests that a company is more likely to control prices at stations with longer hours, greater pump capacities, or convenience stores, and is more likely to delegate price-setting authority to stations that are full-serve and/or have repair bays.[8] Thus, Esso most likely controls prices at Stations 19, 21, and 26. Based on these conclusions, major brand stations are divided in Table 1 by the likely source of price control, where Group A stations' prices are either known or believed to be controlled by the head office of the brand, and Group B stations' prices are not. Independents are also divided by size (i.e., station-specific nozzle counts), where Group C stations have relatively more regular-fuel nozzles than Group D stations.

3.2. Unbalanced Panel of Twice-Daily, Station-Specific Price Data

During these same 103 days, price data were also collected twice per day (noon and midnight) for Guelph from OntarioGasPrices.com. These data include the brand and location of each station, the nickname of each price spotter ("visitor" for non-members), and the time that each price was posted on the site. A total of 2,101 price observations are included in the data set;[9] and visual comparisons of the

---

[7] See "Esso Rebecca Run for SMA" (http://www.rebeccarun.com/esso_dealers/rainbow.html; visited 2007-05-01).

[8] For example, see Shepard (1993), Slade (1998), and Taylor (2000).

[9] One price posted on November 14 is excluded from the data, because the same spotter reported the same price for the same station two minutes earlier. Also, 11 "prank" prices have been omitted from the data, as they are for non-existent (and usually obscene) locations. Finally, the specific station being identified by a spotter is reasonably assumed for 24 price reports, based either on knowledge of the road network or the locations of stations reported concurrently by the same spotter; such ambiguities include identifying a location that applies to two different stations (e.g., Esso on Edinburgh Road N.), identifying a nearby intersection (e.g., Petro-Canada on Willow & Edinburgh instead of Willow & Silvercreek), and identifying the wrong brand for the listed intersection (e.g., identifying the station as Cango instead of Canadian Tire or Can-Op).

city-wide daily mode and mean price series across the two data sets are provided in Figures 2a and 2b, respectively, where it can be seen that the Internet trends coincide quite well with their corresponding balanced panel trends. However, the mode price series appear to overlap more accurately than the mean price series, with the notable exception of certain days when the mode price in one data set rises one day later than the mode price in the other data set.

The reliability of the Internet data depend, in part on the accuracy of the prices reported by each spotter. Thus, each report has been categorized based on its accuracy as follows. Using the collection times reported with Figure 1, each price included in this unbalanced panel is compared to the prices observed in the bi-hourly periods immediately preceding and following the price report. If the reported price equals one of these two prices, then the reported price is considered to be "correct". However, if the spotted price is observed no earlier than 10:00PM on the previous evening, then it is labeled as "outdated".[10] If it is instead observed no later than 8:00AM the next morning, then the post is flagged as being reported "early".[11] It is otherwise labeled "wrong".[12] Based on this methodology, 1,502 (71.5%) prices in the unbalanced panel are considered correct, while 138 (6.6%) are outdated. Another 14 (0.7%) are flagged as early, and 447 (21.3%) are considered wrong. In summary, 1,655 (78.8%)

---

[10] Note that six of these outdated reports are observed between midnight and 1:00AM, but are treated as being posted before midnight to assess their correctness.

[11] This category was created to consider the possibility that market participants might use this site to signal price changes to their competitors, or as a way to advertise their prices to consumers. There does not appear to be any evidence in the data to support this hypothesis.

[12] Note that until its signs were amended to display four-digit prices, Station 13 rounded its $1+ prices down to the nearest cent on its signs; for example, when it charged 101.9 cpl, it posted "101". Similarly, Station 15 always rounded its $1+ prices *up* to the nearest cent, so when it charged 102.5, for example, it posted "103". For these stations, the actual prices being charged were read directly from one of their pumps each period. Both the posted and actual prices were considered when assessing the accuracies of the reported prices for these two stations.

of all 2,101 price reports are either known or believed to have been charged some time during that day.

Next, there is a question of the degree to which prices are wrong. It might be that errors tend to be lower than a cent per liter, because consumers tend to ignore the last digit of the price; they might assume it is zero (for example, see Schindler and Kirby, 1997) or a frequently observed ending, such as "9" or "5" (e.g., Basu, 2004). It is found that of the 461 wrong/early price reports, the integer parts of 274 (59.4%) of them are accurate at some point of the day prior to the price post. Furthermore, the average absolute difference between the reported price and the closest price charged prior to that posting is 1.1 cpl, which is larger than the 0.9 cpl maximum predicted by the focal price explanation.[13] These statistics suggest that while consumers might tend to pay less attention to price endings, there may be other explanations for why errors are made, such as expectations based on prices observed at other stations or simple carelessness on the part of certain price spotters.

Finally, the Guelph site is dominated by a few spotters, and therefore any biases held by these spotters might influence the degree to which the overall data are biased. In particular, there are 81 specific price spotters in the data who post 2,033 of these prices (the other 68 prices are reported by "visitor"), the majority (50.8%) of all 2,101 price reports are made by four spotters, and two particular spotters are responsible for 858 (40.8%) of all price reports. Also, 292 (63.3%) of all 461 wrong/early price reports are made by these two spotters. This disproportionately high number of errors suggests that regular spotters are not necessarily more conscientious of the accuracy in their price reports.

---

[13] Note that this proportion includes two "visitor" posts on September 13 that seem to be pranks; they are for real stations (Stations 21 and 24), but are approximately 30 and 60 cpl lower than all other prices observed in both data sets. If these two posts are omitted from the analysis, then this bias falls to 0.9 cpl. Thus, a researcher using Internet price data should examine the data for any such outliers.

4. The Randomness of the Sample

The purpose of this section is to examine the extent to which prices reported on this site are not simply a random sample, and then identify those factors which make a particular station's price more likely to be reported. As argued in Section 2.1, stations might be more or less likely to be spotted based on spatial and product differentiation, as well as relative prices. Therefore, two potential sources of bias will be examined in this section: biases based on groups of stations and biases based on price differences.

4.1. Biases Across Groups of Stations

It was argued in Section 2.1 that Groups A and C stations might be more likely to be reported in Internet data than other stations; if this is true, then a researcher likely should not use these data to study issues that require individual Group D station prices, such as the speed with which these stations respond to rival price movements, or whether they tend to set the minimum price in the market on a daily basis.

An examination of this prediction uses Table 2, which lists the average number (and proportion) of days that each station's price is observed in the Internet data (averaged by group), both before and after omitting price reports for the two most frequent spotters. In both cases, the counts for Groups A and C stations are higher than for other stations, especially Group D stations. This conclusion is strengthened statistically by constructing a variable ($GROUP_i$) that equals "1" for Group A stations, "2" for Group C stations, "3" for Group B stations, and "4" for Group D stations; the correlation coefficient between $GROUP_i$ and the number of days (out of 103) that Station i is spotted ($COUNT_i$) is -0.75 if all spotters are included in the data, and -0.63 after excluding the two most frequent spotters.[14]

It was also argued in Section 2.1 that spatial location might be an important determinant of

[14] All correlation coefficients reported in this paper are Pearson correlations.

whether a station's price is observed by consumers. To examine this prediction graphically, Figure 3 differentiates stations based on how many days (out of 103) they are observed in the Internet data (including all spotters). This figure suggests that a spatial bias does exist in regard to which stations tend to be spotted; while stations in the northwest portion of the city (Stations 18 to 23) are all spotted on at least 80% of the 103 days in the sample, northeastern stations (Stations 1 to 11) are spotted less frequently, particularly Group D stations. If the two most frequent spotters are excluded from the data (not shown in the map), then all 11 northeastern stations are reported less than 20% of the time. Thus, it seems that whether a specific station's price is observed can depend largely on a single spotter.

In order to quantify spatial differences across stations, the estimated number of cars that pass each station every year have been calculated, and are summarized by group in Table 2. It can be seen that in addition to being spotted more often, Groups A and C stations tend to have higher traffic flows than Group B stations, and especially Group D stations. Statistically, the correlation coefficient between a station's traffic flows ($TRAFFIC_i$) and the number of days that it is spotted is 0.53, which suggests that traffic flows might influence whether a station's price is spotted in the Internet data.

Therefore, based on the statistical analysis conducted in this section, it is concluded that there does appear to be a bias in the data toward Groups A and C stations, which tend to not only have other non-gasoline operations which can act as demand complementarities, but also tend to be located along relatively high-traffic routes where they are likely visible to more consumers than other stations.[15]

[15] While an econometric model could be formulated where $COUNT_i$ is regressed on three group dummies (Groups A to C) and $TRAFFIC_i$, there is a potential endogeneity problem with this regression, because a refiner brand might separate its stations into Groups A and B based on traffic flows at each station. Consistent with this argument, the correlation coefficient between $GROUP_i$ and $TRAFFIC_i$ is -0.64. However, this model was estimated both with and without $TRAFFIC_i$ using GeoDa™ (available at https://www.geoda.uiuc.edu; visited 2007-05-01), which controls for spatial autocorrelation; the results are consistent with the conclusions made above, which are that Groups A and C stations are reported statistically significantly more often than Groups B and D stations.

4.2. Biases Based on Relative Prices

As argued in Section 2.1, a consumer might be more likely to report a station's price if it falls relative to the prices of competing stations. Consistent with this prediction, the correlation coefficient between the mean retail-rack margins (as summarized by group in Table 2) and $COUNT_i$ is -0.65 using all observations, and -0.67 after excluding the two most frequent spotters, implying that stations with relatively high margins are reported less often than stations with lower margins.[16] If consumers are indeed more likely to report a station's price as its relative price falls (rises), then not only will station-specific price data be biased (as concluded in Section 4.1), but market mode and mean prices will also tend to be too low (high).

Thus, an important question to address is whether a station's price is more likely to be reported as its price falls relative to the market mode; if this null hypothesis is rejected, then a possible alternative explanation is that consumers tend to report prices based on other factors, such as spatial and product differentiation. To econometrically test this null hypothesis, the following probit regression has been estimated using a balanced panel of 2,781 observations (one per day for each station):

$$COUNT_{it} = \alpha + \beta DIFF_{it} + \gamma DAY0_t{*}DIFF_{it} + \delta NOPOST_{it} + \lambda STATION_i + \sigma DAY_t + v_{it}$$

where $COUNT_{it}$ is a dummy that equals one when Station i's price is observed in the Internet data on Day t, $DIFF_{it}$ equals the difference between Station i's mode price and the market mode price on Day t, $DAY0_t$ is a dummy that equals one on a day when a restoration attempt is observed in the balanced panel, $NOPOST_{it}$ is a dummy that equals one if Station i's price is not posted on its pricing sign on at least four periods on Day t, $STATION_i$ is a vector of 26 station-specific dummies, DAY_t is a vector of

[16] The retail-rack margin is calculated to be the retail price minus the current London rack price, the Ontario provincial gas tax (14.7 cpl), the federal gas tax (10.0 cpl), and the federal Goods and Services Tax (7%), which is levied on both the price and the excise taxes.

102 daily dummies, and v_{it} is a random error term.

The model specification is based on the literature review in Section 2.1. Specifically, modes are chosen to define $DIFF_{it}$ because these are the prices that are most often observed on a given day. It might be predicted that β will be negative, because consumers may be more likely to report a station's price as it falls relative to the market mode; on the other hand, consumers might be more likely to report a station's price as it *rises* relative to the market mode, in which case β would be positive. Furthermore, to control for the possibility that consumers might change their focus from low prices to high prices on days when restorations are attempted, the interaction term, $DAY0_t$*$DIFF_{it}$ is included in the regression, where a restoration day (Day 0) is defined to be a day when the bi-hourly market mode price rises in the balanced panel after at least one station raises its price to this mode.[17]

Next, $NOPOST_{it}$ is included to control for days when a station did not post its price for at least four periods (half the day); δ is predicted to be negative. Finally, 26 station-specific dummies are included to control for differences across stations, such as product and spatial differentiation, while 102 daily dummies are included to control for unusual circumstances on specific days, such as bad weather or temporary city-wide demand shocks. These daily dummies indirectly control for day-of-the-week effects, as well, such as fewer reports on weekends when consumers are not commuting to work.[18]

The results of this regression are provided in the second and third columns of Table 3. First, it

[17] Note that while this definition differs from the one used in Section 5 (which is based on *daily* mode price increases to permit direct comparison between the two data sets), it is appropriate for this analysis because consumers are expected to notice when more than 50% of the stations in the market raise their prices by several cents per liter during the afternoon/evening of a single day, even if these increases are not enough to raise the daily market mode price.

[18] The excluded dummies are $STATION19_i$ (because it is one of the stations that often did not post a price) and $DAY96_t$ (which is considered to be a typical day since no restoration attempts were observed, and no unusual conditions were evident with respect to traffic, weather, queues, etc.).

is shown that the LR test statistic for the overall significance of the model is 1,144.4624, so the null hypothesis that all 131 coefficients simultaneously equal zero is rejected at the 1% level of significance. Also, the estimated coefficients were used to obtain the predicted probability that a station's price will be spotted. If it is predicted that Station i's price will be reported if this probability exceeds 0.5, then the probit correctly predicts the dependent variable 76.7% of the time (66.9% for $COUNT_{it} = 0$ and 83.6% for $COUNT_{it} = 1$). Compared to the success rate of the naive prediction that every station's price will be reported every day (58.4%), the model appears to fit the data reasonably well.

With respect to the estimated coefficients, β and γ are positive but not statistically significant from zero at the 5% level of significance, suggesting that a station's price is not significantly more likely to be reported if its price falls relative to the market mode price. On the other hand, δ is negative and significant at the 5% level, indicating that Station 19 is significantly less likely to be reported when it does not post its price on its pricing sign.

Since δ is statistically significant, the partial effects of a change from Station i posting its price to not posting it on the probability of that station's price being reported have been calculated for Stations 10 and 19; they are -13.4% and -5.7%, respectively.[19] An interpretation of these effects is that Group C stations tend to be reported less often than Group A stations, as found in Section 4.1, and thus might be more likely to be omitted from the Internet data when they do not post their prices; an alternative explanation is that stations in the northeastern section of the city tend to be reported less frequently than stations in the northwestern section of the city, as shown in Figure 3.[20]

---

[19] The partial effect for Station 19 is calculated by initially setting $NOPOST_{it}$ equal to zero, all other dummies always equal to zero, and $DIFF_{it}$ equal to its sample mean (= -0.0798993); for Station 10's partial effect, $STATION10_i$ is also always set equal to one (its coefficient is -0.64423).

[20] Several alternative model specifications have been estimated to examine the robustness of these general results, i.e., that α is positive and significant at the 1% level, δ is negative and significant

Finally, to examine whether the results are robust to different measurements of price differences across stations, this probit has been re-estimated using daily mode price rankings instead of differences between a station's daily mode price and the daily market mode price. Specifically, stations are ranked on a 27-point scale, where a higher ranking indicates a higher mode price for the station in question; if more than one station has the same daily mode price, then they are all ranked equally. For example, if two stations are tied for the lowest mode price, then they are both ranked number one, and the next-lowest priced station is ranked third.

The results of this alternative regression are displayed in the last two columns of Table 3; as with the first probit regression, the constant is positive and statistically significant at the 1% level of significance, while the coefficient for $NOPOST_{it}$ is negative and significant at the 5% level. However, an important difference between the results in the two models is that the variable used to compare prices across stations ($MODERANK_{it}$) is negative and statistically significant at the 5% level, meaning a station's price is significantly more likely to be reported as its ranking falls along the 27-point scale.

However, it appears that the methodology used in the first regression is more appropriate than in the second regression, because the ranking system can overestimate price differences across stations. To demonstrate this possibility, in the balanced panel, seven to eight stations have the same daily mode price on an average day. Thus, since the ranking system assigns all of these stations the same rank, if Station x's daily mode price rises by as little as 0.1 cpl in one day, while all other station's daily mode

---

at the 5% level, and neither β nor γ are significantly different from zero at the 10% level of significance. First, six alternative definitions of $DIFF_{it}$ are the differences between Station i's daily mean/minimum/ maximum price and the city-wide daily mode/mean price; a seventh is the difference between Station i's daily mode price and the city-wide daily mean price. Despite which definition is used, none of the main conclusions of this analysis change. These conclusions also remain unchanged if $NOPOST_{it}$ is modified such that it equals one on Day t iff Station i's price is not posted for at least j periods during that day (j = 2, ..., 8). If j = 1, then δ is significant at the 1% level for all eight definitions of $DIFF_{it}$.

prices remain unchanged, then Station x's ranking will jump by six to seven points. In other words, the significance of the ranking variable in Table 3 might be due to the ranking system, rather than because a station's price is more likely to be reported as its daily mode price falls relative to its rivals.[21]

5. Implications of Sample Selection Biases for Empirical Research

In Section 4, it was demonstrated econometrically that while it does not appear that a particular station's price is reported more often as its mode price falls relative to the market mode price, there appears to be a sample selection bias with respect to which types of stations are reported: Groups A and C stations are reported significantly more often than Groups B and D stations. These results suggest that these data might be reliable if a researcher wants to analyze retail gasoline markets using daily mode prices (assuming a clear mode price exists in the market). However, they likely cannot provide clear answers to questions that rely on the pricing dynamics of particular stations, especially Groups B and D stations.

The purpose of this section is to explore these expectations further, by examining some basic predictions of the Edgeworth cycle theory. Consistent with the approach of Eckert and West (2004b) who examine mean and mode price movements in their Internet data for Vancouver, duplicate price reports for the same station on the same day are excluded from the data, leaving a total of 2,005 price observations; these deletions are made to avoid double-counting a station's price.

---

[21] If rankings are based on mean daily prices, which are less likely to be identical across stations (three to four per day, on average), and therefore less prone to large jumps in a station's ranking, then the rank variable is *not* statistically significant at the 5% level of significance. This result is consistent with the argument that a change in a station's price relative to other stations will not significantly influence the probability that it will be spotted. Similarly, if rankings are based on maximum daily prices, then the rank variable is not significant at the 5% level. Finally, while the rank variable is significant at the 5% level if minimum daily prices are used, minimum price rankings have a similar drawback to mode rankings: on average, seven to eight stations have the same daily minimum price.

5.1. The Timing of Restorations

In Section 4.2, a Day 0 was defined to be a day when the periodic (bi-hourly) market mode price rises after at least one station raises its price to this mode price; the next day is labeled Day 1, and so on. However, a problem with applying this definition to the Internet data is that every station's price is not reported in every 12-hour period, and therefore determining whether a station increases or decreases its price is not always clear. Therefore, for this section and those that follow, a modified version of the methodology of Eckert and West (2004b) will be used to define Days 0: first, identify every day when the daily city-wide mode price rises by at least 4.0 cpl. If the first station-specific occurrence of this mode price is observed on this day, then this is Day 0; if it is the previous day, then that is Day 0. Using this methodology, both data sets identify 13 Days 0.

Comparing the 13 dates identified in each data set, it appears that the Internet data reasonably approximate the days on which restoration attempts are made, provided they are not "false starts",[22] and as long as restoration attempts are not made on consecutive days. Specifically, 12 of the 13 restorations identified in each data set are identified in both data sets, but five are identified one day apart. This is because it typically takes 24 hours for all stations in Guelph to raise their prices to their cycle peaks in the balanced panel, so it might take an additional day to observe an increase in the daily mode price in either data set. Given this lag, if a restoration attempt is abandoned on Day 1 (implying it is a false start), then a daily mode price increase might never be identified in the Internet data. Thus, while the balanced panel identifies a restoration attempt on October 18, the Internet data never identify it.

It should be stressed that while the Internet data might identify restoration attempts with a lag of one day, the same problem is faced by the balanced panel when restorations are identified based on

[22] See Noel (2006), *supra* note 3.

daily mode price changes. For example, during the days following the impact of Hurricane Katrina on the United States, three daily mode price increases are observed in the Internet data over four days (August 31 to September 3), but since the first two mode price increases are observed on consecutive days, the balanced panel identifies one big restoration instead. These results indicate that while the Internet data can overlook false starts and restorations that occur on consecutive days, this problem is not unique to this type of data. Furthermore, unless the market being studied follows daily cycles, the number of restoration attempts missed by the Internet data are likely to be relatively rare anomalies.[23]

A second implication of these results is also relevant for more general applications of these data. If the market being studied is characterized by such highly asymmetric cyclical pricing patterns, then even if mode and/or mean price movements tend to be accurately portrayed by the Internet data, they might be highly underestimated on days when restorations are initiated. Thus, a researcher should take this possibility into account when using mode or mean prices to study retail gasoline markets where prices cycle. The biases that can arise from the inclusion of Days 0 will be further demonstrated below.

5.2. Basic Patterns in Retail Price Changes

This subsection examines whether basic patterns in retail price changes, in terms of daily mode and mean prices, can be reasonably approximated using data collected from the Internet. The results of this section will be used to evaluate arguments made in previous sections that while one might be advised

---

[23] Note that the definition of Days 0 used in Section 4.2 identifies 16 Days 0, which include the 14 days identified between the two data sets in the current section. The other two are identified on September 12 and October 12, during which the market mode prices in the balanced panel are observed to increase by no more than 2.7 cpl, at least 11 stations are not observed to raise their prices at all, and most price increases are completely reversed by the end of Day 1. The finding that these two restoration attempts are not identified in the Internet data (using the 4.0 cpl threshold) is consistent with the conclusion that Internet data can reasonably approximate the timing of restoration attempts as long as the price increases are sustained for longer than a day.

to not base their analysis on Internet data for particular stations, results based on mode (and possibly also mean) price movements might be reasonable. The results are summarized in Table 4, and when appropriate, are divided into those for the Balance Panel ("BP") and the Internet data.

It is first found that the Internet data more accurately characterize "true" price movements if mode prices are studied instead of mean prices. Specifically, the daily city-wide average price decreases 2.5 (3.1) times more often than it increases in the Internet (BP) data, while the city-wide daily mode price decreases 4.2 (4.3) times more often in the Internet (BP) data. Similarly, city-wide average price increases are 2.4 (2.9) times greater in magnitude than decreases using the Internet (BP) data, while city-wide mode price increases are 4.1 times greater in magnitude than decreases, regardless of which data set is used.

Further evidence regarding the relative accuracy of the mode price in the Internet data is found by examining the proportion of times that the daily mode and mean prices are exactly identical across data sets. Table 4 shows that while the mode price is identical across data sets on more than 50% of all 103 days, the mean prices are identical less than 10% of the time. One potential explanation for the relative accuracy demonstrated in the mode series is that Group D stations are under-represented in the Internet data. On average, the daily mode price for a Group D station equals the daily city-wide mode price in the balanced panel on 5.7% of all 103 days, so the under-representation of Group D stations is unlikely to affect the mode price. Also, it was statistically demonstrated in Section 4.1 that Group D stations tend to have higher margins than Group A stations. Thus, the low number of price reports for Group D stations appear to contribute to the Internet data's frequent underestimation of the mean daily city-wide price (65.0% of all 103 days), but are less likely to affect the calculated mode price.

However, while daily mean prices do not tend to be exactly accurate, they do tend to be *almost* accurate. According to summary statistics and correlations provided in Table 4, including all 103 days in the sample, the Internet data underestimate the "true" daily mean price by 0.3 cpl, on average, while the correlation coefficient between the two daily mean price series is quite high at 0.998. On the other hand, if the mean sample margins are calculated *for each station* using both data sets, then the correlation coefficient between these two mean series is 0.640. Therefore, if the questions being studied by a researcher do not require exact average prices, then the Internet data might be relevant for these purposes, even though one likely should not base any conclusions on station-specific prices.

Next, an interesting observation made from the statistics in Table 4 is that even though the daily mode price tends to be exactly accurate more than 50% of the time, overall, it is still less accurate than the average price series; specifically, the Internet data underestimate the true daily mode price by 0.6 cpl, on average, while the correlation coefficient between the two mode price series is 0.973. A possible explanation for this result goes back to Section 5.1, where it was argued that the mode and mean prices might be less accurate on days when restoration attempts are made; it typically takes 24 hours for all stations to raise their prices, and therefore the mode price increase might be observed with a lag. To examine this possibility, the statistics in the previous paragraph have been re-computed after excluding all 14 Days 0 identified in Section 5.1.[24] It is found that both the mean and mode price series become more accurate after the exclusion of these 14 days, particularly the mode series: the Internet data underestimate the daily mode and mean prices by 0.1 and 0.2 cpl, respectively, while the correlation coefficients between the two data sets are 0.994 for the modes and 0.999 for the means.

The findings of this subsection can be summarized visually using Figures 2a and 2b, where the daily city-wide mode and mean prices are plotted for each data set. Both the mode and mean prices computed from the Internet data roughly overlap the corresponding series computed using the balanced

---

[24] When one data set identifies a Day 0 one day later than the other data set, only the first day is counted for these computations.

panel. However, the Internet data tend to approximate daily mode prices more accurately than daily mean prices, with the primary exception of restoration days when the price increases can take a day to be reflected in the mode price. The overall results are consistent with conclusions made in Section 4, which are that a researcher can likely generate reliable results using mode and mean prices that are calculated from Internet data, but one should take into account the possibility that these prices might be highly underestimated on days when restorations are attempted.

5.3. Price Leadership

The purpose this section is to demonstrate that empirical analyses that rely on price data for specific stations likely should not be conducted using Internet data. One reason for this argument is that only a subset of the stations in a market are sampled each period, which can change every period. Also, there are some stations, particularly those in Group D, that are sampled as infrequently as once every few weeks; an extreme example is Station 8, which is spotted in the data nine times over the entire 103 days. Finally, stations' prices are unlikely to be observed in the order that they actually change.

To demonstrate the problems that can arise from relying on prices for particular stations, the leaders of price increases are defined in both data sets to be the first stations observed to raise their prices to the mode peak price on Day 0. The balanced panel identifies six specific major brand stations as leaders at least twice as often as the other 21 stations in the city: Petro-Canada Stations 5, 18 and 25 are identified as leaders of five to six restorations, while Esso Stations 17, 19, and 26 are identified as leaders four to five times each; the remaining 21 stations are identified as leaders no more than twice each. On the other hand, the Internet data identify four of these major brand stations as leaders no more often than two Group C stations: Pioneer Station 23 and Canadian Tire Station 24.

Similar conclusions are made regarding price decreases, in general. Using each data set, a station is identified as setting the minimum daily price in Guelph if it is observed to set this price before the other 26 stations in the city. At first glance, the Internet data seem to be quite reliable, as Pioneer Station 23 is consistently observed to set the daily minimum price much more frequently than all other stations. In the balanced panel, it sets the minimum 22 times, which is at least 2.7 times more often than all other stations; in the Internet data, Pioneer sets the minimum price 20 times, or at least four times more often than all other stations. However, the results are less accurate for the remaining 26 stations. Specifically, the balanced panel identifies 10 other stations are setting the market minimum at least once, while the Internet data identify 18 stations after Pioneer. Furthermore, seven of these new stations are identified at least as often as eight of the stations identified in the balanced panel.

Nonetheless, one can still likely use these data to determine whether prices tend to decrease in certain areas of a market, or whether certain brands might be more or less aggressive than other brands. For example, Eckert and West (2004a) compare the prices of each sampled station in Vancouver and Ottawa to the mode market price for each day; they find evidence that is consistent with the argument that certain brands in each city (such as Pioneer) are mavericks, because they tend to price below the daily market mode price. To examine whether these same conclusions might have been made if the authors used balanced panel data for each city, daily mode prices are calculated for Guelph using both data sets. Next, for each data set, the mode price for a given day is subtracted from all prices observed on that same day. The resulting differentials are then averaged across the entire sample for each brand, and displayed in the top section of Table 5.

According to this table, regardless of which data set is used, two key conclusions are made. First, Pioneer sets its price equal to the mode price with a relatively low frequency, and also tends to undercut the market mode by more than other brands; this is consistent with the argument that Pioneer is a maverick brand in Guelph. Second, the Group D stations also rarely price at the mode, but tend to price well above the mode price, suggesting that they do not price as aggressively as one might expect.

In the last column of Table 5, the amounts by which the Internet data overestimate the mean brand differentials are listed, which range between 0.1 and 1.0 cpl. However, as was argued in previous sections, these differentials are likely to be smaller if restoration days are excluded from the calculations, because the mode price can be highly underestimated on these days. Thus, the numbers provided in Table 5 have been re-computed after omitting the 14 Days 0 identified in Section 5.1. The results, which are displayed in the bottom section of Table 5 are similar to those in the top section, i.e., that Pioneer appears to be a maverick in the market, while the Group D stations tend to price above the mode price. Finally, the numbers in the last column demonstrate that the amounts by which the Internet data overestimate the mean brand differentials are all smaller than when all 103 days are included in the analysis. The absolute differential is also smaller for every brand, except Canadian Tire which is 0.2 cpl larger than when all 103 days are included in the calculations.

6. Conclusions

While the benefits of using high frequency, station-specific data to study retail gasoline price competition are not in dispute, data of such quality can usually only be obtained at a high cost, both in terms of time and money. Furthermore, while station-specific data are available from Internet gasoline pricing sites such as GasBuddy.com, no studies appear to exist which examine the conditions under which such data can be reliably used by a researcher. Thus, two data sets were constructed for stations in Guelph, Ontario to fill this gap in the literature: one consisting of a balanced panel of data collected bi-hourly from 27 stations, eight times per day for 103 days; and the other including prices that were collected every 12 hours from OntarioGasPrices.com during the same 103 days. These data were used to first identify certain sample selection biases that might arise in these Internet data sets, and the extent to which they can influence a researcher's results. These biases were then explored in more detail by examining some basic predictions of the Edgeworth cycle theory, using both data sets.

The main results of this paper are as follows. First, consistent with predictions in the relevant literature, it appears that consumers tend to be relatively motivated to report the prices of major brand stations that are price controlled by their head offices; they are also relatively unlikely to report prices for small independent stations. Potential explanations for these biases include spatial and product differentiation between stations. However, it seems that consumers do not tend to report a station's price more often if its price rises or falls relative to other stations' prices in the city. Thus, the Internet data tend to accurately identify features of cycles that can be distinguished using company-operated major brand station prices (e.g., the existence of a cycle, its period, height, approximate starting days, asymmetry, etc.), while features that require individual independent station data or very high frequency data might not be well-identified (e.g., the identities of price leaders and the order in which stations change their prices).

Furthermore, for markets such as Guelph where a clear city-wide mode price tends to exist, daily mode prices tend to be more accurately measured in the Internet data than average prices, in part because small non-branded independents that tend to price above the city-wide mode price are under-represented in these data for Guelph. However, mean prices still appear to be approximated with reasonable accuracy by the Internet data. Finally, the mode and mean price series are found to be less accurately measured by the Internet data on days when restorations are initiated, because it typically takes two days for these price increases to be fully reflected in the data.

These results have important implications for economic researchers, because depending on the market and the questions being asked, they suggest that these data can be reliably used to examine retail gasoline price competition under a number of different scenarios. In particular, when combined with station characteristic data, including capacities, traffic counts, and other non-gasoline operations, these data can be used to examine spatial price competition, both within and across markets. Such studies are not usually possible using data that are publicly available, because these data tend to either be averaged across stations, or for a limited number of stations in a market. Furthermore, these data can be collected for free from the Internet, and are therefore widely available to researchers regardless of available funding. In summary, the results in this paper suggest that the availability of this relatively new source of data can open up new avenues for research that were previously unavailable due to data restrictions.

References

Basu, Kaushik (2004), "Consumer Cognition and the Pricing in the 9's in Oligopolistic Markets," Harvard Institute of Economic Research Discussion Paper Number 2053.

Eckert, Andrew (2002), "Retail Price Cycles and the Response Asymmetry," *Canadian Journal of Economics* 35, 52-77.

----- (2003), "Retail Price Cycles and the Presence of Small Firms," *International Journal of Industrial Organization* 21, 151-70.

Eckert, Andrew and Douglas S. West (2004a), "A Tale of Two Cities: Price Uniformity and Price Volatility in Gasoline Retailing," *Annals of Regional Science* 38, 25-46.

----- (2004b), "Retail Gasoline Prices Cycles Across Spatially Dispersed Gasoline Stations," *Journal of Law & Economics* 47, 245-73.

----- (2005), "Price Uniformity and Competition in a Retail Gasoline Market," *Journal of Economic Behavior & Organization* 56, 219-37.

Kleit, Andrew N. (2003), "The Economics of Gasoline Retailing: Petroleum Distribution and Retailing Issues in the U.S.," Available at http://www.api.org (December 2003).

Lewis, Matthew S. and Howard P. Marvel (2007), "When Do Consumers Search?," Ohio State University Working Paper (March 2007).

Maskin, Eric and Jean Tirole (1988), "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica* 56, 571-99.

Noel, Michael (2006), "Edgeworth Cycles and Focal Prices: Computational Dynamic Markov Equilibria," UCSD Working Paper (April 10, 2006).

----- (2007a), "Edgeworth Price Cycles, Cost-Based Pricing and Sticky Pricing in Retail Gasoline Markets," *Review of Economics and Statistics* 89, 324-34.

----- (2007b), "Edgeworth Price Cycles: Evidence from the Toronto Retail Gasoline Market," *Journal of Industrial Economics* 55, 69-92.

Plummer, Paul, Robert P. Haining, and Eric Sheppard (1998), "Spatial Pricing in Interdependent Markets: Testing Assumptions and Modeling Price Variation. A Case Study of Gasoline Retailing in St. Cloud, Minnesota," *Environment and Planning A* 30, 67-84.

Schindler, Robert M. and Patrick N. Kirby (1997), "Patterns of Rightmost Digits Used in Advertised Prices: Implications for Nine-Ending Effects," *Journal of Consumer Research* 24, 192-201.

Shepard, Andrea (1993), "Contractual Form, Retail Price, and Asset Characteristics in Gasoline

Retailing," *RAND Journal of Economics* 24, 58-77.

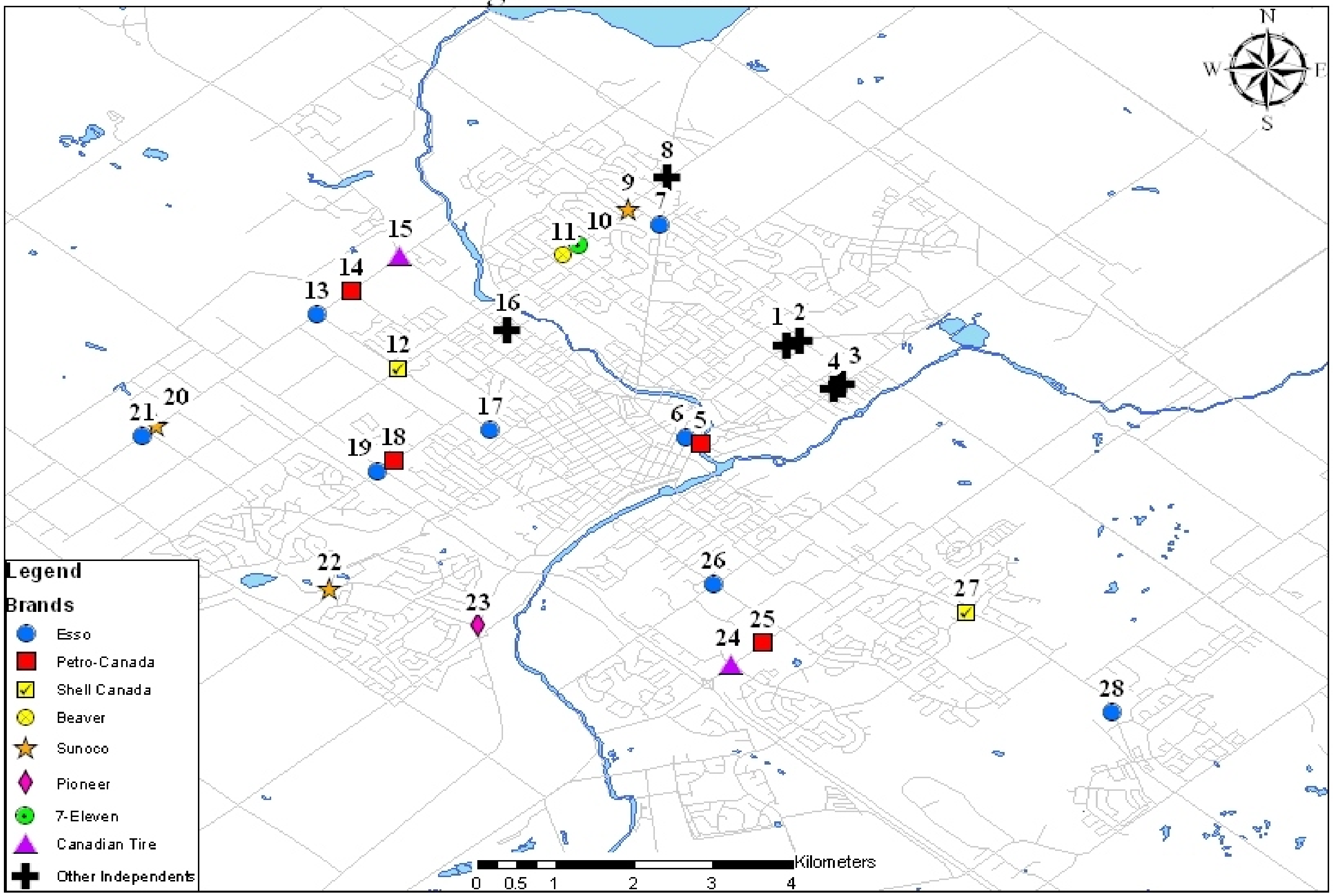
Sheppard, Eric, Robert P. Haining, and Paul Plummer (1992), "Spatial Pricing in Interdependent Markets," *Journal of Regional Science* 32, 55-75.

Slade, Margaret E. (1998), "Strategic Motives for Vertical Separation: Evidence from Retail Gasoline Markets," *Journal of Law, Economics & Organization* 14, 84-113.

Taylor, Beck A. (2000), "Retail Characteristics and Ownership Structure," *Small Business Economics* 14, 157-64.

Wang, Zhongmin (2006), "Endogenous Timing in Oligopoly Pricing: Evidence from a Repeated Game in the Field," Northeastern University Working Paper (October 2006).

Figure 1: Station Locations



Approximate Collection Times	
1	:00
2	:00
3	:01
4	:01
5	:03
6	:03
7	:08
8	:09
9	:11
10	:12
11	:12
12	:14
13	:15
14	:16
15	:17
16	:21
17	:24
18	:26
19	:26
20	:30
21	:30
22	:35
23	:36
24	:40
25	:40
26	:42
27	:45
28	N/A

Figure 2a

Daily Mode Retail Prices for Guelph, Ontario

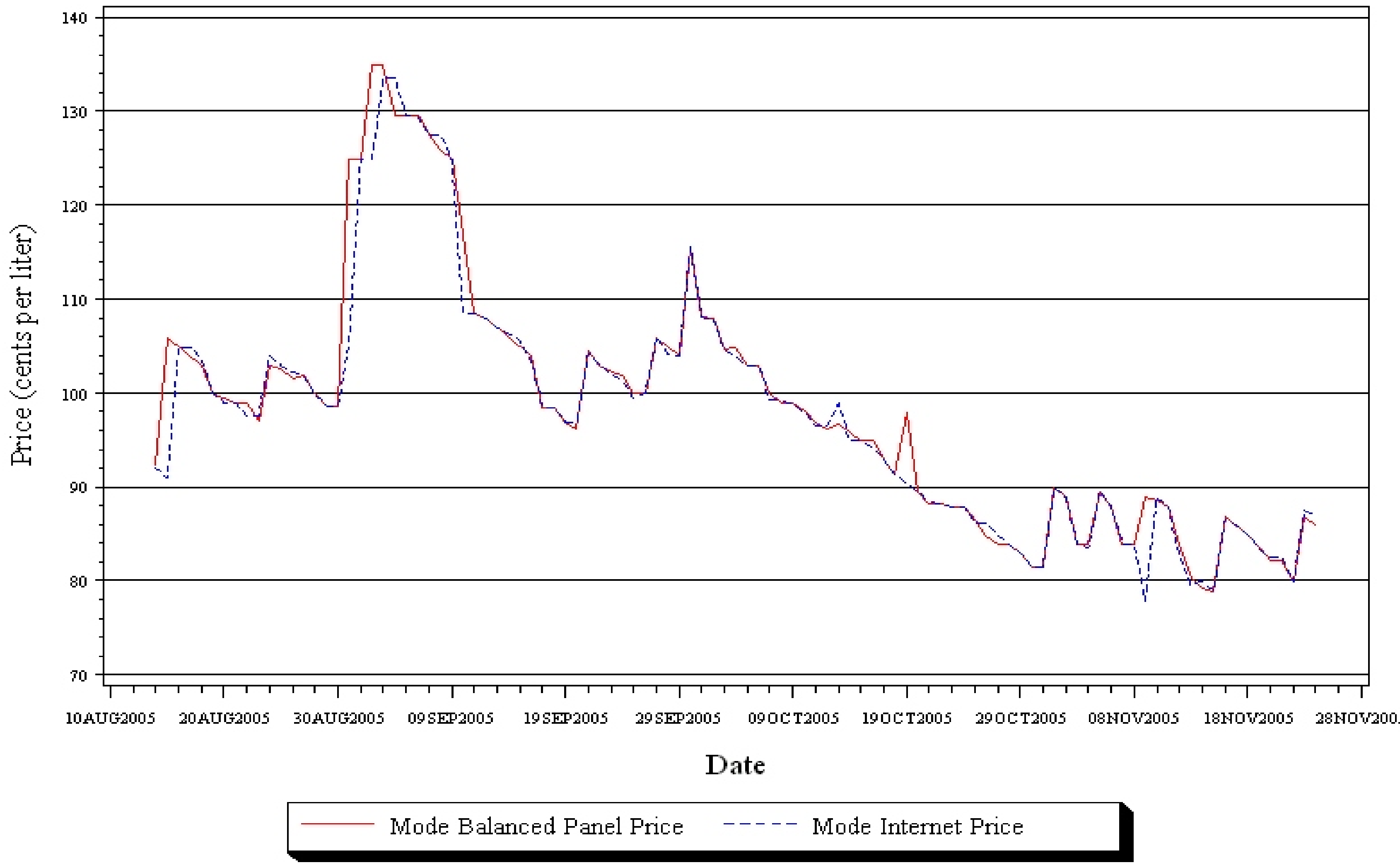


Figure 2b

Daily Mean Retail Prices for Guelph, Ontario

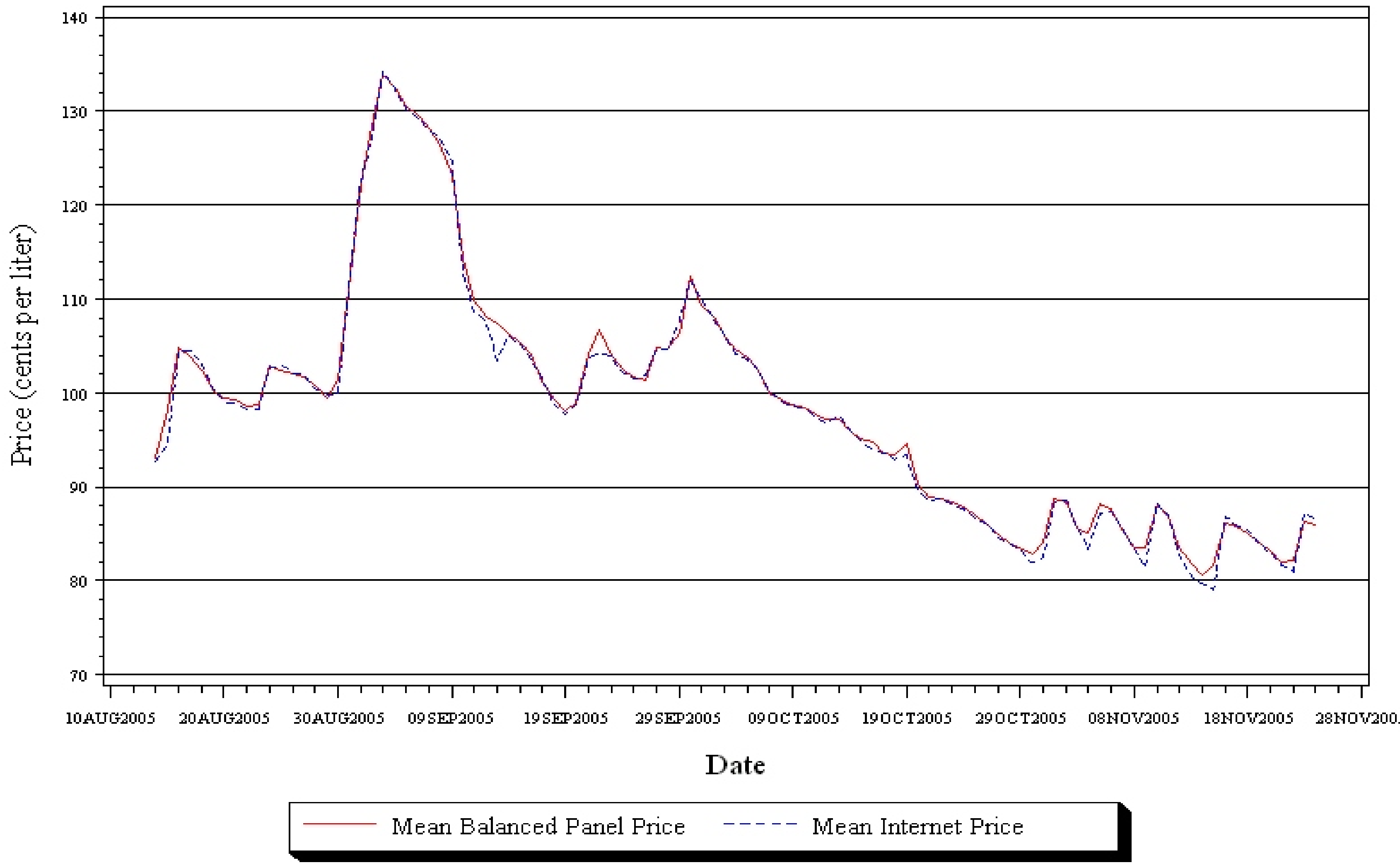


Figure 3: Proportion of Dates Each Station is Reported (N = 103)

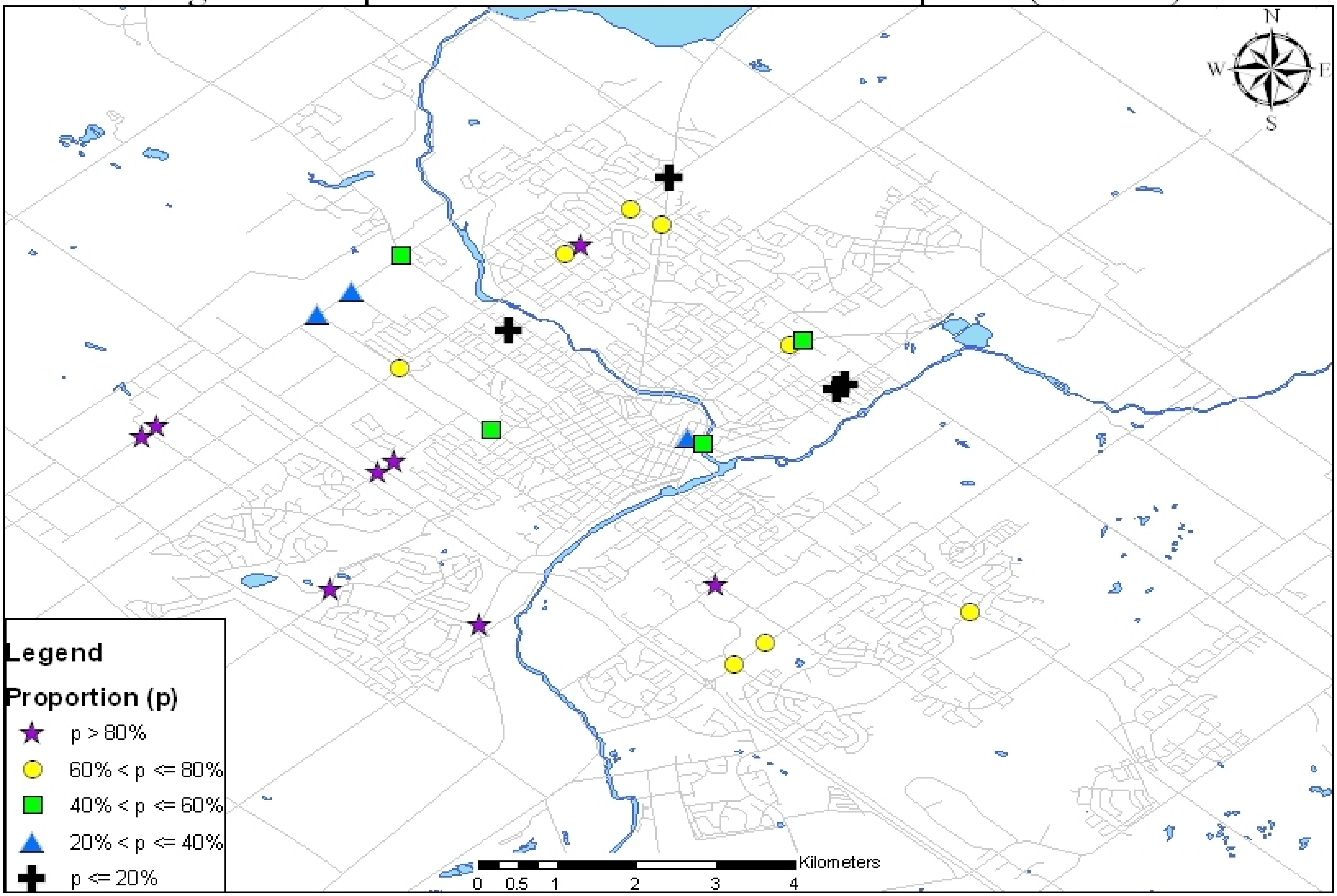


Table 1: Selected Station Characteristics*

Brand	ID	Nozzle Count	Open 24 Hrs	Self-Serve	Store	Car Wash	Auto Repair
Group A							
Esso	19	12	✓	✓	✓	✓	
Esso	21	10	✓	✓	✓	✓	
Esso	26	8	✓	✓	✓		
Petro-Canada	5	8		✓	✓		
Petro-Canada	18	8	✓	✓	✓	✓	
Petro-Canada	25	8	✓	✓	✓	✓	
Shell	12	8	✓	✓	✓		
Shell (Mac's)	27	8	✓	✓	✓		
Shell (Beaver)	11	6					✓
Sunoco	9	6				✓	
Sunoco	20	4					
Sunoco	22	8	✓	✓	✓	✓	
Group B							
Esso (Norm's Garage)	6	4					✓
Esso (7-Eleven)	7	8	✓	✓	✓		
Esso (Gas-Up Carwash)	13	4		✓		✓	✓
Esso (Rainbow)	17	6	✓	✓	✓	✓	
Petro-Canada	14	8		✓	✓		
Group C							
7-Eleven	10	8	✓	✓	✓		
Canadian Tire	15	8		✓	✓		✓
Canadian Tire	24	6	✓	✓		✓	
Pioneer	23	8	✓	✓	✓		
Group D							
Amco	1	4					✓
Cango	2	3					✓
Hilton Group Gas	3	4					✓
Maple Leaf Gas and Fuels	4	4					✓
CAN-OP	8	2					✓
Quik-N E-Zee Gas & Snacks	16	2					✓

* Petro-Canada Station 18 is the only station with both full- and self-serve pumps; since the self-serve price is observed, only those pumps are counted. Also, a "store" means a convenience store, except for Canadian Tire Station 15 where it is a Canadian Tire department store.

Table 2: Selected Statistics, Averaged Across Stations by Group

Group	Number (%) of Days "Spotted": Including All Spotters		Number (%) of Days "Spotted": Excluding Two Most Frequent Spotters		Mean Sample Margin (cpl)*	Annual Traffic Flows (Cars, '000s)
A	79	(76.7%)	50	(48.5%)	4.8	25.664
B	45	(43.7%)	32	(31.1%)	5.0	20.661
C	70	(68.0%)	47	(45.6%)	4.5	24.196
D	29	(28.2%)	10	(9.7%)	6.0	10.627

* The mean sample margin for each station is calculated using the balanced panel data set.

Table 3: Probit Regressions (Dependent Variable: $COUNT_{it}$; N = 2,781)

Variable†	Price Differentials: Coefficient		Price Differentials: χ^2 Stat	Price Rankings: Coefficient		Price Rankings: χ^2 Stat
$MODEDIFF_{it}$	0.00012		0.00	-----		-----
$MODERANK_{it}$	-----		-----	-0.00982	**	4.16
$DAY0_t$*$MODEDIFF_{it}$	0.00261		0.01	-----		-----
$DAY0_t$*$MODERANK_{it}$	-----		-----	0.00447		0.19
$NOPOST_{it}$	-0.54350	**	5.74	-0.53544	**	5.57
CONSTANT	1.90989	*	28.29	1.97055	*	30.03
LR Test Statistic (df = 131)	1,144.4624			1,148.7847		

† Results for the 26 station- and 102 daily-dummies are not reported for presentation purposes.
* Statistically significant at the 1% level of significance (two-tail)
** Statistically significant at the 5% level of significance (two-tail)

Table 4: Inter-Data Set Comparisons of Prices and Price Correlations*

Inter-Data Comparisons of Cycle Characteristics	Internet Data	Balanced Panel
Average increase in city-wide average price	3.6 cpl	3.8 cpl
Average increase in city-wide mode price	7.8 cpl	7.7 cpl
Average decrease in city-wide average price	1.5 cpl	1.3 cpl
Average decrease in city-wide mode price	1.9 cpl	1.9 cpl
Number of increases (average price)	29	25
Number of increases (mode price)	16	16
Number of decreases (average price)	73	77
Number of decreases (mode price)	67	68
Average number of days between attempted restorations	7.5	7.8
Daily Price Comparisons		
Number (%) of days when mode daily prices are equal		52 (50.5%)
Number (%) of days when mean daily prices are equal		8 (7.8%)
Average difference between Balanced and Internet mode prices (all days)		0.6 cpl
Average difference between Balanced and Internet mean prices (all days)		0.3 cpl
Average difference between Balanced and Internet mode prices (non-Days 0)		0.1 cpl
Average difference between Balanced and Internet mean prices (non-Days 0)		0.2 cpl
Pearson Price Correlations		
Between daily city-wide mode prices (all days)		0.973
Between daily city-wide mean prices (all days)		0.998
Between daily city-wide mode prices (non-Days 0)		0.994
Between daily city-wide mean prices (non-Days 0)		0.999
Between station-specific sample mean prices (all days)		0.640

* Note that the mean price rises more frequently and by smaller amounts than the mode price, because it typically takes two days for all stations in the city to raise their prices during restorations. Also, a Day 0 is a day when a restoration attempt is observed; if one data set identifies a restoration one day earlier than the other, then only the first day is counted for the purpose of the bottom two sections of this table.

Table 5: Comparisons of Station Prices to Daily Market Mode by Brand

Brand	Internet Data		Balanced Panel		Internet - Balanced Mean Diff
	Mean Diff	% Diff = 0	Mean Diff	% Diff = 0	
Including All 103 Days					
Esso	0.3	36.5%	-0.1	32.2%	0.4
Petro-Canada	0.7	39.1%	-0.3	34.5%	1.0
Shell	0.9	25.8%	0.3	28.9%	0.6
Sunoco	0.3	23.2%	-0.3	23.5%	0.6
7-Eleven	0.3	33.3%	-0.2	36.2%	0.5
Canadian Tire	-0.3	26.2%	-0.4	31.1%	0.1
Pioneer	-0.6	17.2%	-1.1	13.4%	0.5
Other	1.5	4.6%	1.1	5.5%	0.4
Excluding Days 0					
Esso	0.1	35.9%	0.1	32.9%	0.0
Petro-Canada	0.1	41.0%	-0.3	35.8%	0.4
Shell	0.5	26.2%	0.4	29.8%	0.1
Sunoco	0.2	22.3%	-0.1	23.8%	0.3
7-Eleven	0.2	32.3%	0.3	37.4%	-0.1
Canadian Tire	-0.6	27.0%	-0.3	30.5%	-0.3
Pioneer	-0.7	16.5%	-0.9	15.4%	0.2
Other	1.6	4.9%	1.4	5.6%	0.2