The Role of Information in Retail Gasoline Price Dispersion

Michael D. Noel* and Hongjie Qiang†

November 1, 2018

Abstract

Collecting information on prices is a costly endeavor. The cost depends on the relative ease with which those prices can be collected, and in many retail gasoline markets, there is a substantial divide in the ease of collecting information with regular grade gasoline on one side of the divide and midgrade and premium grade gasoline on the other. Regular grade prices are prominently displayed on large signboards in front of gasoline stations while the prices of higher octane grades, except where required by law, are rarely displayed. In this article, the effects of differential-by-grade price information on search and gasoline price dispersion are examined. A rank reversal test is used to test whether the observed grade-specific price dispersion is consistent with search or non-search related causes and, finding the former, a series of tests are presented to test for the effect of price information and other leading hypotheses. A significantly concave curvature in the price dispersion coefficients across the three grades supports a price information hypothesis. Detailed socioeconomic data on consumers, spatially matched to the stations they are most likely to patronize, shows that income is a secondary factor. Implications for policy are discussed.

JEL Classification Codes: L11, L15, L81, L91, Q31

Keywords: Information, Price Dispersion, Search, Search Costs, Income, Retail, Gasoline

1 Introduction

As Sorensen (2000) notes in his study of price dispersion in the pharmaceutical industry, the "law of one equilibrium price" rarely if ever holds in real-world markets. Markets exhibit various degrees of price dispersion in equilibrium, even in largely-homogeneous goods markets where classic textbook models of competition would have predicted a uniform price. The puzzle of why price dispersion is so pervasive across industries has spawned a large game-theoretic literature that seeks to explain its causes. A key finding of this literature is that price dispersion largely depends on how intensely

*Texas Tech University, Lubbock, Texas. michael.noel@ttu.edu.
†Texas Tech University, Lubbock, Texas. hongjie.qiang@ttu.edu.

consumers search for low prices (Diamond (1971), Varian (1980), Stahl (1989) and others). This in turn depends on the search costs consumers bear, and the potential benefits they expect to receive, from conducting a search (Stigler (1961)). One particularly important component of the equation is the ease with which price information can be gathered. When it is easier for consumers to collect price information, search costs will be lower, and we expect consumer search to be higher. Consumers will be more informed, and competitive forces will drive equilibrium price dispersion lower.[1] In this article, we examine the relationship between consumer search and price dispersion in an empirical setting. In particular, we examine how the relative ease of gathering price information in a certain industry affects the equilibrium level of price dispersion, through its intermediate effects on consumer search.

Our application focuses on the retail gasoline industry, for three reasons. First, the retail gasoline industry is an important industry to study in its own right, since gasoline expenditures make up a considerable expense for most households. According to the Energy Information Administration, Americans consume approximately 400 million gallons per day, at a cost of about one billion dollars per day, and the annual average expenditure per licensed driver is a little under two thousand dollars.[2] Consumers are very inelastic to gasoline in the short run (with short run elasticity estimates close to zero (Hughes et al. (2008)), and are vulnerable to large negative income effects each time crude prices rise unexpectedly (Gicheva et al. (2010)). Thus, understanding the nature of price information in an industry such as gasoline, and how consumers use that price information to search for gasoline prices, has important welfare considerations.

Second, our analysis is a first step in understanding the potential effects of so-called "price transparency regulations", a recent hot topic among industry regulators. These regulations are designed to make gasoline prices even more transparent than they already are, and range from requiring firms to post a complete set of grade-specific prices on large streetside signboards (e.g. California) to requiring firms to continually post real-time prices to a centralized government website for consumer reference (e.g. New South Wales, Australia). The stated purpose of these efforts is

[1] While the relationship between search and price dispersion is non-monotonic in general (Chandra and Tappata (2011)), it is negative in the context of competitive markets. In these markets, a shock to search costs or benefits that leads to a higher level of consumer search can be expected to reduce price dispersion.

[2] There were 214 million licensed drivers in the U.S. in 2014.

to enhance competition and improve consumer outcomes, yet surprisingly little research has been done to actually understand the effects of increased price information on consumer outcomes. This article is a first step in that direction.

Third, and importantly, the retail gasoline industry offers an excellent laboratory for studying the effects of price information, due to a unique historical artifact that has largely been ignored in the literature. The artifact is that, except where required by law, it is standard practice for gasoline stations to provide especially prominent and easily observable information on the price of regular grade gasoline, but not as much on the prices of midgrade and premium grade gasoline. In the study market we examine, the price of regular grade gasoline is virtually always displayed on large streetside signboards visible to consumers driving past the station, whereas the prices of midgrade and premium grade gasoline are almost never displayed on signboards and are observable at the station only once one drives up to the pump.[3] We exploit this differential-by-grade signboard price information to identify the relationship between price information and price dispersion.

In short, the signboards create a discrete and substantial divide in the ease of collecting price information across grades – with regular grade gasoline on one side of the divide and midgrade and premium grade gasoline on the other – which in turn predicts a similarly substantial divide in price dispersion by grade along that same boundary. This non-linear differential-by-grade shock can be used to test for the impact of signboard price information. We design a series of concavity tests that separate our price information hypothesis from the other usual suspects such as income differences or cost variation that could also cause differential degrees of price dispersion across grades. We show that price dispersion due to a price information story will lead to a concave pattern in our grade-specific price dispersion coefficients, whereas price dispersion due to income or cost based factors will lead to a linear or convex pattern instead. We test for concavity and find that price transparency indeed matters – more transparent gasoline prices (as displayed on large streetside signboards) leads to lower levels of gasoline price dispersion, all else equal. Our results are consistent with the theoretical predictions from the literature.

There are several novel features of our study. First and foremost, to our knowledge, ours is the

[3] Midgrade and premium grade gasoline prices are usually displayed at the station in one to two-inch digital display windows. One may also research higher octane gasoline prices on a gasoline price reporting website (at an additional search cost), but even here premium grade and midgrade gasoline prices are more sparsely reported.

first to consider the informational content contained on individual streetside signboards, and the first to examine the effects of differential signboard price information on price dispersion. The lack of research may seem surprising, given the potential importance of signboard price information on gasoline competition, but we expect the reason is a logistical one. Signboard information is not available in standard datasets on the retail gasoline industry and gathering this information necessitates an extensive primary data collection exercise. We do that here. Second, our study is among the few, whether related to consumer search or other topics, that examine midgrade and premium grade gasoline prices in their own right. The vast majority of studies relating to gasoline prices focus only on regular grade gasoline and, in framing their conclusions about "gasoline" generally, implicitly assume that their results also apply to higher octane grades. Our study shows that this generalization is premature.[4] Third, our study is among the first to provide evidence on the potential effects of gasoline price transparency regulations, a current hot topic among gasoline industry regulators in many countries.

Several other studies examine search and price dispersion in the context of the retail gasoline industry, but the only other study that examines price dispersion across the different grades of gasoline, to our knowledge, is Chandra and Tappata (2011). That study examines a handful of U.S. States and finds, as we do here, that price dispersion tends to be higher for higher octane grades than for regular grade gasoline. The authors postulate that their finding could result from differences in the incomes of consumers that typically purchase the different grades of gasoline. Such a postulate is surely reasonable. However, they do not explicitly test for this or other potential underlying causes of the differential-by-grade price dispersion they observe and, in fact, rule out any role for price information, stating that "the phenomenon of gas stations prominently displaying their prices [...] controls for imperfect information." As it turns out, however, it is actually very rare for stations to display midgrade and premium grade gasoline prices where not required, and, in all but one of the states they examine, there is no state requirement to do so and very few stations do.[5] Signboard information thus remains a potentially important factor for any differential-by-grade

[4] While the usual justification is that regular grade gasoline accounts for the majority of sales, still almost one in five purchase a higher octane grade. Approximately 82% of all gasoline sold is regular grade, 7% is midgrade, and 11% is premium, according to the Energy Information Administration.

[5] The states that they examine without requirements are Texas, New Jersey, and Florida. In their dataset, only California (and some individual counties, e.g. Miami-Dade in Florida) have a requirement to display all three prices

price dispersion that is observed, and one that has been largely unexplored in the literature. This is where our study comes in. We contribute to the literature by exploring the potential causes of differential-by-grade price dispersion and ultimately highlight the importance of price information in generating differential levels of price dispersion across grades.

Our analysis proceeds in two stages. First, we provide evidence, independent of signboard information, that price dispersion results from a search cost story in general rather than a pure product or station differentiation story. Chandra and Tappata (2011) introduce a novel rank reversal test that can distinguish between these competing theories, based on the observation that search models tend to predict frequent rank reversals, i.e. frequent changes in which stations have higher or lower prices over time. A pure product or station differentiation story has no role for rank reversals. We search for evidence of rank reversals and find that they are very frequent within each grade, consistent with a search-cost hypothesis.

With a search-based explanation in hand, our second stage is to explore the potential sources of search costs, including the role of price information, that could lead to the differential-by-grade price dispersion we observe. Our identification comes from the fact that the discrete and substantial divide in price information posted on signboards predicts a similarly discrete and substantial divide in price dispersion along that exact same boundary. This implies a strong concavity in our grade-specific price dispersion coefficients that separates a price information story from other leading hypotheses. Other hypotheses such as wholesale price dispersion or income differences lead to linear or convex patterns instead. In the end, we find significantly concave patterns in our coefficients and conclude that signboard price information is a primary source of the differential-by-grade price dispersion we observe.

Given the special attention in the literature to income as a potential source of search costs, we conduct an additional analysis in which we collect detailed information on consumer income and other socioeconomic factors at the census block level and directly control for area income effects. The unique spatial distribution of income in our study market and its unique traffic patterns improves the match between consumers and the stations they are most likely to patronize. This in turn improves the identification of our results. After controlling for income and other factors

on streetside signboards.

explicitly, our main result – that signboard price information is an important driver of differential-by-grade price dispersion – carries through as strongly. We conclude with a discussion of the potential policy implications of our results.

2 Literature and Background

There is a large theoretical literature on equilibrium consumer search and price dispersion. In an early article, Diamond (1971) finds that if all consumers have positive search costs, prices necessarily rise to monopoly levels, price dispersion falls to zero and, absent any benefit to search in a world of identical prices, there is no search in equilibrium. Varian (1980) and Stahl (1989) develop models where some consumers have zero or negative search costs and find in these cases that price dispersion and search do occur in equilibrium. A general result from this literature is that price dispersion and search are jointly determined - an exogenous shock to one impacts the other.

Chandra and Tappata (2011) show that the relationship is a non-monotonic one, depending on the degree of competition. When no consumers search, these models reproduce the Diamond result where all prices are equal to each other and equal to monopoly prices. When all consumers search, the models reproduce the perfectly competitive result where all prices are equal to each other and equal to marginal cost. There is no price dispersion in either case. For intermediate cases, there is positive consumer search and positive price dispersion in equilibrium.[6] For highly monopolistic markets with prices relatively close to monopoly levels, the predicted effect of search on price dispersion is positive, and for relatively competitive markets the predicted effect is negative.[7] Given the generally competitive nature of gasoline retailing, we would expect to see a negative relationship between search and price dispersion in our setting.

Numerous empirical studies have tested the consumer search vis-a-vis price dispersion relationship in various industries and generally find it to be negative, consistent with competitive markets. Search is difficult to measure directly, so these studies generally use a proxy for consumer search, often a variable correlated with search costs or search benefits. Sorensen (2000), for example, finds

[6] See the appendix for a more complete treatment.

[7] See also Pennerstorfer et al. (2016) and Janssen et al. (2011).

that price dispersion in medications used to treat chronic illnesses is lower than in medications used to treat acute illnesses, noting that in the former case, repeat purchases should make the benefits of search, and therefore search, greater. Brynjofsson and Smith (2000) find that price dispersion is lower for goods sold on-line, where search costs are likely to be lower than in brick and mortar stores. Examples of other studies examining the effect of consumer search on price dispersion include Dahlby and West (1986) on auto insurance, Walsh and Whelan (1999), Zhao (2006), Dubois and Perrone (2015) and Sherman and Weiss (2017) on groceries, Baye et al. (2003) and Tang et al. (2010) for on-line purchases via shopbots, Brown and Goolsbee (2002) for online purchases of life insurance, Hortacsu and Syverson (2004) on new stock market investors, Milyo and Waldfogel (1999) on liquor, Orlov (2011) on airlines, and others.

In the context of retail gasoline, Barron et al. (2004) find that price dispersion is lower with more densely located stations and Lewis (2008) finds that price dispersion is lower when there are more competing stations of a similar brand type. Lewis and Marvel (2011) find that price dispersion is higher (lower) when prices are falling (rising) and consumer search is low (high). Pennerstorfer et al. (2016) find that price dispersion depends on the percentage of commuter traffic present in an area, noting that commuters tend to be better informed about prices. Hosken et al. (2008) opine that the degree of price dispersion across stations in Washington, D.C. is relatively high considering the homogeneity of the product. Byrne et al. (2015), Byrne and de Roos (2015) and Noel (forthcoming) examine the reverse relationship of how shocks to price dispersion affect consumer search in the context of markets with gasoline price cycles (Noel (2007)). None of these studies, however, consider different grades of gasoline or the effects of differential-by-grade signboard price information.

A different but related branch of the literature examines how consumer search following general cost shocks can lead to the asymmetric passthrough of costs into prices, the so-called "rockets and feathers" effect. Rockets and feathers refers to a situation where prices tend to rise quickly after a cost increase but fall more slowly after a cost decrease, and is often seen in retail gasoline and other markets (Borenstein et al. (1997), Peltzman (2000)).[8] While our interest and focus

[8] There are very many such studies, including (for the retail gasoline industry): Bacon (1991), Godby et al. (2000), Bachmeier and Griffin (2003), Galeotti et al. (2003), Radchenko (2005), Deltas (2008), Verlinda (2008), Noel (2009), Kristoufek and Lunachova (2015), Li and Stock (2018), and Eleftheriou et al. (2018). See Eckert (2013) and Noel (2016) for surveys.

is fundamentally different than that of an asymmetric passthrough study (as described in the footnote),[9] these studies highlight the importance of consumer search in generating equilibrium price distributions. Lewis (2011) offers a theoretical explanation for rockets and feathers using a reference price model in which consumers have expectations of the price distribution based on past price observations. He finds that consumers ultimately search more (less) when they observe relatively high (low) prices compared to their expectation, generating asymmetric passthrough in which search tends to be high (low) and price dispersion low (high) when costs are rising (falling). Yang and Ye (2008), Tappata (2009) and Cabral and Fishman (2012) all develop more formal dynamic models of consumer search that produce asymmetric passthrough in equilibrium. While assumptions of the models vary, the general framework is one of competing firms with fluctuating marginal costs that follow a Markov process. In these models, changes in consumer search due to aggregate cost shocks naturally lead to asymmetric passthrough in prices, with short term changes in price dispersion as an intermediate step in the process.

Most closely related to this article is Chandra and Tappata (2011) who examine price dispersion across several states and find that price dispersion is lower between stations at the same intersection than at stations further out. They also consider multiple grades of gasoline and find that price dispersion is higher for higher octane grades. The authors do not test for the source of the differential-by-grade price dispersion or the impact of signboard information, however, and we contribute to the literature by doing that explicitly here.

[9] The shocks to search in these studies stem from shocks to wholesale costs over time, which is of little issue in our short 33 day sample with relatively flat wholesale prices. Even still, our approach and identification is very different from a rockets and feathers study. We are not interested in rockets and feathers in its own right, or in aggregate shocks to costs over time, but rather in the differential degrees of search and price dispersion across the three different grades of gasoline at a given point in time. Since wholesale gasoline prices across the three grades are, by historical convention, perfectly serially correlated, general shock to costs over time is simply differenced out of our differential-by-grade comparative analysis. In other words, the first differences in the cost differentials are almost always uniformly zero in retail gasoline markets. Nonetheless, we include daily fixed effects to account for any general changes over time, including cost shocks, and find no differences in our results whether or not these are included. We also design a concavity test to examine the potential effect of cost variation on the differential-by-grade price dispersion we observe and can easily rule this out as a cause as well.

3 Data and Methodology

We collect station-specific daily price information for the three primary grades of gasoline – regular, midgrade, and premium – for retail gasoline stations in and around the city of Lubbock, Texas, from November 16, 2016 to December 18, 2016. The price information comes from GasBuddy.com, a popular crowd-sourced gasoline price information website, in which drivers report prices to the GasBuddy website to help inform other drivers of current prices in their areas. We have information on virtually all of the stations in the Lubbock area, 104 stations in all, and a total of 7133 price observations over the 33 day sample period. Through a combination of satellite imagery and on-site data collection, we are able to determine that 101 of the 104 stations display the price of regular gasoline but not the price of midgrade or premium grade gasoline on their signboards.

The GasBuddy source of gasoline price data used in this study has also been used in a number of previous studies for other areas (e.g. Lewis and Marvel (2011), Atkinson et al. (2014), Noel (forthcoming)). An advantage of the GasBuddy data is that it is of a higher frequency and has broader station coverage than most other gasoline price surveys. The concern is that, as a crowd-sourced database, there is potential for error, especially if drivers disproportionately report unusually low or "surprising" prices. Atkinson (2008) and Atkinson et al. (2014) test this directly (in the case of markets exhibiting and not exhibiting price cycles, respectively) and, fortunately, find it not to be the case. Comparing GasBuddy price data for a given area to similar data from traditionally taken surveys, they find that the GasBuddy price data does a good job representing the complete distribution of gasoline prices in an area, and prices are not selectively chosen from the lower tail of the distribution or systematically biased in other ways.

To calculate price dispersion within groups of stations, we must first group stations together with its nearby competitors. We obtain geolocation information for each station and assign them to one of nine subregions of the city (Central East, Central West, West, Southwest, South, Southeast, Northeast, Northwest and North).[10] The subregions are chosen so that each has a major highway running roughly through its center, and each is about two miles wide.[11] We then calculate various

[10] There is no "East region" as directly east of the city is parks and farmland with no gasoline stations.

[11] The major highways are the north-south I-27, the east-west U.S. 62/82 and the loop highway around the city, Texas Loop 289.

measures of price dispersion over the sets of stations within each subregion, for each grade of gasoline and at a daily frequency. We experimented with alternative definitions of subregions and our results are unaffected by this choice.

Because of the special attention given to income in understanding search costs and price dispersion, following our main concavity tests, we conduct an additional analysis in which we directly incorporate measures of income, along with other pertinent demographic information, into our model. We use median household income data from the U.S. Census Bureau at the block-group level and aggregate it up to the subregional level. In addition to income, we include demographic variables of gender and race.[12] Consistent with the literature, we implicitly assume that the probability that a consumer visits a given station in a given subregion is higher when that consumer lives and regularly drives in that subregion. The unique spatial distribution of income in our sample city – with incomes rising substantially as one moves from east to west – and the more compartmentalized nature of the city in the absence of a high-density downtown core and the absence of long commutes – improves the match. We report summary statistics for our variables in Table 1.

There are several advantages to conducting a single market study for a city like the one we examine here. One advantage is that Lubbock, a midsized city of a quarter of a million people and over a hundred gasoline stations, is relatively isolated from other cities. The next nearest midsized city is a two hour drive away and the nearest large city (Dallas) is a five hour drive away. The relative isolation means that its residents reasonably make up the entire market for gasoline retailers in the city, and it is substantially easier to spatially match the demographic and socioeconomic information of consumers to the gasoline stations they are most likely to patronize. This is contrast to larger, interconnected metropolitan areas with many adjoining cities and suburbs, long commutes and a large degree of adjoining-city through-traffic, where matching stations to potential consumers is especially difficult.[13] A second advantage of our sample market is that it has a useful spatial distribution of income – with substantially lower incomes in the eastern portions of the city and higher incomes in the west – that helps identify effects. A third advantage is that its traffic patterns

[12] Other socioeconomic variables such as average housing prices or rents, unemployment rates, and percentage of residents below the poverty line tend to be largely collinear with income and we exclude these.

[13] See Cooper and Jones (2007), Houde (2012), and Pennerstorfer et al. (2016) for early efforts and a discussion of the difficulties in incorporating large-city commuting patterns into the analysis.

are more localized – absent a high density downtown, commutes in the sample city are short, and this improves the correlation between where consumers purchase gasoline and where they live. A final practical advantage of our single market approach is that its census of just over one hundred gasoline stations feasibly enables us to visually observe every station in the marketplace, including the content of the streetside signboards on display. Signboard price information is not found in publicly available datasets on the gasoline industry and is a likely reason why, in spite of its importance, it has received little attention in the literature.

The main disadvantage of a single market study is the usual one – that its results are specific to the city and may or may not fully carry over to other areas. We accept this trade-off as necessary for gathering the necessary signboard information required to address the question of interest. Fortunately, we do not expect representation issues to be a large concern. Gasoline products offered from one station to the next are generally homogeneous, and gasoline stations from one city to the next are also largely homogenous. The same major wholesalers and retailers reappear in many other markets, many nationally. Market shares, station formats, ingress and egress patterns, and quality standards in the sample market are all typical and gasoline prices are typical for the Gulf Region. Other than its beneficial spatial geography, the sample city does not appear to be unusual and its consumers are not different in any apparent way. In future research, we plan to expand our data collection effort and analysis to additional geographies with varying signboard conventions and test whether the intuitive results we find here do or do not carry over as strongly to other areas.

Our analysis proceeds in two stages. We first examine whether the observed price dispersion is attributable to a search cost story in general, by testing for the presence of rank reversals. The rank reversal measure, introduced by Chandra and Tappata (2011), measures how often the relative ranking of prices at two nearby stations change. Search models predict rank reversals as a consequence of stations' practice of mixing among a distribution of prices, whereas there is no role for them under a pure station differentiation story (in which relative prices would be different but stable). At this stage, we can remain agnostic about the source of any search costs.

To test for the presence of a non-negligible amount of price rank reversals among nearby stations, we calculate rank reversals between each pair of stations i and j within the same subregion s, as:

$$RANKREVERSAL_{ijs} = \sum_{t=1}^{T} \frac{I(p_{ist} > p_{jst} \mid [\sum_t I(p_{ist} \geq p_{jst}) < \sum_t I(p_{ist} \leq p_{jst})])}{T} \quad (1)$$

where $I(\cdot)$ is an indicator function equal to one when the internal condition holds. Basically, it measures the fraction of days that the price of station i exceeds that of station j, conditional on station j being the "usually" higher priced station. By "usually", we mean that station j has a higher price than station i more often than it does not. The measure ranges from a minimum of zero (with no rank reversals) to a maximum of one-half (where no station is "usually" the higher priced one.)

Once a search based explanation is established, we move to the second stage of analysis to establish the differential degrees of price dispersion across the three grades of gasoline and explore their causes. The basic estimating equation is:

$$\begin{aligned} PRICEDISPERSION_{gst} &= \beta_0 + \beta_1 MIDGRADE_g + \beta_2 PREMIUM_g \\ &+ \sum_{t=2}^{\overline{T}} \phi_t T_t + \sum_{r=2}^{\overline{R}} \theta_r R_r + X_{srt} B + \eta_{gst} \end{aligned} \quad (2)$$

where $PRICEDISPERSION_{gst}$ is calculated as either the standard deviation of prices or the difference between the maximum and the minimum price (the "max-min range"), for gasoline grade g within a given subregion s of the city at time t.[14] The variable $MIDGRADE_g$ is an indicator variable equal to one when the grade is midgrade and $PREMIUM_g$ is an indicator variable equal to one when the grade is premium. The T_t are a set of daily fixed effects and the R_r are a set of regional fixed effects for each of the three major regions of the city (West, Central, and East). The value of $\overline{T}$ is 33 (the number of days in the sample) and the value of $\overline{R}$ is 3. The explanatory variable matrix X contains additional controls, including a market structure variable (discussed below) and, in later specifications, socioeconomic variables including consumer income. We also perform alternate regressions in which we interact our spatial measure of income with the

[14] Later we consider other price dispersion metrics such as interquartile ranges and Gini coefficients.

$MIDGRADE$ and $PREMIUM$ coefficients to test for the effect of income on differential-by-grade price dispersion directly. We assume the η_{gst} are normally distributed error terms. Because prices at a given station are unlikely to be independent over time, and measures of price dispersion are thus unlikely to be independent within a given subregion over time, we calculate robust standard errors clustered by subregion to account for the potential non-independence.

Our identification is based on the realization that differential-by-grade price dispersion from signboard price information will create a strong and concave pattern in the grade specific coefficients. (Including the "zero" coefficient on the omitted regular grade gasoline indicator variable, the grade-specific coefficients of interest are 0, β_1, and β_2.) We will show that price dispersion from other competing hypotheses like wholesale price variation or income differences would create a linear or convex pattern instead. We thus take a significant concave relationship in the grade-specific coefficients as evidence of the impact of signboard price information on price dispersion. More specifically, we simultaneously test whether the $MIDGRADE$ and $PREMIUM$ coefficients are each statistically significantly greater than zero, and whether the premium grade coefficient is less than double the midgrade coefficient (i.e. concavity). We refer to this as a concavity test or, alternately (in reference to the null), a linearity test:

$$H_0 : \beta_2 = 2\beta_1 \tag{3}$$

To understand the concavity/linearity test, begin by noting that an even stronger test for the impact of differential signboard price information would be to test for the significance and *equality* of the $MIDGRADE$ and $PREMIUM$ coefficients, i.e. an "equality test". The logic behind the stronger equality test is that, if the data were complete and all else were held equal, the $MIDGRADE$ and $PREMIUM$ coefficients should be significant and equal because they receive the same treatment effect of no signboard price information. Signboard price information for both higher octane grades is equally absent and its impact on price dispersion should be equal as well, leading to equal $MIDGRADE$ and $PREMIUM$ coefficients. This yields an extreme form of concavity that takes the shape of a right-slanted capital Γ in a plot of price dispersion against the three individual and ordered grades of gasoline – with the bottom left corner of the Γ representing the omitted regular grade "coefficient" of zero, and the upper left and upper right corners of the Γ corresponding to the higher but equal $MIDGRADE$ and $PREMIUM$ grade coefficients. The equality test is essentially a test of an extreme form of Leontief-style concavity, while the linearity test is a test of general concavity.

While we could focus on the equality test, there is an important reason for using the more flexible concavity test instead. As discussed below, other leading competing hypotheses tend to work towards either a linear or convex relationship instead of a concave one. We consider a cost dispersion explanation which mathematically requires a linear or convex pattern in the coefficients because of the nature under which midgrade gasoline is dispensed at the pump. We also consider income-based explanations which lead to a convex relationship in the coefficients because of the nature of buyers of different grades. When we reject linear patterns and convex patterns in our coefficients, we can rule out these competing hypotheses as leading causes. But it does not mean that these hypotheses must have no effect. An effect can still be present to a lesser degree and, by working toward a linear or convex pattern, potentially offset the extreme-concave, or perfect, Γ shape. For example, to the extent that there is *some* degree of cost dispersion present, it can lead to some linearization or convexation of the coefficients that would bend the coefficients away from the perfect Γ shape. Similarly, and putting aside the fact that we later use direct information on income (which serve to confirm our results), the presence of income effects can lead to *some* convexation of the coefficients that would again bend the coefficients away from the perfect Γ shape. In each case, the equality test could fail but we should still see concavity on net if signboard price information were the dominant effect.

The stronger equality test is thus consistent with a test of whether signboard price information strongly dominates to the exclusion of everything else, while the weaker concavity test is consistent with a test of whether signboard information is simply dominant on net. For these reasons, we think that the concavity test is the more relevant benchmark. Having said that, it is worth noting at this point that our results actually do satisfy the stronger equality test as well as the more flexible concavity test, and we interpret this as evidence of a high importance of differential signboard price information on differential-by-grade price dispersion.[15]

[15] A common alternate method of testing for a concave relationship is to use panel threshold techniques (e.g. Hansen

We consider two leading alternate hypotheses, both of which produce a linear or convex pattern in our grade-specific price dispersion coefficients, instead of a concave one. The first is cost dispersion, or alternately, wholesale price dispersion. In spite of the fact that we cannot measure wholesale price dispersion directly (as this depends on the terms of confidential retailer-wholesaler contracts), we can easily test this hypothesis by relying on a simple accounting identity. The identity is that midgrade gasoline (which generally carries an octane rating of 89) is an exact 50-50 mix of regular grade gasoline (which generally carries an octane rating of 87) and premium grade gasoline (which generally carries an octane rating of 91). Midgrade gasoline itself is virtually never sold at the wholesale level separately, but rather is mixed "on the fly" at gasoline stations by pumping equal amounts of regular grade and premium grade gasoline from underground storage tanks into the dispenser hose and the consumer's car.[16] In other words, the marginal cost of midgrade is by construction the average of regular grade and premium grade marginal costs.

The identity implies that if wholesale price dispersion were the cause of retail price dispersion, the pattern of grade-specific coefficients must be linear or convex. The coefficient on midgrade gasoline (relative to regular grade gasoline) must be no more than half that of premium grade gasoline (relative to regular), and this remains true whether price dispersion is measured as standard deviations or as max-min ranges. Said another way, the standard deviation of midgrade gasoline must be less than or equal to the average of the regular grade and premium grade gasoline standard deviations. It is "equal to" (creating a linear pattern in the coefficients) in the limiting case when the regular and premium grade prices are perfectly correlated, and "less than" (generating a convex pattern in the coefficients) in all other circumstances. Specifically, given that $c_{Mj} = (c_{Rj} + c_{Pj})/2$ for station

(1999)). For example, the Hansen model posits that $y_{it} = \alpha_{it} + \beta_1 x_{it}$ if $z_{it} \leq \gamma$, and $y_{it} = \alpha_{it} + \beta_2 x_{it}$ if $z_{it} > \gamma$, and tests whether $\beta_1 = \beta_2$. The model as written is most applicable when x is a continuous variable and the threshold parameter γ (that determines the "kink" point) is unknown to the researcher. In our setting, this approach is less applicable for two reasons. First, we know the value of γ. Signboard information is present for regular grade prices and absent for midgrade and premium grade prices, so any break in price dispersion due to differential signboard information must be along grade-specific lines. One can thus simply test the freely-estimated grade-specific coefficients for concavity directly, which is what we do. Second, when the right hand side variable x and the threshold variable z are the same, the model is only estimable as written when the x variable is not dichotomous. (If dichotomous, x would uniformly be equal to zero in the first equation and uniformly equal to one in the second (or vice versa) and perfectly correlated with the constant.) In our setting, x and z are indeed one and the same (each represents the same set of grade-specific indicator variables) and x is indeed dichotomous. With γ known, the model boils down to nothing more than a comparison of price dispersion y for each grade of gasoline x, which is of course what we do. Concavity can still be tested by a comparison of the freely-estimated grade-specific coefficients, which is again what we do.

[16] This eliminates the need for an expensive and redundant third tank.

j, where c is marginal cost and R, M, and P denote regular, midgrade, and premium respectively, it is straightforward to show that $stddev(c_R) + stddev(c_P) - 2stddev(c_M) \geq 0$, i.e. convexity (or linearity when the equality holds), if and only if $corr(c_R, c_P) = cov(c_R, c_P)/(stddev(c_R)stddev(c_P)) \leq 1$. The latter is necessarily true because $corr(c_R, c_P) \in [-1, 1]$.

Similarly, the max-min range of midgrade gasoline must be less than or equal to the average of the max-min ranges of regular grade gasoline and premium grade gasoline. It is "equal to" (creating a linear pattern in the coefficients) when the maximum of both regular grade and premium grade prices always occurs at the same station and the minimum of both regular grade and premium grade prices also always occurs at the same station (the max-min equivalent of perfect correlation). It is "less than" (creating a convex pattern in the coefficients) in all other circumstances. Specifically, when the maximums (minimums) occur at different stations, the maximum (minimum) midgrade cost is necessarily lower (higher) than the average of the maximum (minimum) regular grade and premium grade costs, so that the max-min range for midgrade is necessarily smaller than the average of regular grade and premium grade max-min ranges, yielding convexity. That is, given that $c_{Mj} = (c_{Rj} + c_{Pj})/2$, it follows that $\max(c_M) \leq (\max(c_R) + \max(c_P))/2$ and $\min(c_M) \geq (\min(c_R) + \min(c_P))/2$, and therefore $\max(c_M) - \min(c_M) \leq (\max(c_R) - \min(c_R))/2 + (\max(c_R) - \min(c_P))/2$, i.e. convexity (or linearity if equal). We can test for the impact of wholesale price dispersion using the aforementioned linearity test, comparing the premium grade coefficient with double that of the midgrade coefficient. A finding of concavity would rule out wholesale price dispersion as a leading cause.[17]

Note that this discussion relates to unobserved wholesale price variation arising from the different terms of individual contracts. Changes in overall daily wholesale prices over time - i.e. the type of cost shocks that are the focus of asymmetric passthrough studies - are less relevant to our analysis. The first reason is that wholesale prices across the three different grades are perfectly correlated over time - premium and midgrade gasoline wholesale prices are, by convention, set as fixed price differentials above regular grade gasoline.[18] This means that any shock to wholesale

[17] We gently point out that, even if we find wholesale price dispersion to be a cause of retail price dispersion, it would be no less interesting. It would simply suggest that wholesalers were strategically setting wholesale prices cognizant of end consumers' search costs, either instead of or in addition to retailers.

[18] This is an often overlooked fact. The differentials vary by wholesaler, but are generally about sixteen cents per gallon for premium grade gasoline and eight cents per gallon for midgrade gasoline, up and above the price of regular

costs over time is actually a simultaneous shock to the cost of all three grades of gasoline, and its effect is differenced out of our differential-by-grade comparative analysis (i.e. the first differences in the cost differentials are always uniformly equal to zero). The second reason is that we have a short 33 day sample at a time of relatively flat wholesale prices and we would not expect cost shocks to play an important role anyway. In these ways, our study is orthogonal in both methodology and focus to asymmetric passthrough studies. Asymmetric passthrough studies examine general cost shocks over time, almost always for a single grade of gasoline, and how it affects the speed of price movements over time. In contrast, we examine *differential-by-grade* search costs and how it may affect *differential-by-grade* price dispersion, at a given point in time. (We also include daily fixed effects to account for the transmission of cost or other shocks over time, though their inclusion has no meaningful effect on results.)

We next turn to a discussion of the effect of income on differential-by-grade price dispersion. While we explicitly control for income later in our analysis, even without direct data, we can still test this hypothesis since an income story again leads to a convex pattern in the grade-specific coefficients. To see this, we again appeal to the aforementioned accounting identity, noting that midgrade gasoline is a perfect physical substitute for a 50-50 mix of regular grade gasoline and premium grade gasoline. We can thus model product choice not as a choice among three distinct products but rather as a choice of the average probability p that a consumer of a given income chooses premium grade gasoline. A consumer that always purchases regular grade gasoline has $p = 0$. A consumer that always purchases premium grade gasoline has $p = 1$. A consumer that always purchases midgrade, or equivalently alternates equally between premium grade and regular grade gasoline has $p = 0.5$. Now imagine an Engel Curve with income on the vertical axis and the probability p of purchasing premium grade gasoline (relative to regular grade gasoline) on the horizontal axis. The curve is convex, intuitively, because the gap in the incomes of consumers who purchase regular grade and midgrade gasoline tends to be smaller than the gap in incomes between those who purchase midgrade and premium grade gasoline.

It is easy to understand the convexity by noting that there are virtually no vehicles that require midgrade gasoline, so that midgrade gasoline is generally purchased, and often only occasionally,

(and recognizing that midgrade is very rarely sold at wholesale).

by consumers whose vehicles only require regular grade gasoline. These consumers have made relatively similar vehicle choices as consumers of regular grade gasoline, and are often one and the same consumer. In contrast, there are numerous vehicles – luxury vehicles, high-end sports cars, and other high performance vehicles – that require premium grade gasoline exclusively and these are disproportionately owned by the wealthiest group of consumers in a market. While it is well known that some consumers with "regular" cars may sometimes purchase premium grade gasoline (though unnecessarily so, since there is no mechanical benefit to a car that does not require it), premium grade gasoline is disproportionately purchased by the wealthiest individuals whose high-end vehicle choices always require it. The result is a convex Engel Curve. The convexity in the Engel Curve has been empirically confirmed by Coats et al. (2005), who estimate income elasticities for different grades of gasoline and whose results imply a convex Engel Curve when drawn this way.[19]

Next, since search costs are essentially the opportunity cost of time spent searching, the corresponding plot of search costs against p (instead of income against p) is also convex, as is the plot of price dispersion against p if search costs matter as predicted. In essence, and in contrast to differential signboard price information which creates a strong divide between regular grade gasoline on one hand and midgrade and premium grade gasoline on the other, leading to concavity, income differences create a divide between regular and midgrade on one hand and premium grade on the other, leading to convexity. Even before incorporating direct income information into our analysis, we can test for the income hypothesis using our linearity test, with the alternate hypothesis that the test statistic will be significantly positive (convexity) instead of negative (concavity).

Given the interest in income as a potential cause, we conclude our results section with two additional analyses. In the first, we conduct our analysis separately for each of the three main regions in the city - the West, Central, and East regions - noting the strong gradient in income from east to west. We compare the patterns of coefficients across the income-based regions in each case and test for continued concavity in our grade specific price dispersion coefficients within each subregion. In the second, we directly incorporate measures of income, gender, and race in the

[19] The authors calculate income elasticities for regular and premium grade gasoline and find positive income elasticities, higher for premium grade than regular grade. Their results imply that the elasticity of p (the probability of choosing premium) with respect to Y is positive, smaller than one, and concave in Y. Thus, after reversing the axes by putting Y on the vertical and p on the horizontal, the Engel curve is seen to be convex.

model, by subregion. We then estimate our full model - including interactions between income and grade-specific measures of price dispersion - and test for continued concavity in our grade-specific price dispersion coefficients.

Throughout our analysis, we include a measure of market structure as well. We use the number of stations in a given subregion upon which each calculation of price dispersion is based, though our main results are unaffected if we use the total number of stations in a given subregion instead. The variable potentially helps control for two things. First, it helps control for station density and the degree of spatial competition given that the subregions are of roughly equal size. The effects of competition and density on price dispersion has been widely explored in the literature (e.g. Lewis (2008), Deltas (2008)). Secondly, it controls for the fact that measures of price dispersion can mechanically depend on the number of observations used to calculate it, especially for smaller numbers of observations. The overall effect of the market structure variable on search costs and price dispersion is in general ambiguous, and empirical results have varied (Borenstein and Rose (1994), Barron et al. (2004), Dai et al. (2014)).[20] In our particular setting, we refrain from claiming a causal interpretation of our market structure variable for three reasons. First, market structure is itself an outcome variable and endogenously determined. Second, and on a related note, in our sample period of 33 days, there are no identifying shocks to market structure that would allow us to separate the marginal effect of an additional station vis-a-vis the underlying factors that would have led to that station's entry in the first place. Third, all three grades are sold by virtually all stations in the sample, so there is no differential variation in market structure across the three grades that would be relevant for our cross-grade comparisons. We include the variable as an extra control and to be complete, and our main results on the difference in price dispersion across the three grades are unaffected by its inclusion or exclusion. Our results are also unaffected by the particular definition of this variable (e.g. counts, density measures, etc.) or the definition of subregions over which it is calculated.

[20] If all firms and consumers are perfectly homogeneous, we expect a negative relationship. However, if consumers are heterogeneous and firms respond by targeting those different consumer niches, increased competition can increase price dispersion.

4 Results

We begin by testing for the presence of rank reversals. As discussed in Chandra and Tappata (2011), if price dispersion were caused by consumer search, we should observe a non-negligible amount of rank reversals among nearby stations. If price dispersion were instead caused by product or station differentiation alone, then relative prices would be stable and we should not.[21] We compare prices within the same grade of gasoline for all pairs of stations located within the same subregion, for all nine subregions in the city. We compare 766 unique station-pairs in total.

The results are reported in Table 2 and show that rank reversals are indeed very frequent, supporting a search-based explanation for the observed price dispersion. Rank reversals occurred for approximately 73.5% of the 766 station-pairs, and the mean value of the rank reversal measure conditional on a pair having at least one rank reversal was 0.23 (out of a maximum of 0.5). We find that the percentage of rank reversals and the value of the conditional mean are high for all grades and, as expected, highest for regular grade gasoline (for which the price distribution is especially tight). Similar to the findings of Chandra and Tappata (2011), the high percentage of rank reversals is consistent with a search cost based explanation and inconsistent with a pure station differentiation story.

With evidence of a search-cost related cause in hand, we turn to an analysis of the potential sources of search costs that could lead to observed price dispersion. We focus on the factors that could cause the differential-by-grade price dispersion we observe (rather than common changes over time), and begin with differential signboard price information. The shock provided by the differential provision of signboard price information creates a substantial divide in the ease of collecting price information, with regular grade gasoline on one side of the divide and midgrade and premium grade gasoline on the other. We expect it to lead to a similarly strong divide in the

[21] Two other possible causes for a presence of rank reversals, aside from search, would be the presence of Edgeworth price cycles or the practice of intertemporal price discrimination (Noel (2012)). These are easily ruled out, however. First, there are no Edgeworth price cycles in the study market so this cause is easily ruled out. Second, and unlike supermarket goods or some other consumer goods, the three grades of gasoline taken together do not go on frequent and regular "sales" in the absence of Edgeworth price cycles. There is no day-of-the-week or other predictable intertemporal pattern of prices that consumers could use to shift future purchases to periods of lower prices, and there is extremely limited ability to stockpile if gasoline ever were to go on sale unexpectedly. While gasoline prices do rise and fall with crude oil prices, it is generally at a much slower frequency than the frequency with which consumers fill their tanks, limiting the scope for intertemporal switching in the absence of Edgeworth price cycles.

degree of price dispersion along that same boundary, i.e. a concave relationship in our estimated price dispersion coefficients.

Table 3 presents results for two different and common measures of price dispersion (e.g. Sorensen (2000)). The first set, in Specifications (1) through (3), uses standard deviation as the measure of price dispersion, and the second set, in Specifications (4) through (6), uses the max-min range. The right hand side variables of interest are the grade indicator variables, $MIDGRADE$ and $PREMIUM$.

The first specification in each set (Specifications (1) and (4)) is based on the complete dataset and includes an indicator variable for each grade (with regular grade gasoline being the omitted variable), plus daily and regional indicator variables. The daily indicator variables control for any station-invariant, grade-invariant but time-variant factors, including daily changes in wholesale price levels (which are the focus of asymmetric passthrough studies). The second specification in each set (Specifications (2) and (5)) is similar but includes our measure of market structure as well.

The third specification in each set (Specifications (3) and (6)) is similar to the second except that it restricts the price data to only those stations for which there is complete price information across grades at a point in time, i.e. for which the prices of all three grades of gasoline at a given station are recorded on a given day. These are our preferred specifications, as they eliminate any potential composition bias that could arise when price dispersion for one grade of gasoline is calculated over one group of stations and price dispersion for another is calculated over a different albeit overlapping group of stations. For example, imagine that a station was very expensive relative to other stations generally. Its relatively high regular, midgrade, and premium grade prices should lead to higher price dispersion for each grade if they are observed. However, if its premium grade (midgrade) price were not observed on a given day, while its regular grade price were, then price dispersion in premium grade (midgrade) gasoline would be underestimated and the $PREMIUM$ ($MIDGRADE$) coefficient would be biased downward. Specifications (3) and (6) avoid this issue.

The results of Table 3 provide evidence that, as hypothesized, signboard price information plays an important role in differential-by-grade price dispersion. First, we find a substantial difference in the degree of price dispersion between regular grade gasoline, whose prices are posted on signboards, and midgrade and premium grade gasoline, whose prices are not. The $MIDGRADE$ and

PREMIUM coefficients are significantly different from zero at better than the 1% level in all six specifications. Second, we find a substantial concavity in the grade-specific coefficients, consistent with the discrete divide in search costs created by the differential signboard price information that divides regular grade gasoline from higher octane grades. The premium grade coefficient is significantly less than double the midgrade coefficient in four specifications, and barely shy of significance in the remaining two (with p-values of 0.11 in each case). In our preferred Specifications (3) and (6), which remove any composition bias, the concavity is statistically significant at better than the 1% level in all cases.

Taking Specification (3) as an example for discussion, we find that the standard deviation of midgrade gasoline prices is 3.9 cents per gallon higher than that of regular, and the standard deviation of premium grade prices is 4.7 cents per gallon higher, both statistically significant at the 1% level. The difference between the standard deviation of premium grade gasoline and twice the standard deviation of midgrade gasoline, as seen in the "Linearity of Coefficients Test" near the bottom of the table, is also statistically significantly different from zero. The test statistic is negative, at -3.2 cents per gallon, meaning that the coefficient pattern is significantly concave as predicted. In fact, the "Equality of Coefficients Test" near the bottom of the table shows that the *MIDGRADE* and *PREMIUM* coefficients are statistically indistinguishable from each other, consistent with the extreme case of perfect Γ concavity. Our finding of concavity separates our price information hypothesis from competing hypotheses and identifies differential signboard price information as a significant cause underlying the differential-by-grade price dispersion that we observe.

Specification (6), which repeats Specification (3) using the max-min range instead, shows a similar pattern. The range of midgrade gasoline prices is 10.9 cents per gallon higher than that for regular grade gasoline, and the range of premium grade gasoline prices is 12.9 cents per gallon higher, both statistically significant at the 1% level. The Linearity of Coefficients Test yields a test statistic of -8.8, statistically significant at the 1% level. The large and negative statistic shows a strong concavity in the grade-specific coefficients. In fact, the Equality of Coefficients Test shows that the *MIDGRADE* and *PREMIUM* coefficients are statistically indistinguishable from one another, again consistent with the extreme case of perfect Γ concavity.

In terms of our market structure variable, included in the last two specifications of each set (Specifications (2), (3), (5), and (6)), we find a positive effect in Specifications (2) and (5) (ignoring potential composition bias), but the effect becomes insignificant when controlling for composition bias in Specifications (3) and (6). While the insignificant result differs from studies that try to identify a causal effect of market structure on price dispersion (e.g. Barron et al. (2004), Lewis (2008)), recall that we cannot offer a causal interpretation of our variable due to the nature of our sample and the lack of an exogenous shock to identify any effect. Our results are consistent with the limited amount of market structure variation across the subregions of the city and the absence of market structure variation across the three grades.

To this point we have performed regressions where price dispersion is calculated on the basis of raw prices. A common alternate approach in the literature is to first regress price dispersion on product and store fixed effects and then construct measures of price dispersion based on the residuals of this regression instead of raw prices (e.g. Sorensen (2000)). We do that here as well. The first stage regression that produces the residuals is given by:

$$\begin{aligned} PRICE_{git} &= \gamma_0 + \gamma_1 MIDGRADE_g + \gamma_2 PREMIUM_g + \sum_{i=2}^{\bar{I}} \zeta_i I_i \\ & \sum_{t=2}^{\bar{I}} \zeta_{im} * I_i * MIDGRADE_g + \sum_{t=2}^{I} \zeta_{ip} * I_i * PREMIUM_g + \nu_{git} \end{aligned} \tag{4}$$

where the I_i are station-specific indicator functions and $\bar{I}$ is the number of stations. Price dispersion is then calculated across the ν_{git} instead of the $PRICE_{git}$. Our new measures of price dispersion reflect the extent to which individual station operators vary their own midgrade and premium grade prices more than their own regular grade price, relative to their own station-specific and product-specific means. In other words, these measures remove any common or stable component in the differential between regular grade and premium grade (midgrade) prices for a given station and focus only on unpredictable "surprises" in that differential. For example, imagine a station sets its premium grade price twenty cents per gallon above its regular grade price on average. That twenty-cent-on-average differential could be used by consumers to help predict the price of premium grade gasoline from the price of regular grade gasoline, even when the premium grade price is not posted. In these residual-based specifications, we remove that stable component (the station-grade specific constant for that station, ζ_{ip}, would be twenty) and the residuals contain only the leftover surprises in the differential above or below the twenty cent average.[22]

We report the results under this approach in Specifications (7) to (12) in Table 4. The six new specifications correspond to the earlier six specifications in Table 3 based on raw prices. The preferred specifications that remove potential composition bias are Specifications (9) and (12).

In short, we find that our concavity results continue to hold and are, in fact, stronger than before. The estimates on the $MIDGRADE$ and $PREMIUM$ coefficients are positive and statistically significantly different from zero across all six specifications, at the 1% level. The Linearity of Coefficients Test strongly rejects linearity in favor of concavity at the 1% level in all cases. The $MIDGRADE$ and $PREMIUM$ coefficients are now almost identical as well, and the Equality of Coefficients Test confirms that they are indistinguishable from one another in all cases, consistent not only with concavity but with perfect Γ concavity.

Taking Specification (9) as an example for discussion, we find that the standard deviation of midgrade gasoline residuals is 0.99 cents per gallon higher than that of regular, and the standard deviation of premium grade residuals is 0.96 cents per gallon higher, both statistically significant at better than the 1% level. The Linearity of Coefficients Test shows that the difference between the standard deviation of premium grade gasoline and twice the standard deviation of midgrade gasoline is -1.0, statistically significantly negative at the 1% level, and showing concavity. The Equality of Coefficients Test yields a test statistic of -0.03 (with a small standard error of 0.14), showing the $MIDGRADE$ and $PREMIUM$ coefficients are extremely close and statistically indistinguishable from each other, consistent with perfect Γ concavity.

In Specification (12), using the max-min range instead, the range in the midgrade gasoline

[22] It is worth noting up front that the differential between regular and premium grade prices is largely unpredictable from the regular grade price alone. This is evident from the results of the residual regressions, but it can be seen another way as well. We calculate the variance in the regular-premium grade differential as a fraction of the variance in the premium grade price for each station. A value of zero means a perfectly stable differential. A value of 100% means the differential is as volatile as the premium grade price itself. We find a high median value of 40%, showing that the differential between a station's posted regular grade price and its unposted premium grade price, and thus a station's premium grade price ranking, is difficult to infer from its posted regular grade price alone. We perform the same analysis on the midgrade-regular grade differential and find a median value of 49% in that case.

residuals is 2.9 cents per gallon higher than that for regular grade gasoline, and the range in the premium grade gasoline residuals is 2.6 cents per gallon higher, both statistically significant at the 1% level. The Linearity of Coefficients Test yields a test statistic of −3.2, statistically significant at better than the 1% level, and showing concavity. The Equality of Coefficients Test yields a test statistic of −0.3, statistically insignificantly different from zero and showing not only concavity but perfect Γ concavity.

The results of Tables 3 and 4 taken together support the hypothesis that differential signboard information is an important driver of the differential degrees of price dispersion across grades. The divide in signboard price information with midgrade and premium grade gasoline on one side of the divide and regular grade gasoline on the other, predicts a similarly substantial divide in price dispersion along this same boundary, and that is what we observe.[23]

We now consider alternative explanations and begin with the hypothesis that the differential-by-grade retail price dispersion could be the result of differential wholesale price dispersion. Under this hypothesis, the pattern of grade-specific coefficients is predicted to be convex, or linear in the limiting case, as discussed above. We test this using linearity tests and summarize the relevant results in Table 5. In the first row of Table 5, we summarize the Linearity of Coefficients Tests when comparing price dispersion based on raw prices (from Specifications (1) through (6) in Table 3) and in the second row of the table, we report the Linearity of Coefficients Tests using price dispersion based on price residuals instead (from Specifications (7) through (12) in Table 4).

The results of Table 5 reject the wholesale price dispersion hypothesis. Beginning with raw price dispersion in the first rows, we reject linearity and convexity in favor of concavity in four of six specifications, with p-values of 0.11 in the remaining two. In our preferred Specifications (9) and (12), we reject linearity and convexity at the 1% level in both cases. Using residual dispersion instead of raw price dispersion, we now reject linearity and convexity at the 1% level

[23] Even though the sample period is only 33 days long and there is no change in the formulation of the three different grades or in the presence or absence of signboards, we test for the possibility of a structural break in the residual price dispersion series, for each grade, using a series of Wald tests. We allow for the possibility of a break starting on any given day. In short, we find no meaningful structural breaks in the data. We do get on occasion a significant Wald test when assuming a break in the first few or last few days of the sample, which often occurs under this method when the first or last few observations are meaningfully above or below average. However, dropping the first or last few days accordingly and reperforming the analysis on the remaining data has no meaningful impact on our results or conclusions.

in all specifications. Since a wholesale price dispersion story requires linearity or convexity in the grade coefficients, we reject that wholesale price dispersion hypothesis is a leading cause.

The same Table 5 can also be used to test whether a pure income story is responsible for the differential-by-grade price dispersion we observe. Under a pure income story, the coefficients would again take on a convex pattern because of the nature of the consumers that buy different grades, as discussed above. For the same reason we can reject wholesale price dispersion as a leading cause, we can also reject a pure income story as the leading cause of the differential-by-grade price dispersion we observe.

In summary, the net concavity we find in the grade specific coefficients is consistent with differential signboard price information as a leading cause of the differential-by-grade price dispersion in gasoline prices. It is inconsistent with other leading hypotheses, such as wholesale price dispersion or income differences, which favor a linear or convex pattern instead.

This is not to say that wholesale price dispersion or income differences can have no effect and we suspect this is unlikely, at least for income. Our finding of concavity *on net* simply reflects the relatively larger impact of differential signboard price information on price dispersion compared to other contributors.

Even if not primary, it would be interesting to further explore the potential effect that income has on price dispersion, and we do this in two ways. First, we divide the city into three equally sized regions – the East, Central, and West Regions – based on its steep east-to-west income gradient, and re-estimate our model separately for each income-based region in search of differential patterns. Second, we directly incorporate income and socioeconomic data into our main analysis and explicitly estimate the effect of these variables on price dispersion.

We begin by performing region-specific versions of our main model and report the results in Table 6. Specifications (13) to (15) are region-specific versions of Specification (9), using the standard deviation of price residuals. Specifications (16) to (18) are region-specific versions of Specification (12), using the max-min range of residuals. The regions exhibit a large gradient in income – the median income is \$32,371 in the East Region, \$38,497 in the Central Region, and \$55,671 in the West Region.[24] Because of the finer split, sample sizes are smaller than in the

[24] We have performed various alternate splits of the city with similar results.

main analysis and standard errors are expected to be higher. In Specifications (13) to (15), we find that, even with fewer observations, the $MIDGRADE$ and $PREMIUM$ coefficients continue to be positive and statistically significant in each region. The concavity in the point estimates in each region is again clear and consistent with earlier results, but the higher standard errors do weaken significance some. The Linearity of Coefficients Test is unable to reject linearity in the Central region (unable to statistically show concavity comparing the regular grade "coefficient" of zero to the midgrade and premium grade coefficients of 0.97 and 1.28 respectively) but does reject linearity in favor of concavity at the 1% level in the West Region. It is marginally shy of significance in the East Region (with a p-value = 0.106). The Equality of Coefficients Test, as before, cannot reject perfect Γ concavity in any region, i.e. that the grade specific coefficients are equal to one another. In Specifications (16) to (18), we find similar and more significant results, with statistically significant concavity in both the West and East Regions, and point estimates consistent with concavity in the Central Region, but just short of significance with its smaller sample size. We performed region-specific versions of all other specifications from Tables 3 and 4 as well, and continue to find concavity in all regions, statistically significantly so in the East and the West in particular.

Examining the magnitudes of the $MIDGRADE$ and $PREMIUM$ coefficients across regions, they tend not to be too different from one region to the next and all differences between them are insignificant. Moving from the left to the right across the first three columns, the $MIDGRADE$ coefficients are modestly and insignificantly higher as income increases while the $PREMIUM$ coefficients show no particular directional relationship. In the last three columns, the $PREMIUM$ coefficients are modestly and insignificantly higher while the $MIDGRADE$ coefficients show no particular directional relationship. While possibly impacting price dispersion, it does not appear that income is the dominant force behind the differential price dispersion across grades.

We explore the potential effects of income further by explicitly adding income and socioeconomic variables to the model directly. The key variables are income ("Income", in thousands of dollars), race ("Black" and "Hispanic", with whites and others omitted), and gender ("Female" with males omitted), aggregated to the subregional level. While we cannot observe transactional data and the identity of the consumers that purchase each grade of gasoline at each station, we expect that, if income plays an important role, we should observe particularly higher price dispersion on the higher grades of gasoline where wealthier consumers live and drive. The steep gradient of income from one side of the city to the other, the short commutes, and the compartmentalized nature of traffic relative to larger cities in the absence of a high-density commercial downtown, improves the match. We report results in Table 7. Specifications (19) to (21) use the standard deviation of prices, Specifications (22) to (24) the standard deviation of residuals. Results using max-min ranges in place of standard deviations are similar.

In a preliminary step, in Specifications (19) and (20), we first test whether income affects overall price dispersion independent of grade. We add income as an explanatory variable to both specifications and add demographic variables to Specification (20) as well. We find that the results relating to income are mixed. We find a positive, statistically insignificant coefficient on income in Specification (19), but a larger and statistically significant coefficient on income in Specification (20).[25] When we use residuals instead of raw prices in Specifications (22) and (23), we find insignificant coefficients in both cases.[26]

We now turn to the question of whether income has an effect on price dispersion differentially by grade, even if not dominant. Results are presented in Specifications (21) and (24). Specification (21) is based on dispersion in raw prices and Specification (24) is based on dispersion in price residuals. Specifically, we are testing whether differential-by-grade price dispersion is more pronounced in wealthier areas than in poorer areas, using income data directly instead of a regional proxy. We thus add interaction terms $MIDGRADE * INCOME$ and $PREMIUM * INCOME$ to the analysis, in addition to income and demographic main effects. We again find that the results on income are mixed and marginal. Beginning with Specification (21), we find a statistically significant interaction term for midgrade suggesting that midgrade price dispersion, relative to regular grade price dispersion, tends to be a little higher in wealthier areas. The estimated coefficient on the premium-income interaction term is also positive and higher still, but statistically insignificant. In Specification (24), the interaction term for midgrade is again marginally statistically significant and the interaction term for premium grade is insignificant. In both cases, we cannot rule out the

[25] The $MIDGRADE$ and $PREMIUM$ coefficients do not change across the two specifications since all consumer-specific variables are grade-invariant.

[26] Similarly, when adding the square of income, the income coefficients are insignificant.

expected convex relationship among the income *interaction* variables, as expected.

Even in these richer specifications, our main results on signboard price information continue to hold. We are able to rule out linearity in favor of a concave pattern in the grade-specific coefficients, $MIDGRADE$ and $PREMIUM$, even after holding income and other X matrix variables fixed at their means.[27]

So relatively speaking, how important is the estimated effect of income relative to the estimated effect of signboard price information in the coefficient patterns? The point estimates on the income interaction terms, even if one were to excuse the statistical insignificance for premium, turn out to be relatively modest. Using Specification (21) as an example for discussion, and optimistically taking the point estimates as precise, we find that a 10% increase in income would increase the standard deviation of midgrade prices only from 3.96 to 4.24, and the standard deviation of premium grade prices only from 4.73 to 5.09. From the poorest area to the wealthiest area in the data, the standard deviation of midgrade gasoline (up and above that of regular grade) would increase only from 3.18 to 4.71 and the standard deviation of premium grade gasoline (up and above regular) would increase only from 3.74 to 5.70. In other words, even in the poorest areas, there are meaningfully higher degrees of price dispersion for midgrade and premium grade gasoline than for regular grade, even as income is largely taken out of the equation. Midgrade price dispersion is still higher by 3.18 cents per gallon and premium grade price dispersion is still higher by 3.74 cents per gallon. In fact, just 19.5% or 20.8% of the mean price dispersion, for midgrade and premium grade gasoline respectively, is potentially attributable to the existence of differential incomes above or below the mean. Using Specification (24) instead, and again excusing the statistical insignificance in the interaction term for premium, we calculate only 37.4% or 29.6%, respectively, of the mean price dispersion is potentially attributable to differential incomes above and below the mean. We conclude that, even though income might be assumed to be the obvious candidate for differential-by-grade price dispersion, in our application its effect turns out to be secondary.

The primary driver of the higher degrees of price dispersion on higher octane gasolines in our

[27]The $MIDGRADE$ and $PREMIUM$ coefficients are no longer easily interpretable on their own, as they correspond to the differential degrees of price dispersion that would occur in a non-existent neighborhood with a median income of zero. The appropriate measure for determining concavity is the sum of each coefficient and its interaction term, evaluated at the mean income.

sample market is the substantial divide in signboard price information, with regular grade gasoline on one side of that divide and midgrade and premium grade gasoline on the other. It implies a similar divide in differential-by-grade price dispersion along those same lines, and that is indeed what we observe.

Finally, we conduct additional robustness checks to confirm the robustness of our earlier results and report them in Table 8. In Specifications (25) through (27), we use the interquartile range as the measure of price dispersion instead of standard deviations or max-min ranges, and find that the results carry through. These specifications are based on dispersion in the price residuals; specifications based on raw prices produce similar results. We also re-estimate our model using the Gini coefficient as the measure of price dispersion (not shown) and again all results carry through, whether using raw prices or price residuals. Specifications (28) to (30) use subregion indicator variables to control for subregional differences in place of region indicator variables and again all our main results carry through. (The station count variable cannot be included in these specifications because it is absorbed into the subregion indicator variables.) Specification (28) uses standard deviations, Specification (29) uses max-min ranges, and Specification (30) uses interquartile ranges. In all cases, the $MIDGRADE$ and $PREMIUM$ coefficients are very close to one another, the Linearity of Coefficients Tests easily reject linearity and convexity in favor of concavity, and the Equality of Coefficients cannot reject perfect Γ concavity.

5 Discussion and Conclusion

Information plays a central role in market economies. It is compiled and used by consumers and competitors everyday to help them in their decisionmaking. In this article, we examine the relationship between information, search costs, and gasoline price dispersion in the retail gasoline industry. Our study exploits a unique historical artifact – that in most areas, the price of regular grade gasoline is prominently displayed on large streetside signboards that are easily viewable by consumers, while the prices of midgrade and premium grade gasolines are generally not. This creates a divide in the relative ease of gathering gasoline price information with regular grade gasoline on one side of the divide and midgrade and premium grade gasoline on the other. It implies a particular concave curvature in the distribution of our grade-specific coefficients, and we do indeed find evidence of such a curvature. We consider alternate hypotheses as well, such as cost dispersion and income differences and, while we find mixed support for a secondary effect of income, we conclude that the differential provision of price information is an important cause.

Differential signboard price information by grade is an interesting feature of retail gasoline markets that has received little attention in the literature. We are not aware of other studies that explicitly collect and use it, likely because standard datasets do not contain this information and its collection on a large scale would require a substantial effort on the part of researchers. Yet with all the interest in price transparency regulations, understanding how consumers search for and use signboard price information is an important issue with clear policy implications. Our study is the first, to our knowledge, to consider the effects of signboard information. In future work, we plan to expand our own effort, to extend the current study to multiple cities with local variation in the content of signboard information.

As discussed above, there are numerous other papers that examine the effect of various types of search costs and benefits on gasoline price dispersion, including differences in market structure and station density (e.g. Barron et al. (2004)), the prevalence of commuting traffic (e.g. Pennerstorfer (2016)), and volatility in wholesale prices over time (e.g. the numerous asymmetric passthrough papers cited above). Our focus on signboard price information is uniquely different. While the literature has generally focused on regular grade gasoline only, the question of how signboard information affects price dispersion cannot be properly addressed using regular grade price data only, since stations almost universally post their regular grade prices on signboards.[28] Therein lies a main contribution of the current paper. By expanding the analysis to all three grades of gasoline, including higher octane grades for which there is often no price information on signboards, we are able to address the question of how price information affects price dispersion in a way not before possible.

The only other paper we are aware of that examines differential-by-grade gasoline is Chandra and Tappata (2011). While the primary goal of that paper is to establish the non-monotonicity of

[28] If an isolated station were to post no signboard information at all (where allowed), presumably there would be a reason for doing so, meaning that the choice of signboard information would be endogenous to price setting, and more difficult to disentangle.

the consumer-search-price-dispersion relationship and show that it is negative in the case of retail gasoline, they include an analysis of price dispersion by grade and find that price dispersion on premium grade and midgrade gasoline tends to be higher than that for regular grade gasoline. They postulate that the reason behind the differential-by-grade price dispersion may be income-based, ruling out imperfect information due to the common practice of posting price information on large signboards. However, in many markets including our study market, this practice does not apply to all grades of gasoline. Only the price of regular grade gasoline is posted, while those of higher octane grades are not, causing a discrete divide in the ease of collecting price information along this boundary. In this study, we explore a series of possible causes for differential-by-grade price information and find that signboard information does indeed matter.

It would be easy to interpret our findings as a call for increased transparency regulation, but this would also be premature. Our study is only the first step in this discussion and additional research is needed. Our analysis allows us to identify effects on price information on price dispersion and conclude that signboard information matters to consumers. We cannot yet say to what extent this would translate into higher elasticities on high octane grades or improve market efficiency post-regulation. Since regulation carries compliance and enforcement costs, and one must compare the costs to the benefits, there is no assurance ex ante that the *net* benefit will necessarily be positive. In fact, there is no assurance ex ante that even the *gross* benefit will be positive. The hypothesized benefits of expanding signboard information would presumably accrue to the purchasers of higher grades of gasoline, which are generally higher income individuals. Since the retail gasoline industry is a competitive industry in which firms sell a portfolio of products in an attempt to recover its fixed costs and earn a competitive rate of return, higher markups on higher octane grades can potentially allow firms to be more aggressive on regular grade gasoline prices, while still maintaining overall profitability. If a price transparency regulation were to lead to lower prices on higher octane grades, it could force firms to compensate by increasing prices on regular grade gasoline, which is generally bought by lower income individuals instead. There are many interconnected parts, and additional research is needed to consider each of the costs and benefits of any proposed transparency regulation. Indeed, little research has been done to understand the costs and benefits of greater gasoline price transparency, and this article is an early effort in that direction.

Finally, it is interesting to note that the retail gasoline industry would seem an unusual target for price transparency laws at first glance. In the absence of price regulation, there are few industries where prices are already as transparent as they are in the retail gasoline industry (including premium prices). Prices are always clearly visible at the pump (if not also on the signboard), they are a rare example of tax-inclusive prices, they do not contain hidden fees or surcharges that are commonplace in some other industries (e.g. hotels, rental cars), and are far from the opaque nature of pricing in health care markets, for example. The unique regulatory attention surrounding the retail gasoline industry and the question of price transparency stems in large part from the heightened importance consumers attach to the industry. This is in itself an interesting phenomenon and another potential avenue for future research.

6 References

Atkinson, B. (2008), "On Retail Gasoline Pricing Websites: Potential Sample Selection Biases and their Implications for Empirical Research", Review of Industrial Organization, 33(2), 161-175.

Atkinson, B., A. Eckert and D.S. West (2014), "Daily Price Cycles and Constant Margins: Recent Events in Canadian Gasoline Retailing", Energy Journal, 35(3), 47-69.

Bachmeier, L.J. and J.M. Griffin (2003), "New Evidence on Asymmetric Gasoline Price Responses", Review of Economics and Statistics, 85(3), 772-776.

Bacon, R.W. (1991), "Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes", Energy Economics 13(3), 211-218.

Barron, J.M., B.A. Taylor and J.R. Umbeck (2004), "Number of Sellers, Average Prices, and Price Dispersion", International Journal of Industrial Organization 22(8), 1041-1066.

Baye, M.R., J. Morgan and P. Scholten (2003), "The Value of Information in an Online Consumer Electronics Market", Journal of Public Policy and Marketing 22(1), 17-25.

Borenstein, S., A.C. Cameron, and R. Gilbert (1997), "Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?", Quarterly Journal of Economics, 112(1), 305-339.

Borenstein, S., and N. Rose (1994), "Competition and Price Dispersion in the U.S. Airline Industry", Journal of Political Economy 102(4), 653-683

Brown, J.R. and A. Goolsbee (2002), "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry", Journal of Political Economy 110, 481–507.

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

Byrne, D.P., G. Leslie and R. Ware (2015), "How do Consumers Respond to Gasoline Price Cycles?", Energy Journal 36(1).

Byrne, D.P. and N. de Roos (2015), "Consumer Search in Retail Gasoline Markets", Journal of Industrial Economics 65(1), 183-193.

Cabral, L. and A. Fishman. (2012), "Business as Usual: A Consumer Search Theory of Sticky Prices and Asymmetric Price Adjustment", International Journal of Industrial Organization, 30(4), 371-376.

Chandra, A. and M. Tappata (2011), "Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets", RAND Journal of Economics, 42(4), 681-704.

Coats, R.M., G M. Pecquet and L. Taylor (2005), "The pricing of gasoline grades and the third law of demand", Microeconomics 0506006, EconWPA, 1-17.

Cooper, T.E. and J.T. Jones (2007), "Asymmetric Competition on Commuter Routes: The Case of Gasoline Retailing", Southern Economic Journal, 74(2), 483-504.

Dahlby, B. and D.S. West (1986), "Price Dispersion in an Automobile Insurance Market", Journal of Political Economy 94(2), 418-438.

Deltas, G. (2008), "Retail Gasoline Price Dynamics and Local Market Power", Journal of Industrial Economics, 56(3), 613-629.

Dai, M., Q. Liu, and K. Serfes (2014), "Is the Effect of Competition on Price Dispersion Nonmonotonic? Evidence from the U.S. Airline Industry", Review of Economics and Statistics 96(1), 161-170.

Diamond, P. (1971), "A Model of Price Adjustment", Journal of Economic Theory 3(2), 156-168.

Dubois, P. and H. Perrone (2015), "Price Dispersion and Informational Frictions", working paper.

Eckert, A. (2013), "Empirical Studies of Gasoline Retailing: A Guide to the Literature" 27(1), 140-166.

Eleftheriou, K., P. Nijkamp, and M. Polemis (2018), "Asymmetric Price Adjustments in US Gasoline Markets: Impacts of Spatial Dependence on the Rockets and Feathers Hypothesis", Regional Studies, DOI: 10.0180/00343404.2018.1463093

Galeotti, M., A. Lanza and M. Manera (2003), "Rockets and Feathers Revisited: an International Comparison on European Gasoline Markets", Energy Economics, 25(2), 175-190.

Gicheva, D., J. Hastings, and S. Villas-Boas (2010), "Investigating Income Effects in Scanner Data: Do Gasoline Prices Affect Grocery Purchases?", American Economic Review Papers and Proceedings, 100 (2), 480-484.

Godby, R., A. Lintner, B. Wandschneider and T. Stengos (2000), "Testing for Asymmetric Pricing in the Canadian Gasoline Market", Energy Economics, 22(3), 349-368.

Hansen, B.E. (1999), "Threshold Effects in Non-Dynamic Panels: Estimation, Testing, and Inference", Journal of Econometrics 93(2), 345-368.

Hosken, D.S., R.S. McMillan, C.T. Taylor (2008), "Retail Gasoline Pricing: What Do We Know?", International Journal of Industrial Organization, 26(6), 1425-1436.

Hortacsu, A. and C. Syverson (2004), "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P Index Funds", Quarterly Journal of Economics 119(2), 403-456.

Houde, J.-F. (2012), "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline", American Economic Review, 102(5), 2147-2182.

Hughes, J.E., C.R. Knittel, and D. Sperling (2008), "Evidence of a Shift in the Short-Run Price

Elasticity of Gasoline Demand", Energy Journal, 29(1), 113-134.

Janssen, M., P. Pichler, and S. Weidenholzer (2011), "Oligopolistic Markets with Sequential Search and Production Cost Uncertainty", RAND Journal of Economics, 42(3), 444-470.

Kristoufek, L. and P. Lunackova (2015), "Rockets and Feathers Meet Joseph: Reinvestigating the Oil-Gasoline Asymmetry on the International Markets", Energy Economics 49, 1-8.

Li, J. and J.H. Stock (2018), "Cost Pass-through to Higher Ethanol Blends at the Pump: Evidence from Minnesota Gas Station Data", Journal of Environmental Economics and Management, https://doi.org/10.1016/j.jeem.2018.08.003

Lewis, M. (2008), "Price Dispersion and Competition with Differentiated Sellers", Journal of Industrial Economics, 56(3), 654-678.

Lewis, M. (2011), "Asymmetric Price Adjustment and Consumer Search: An Examination of the Retail Gasoline Market", Journal of Economics and Management Strategy, 20(2), 409-449.

Lewis, M. and H. Marvel (2011), "When Do Consumers Search?", Journal of Industrial Economics 59(3), 457-483.

Milyo, J. and J. Waldfogel (1999), "The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart", American Economic Review 89(5), 1081-1096.

Noel, M.D. (2007), "Edgeworth Price Cycles, Cost-Based Pricing, and Sticky Pricing in Retail Gasoline Markets", Review of Economics and Statistics, 89(2), 324-334.

Noel, M.D. (2009), "Do Retail Gasoline Prices Respond Asymmetrically to Cost Shocks? The Influence of Edgeworth Cycles", RAND Journal of Economics, 40(3), 582-595.

Noel, M.D. (2012), "Edgeworth Price Cycles and Intertemporal Price Discrimination", Energy Economics, 34(4), 942-954.

Noel, M.D. (2016), "Retail Gasoline Markets" in the Handbook on the Economics of Retail and Distribution, E. Basker (ed.), Edward Edgar Publishing.

Noel, M.D. (forthcoming), "Gasoline Price Dispersion and Consumer Search: Evidence from a

Natural Experiment", Journal of Industrial Economics.

Orlov, E. (2011), "How Does the Internet Influence Price Dispersion? Evidence from the Airline Industry", Journal of Industrial Economics 59(1), 21-37.

Peltzman, S. (2000), "Prices Rise Faster Than They Fall", Journal of Political Economy, 108(3), 466–502.

Pennerstorfer, D., P. Schmidt-Dengler, N. Schutz, C. Weiss and B. Yontcheva (2016), "Information and Price Dispersion: Evidence from Retail Gasoline", working paper.

Radchenko, S. (2005), "Oil Price Volatility and the Asymmetric Response of Gasoline Prices to Oil Price Increases and Decreases", Energy Economics, 27(5), 708-730.

Sherman, J. and A. Weiss (2017), "On Fruitful and Futile Tests of the Relationship between Search and Price Dispersion", Economic Inquiry 55(4), 1898-1918.

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

Stigler, G.J., 1961, "The Economics of Information", *Journal of Political Economy* 69(3), pp. 213-225.

Stahl, D. O. (1989), "Oligopolistic Pricing with Sequential Search", American Economic Review 79(4), 700-712.

Tang, Z., M. Smith and A. Montgomery (2010), "The Impact of Shopbot Use of Prices and Price Dispersion: Evidence from Online Book Retailing", International Journal of Industrial Organization 28(6), 579-590.

Tappata, M. (2009), "Rockets and Feathers: Understanding Asymmetric Pricing", RAND Journal of Economics, 40(4), 673-687.

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

Verlinda, J.A. (2008), "Do Rockets Rise Faster and Feathers Fall Slower in an Atmosphere of Local Market Power? Evidence from the Retail Gasoline Market", Journal of Industrial Economics, 56(3), 581-612.

Walsh, P.P. and C. Whelan (1999), "Modelling Price Dispersion as an Outcome of Competition in the Irish Grocery Market", Journal of Industrial Economics 47(3), 325-343.

Yang, H. and L. Ye (2008), "Search with Learning: Understanding Asymmetric Price Adjustments", RAND Journal of Economics, 39(2), 547-564.

Zhao, Y. (2006), "Price Dispersion in the Grocery Market", Journal of Business 79(3), 1175-1192.

7 Appendix

There is a large theoretical literature on equilibrium consumer search and price dispersion which generally finds a non-monotonic relationship. In an early article, Diamond (1971) finds that if all consumers have positive search costs, prices necessarily rise to monopoly levels, price dispersion falls to zero and, absent any benefit to search in a world of identical prices, there is no search in equilibrium. Varian (1980) and Stahl (1989) develop models that allow some consumers to have zero or negative search costs and find in these cases that price dispersion and search do occur in equilibrium.

These and similar models imply a non-monotonic relationship between price dispersion and search. Following the Janssen et al. (2011) extension of the Stahl model, imagine there are N homogeneous goods firms competing for sets of informed and uniformed consumers. A fraction μ of consumers are informed and pay no cost to learn the entire distribution of prices. The remaining consumers are uninformed and must pay s per search. In the case of non-sequential search, the equilibrium price distribution is given by:

$$F(p, \mu, N, c, \overline{p}) = 1 - \left(\frac{1}{N} \frac{1-\mu}{\mu} \frac{\overline{p}-p}{p-c} \right)^{\frac{1}{N-1}} \tag{5}$$

where $p \in [\underline{p}, \overline{p}]$ is the price and c is marginal cost. The support of the distribution is given by

$$[\underline{p}, \overline{p}] = \left[c + \frac{(1-\mu)(\overline{p}-c)}{(1-\mu)+\mu N}, \ \min\left(c + \frac{s}{1-\alpha}, v \right) \right] \tag{6}$$

where v is the consumer value of the good,[29] and the parameter α is given by:

$$\alpha = \int_0^1 \frac{dx}{1 + \frac{\mu}{1-\mu} N x^{N-1}} \tag{7}$$

Pennerstorfer et al. (2016) show that the difference between the expected price and the minimum

[29] In the sequential search case, simply $\overline{p} = v$.

price in the market, a measure of price dispersion, is given by:

$$E(p - p_{\min}) = \int_{\underline{p}}^{\overline{p}} p\left[1 - N(1 - F(p))^{N-1}\right] dF(p) \quad (8)$$

$$= (\overline{p} - c)\left(\alpha - \frac{1-\mu}{\mu}(1-\alpha)\right) \quad (9)$$

It is easy to see that when $\mu = 0$ (there are no informed consumers), $\underline{p} = \overline{p} = v$ (i.e. monopoly prices) and price dispersion is zero, $E(p - p_{\min}) = 0$. Similarly when $\mu = 1$ (all consumers are informed), $\underline{p} = \overline{p} = c$ (i.e. perfectly competitive prices) and price dispersion is again zero. Combined with the Tappata (2009) result that $E(p - p_{\min})$ is strictly concave in μ, price dispersion and search intensity necessarily exhibit a non-monotonic relationship. When no consumers search, the model reproduces the Diamond result where all prices are equal and equal to monopoly prices. When all consumers search, the model reproduces perfect competition where prices are all equal and equal to marginal cost. Zero price dispersion results in these cases. For intermediate situations with partial search, there is positive price dispersion in equilibrium.

Table 1. Summary Statistics

	Num. Obs.	Mean	Std. Dev.	Minimum	Maximum
Price	6489	202.97	22.57	165.0	259.0
Regular Grade	2706	182.69	10.09	165.0	210.0
Midgrade	1860	205.25	11.26	169.0	239.0
Premium Grade	1923	229.30	13.29	195.0	259.0
Income	6489	47.007	11.936	28.936	62.345
Race/Ethnicity:					
White	6489	0.619	0.150	0.228	0.809
Black	6489	0.070	0.036	0.029	0.127
Hispanic	6489	0.275	0.124	0.125	0.728
Gender:					
Female	6489	0.514	0.016	0.457	0.536

Prices in cents per gallon. Income in thousands of 2016 dollars. Other variables are dichotomous.

Table 2. Rank Reversals

	All Gasoline	Regular	Midgrade	Premium
Unique Within-Region Station Pairs	766	758	484	514
Fraction Rank Reversals	0.73	0.63	0.50	0.41
Mean Rank Reversal Measure	0.23	0.22	0.25	0.24
Std. Dev. Of Rank Reversal Measure	0.14	0.13	0.12	0.13

Rank Reversal is the fraction of days that the most-often lower priced station of a given pair has the highest price.

Table 3. Price Dispersion Across Stations by Grade of Gasoline

Dep. Var.: Price Dispersion	(1)	(2)	(3)	(4)	(5)	(6)
	Standard Deviation	Standard Deviation	Standard Deviation	Max-Min Range	Max-Min Range	Max-Min Range
Midgrade	3.555***	3.489***	3.961***	8.797***	11.023***	10.873***
	(0.755)	(0.696)	(0.930)	(2.023)	(2.003)	(2.221)
Premium	4.990***	4.932***	4.729***	12.325***	14.317***	12.952***
	(1.235)	(1.125)	(1.209)	(3.175)	(2.942)	(3.162)
Station Count		3.389*	-0.017		0.619*	1.011
		(1.863)	(0.163)		(0.334)	(0.718)
Daily Indicator Variables	Y	Y	Y	Y	Y	Y
Region Indicator Variables	Y	Y	Y	Y	Y	Y
Only with All Grades Reported	N	N	Y	N	N	Y
Equality of Coefficients Test:	1.435	1.442	0.768	3.528	3.294	2.079
Midgrade=Premium	(0.974)	(0.986)	(0.744)	(2.525)	(2.261)	(2.112)
Linearity of Coefficients Test:	-2.120*	-2.047	-3.193***	-5.269	-7.730***	-8.795***
Premium=2*Midgrade	(1.230)	(1.283)	(1.064)	(3.296)	(3.098)	(2.860)
R-Squared	0.472	0.472	0.423	0.510	0.558	0.534
Num. Obs.	713	713	687	713	713	687

*** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Robust standard errors in parentheses.

Table 4. Price Dispersion Within Stations by Grade of Gasoline

Dep. Var.: Residual Dispersion	(7)	(8)	(9)	(10)	(11)	(12)
	Std. Dev. Residuals	Std. Dev. Residuals	Std. Dev. Residuals	Max-Min Residuals	Max-Min Residuals	Max-Min Residuals
Midgrade	1.384***	1.699***	0.985***	3.267***	5.562***	2.917***
	(0.337)	(0.363)	(0.253)	(1.212)	(1.016)	(0.854)
Premium	1.461***	1.743***	0.959***	3.532***	5.585***	2.615***
	(0.440)	(0.349)	(0.321)	(1.360)	(0.953)	(0.806)
Station Count		0.084	0.147**		0.627**	0.935**
		(0.051)	(0.061)		(0.191)	(0.216)
Daily Indicator Variables	Y	Y	Y	Y	Y	Y
Region Indicator Variables	Y	Y	Y	Y	Y	Y
Only with All Grades Reported	N	N	Y	N	N	Y
Equality of Coefficients Test:	0.077	0.044	-0.026	0.265	0.023	-0.303
Midgrade=Premium	(0.222)	(0.253)	(0.142)	(0.571)	(0.698)	(0.345)
Linearity of Coefficients Test:	-1.307***	-1.655***	-1.011***	-3.002***	-5.538***	-3.220***
Premium=2*Midgrade	(0.363)	(0.519)	(0.247)	(1.320)	(1.459)	(0.951)
R-Squared	0.227	0.264	0.235	0.262	0.458	0.517
Num. Obs.	713	713	687	713	713	687

*** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Robust standard errors in parentheses.

Table 5. Linearity of Coefficients Tests Relevant to Alternative Hypotheses

	(1),(7)	(2),(8)	(3),(9)	(4),(10)	(5),(11)	(6),(12)
	Standard Deviation	Standard Deviation	Standard Deviation	Max-Min Range	Max-Min Range	Max-Min Range
Testing Dispersion in Prices:	-2.120*	-2.047	-3.193***	-5.269	-7.730***	-8.795***
Premium=2*Midgrade	(1.230)	(1.283)	(1.064)	(3.296)	(3.098)	(2.860)
Testing Dispersion in Residuals:	-1.307***	-1.655***	-1.011***	-3.002***	-5.538***	-3.220***
Premium=2*Midgrade	(0.363)	(0.519)	(0.247)	(1.320)	(1.459)	(0.951)
Daily Indicator Variables	Y	Y	Y	Y	Y	Y
Region Indicator Variables	Y	Y	Y	Y	Y	Y
Only with All Grades Reported	N	N	Y	N	N	Y
Num. Obs.	713	713	687	713	713	687

*** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Robust standard errors in parentheses.

Table 6. By Region: A Test of Income Based Price Dispersion

Dep. Var.: Price Dispersion	(13)	(14)	(15)	(16)	(17)	(18)
	Std. Dev. Residuals East	Std. Dev. Residuals Central	Std. Dev. Residuals West	Max-Min Residuals East	Max-Min Residuals Central	Max-Min Residuals West
Midgrade	0.714***	0.967***	1.182***	2.200***	1.838***	4.113***
	(0.230)	(0.312)	(0.238)	(0.612)	(0.602)	(0.721)
Premium	0.791***	1.283***	0.862***	2.230***	2.631***	2.868***
	(0.243)	(0.293)	(0.234)	(0.630)	(0.587)	(0.669)
Station Count	-1.515**	-0.015	0.172**	0.784**	1.009**	1.028**
	(0.636)	(0.058)	(0.025)	(0.151)	(0.113)	(0.073)
Daily Indicator Variables	Y	Y	Y	Y	Y	Y
Region Indicator Variables	Y	Y	Y	Y	Y	Y
Only with All Grades Reported	Y	Y	Y	Y	Y	Y
Equality of Coefficients Test:	0.077	0.316	-0.320	0.030	0.793	-1.245
Midgrade=Premium	(0.231)	(0.330)	(0.236)	(0.631)	(0.761)	(0.953)
Linearity of Coefficients Test:	-0.637	-0.651	-1.502***	-2.169*	-1.045	-5.359***
Premium=2*Midgrade	(0.392)	(0.571)	(0.408)	(1.117)	(1.302)	(1.631)
R-Squared	0.227	0.240	0.336	0.394	0.484	0.540
Num. Obs.	201	192	294	201	192	294

*** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Robust standard errors in parentheses.

Table 7. A Direct Test of Income Based Price Dispersion

Dep. Var.: Price Dispersion	(19)	(20)	(21)	(22)	(23)	(24)
	Std. Dev. Prices	Std. Dev. Prices	Std. Dev. Prices	Std. Dev. Residuals	Std. Dev. Residuals	Std. Dev. Residuals
Midgrade * Income			0.066**			0.032*
			(0.029)			(0.018)
Premium * Income			0.085			0.024
			(0.052)			(0.020)
Midgrade	3.961***	3.961***	1.029	0.985***	0.985***	-0.439
	(0.922)	(0.419)	(1.369)	(0.243)	(0.134)	(0.722)
Premium	4.729***	4.729***	0.963	0.959***	0.959***	-0.119
	(1.228)	(0.546)	(2.095)	(0.307)	(0.155)	(0.806)
Station Count	-0.228	0.068	0.068	0.067	0.121**	0.121**
	(0.210)	(0.071)	(0.079)	(0.100)	(0.046)	(0.047)
Income	0.151	1.098**	1.048**	0.060	0.390	0.371
	(0.276)	(0.436)	(0.526)	(0.117)	(0.466)	(0.591)
Black		0.281**	0.281*		0.095	0.095
		(0.132)	(0.163)		(0.141)	(0.175)
Hispanic		0.046	0.046		0.020	0.020
		(0.029)	(0.036)		(0.032)	(0.039)
Female		-0.067**	-0.067*		0.013	0.013
		(0.028)	(0.038)		(0.029)	(0.036)
Daily Indicator Variables	Y	Y	Y	Y	Y	Y
Region Indicator Variables	Y	Y	Y	Y	Y	Y
Only with All Grades Reported	Y	Y	Y	Y	Y	Y
Equality of Coefficients Test: Midgrade=Premium						
Linearity of Coefficients Test: Premium=2*Midgrade						
R-Squared	0.443	0.508	0.516	0.251	0.292	0.296
Num. Obs.	687	687	687	687	687	687

*** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Robust standard errors in parentheses.

Table 8. Additional Robustness Checks

	(25) Interquartile Range	(26) Interquartile Range	(27) Interquartile Range	(28) Std. Dev. Subregions	(29) Max-Min Subregions	(30) Interquartile Subregions
Midgrade	1.947***	2.239***	1.645***	0.985***	1.645***	2.917***
	(0.608)	(0.617)	(0.485)	(0.149)	(0.267)	(0.397)
Premium	2.037***	2.298***	1.761***	0.959***	1.761***	2.615***
	(0.696)	(0.585)	(0.619)	(0.147)	(0.276)	(0.379)
Station Count		0.093	0.168*			
		(0.079)	(0.097)			
Daily Indicator Variables	Y	Y	Y	Y	Y	Y
Region Indicator Variables	Y	Y	Y	Y	Y	Y
Subregion Indicator Variables	N	N	N	Y	Y	Y
Only with All Grades Reported	N	N	Y	Y	Y	Y
Equality of Coefficients Test:	0.089	0.059	0.116	-0.026	-0.303	0.116
Midgrade=Premium	(0.298)	(0.332)	(0.276)	(0.138)	(0.367)	(0.276)
Linearity of Coefficients Test:	-1.858***	-2.180***	-1.528***	-1.011***	-3.220***	-1.528***
Premium=2*Midgrade	(0.659)	(0.799)	(0.489)	(0.247)	(0.664)	(0.509)
R-Squared	0.160	0.170	0.158	0.297	0.232	0.543
Num. Obs.	713	713	687	687	687	687

*** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. Robust standard errors in parentheses.