Learning Competitors' Identities from the Timing of Pricing Decisions: An Application to Retail Gasoline

Richard Grice*

2024

Abstract

Digitisation is increasing the frequency at which retailers can adjust their prices, increasing the importance of being able to identify who it is they are monitoring and responding to when they make these adjustments. In this paper I develop a general approach for solving this identification problem. The foundation of the method is a continuous-time, discrete-game model of a retailer's strategic decision-making. I derive from this structural foundation a 'reduced-form' expression for the hazard rate of a retailer's price adjustments as a function of their competitors' prices. This hazard rate takes the form of a generalised linear model. Consequently, its estimation can be combined with $l1$-norm regularisation to exploit the consistent model-selection properties of the LASSO and identify the competitors whose prices define the payoff-relevant states of the retailer. Further, in implementing the method we can exploit the simplicity and efficiency of standard, highly-optimised machine learning packages. I demonstrate the method by using 30 months of minute-to-minute price data to estimate the competitive ties between gasoline stations on the southern periphery of Sydney, Australia. I find these ties connect all the stations into a single networked competition structure, which spans several geographical areas that have previously been treated as separate markets by the local competition authority in its investigations and merger authorisations.

*Department of Marketing, NUS Business School

This project was developed from the second chapter of my doctoral dissertation, and I am are grateful to my advisers Maria Ana Vitorino and Paulo Albuquerque for their guidance and support. I also thank Ahmed Gueciouеur and Yufeng Huang, seminar participants at Cátolica Lisbon, HEC Paris, INSEAD, the National University of Singapore, Universidad de Chile, University College London, and in the Hong Kong Quant Marketing and EMAC Pricing SIG virtual seminars, and conference participants at the 9th NSF Network Science & Economics and 2024 Marketing Dynamics Conferences for their helpful comments. Finally, this project was begun while I was a Visiting Researcher at Endgame Economics Pty Ltd., and I thank Oliver Nunn and the Endgame team for their generous support.

1 Introduction

It is becoming clear that it is not just consumers who limit their consideration when faced with broad product assortments. In product categories with large numbers of substitutable products or vendors, managers also restrict their attention to narrow subsets of potential competitors (e.g, Israeli and Anderson, 2023). How then can we identify who it is that managers actually treat as their competitors, and who thereby influence their strategic decisions?

The prevailing approach to this problem is to estimate a demand system and assume substitutes compete. However this assumed equivalence may not hold. For one, accurately estimating substitution patterns between a large number of products is itself a difficult marketing research problem (e.g, J. Q. Li et al., 2015). And even if managers can accurately identify the products or vendors their customers consider substitutes, they may still restrict their attention to the closest or most threatening of these. Thus there can exist a wedge between the substitution patterns captured by a demand system and who managers are actually considering and treating as their competitors. To the extent that such a wedge does exist, it is the latter that influences and constrains managers' strategic decisions.

Therefore, in this paper, I propose a method for identifying who it is managers are actually treating as their competitors when they are making pricing decisions. The crux of the method is to exploit information encoded in the *timing* of managers' price-adjustment decisions about the competitors to whom they are responding when they adjust their prices. The primary advantage of the method is that – to the extent managers restrict their attention when setting prices – it identifies who managers *do treat* as their competitors, not who they *ought to treat* as competitors given the substitution patterns of their customers. A secondary advantage is that the method mainly requires high-frequency price data for all potential competitors. Prices are publicly observed and thus often more easily obtained than the sales data that would be required for estimating a demand system, especially in online markets.

The method I propose combines methods developed in computational neuroscience for the analogous problem of identifying connections between neurons from the timing of their firings (e.g, S. Kim et al., 2011; Sheikhattar et al., 2018; Truccolo et al., 2005) with modern model-selection tools from machine learning (e.g, Bühlmann and van de Geer, 2011). In addition, I found the method in an economic model of the timing of pricing decisions to clarify the behavioural assumptions underpinning econometric identification. This theoretical foundation is provided by a structural model of retail competition as a discrete game in continuous time (Arcidiacono et al., 2016; Doraszelski & Judd, 2012). In such a model, decision points for a manager randomly arrive in continuous time, representing the occasions at which the manager attends to the current optimality of their prices vis-à-vis their competitors. The manager then decides whether to adjust prices at these points based upon a Markovian strategy (Ericson &

Pakes, 1995). Such a decision process makes price adjustments a conditional Markov process, with price adjustments timings characterised by a point process conditioning on the current prices of competitors. The hazard rate of this point process is the objective of estimation, and the connection to the computational neuroscience methods.

Following the computational neuroscience literature, I derive a discrete time expression for the price-adjustment hazard rate, and show that the likelihood of observing price adjustments in discrete intervals of time under this hazard rate is equivalent to the likelihood of a generalised linear model ('GLM') with Bernoulli distribution and logistic link function (McCullagh & Nelder, 1989). That is, I derive a reduced-form GLM expression for the hazard rate that can be estimated using optimised, off-the-shelf machine learning packages. This derivation reduces the problem of identifying the competitors whose prices are conditioning the price-adjustment point process to a model selection problem. Consequently, by incorporating $l1$-norm regularisation into the GLM estimation, we can exploit the consistent model selection properties of the least absolute shrinkage and selection operator ('LASSO') (Meinshausen & Bühlmann, 2006; Zhao & Yu, 2006) to identify the competitors influencing a manager's price-adjustment decisions.

Having derived this method for identifying competitors, I demonstrate it with an application to gasoline stations in Australia. This setting is well suited to testing the method for at least three reasons. First, the setting is relatively simple, with vendors differentiated primarily by their location and brand. This simplicity makes it easier to select potential competitors to feed into the estimation, and to 'sense check' results. Second, this is a setting that frequently sees antitrust reviews of proposed acquisitions and mergers, which would benefit from an accurate characterisation of the competitive relationships mediating price competition. To take an example of one such review, in 2017 the Australian Competition and Consumer Commission (i.e, the 'ACCC', Australia's antitrust regulator) blocked a major acquisition of gasoline stations on the basis it would lessen competition by concentrating local markets, defined by a 3km radius around acquiree stations in urban areas and 10km in rural area (ACCC, 2018b). This characterisation of competition presupposes every station within 3km is being treated as a competitor by the acquiree and constraining their pricing decisions, and that their pricing decisions are entirely independent of any competitive constraints being imposed by the broader network of stations beyond 3km. Such a presupposition is liable to bias conclusions about the extent and influence competition, and thus a more accurate characterisation of the setting's competitive structure would be valuable.

And third, retail gasoline is a setting with several rich, publicly available pricing datasets (e.g, Byrne et al., 2018). The data I use in this application comes from the FuelCheck price monitoring program introduced in 2016 in the Australian state of New South Wales ('NSW'), which includes Sydney. The data record both the price of gasoline at the stations of all the competing brands in NSW, and the time to the minute at which each station adjusts its price.[1] Thus the data are high-frequency enough for estimating the hazard rate of price adjustments. I combine this retail price data with data recording the wholesale price of gasoline available in Sydney – which provide a measure of movements in stations' marginal costs and allows me to separate competitive price responses from price adjustments due to cost movements – and measures of the pairwise distance between all stations through the road network.

Applying the method to the FuelCheck price data identifies the sets of competitors whose prices stations in NSW are monitoring and responding to when they adjust their own prices. These identified competitor sets uncover several novel features of retail competition in general, and gasoline station competition in particular. First, stations do indeed monitor and respond to a narrow set of competitors – on average just eight competitors, and at most only 16.[2] Particularly interesting is the finding that this heterogeneity in the numbers and identities of monitored competitors exists even for adjacent stations. Second, in addition to this heterogeneity in the numbers of competitors, there is also asymmetry. That is, one station monitoring and responding to the prices of another station does not imply the reverse. And finally, partial overlaps of stations' heterogeneous competitor sets connect all the stations into a single market with a networked competitive structure. Moreover, this network spans a wide geographic area, indirectly connecting stations in Sydney and stations in regional NSW into the same market. Consequently, the competitive structure looks markedly different to the distinct local markets defined in the above antitrust example, and stations' prices are indirectly affected by a much broader network of competitors than just those with whom they directly compete. Further, applying appropriate network clustering methods produces a collection of market-like clusters that are not ex ante obvious, and very different from the ACCC's 3km and 10km definitions would have produced.

These findings have deep implications. Most notably, they imply that a retail product or vendor's price is affected by a broader set of competitors than just those revealed by pairwise measures of price monitoring, substitution or demand diversion. The networked structure of competitive interactions revealed by these findings is qualitatively different from the characterisation assumed – even if implicitly so – in most analyses of retail competition. That is, the characterisation assuming a distinct product category or geographic market, defined by a set of products or vendors who all directly compete with each other, and only with each

[1]This comprehensiveness – which allows me to study station-level price responses – provides a substantial improvement over data used in previous studies of retail gasoline competition and price dynamics. For example, the recent studies of Byrne (2019), Foros and Steen (2013), and Remer (2012) all use data with daily, station-level observations, which is itself an improvement relative to earlier studies.

[2]This finding is consistent with earlier heuristic findings about price monitoring by gasoline stations (Atkinson et al., 2009), and of recent findings about the price monitoring of competing products by a large online retailer (Israeli & Anderson, 2023).

other. Thus, this implication cautions against the accuracy of market shares or concentration as measures of the competitiveness – or otherwise – of retail markets, and motivates the need for formal analyses of networked markets. Further, this implication suggests the dynamics of price competition could be much more complex – and result in more complex equilibrium pricing patterns – than allowed for by a Bertrand-like characterisation of price competition. More research is needed to explore these possibilities.

The rest of this paper proceeds as follows. In Section 2 I summarise related literature and highlight the contributions made by this paper. I then introduce the empirical context and data used in my application in Section 3, which also provides a concrete example when it comes to modelling the timing of pricing decisions. I introduce the theoretical foundation in Section 4 and derive from it a method for identifying the competitors retailers are monitoring and responding to when adjusting prices. In Section 5 I apply this method to my data and analyse the results. I conclude in Section 6.

2 Related literature & contributions

Methodologically this paper contributes to the diverse literature devising methods for estimating the competitive structure of markets. In then applying the proposed method to the retail gasoline industry and uncovering a networked competitive structure, the paper also contributes to the much younger literature conceptualising competition in terms of networks.

2.1 Competition structure

There is a sizeable literature in marketing proposing approaches for identifying or characterising market structure using data recording consumer actions like purchases and brand switching (Erdem, 1996; Kannan & Sanchez, 1994; Novak, 1993; Novak & Stangor, 1987; Shugan, 1987; Urban et al., 1984), click streams and search paths (J. B. Kim et al., 2011; Moe, 2006), consideration (Ringel & Skiera, 2016), or brand engagement (Hyoryung et al., 2017; Lee & Bradlow, 2011; Netzer et al., 2012; Yang et al., 2021). My notion of competitive structure differs from much of this literature. In general, this literature is focused on identifying substitution patterns on the demand side of a market to inform brand and product managers about which entities they should consider their competitors. This paper is focused on identifying the entities that retailers already believe to be their competitors, and thus whose prices the retailers are monitoring and responding to when they make strategic pricing decisions. This set of entities retailers consider their competitors could differ from the competitors implied by demand substitution for at least two reasons. One, retailers may not be sufficiently aware of demand drivers to know the identities of all the entities to whom their customers might substitute. Indeed, this possibility motivates many of the prescriptive methods introduced in the marketing literature on market structure. And two, retailers may not monitor and respond to the prices of all the entities with whom their products are substitutable. Instead, they may focus on only those entities who represent a sufficiently big threat to their customer base, or who act as price leaders within their market, making the entity's prices a 'sufficient statistic' for maintaining competitive prices. To the extent that this last reason is true, I am interested in identifying the competitors to whom retailers respond with their pricing decisions, and thus the competitive structure mediating the propagation of prices across a market.

The method I propose estimates the competitor sets of retailers – and thus the structure of retail competition – by estimating the hazard rate of competitive price adjustments from high-frequency price data and exploiting the consistent model selection properties of the LASSO. This approach is inspired by methods proposed for the analogous problem of estimating connections between neurons from spike train data (e.g, S. Kim et al., 2011; Sheikhattar et al., 2018; Truccolo et al., 2005). The most similar method for estimating competitive structure of which I am aware is that proposed by J. Li et al. (2018) for estimating price competition between hotels from price and click-stream data. Their approach also uses model selection methods to learn the identities of the competitors to which hotels are responding when they adjust their rates, but differs from the approach I propose by focusing on price levels rather than the timing of price adjustments. The demand-side data requirements of the approach in J. Li et al., 2018 also make it difficult to apply the approach to other industry contexts, whereas the approach I propose primarily uses high-frequency price data, which is common in many online retail contexts.

In applying my proposed method to identify the competitive structure of a gasoline market from price data in the absence of demand-side data, I follow the precedent of Chan et al. (2007), though they take a substantially different approach. Using data from Singaporean gasoline stations, they specify and estimate an entry model of stations' locations – as decided by a welfare maximising Singaporean policy maker – followed by a Nash-in-prices model of gasoline brand competition.[3] They are motivated in the latter of these model specifications by the institutional fact that station prices are set by a brand-level pricing manager. Stations prices for the major brands in my setting are also set by brand-level pricing managers (I discuss these institutional details further in Section 3.1). However, because I observe the to-the-minute timing of stations' price adjustments, as compared to the three-wave repeated cross sections collected by Chan et al. (2007), I am able to specify a model of the timing of price adjustment decisions at the station level, rather than having to rely on a static equilibrium assumption

[3]The spatial competition model of gasoline brand pricing decisions is an adaptation of Thomadsen (2005), who introduces an approach for estimating the structure of price competition between retailers from price and location data by assuming observed prices are the result of a static equilibrium in Nash-in-prices competition.

for identification of competitive structure. This distinction in the level at which competitive decisions are focused also connects this paper to the semi-parametric method of Pinkse et al. (2002) for distinguishing between global and local decisions, which is demonstrated using data from American wholesale gasoline markets.

2.2 Networked pricing games

Finally, my application finds that retail gasoline competition has a network structure, with directed ties between station-nodes representing competitive monitoring and responding to pricing decisions. This finding contributes to the nascent literature modelling price competition as a game played over a network. While games-on-networks are an established tool for modelling a range of market interactions – including consumer word-of-mouth (Campbell, 2013), R&D collaborations (Goyal & Moraga-González, 2001), and B2B quantity competition (Bimpikis et al., 2019) – their use to model price competition is relatively new. Ushchev and Zenou (2018) introduce the first such model, though my findings are more comparable to the model introduced in Grice (2023), which micro-founds networked price competition between differentiated products with a heterogeneous-consideration model of consumer demand. The current paper adds to this nascent literature by providing empirical evidence for these theorised networked competitive pricing interactions, as well as a method for recovering this latent network from data.

3 Institutional background and data

3.1 The NSW retail gasoline industry

To help fix ideas when modelling competitive pricing decisions, and then provide a context in which to demonstrate my proposed method, I focus on the retail gasoline industry in the Australian state of New South Wales ('NSW'), of which Sydney is the capital. At 810 thousand square-kilometres, NSW is a large area over which to supply gasoline – slightly bigger than the combined US states of Texas and Oklahoma, and slightly smaller than the combined European countries of Poland and Ukraine (including Crimea). However 75% of NSW's 8 million residents live in Sydney and the surrounding extra-urban area extending north to south between the satellite cities of Newcastle and Wollongong.

Fuel retailing Fuel is sold to NSW motorists at roadside service stations. Most stations sell a range of fuel types. These types generally include unleaded petroleum ('gasoline') of varying octane ratings, diesel, and potentially liquid petroleum gas ('LPG'). The octane rating of petroleum refers to its ability to withstand compression before combusting, which can reduce wear and increase the potential power output of an engine. Thus petroleum is considered more premium the higher its octane rating. The lowest octane petroleum offered by stations has a rating of 91 ('U91'), and the premium petroleums offered have ratings 95 and 98 ('P95' and 'P98'). In addition, many stations sell a 90% petroleum - 10% ethanol blend ('E10'), which is considered a close substitute for U91.[4] More generally, a motorist can substitute between any of the petroleums at the point of purchase, but cannot substitute between petroleum, diesel and LPG, which all require different engine technologies to be used as fuel.

The majority of service stations are branded under one of eight major brands operating in NSW. These major brands are the oil majors BP and Shell, Caltex, which is a retail brand of Chevron, the major Australian supermarkets Woolworths and Coles, the convenience store 7-Eleven, and the domestic fuel brands United and Metro. Branded stations may be owned and operated by their brand ('company-owned, company-operated' or 'COCO' stations), or owned and operated by a second party under a franchise or other licensing agreement with the brand ('dealer-owned, dealer-operated' or 'DODO' stations). Whether a particular station is a COCO or DODO is generally not observable.[5] In addition, stations' brands may differ from that of the fuel they sell. In particular, over the period studied in this paper, Woolworths stations sell Caltex-branded fuel and Coles stations sell Shell-branded fuels. Finally, the vast majority of stations operate a convenience store alongside their fuel pumps, and a minority offer other services, such as garage services. Note that these convenience offerings are distinct to the station and its brand – Australia does not have the population density to support larger motorway services with separately-branded fast food restaurants as are common in Europe and North America. Figure 1 depicts a typical highway service station in NSW.

Fuel wholesaling The wholesale supply of fuel to stations in NSW is from import terminals, which are located in Sydney and Newcastle and operated by half a dozen competing companies (ACIL-Tasman, 2009).[6] Fuel is supplied in bulk from these terminals to any certified purchaser at a listed terminal gate price ('TGP'), which is updated daily and closely tracks the Singapore

[4]The sale of E10 by service stations has been mandated in NSW since 2017. Service stations fall under the mandate if they offer three or more types of petroleum or diesel, or if they sell a sizeable volume of either petroleum or diesel. The mandate has applied to these stations since either January or April 2017, depending on the station's owner. See www.nsw.gov.au/driving-boating-and-transport/e10-fuel/e10-history. Roughly 1,400 stations in the FuelCheck dataset sell E10.

[5]One prominent exception is BP's COCO stations, which in 2017 were identified in a public filing during a review of a proposed merger (ACCC, 2017). While the distinction between COCOs and DODOs does not feature in this paper, it is part of a plan for future work to analyse the impact of this distinction on price setting.

[6]Stations located in the north of NSW may also be supplied from import terminals located across the NSW-Queensland border in Brisbane. NSW has supplied all its gasoline demand with imports since the Clyde and Kurnell oil refineries in Sydney were closed and converted to import terminals in 2013 and 2014, which preceded the period covered by the FuelCheck dataset. Nationally Australia has only four refineries, which supply about 65% of the country's gasoline demand. The remainder is supplied by imports, primarily from refineries across Asia via Singapore.

Figure 1: Coles Express, Wagga Wagga The Coles Express service station on Sturt Highway in Wagga Wagga, a regional centre south-west of Sydney. The station is typical in its branding, price display and convenience offering.

MOGAS 95 petroleum price index. While some wholesale customers will purchase bulk fuel under long-term supply contracts, regulatory analysis of transactions shows paid wholesale prices vary very little from TGPs (AIP, 2019). Consequently, movements in TGPs provide a good proxy for movements in stations' marginal cost of fuel, which differs from the TGP primarily due to the fixed cost of transporting fuel between terminals and stations.

Station price setting Service stations in NSW are free to change their price at any time. They display their current fuel prices both at the pump and on a large roadside sign at the station entrance.[7] Conversation with industry participants reveals the major brands employ brand price managers or pricing teams to centrally set the prices of their COCO stations. These price managers are responsible for setting prices at all stations in the state, and in some cases nationally. They do so for each station by reference to a real-time data feed displaying the current prices of a set of relevant, nearby stations. They refer to these stations as *markers*. Conversations with industry participants also suggests the number of markers for a station can range from one up to roughly a dozen, with the managers for higher-priced brands tending to monitor fewer markers. Finally, conversations with industry participants reveals one major brand had partially automated their price setting during the period used in our application, though not the identity of that brand.

[7]During the period covered by our data, Australia did not have highway price noticeboards of the sort studied by Chintagunta and Rossi (2016, 2018).

Price monitoring programs A notable feature of NSW's retail gasoline industry is its price monitoring program, FuelCheck. Launched by the NSW government in July 2016, the program requires every station in the state to upload price changes to its website as they are made. The program then makes these prices available to consumers in real time via a website and associated smartphone application. Consumers are encouraged to report discrepancies between the prices in FuelCheck and at stations, and each violation attracts a fine. This crowd-sourced monitoring provides some guarantee of the timeliness of historical price changes in the FuelCheck database, which is publicly available to researchers.[8] A further and stronger guarantee is provided by the fact that the major brands shared their prices with each other through the third-party data aggregator Informed Sources prior to the introduction of FuelCheck, and thus already had in place the systems to automatically upload changes in their stations' prices.[9] I use the FuelCheck data in this research, and describe it in detail in Section 3.2. An extended discussion of Australia's price monitoring programs and the data they generate can be found in Byrne et al. (2018).

2019 industry developments The NSW retail gasoline industry experienced a period of relative stability from the introduction of FuelCheck until two major acquisitions in 2019. In early March 2019, responsibility for setting the retail prices of fuel sold at Coles' stations was transferred to Viva Energy, the supplier of Shell-branded fuels to Coles' stations. Under the arrangement between Coles and Viva, Viva supplied, branded and priced fuel, while Coles continues to operate the stations and manage their convenience offerings. Separately, in April 2019, Woolworths sold its stations to the Euro Garages Group, a fuel and convenience retailer with operations across Europe and North America (ACCC, 2019b).[10] It is because of these changes in the parties responsible for price setting at a large fraction of NSW gasoline stations that I restrict the sample period for my application in Section 5 to the period preceding March 2019.[11]

[8] See the NSW data portal at https://data.nsw.gov.au/data/dataset/fuel-check.

[9] This data sharing through Informed Sources in part motivated the introduction of FuelCheck. In 2014, the Australian Competition and Consumer Commission (the 'ACCC') commenced legal proceedings against Informed Sources and the major brands alleging that their sharing of price information contravened Australia's competition laws. The ACCC settled these proceedings in late 2015 with the agreement that the brands and Informed Sources would not share prices unless they make the information available at the same time to consumers (for free) and third parties (ACCC, 2015). The NSW Government then launched FuelCheck in 2016.

[10] This sale followed a failed attempt in 2017 to sell its stations to BP. Australia's antitrust regulator, the Australian Competition and Consumer Commission (the 'ACCC'), opposed the merger on the grounds it would substantially lessen competition in retail gasoline markets across Australia, and the brands abandoned the sale.

[11] The average price of gasoline at Coles stations decreased in the months following the change in price setting responsibility, and increased at Woolworths stations (ACCC, 2019a). In future work I hope to explore the impact of the acquisitions on station pricing suggested by these average price changes.

3.2 Data

Price data The primary data for this project comes from the FuelCheck program. As well as requiring stations to upload price changes as they are made, the program provides an API via which station and pricing managers can automatically upload price changes. This facility provides confidence the historical price data in the FueCheck database is accurate, at least for stations managed by the major brands. In particular, it provides the pricing managers for these brands the ability to simultaneously upload to FuelCheck the price data they were previously sharing with a third-party data aggregator prior to July 2016.[12]

The historical prices in the FuelCheck database take the form of station-fuel observations recording the time and price for every price adjustment by a NSW service station. I focus on price adjustments for U91. Because U91 and E10 are close to perfect substitutes, I include E10 price adjustments for those stations selling E10 but not U91.[13] I use a 30 month sample period from 1 September 2016 to 28 February 2019. This sample period is chosen to allow for a "burn-in" period in which stations adapt to the practice of uploading prices to FuelCheck after its launch in July 2016, and to precede the major industry changes beginning in March 2019.

Station location data The FuelCheck historical price database also includes a station's name, brand and street address for each price adjustment observation. However it does not include the latitude and longitude associated with stations' street addresses. I therefore use the API offered by the FuelCheck program for consumer applications to retrieve the latitude and longitude of each station's street address. I then use these location coordinates to create unique station IDs with which to associate all of a station's price adjustments.The resulting dataset contains 2,277 stations, of which 2,135 sell U91 and/or E10.[14] Armed with the coordinates of each station's location, I am also able to calculate the distance between each pair of them. Using the OpeRouteService API, I calculate the distance between each pair of the 2,135 stations in terms of both the driving distance through the road network between

[12] In August 2014, the 'ACCC' commenced legal proceedings against the data aggregator Informed Sources and five major gasoline station brands – BP, Caltex, Coles, Woolworths and 7-Eleven – alleging they had contravened Australia's competition laws by collecting near real-time price data from the brand's stations and then sharing that data amongst the brand's pricing managers. In early December 2015, Coles settled the action by agreeing to terminate its arrangement with Informed Sources. The action against the remaining four brands was settled by Australia's Federal Court later in December 2015 (ACCC, 2015). The four brands continued supplying price data to Informed Sources, but all parties agreed to also make that data available to consumers.

[13] Stations selling both fuels rarely adjust the price of one without adjusting the price of the other.

[14] I code a station as selling a fuel if there are at least 10 recorded price adjustments for that station-fuel combination during the sample period. I apply this filter to remove data artefacts. While the filter may miscode some remote stations who rarely adjust prices as not selling unleaded fuels, such remote stations do not feature in my application.

them, and the duration of this drive.[15]

Cost data Finally, I obtain TGP data from the Australian Institute of Petroleum ('AIP') to serve as a proxy for stations' cost of gasoline. This data records the daily average TGP for gasoline available at terminals in Sydney. It averages the TGPs posted by the companies managing Sydney terminals, which are collected and collated by the AIP every weekday morning. We match this TGP data to the station price data by date, designating 6am as the time of day at which wholesale fuel becomes available at that day's TGP.

3.3 Retail gasoline price cycles

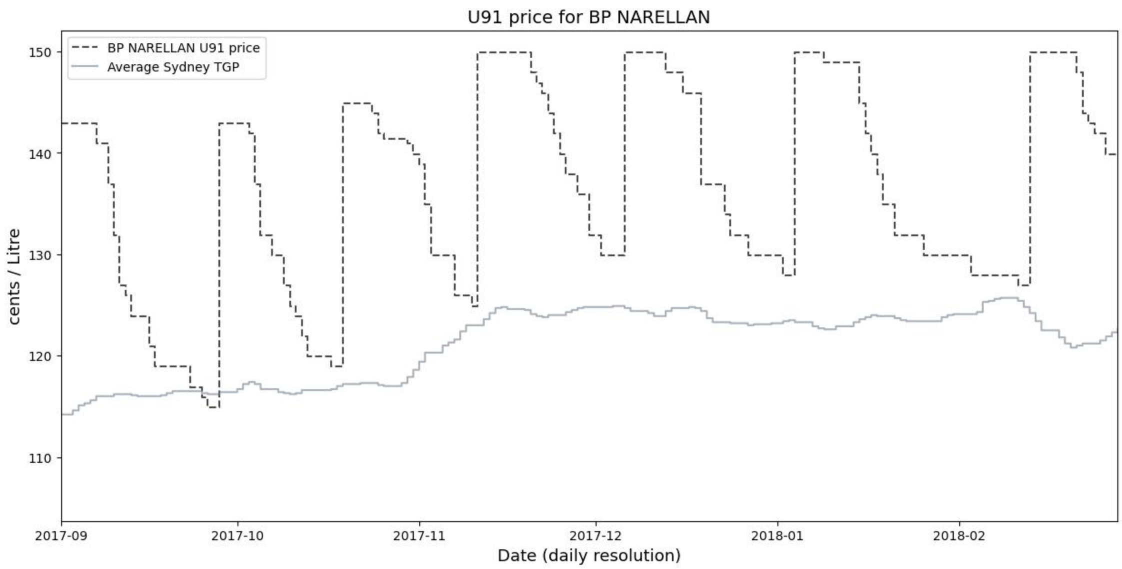
An additional feature of NSW's retail gasoline industry of relevance to my application is gasoline price cycles. These are coordinated cycles in retail prices characterised by frequent decreases in retail prices punctuated by substantial, city-wide increases. A handful of markets globally exhibit gasoline price cycles, and they have been the subject of much research and regulatory attention.[16] Their relevance to my application is that they generate substantial, endogenous variation in retail prices – as competition between stations brings prices close to costs, a substantial increase occurs that restarts the process of competitive price adjustment. In this section I use the FuelCheck data to present the key, relevant features of these price cycles.

Retail vs. wholesale dynamics: An example of these price cycles is depicted in Figure 2. The dashed black line plots the U91 price of a station in the outer Sydney suburb of Narellan. The solid grey line plots the average TGP in Sydney. Roughly every month, when the station's price approaches or reaches the TGP, the station increases its price by 20 to 30 cents per litre within a single day. This action is referred to in the industry as a 'restoration' of the station's price. Following the restoration, the station maintains its price at the higher level for up to a week, before adjusting its price downwards every few days until the next restoration. This longer period between the peak and trough of the cycle is referred to as the 'discounting' or 'undercutting' phase of the cycle.

[15] See the OpenRouteService documentation at https://openrouteservice.org/.

[16] In particular, gasoline price cycles have recently been noted in Australia's five major cities and their surrounding satellite cities; the Canadian cities Calgary, Montreal and Quebec; the US cities Chicago and Indianapolis; and in some locations in Norway (ACCC, 2018a). Researchers have also previously studied price cycles in Australian cities (Byrne & de Roos, 2019; Wang, 2009), multiple cities in Canada (Atkinson, 2009; Eckert & West, 2004; Noel, 2007a, 2007b), the Midwestern states of the US (Doyle et al., 2010; Lewis, 2009; Zimmerman et al., 2013) and, in the early 1970s, California (Castanias & Johnson, 1993), and Norway (Foros & Steen, 2013)

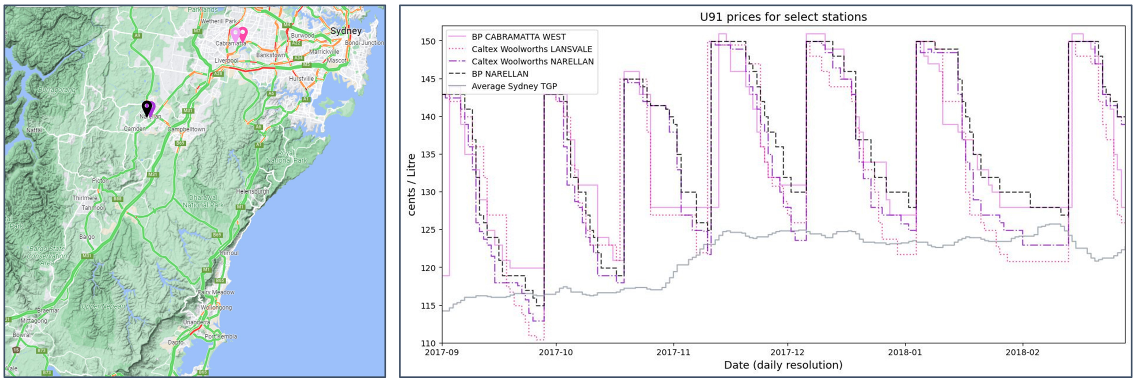
Figure 2: Gasoline price cycles Six months of U91 prices from September 2017 through February 2018 for the BP station located in Narellan, an outer-suburb of Sydney.



Coordination across stations: The coordination of price cycles across stations can be seen in Figure 3. The figure plots the U91 prices of four stations – the BP station from Figure 2 and a nearby Woolworths station, and two other neighbouring BP and Woolworths stations half an hour's drive north in Western Sydney. All four stations have highly correlated prices despite their dispersed locations. They frequently restore their prices on the same day, and adjust prices downward at a similar rate during the undercutting phase. This coordination in the timing of restorations and undercutting phases is representative of stations in Sydney, which all restore their prices within several days of each other.

Figure 3 also shows how station prices are more closely coordinated for neighbouring stations, even if all four stations have prices following the same cyclical trend. The BP and Woolworths stations in Narellan – whose prices are plotted in black and dark violet – restore their prices on the same day and to the same level in each of the plotted restorations. They also adjust their prices at very similar times throughout the undercutting phase, and maintain a relatively stable margin between their prices, though the Woolworths station discounts its price more aggressively at the trough of the cycle. Compare this intra-location coordination to the similarity of each of these two stations' prices to their corresponding brand station in Western Sydney. For both brands, neither the timing of price adjustments nor the stability of the cross-station margin is as coordinated for stations within brand as for stations within location. Thus the coordination through the undercutting phase is more likely the result of competitive price monitoring and response than any coordinated price setting by a common

Figure 3: Coordination of price cycles (Right) Six months of U91 prices from September 2017 through February 2018 for four stations with price cycles located in Sydney. (Left) Locations of stations. Two stations are located in the south-western suburb of Narellan, including the BP station from Figure 2. The other two stations are located in the Cabramatta area of West Sydney. Both locations include a station of the BP and Woolworths brands.



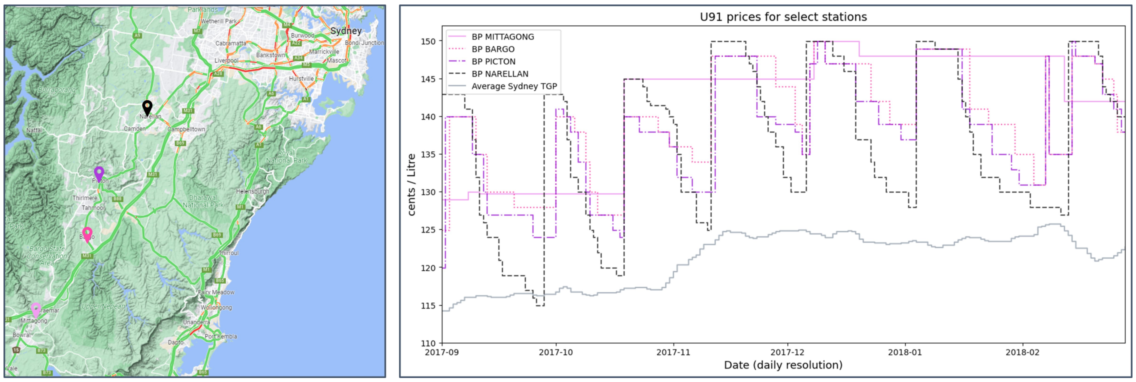
price manager.

Cross-sectional variation: Not all stations' prices cycle with the same robustness as the prices of the stations in Figure 3. Figure 4 illustrates the variation in stations' price dynamics across NSW. Six months of U91 prices are plotted for four BP stations south of Sydney. The first is the BP station in Narellan, whose cyclical prices are the focus of Figure 2. These prices are compared to those of the BP stations 25 minutes drive south along the Old Hume Highway in Picton, 40 minutes drive south in Bargo, and an hour drive south in Mittagong.[17] The prices in Figure 4 illustrate the primary feature by which price dynamics vary across stations: price stickiness. While the stations in Picton and Bargo adjust their prices downwards during the undercutting phase of the price cycle, they do not do so with the frequency of the station in Narellan. As result, they do not discount their prices to as thin a margin at the trough of the cycle before the next restoration.

The BP station in Mittagong displays the sticky-price extreme of cost-based or constant-markup pricing. It rarely adjusts its price –on average three times every two months. However, when it does, it often raises its price at the same time as a restoration by the cycling stations. This coordination implies its pricing is not entirely independent of that of the other stations. Further support for this implication can also be seen in the Mittagong station raising its price during a restoration to the level to which the other stations are restoring their prices, as in

[17]The inner regional towns of Picton and Bargo, with populations in 2016 of 4.8 and 4.4 thousand people, are part of a cluster of towns in the MacArthur region south-west of Sydney. Mittagong, with a 2016 population of six thousand, is located in the Southern Highlands region of NSW in Australia's Great Dividing Range. All four of these stations are included in the sample of stations used in my application in Section 5.

Figure 4: Variability in price dynamics (Right) Six months of U91 prices from September 2017 through February 2018 for four BP stations located south of Sydney. (Left) Locations of stations, from the BP station from Figure 2 in the outer-Sydney suburb of Narellan, down through the inner regional towns of Picton and Bargo on the Old Hume Highway, to the Inner regional town of Mittagong in the Southern Highlands.



December 2017, and in the Bargo and Picton stations temporarily increasing their prices to the level of the Mittagong station in February 2018.

Figure 4 also illustrates how price stickiness increases with the distance of stations from central Sydney. The prices of the BP station in Bargo, which is marginally more regional than Picton, exhibit slightly more stickiness throughout the price cycle than the prices of the Picton station. And the prices of the Mittagong station – which is connected to Sydney by only the Hume Motorway seen in Figure 4 and indirectly via the Macquarie Pass from the southern satellite city of Wollongong to the east – exhibit substantially more stickiness. The progressive increase in price stickiness exhibited by these four BP stations for the southern periphery of Sydney is representative of the variation we observe across stations on Sydney's western and northern peripheries.

4 A method for learning the identities of competitors

The goal of this section is to derive a method for recovering the set of competitors a retailer is monitoring and responding to when adjusting their prices. The institutional details on station price setting summarised in Section 3 imply that this goal is equivalent to recovering the sets of "markers" whose prices brand price managers monitor and use when adjusting their own station's prices. Therefore, my approach is to treat stations' marker sets as latent objects upon which the pricing decisions generating my data depend. Thus, just as demand models recover consumer preferences from purchase data (e.g, Guadagni and Little, 1983), and models of market entry recover firms' profit expectations from store opening data (e.g, Vitorino, 2012), the method I derive will recover marker sets from the timing of price adjustments in the FuelCheck data.

I derive the method from a structural model of retail pricing decisions. The purpose and benefit of this structural foundation is to make explicit the assumptions on behaviour implicit in subsequent statistical assumptions and results. Because there is little economic theory characterising the decision of *when* to change prices to guide this modelling, I follow the structural dynamic games literature by modelling the pricing problem of a pricing manager as a strategic decision to be made at decision points arriving stochastically in continuous time (Arcidiacono et al., 2016; Doraszelski & Judd, 2012). These decision points model the reality that a pricing manager cannot continuously monitor the market conditions for all products or retail locations (e.g, stations) in their portfolio, but at any point in time must necessarily attend to one at the exclusion of others.

From this structural model of pricing, I derive the hazard rate characterising the times at which a station chooses to adjust its price. This reduced-form expression combines the arrival rate of the station's decision points with the pricing rule by which it decides whether its price needs adjusting. Thus the expression characterises the functional dependence of the probability a manager adjusts their price at a point in time on the observed market conditions of the retailer at that time – including the prices of its competitors, as well as costs and demand conditions.

I next derive a discrete time analogue for this continuous time hazard rate. This analogue defines a likelihood of observing price-adjustment timings in discrete periods of time that is equivalent to the likelihood of those observations under a generalised linear model ('GLM'). As a result, estimating the mapping from a retailer's market conditions to the timings of its observed price adjustments can exploit existing methods and optimised software for estimating GLMs.

However, while I have data recording station prices, and proxies for costs and demand, I do not know which stations are the markers whose prices enter into a manager's pricing rule, nor the form of the pricing rule mapping the prices of those markers into a price adjustment decision. I address this challenge with two additional elements. First, I exploit the fact a manager's pricing rule is a mapping from a partition of current market conditions into a 0-1 decision about whether to adjust prices to propose a linear expansion of market conditions forming possible boundaries for this partition. That is, I show we can construct a (potentially large) set of binary variables describing possible combinations of markers and pricing rules, which will contain within it variables capturing the *true* markers and pricing rule used to respond to their prices. This construction transforms the problem of identifying the set of markers into a model selection problem – the highest-likelihood GLM will be that which includes on the right-hand side the true variables.

To efficiently search through the combinatorial space of possible GLM specifications for the one containing the true marker-pricing rule variables, I propose using consistent model selection methods for high-dimensional data from machine learning. Specifically, I exploit the consistent model selection properties of the least absolute shrinkage and selection operator ('LASSO') to learn the boundaries of the pricing-rule partition, and identify as markers those stations whose prices define the partition boundaries.

The rest of this section proceeds as follows. In Section 4.1 I lay out the structural model of retail pricing decisions. I derive the reduced form expression for the hazard rate of price adjustments in Section 4.2, and show its discrete time analogue has the same likelihood as a GLM. Then, in Section 4.3, I define the class of possible partition boundaries and summarise the consistent model selection results allowing us to recover the identities of markers by estimating the GML over this class with LASSO regularisation. I use the language of my retail gasoline application throughout this section, but the method could be just as readily applied to identify competitors in other retail contexts.

4.1 A continuous-time model of price competition between gasoline stations

There is a retail gasoline market with N stations, indexed by $i = 1, ..., N$ and collectively denoted by the set $\mathcal{N}$. The market evolves in continuous time, with the current time indexed by $t \in (0, T]$. The profits flowing to a station i at any point in time t are a function of a set of random variables common to all stations, which account for common movements in wholesale gasoline prices and demand conditions, and the current prices of a subset of all stations. This subset, denoted by $\mathcal{N}_i \subseteq \mathcal{N}$ and indexed by $j = 1, ..., N_i$, is the set of direct competitors or 'markers' whose gasoline is considered substitutable for that of station i by its customers. It is the identity of the stations in $\mathcal{N}_i$, for every i, that we wish to learn.

States and payoffs The price $X_i(t)$ of station i, which at any time t is the current state of i, takes values $k_i \in \mathcal{X}_i$, where $\mathcal{X}_i$ is a finite state space space with K_i elements. This state evolves according to the following decision process. At points in time arriving according to a Poisson process with rate parameter λ_i, which we will call 'decision points', the pricing manager of station i evaluates the optimality of $X_i(t)$ and chooses a price $p_i \in \mathcal{X}_i$. Thus at every decision point the station's action space is equal to $\mathcal{X}_i$, which implies its manager has a (costless) continuation action of not adjusting its price, $p_i = X_i(t)$, and that every action $p_i \in \mathcal{X}_i$ leads to a distinct continuation state $X_i(t_+) = p_i$.

Stations' payoff relevant states at time t also include the values of the random variables common to all stations. These common random variables $X_0(t)$, which we will call 'nature' and index by $i = 0$, take values $k_0 \in \mathcal{X}_0$, where $\mathcal{X}_0$ is a finite state space with K_0 elements.

The state of nature evolves over time according to a homogeneous Markov process on $\mathscr{X}_0$ with a $K_0 \times K_0$ intensity matrix Q_0. The off-diagonal elements of Q_0, $q_{k_0 k_0'}$ for $k_0 \neq k_0'$, are the hazard rates for transitions from state k_0 to state k_0':

$$q_{k_0 k_0'} \equiv \lim_{\Delta \to 0} \frac{\mathbb{P}\big(X_0(t+\Delta) = k_0' \mid X_0(t) = k_0\big)}{\Delta} \in (0, \infty). \tag{1}$$

The leading diagonal elements $q_{k_0 k_0}$ are then the negative overall rate with which nature leaves the state k_0:

$$q_{k_0 k_0} = -q_{k_0} = -\sum_{k_0' \neq k_0} q_{k_0 k_0'}$$

where:

$$q_{k_0} \equiv \lim_{\Delta \to 0} \frac{\mathbb{P}(X_0(t+\Delta) \neq k_0 \mid X_0(t) = k_0)}{\Delta} \in (0, \infty). \tag{2}$$

Therefore, conditional on leaving state k_0, the probability of transitioning to state k_0' is $q_{k_0 k_0'}/q_{k_0}$. Finally, transitions by nature out of state k_0 follow an exponential distribution with rate parameter $-q_{k_0}$ (or, equivalently, nature stays in state k_0 for a period of time that is exponentially distributed with parameter q_{k_0}). For additional details on Markov processes I recommend Siegrist (1997).

The optimality of station i's current price $X_i(t)$ is assessed by reference to the current and discounted expected future flow profits resulting from $X_i(t)$. All stations discount the future at the common rate $\rho \in (0,1)$. The flow profits of a station i at time t are a function of its payoff-relevant state, which comprise the state of nature $X_0(t)$, its own current price $X_i(t)$, and the current prices of its markers, $X_j(t)$ for $j \in \mathscr{N}_i$. Let us denote this state by $\mathbf{X}_{-i}$, where the payoff-relevant state space is:

$$\mathscr{X}_{-i} \equiv \bigtimes_{j \in \{0,i\} \cup \mathscr{N}_i} \mathscr{X}_j. \tag{3}$$

The function mapping from $\mathbf{X}_{-i}$ to station i's flow profits is then:

$$\pi_i : \mathscr{X}_{-i} \longrightarrow \mathbb{R}. \tag{4}$$

Information and beliefs To characterise a station's expectations over future flow profits we must specify the information available to the station at time t and the beliefs it forms conditional on that information. Consistent with the institutional knowledge summarised in Section 3, we will assume station i continuously collects information on the payoff relevant state $\mathbf{X}_{-i}$. Thus station i at time t has the information set:

$$\mathscr{I}_i(t) \equiv \{\mathbf{X}_{-i}(\tau) \mid \tau \in (0, t]\}. \tag{5}$$

Based on this information set, station i forms beliefs about the actions $p_j \in \mathscr{X}_j$ that its markers $j \in \mathscr{N}_i$ will take in any given state. However, an implication of the assumption on the information collected by station i is that it can – and most likely will – have incomplete information about the payoff relevant states of its markers. For any marker $j \in \mathscr{N}_i$ for which $\mathscr{N}_i \neq \mathscr{N}_j$, station i will not fully observe the payoff relevant state $\mathbf{X}_{-j}$ informing j's actions. Thus even if i knows the decision rule used by its markers, it can have only probabilistic beliefs – based on the state $\mathbf{X}_{-i}$ that it observes – about the actions that its markers will take in any observed state. Let us denote these beliefs by:

$$\zeta_i(k'_j \mid \mathbf{k}_{-i}) \equiv \mathbb{P}_i\left(p_j = k'_j \mid \mathbf{X}_{-i} = \mathbf{k}_{-i}\right) \tag{6}$$

for every $\mathbf{k}_{-i} \in \mathscr{X}_{-i}$, and every $k'_j \in \mathscr{X}_j$ for every $j \in \mathscr{N}_i$, where the subscript i explicitly denotes the subjective nature of these probabilities.

Decision objective and strategies We can now define the objective against which a station i assesses the optimality of its current price $X_i(t)$ given a decision point at t. By Bellman's principle of optimality, the value function for station i in state $\mathbf{X}_{-i}(t) = \mathbf{k}_{-i}$ is defined recursively as:

$$\begin{aligned} V_i(\mathbf{k}_{-i}) = & \Big[\pi_i(\mathbf{k}_{-i}) + \sum_{k'_0 \neq k_{-i,0}} q_{k_{-i,0}k'_0} V(\mathbf{k}_{-i} \cup k'_0) \\ & + \sum_{j=1}^{N_i} \lambda_j \sum_{k'_j \in \mathscr{X}_j} \zeta_i(k'_j \mid \mathbf{k}_{-i}) V(\mathbf{k}_{-i} \cup k'_j) + \lambda_i \max_{p_i \in \mathscr{X}_i} V(\mathbf{k}_{-i} \cup p_i)\Big] \\ & \Big/ \Big[\rho + q_{k_{-i,0}} + \sum_{j=1}^{N_i} \lambda_j + \lambda_i\Big] \end{aligned} \tag{7}$$

where $\mathbf{k}_{-i} \cup k_j$ is equal to $\mathbf{k}_{-i}$ except with the jth element swapped for k_j, and $k_{-i,0}$ is equal to the 0th element of $\mathbf{k}_{-i}$. The denominator of this expression is the discount factor plus the sum of the rates of state changes by nature and moves by station i and its markers.[18] The values in the numerator can be understood by looking at each term separately. The first term is the flow profits station i earns each instant the payoff relevant state remains $\mathbf{k}_{-i}$. The second term

[18] Note that we are assuming each station i is aware of the rates at which its markers make moves, λ_j for every $\mathscr{N}_i$.

is the sum over states reached by changes in the state of nature of the value of those states, weighted by the hazard rates for those state transitions. The third term is the sum over states reached by the actions of a marker of the value of those states, weighted by i's beliefs about the probability of those actions, and summed over markers, with each marker weighted by the arrival rate of its decision points. The last term is the value of the state reached by i choosing an optimal action, weighted by the arrival rate of i's decision points.

A pricing rule by which station i chooses prices at decision points to maximise V_i is a strategy. Following the dynamic games literature (e.g, Arcidiacono et al., 2016; Ericson and Pakes, 1995; Fershtman and Pakes, 2012), we focus on pure Markov strategies. A Markov strategy for player i is a mapping assigning a price action $p_i \in \mathscr{X}_i$ for every payoff relevant state $\mathbf{k}_{-i} \in \mathscr{X}_{-i}$. By the Markov property, the choice of action only conditions on the state at time t, not the state at any time prior to t. Given beliefs ζ_i, station i's pricing rule δ_i will a best response if:

$$\delta_i(\mathbf{k}_{-i}) = p_i \iff V(\mathbf{k}_{-i} \cup p_i) \geq V(\mathbf{k}_{-i} \cup p_i') \tag{8}$$

for every $p_i' \in \mathscr{X}_i$. We will assume a best response pricing rule is being used by each station to set prices, and denote these by δ_i^* for every $i \in \mathscr{N}$.

The Markov nature of a station's pricing rule, combined with the Poisson arrival process of its decision points, implies the station's prices evolve as a conditional Markov process. That is, $X_i(t)$ is a non-homogeneous Markov process with hazard rates for state transitions on $\mathscr{X}_i$ that vary with time, but not as a direct function of time (Nodelman et al., 2002). Instead the hazard rates are a function of the current payoff relevant state of the station. Therefore, the $K_i \times K_i$ conditional intensity matrix $Q_{X_i|\mathbf{X}_{-i}}$ describing this process by which station i's prices evolve is:

$$Q_{X_i|\mathbf{X}_{-i}} \equiv \begin{bmatrix} -q_1(\mathbf{X}_{-i}) & q_{12}(\mathbf{X}_{-i}) & \cdots & q_{1K_i}(\mathbf{X}_{-i}) \\ q_{21}(\mathbf{X}_{-i}) & -q_2(\mathbf{X}_{-i}) & \cdots & q_{2K_i}(\mathbf{X}_{-i}) \\ \vdots & \vdots & \ddots & \vdots \\ q_{K_i1}(\mathbf{X}_{-i}) & q_{K_i2}(\mathbf{X}_{-i}) & \cdots & q_{K_iK_i}(\mathbf{X}_{-i}) \end{bmatrix} \tag{9}$$

where:

$$q_{k_ik_i'}(\mathbf{X}_{-i}) = \lambda_i \mathbb{1}_{\{X_{-i,i}=k_i\}} \mathbb{P}\left(\delta_i^*(\mathbf{X}_{-i}) = k_i'\right) \tag{10}$$

and:

$$q_{k_i}(\mathbf{X}_{-i}) = \sum_{k_i' \neq k_i} q_{k_ik_i'}(\mathbf{X}_{-i}). \tag{11}$$

This conditional intensity matrix can be equivalently understood as a set of intensity matrices, one for each state in station i's payoff relevant state space, $\mathbf{k}_{-i} \in \mathscr{X}_{-i}$.

Stationarity and market equilibrium To be able to learn $\mathcal{N}_i$ from data generated by $X_i(t)$, we require both $X_i(t)$ and the payoff relevant process $\mathbf{X}_{-i}(t)$ upon which it is conditioned to be stable across time. We again follow the dynamic games literature to obtain this stability by assuming the beliefs ζ_i and pricing rules δ_i^* of all stations $i \in \mathcal{N}$ constitute a dynamic equilibrium in the market's pricing game. Specifically, for the market state space:

$$\mathscr{X} \equiv \bigtimes_{i=0}^{N} \mathscr{X}_i \tag{12}$$

we will assume there exists an Experience-Based Equilibrium ('EBE'; Fershtman and Pakes, 2012) consisting of a subspace $\mathscr{Y} \subseteq \mathscr{X}$, and equilibrium beliefs $\zeta_i^*(k_j' \mid \mathbf{k}_{-i})$ and pricing rules $\delta_i^*(\mathbf{k}_{-i})$ for every $k_j' \in \mathscr{X}_j$, for every $j \in \mathcal{N}_i$, for every $\mathbf{k}_{-i}$ that is a component of a state $\mathbf{Y} \in \mathscr{Y}$, and for every station $i \in \mathcal{N}$.

The market being in an EBE implies:

1. $\mathscr{Y}$ **is a recurrent class** – With probability 1, the Markov process on $\mathscr{X}$ generated by $X_0(t)$ and $\{\lambda_i, \delta_i^*\}_{i \in \mathcal{N}}$ and starting from any state $\mathbf{Y} \in \mathscr{Y}$ will remain forever in $\mathscr{Y}$, visiting each state in $\mathscr{Y}$ infinitely often.

2. **Pricing rules δ_i^* are optimal on $\mathscr{Y}$** – For every station $i \in \mathcal{N}$, given beliefs ζ_i^*, the pricing rule $\delta_i^*(\mathbf{k}_{-i})$ is a best response for every payoff relevant state $\mathbf{k}_{-i}$ that is a component of a state $\mathbf{Y} \in \mathscr{Y}$.

3. **Beliefs ζ_i^* are consistent on $\mathscr{Y}$** – For every station $i \in \mathcal{N}$, given the pricing rule δ_i^*, the beliefs ζ_i^* satisfy the condition:

$$\zeta_i^*(k_j' \mid \mathbf{k}_{-i}) = \frac{\sum_{\mathbf{k}_{-j} \in \mathscr{X}_{-j}} \mathbb{1}_{\{\mathbf{k}_{-j,\mathcal{N}_i \cap \mathcal{N}_j} = \mathbf{k}_{-i,\mathcal{N}_i \cap \mathcal{N}_j}\}} \mathbb{1}_{\{\delta_j^*(\mathbf{k}_{-j}) = k_j'\}}}{\sum_{\mathbf{k}_{-j} \in \mathscr{X}_{-j}} \mathbb{1}_{\{\mathbf{k}_{-j,\mathcal{N}_i \cap \mathcal{N}_j} = \mathbf{k}_{-i,\mathcal{N}_i \cap \mathcal{N}_j}\}}} \tag{13}$$

   for every $k_j' \in \mathscr{X}_j$, for every $j \in \mathcal{N}_i$, and for every payoff relevant state $\mathbf{k}_{-i}$ that is a component of a state $\mathbf{Y} \in \mathscr{Y}$, where $\mathbf{k}_{-i,\mathcal{N}_i \cap \mathcal{N}_j}$ is the sub-vector of $\mathbf{k}_{-i}$ containing the elements of $\mathbf{k}_{-i}$ for those stations that are markers of both i and j.

Thus the EBE requires only that stations' pricing rules are optimal conditional on the beliefs they hold about their markers' actions, and that their beliefs are consistent with their markers' actual actions for only those states of the market entering their information sets. For market states that have not previously been observed, and thus about which the stations have no historical information – "off equilibrium path" states – an EBE allows stations to hold any beliefs whatsoever. Therefore, assuming EBE ensures stability of the data generating process while making minimal assumptions on the information collection and processing abilities of station price managers. This is preferable to the informational assumptions of Markov Perfect Equilibrium (Maskin & Tirole, 1988, 2001), which would require each station to know both the marker set and the pricing rule of every other station.

4.2 The hazard rate of a station's price adjustments

Having ensured stability of the market's price dynamics, let us return to the pricing decisions of a single station to derive the process characterising the timing of price adjustments. Because these adjustments are events occurring at random points in time, their timings will be characterised by a point process. Therefore, we will begin by describing the essential structure of point processes in general before deriving the process defined by the evolution of a station's prices $X_i(t)$ over time.

Point processes A point process is a set of discrete stochastic events occurring in continuous time. That is, for an observation interval $(0, T]$, the sequence of M times at which events are observed to occur:

$$0 < \tau_1 < \ldots < \tau_m < \ldots < \tau_M \leq T \tag{14}$$

is a point process. Let $M(t)$ denote the number of events in the interval $(0, t]$ for any $t \in (0, T]$. The stochastic structure of a point process is fully characterised by its conditional intensity function ('CIF'):

$$\lambda(t \mid \mathscr{I}(t)) \equiv \lim_{\Delta \to 0} \frac{\mathbb{P}(M(t+\Delta) - M(t) = 1 \mid \mathscr{I}(t))}{\Delta} \tag{15}$$

where $\mathscr{I}(t)$ is a set of information upon which the CIF depends at time t. For small Δ, the CIF gives the approximate probability of an event occurring at time t, with the approximation being better the smaller is Δ. Thus the CIF can be understood as the hazard rate of the process. It can also be understood as the generalisation of the rate parameter of a Poisson process to the event arrival probability of a process that is dependent on an information set $\mathscr{I}(t)$, and only obtains the 'memoryless' property of the Poisson process after conditioning on this information. For additional details on point processes I recommend Daley and Vere-Jones (2003, 2008).

The point process characterising the times at which a station i adjusts its price is a component of the conditional Markov process $X_i(t)$. A Markov process can be equivalently defined by a point process characterising the arrival of transitions out of the current state, and an embedded Markov chain characterising the state to which the process next transitions conditional on it transitioning out of the current state (Daley & Vere-Jones, 2008). The rate at which the point process produces transition events is the overall rate at which the Markov process leaves the current state, the negative of which populates the leading diagonal of the

Markov process's intensity matrix. Therefore, the CIF for the point process characterising the transition times of the conditional Markov process $X_i(t)$ is:

$$\begin{aligned}\lambda_i^*(t \mid \mathbf{X}_{-i}) &= q_{X_{-i,i}(t)}(\mathbf{X}_{-i}(t)) \\ &= \lambda_i \mathbb{P}\left(\delta_i^*(\mathbf{X}_{-i}(t)) \neq X_{-i,i}(t)\right). \end{aligned} \tag{16}$$

where $X_{-i,i}$ is the ith element of $\mathbf{X}_{-i}$. The information set upon which this point process depends at time t is equal to $\mathbf{X}_{-i}(t)$ due to the Markov nature of δ_i^*. Therefore, the instantaneous probability of station i adjusting its price at time t is the product of the instantaneous probability of i's pricing manager evaluating the optimality of its current price, and the probability that i's current price is not a best response to its current payoff relevant state.

Discrete time The CIF for the price-adjustment times produced by $X_i(t)$ can be used to define the likelihood of observing a price adjustment in the interval $(t, t+\Delta]$ as a function of $\mathbf{X}_{-i}(t)$. This likelihood enables estimation of the relationship between a station's payoff relevant state and the times at which it adjusts its prices, which in turn enables learning of $\mathcal{N}_i$ with model selection methods. However, applying these model selection methods requires data recorded in discrete time. Consequently, we require a discrete time representation of the data likelihood and thus of the CIF.

To obtain a discrete time representation, we will choose an integer H and partition the observation interval $(0, T]$ into H subintervals $(t_{h-1}, t_h]_{h=1}^H$ of length $\Delta = T/H$. The chosen H must be sufficiently large to ensure there is at most one price adjustment event per subinterval. Denote the discrete time analogues of the continuous time variables as $M_{i,h} = M_i(t_h)$ and $\mathbf{X}_{-i,h} = \mathbf{X}_{-i}(t_{h-1})$. By choosing H sufficiently large, the differences $\partial M_{i,h} = M_{i,h} - M_{i,h-1}$ for $h = 1, ..., H$ define a timeseries of zero-one indicators of price adjustment events. Lastly, denote the discrete time CIF by $\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)$, where the parameter vector θ_i parameterises the functional relationship between $\mathbf{X}_{-i,h}$ and the probability of an event in the subinterval $(t_{h-1}, t_h]$.

Likelihood Denote a sequence of M_i price adjustment events for station i over the partitioned observation window $(t_0, t_H]$ by:

$$M_{i,1:H} \equiv \left\{t_0 < \tau_1 < ... < \tau_{m_i} < ... < \tau_{M_i} \leq t_H \ \cap \ M_{i,H} = M_i\right\}. \tag{17}$$

By construction of the partition, it must be that:

$$\tau_{m_i} \in (t_{h_{m_i}-1}, t_{h_{m_i}}] \ \ \forall m_i = 1, ..., M_i \tag{18}$$

for a subset of M_i subintervals satisfying $h_1 < ... < h_{m_i} < ... < h_{M_i}$, with the remaining $H - M_i$ subintervals containing no events. Therefore the probability of $M_{i,1:H}$ is the probability of exactly M_i events occurring, and of those events falling in the intervals $(t_{h_{m_i}-1}, t_{h_{m_i}}]_{m_i=1}^{m_i=M_i}$. This probability is given by:

$$\begin{aligned}\mathbb{P}(M_{i,1:H}) &= \mathbb{P}\Big(M_{i,H} = M_i \ \cap \ \tau_{m_i} \in (t_{h_{m_i}-1}, t_{h_{m_i}}] \ \ \forall m_i = 1, ..., M_i\Big) \\ &= \prod_{h=1}^{H} \mathbb{P}(\mathscr{A}_{i,h})^{\partial M_{i,h}} \mathbb{P}(\mathscr{A}_{i,h}^c)^{1-\partial M_{i,h}} \end{aligned} \tag{19}$$

where

$$\mathscr{A}_{i,h} \equiv \big\{\exists m_i \text{ s.t } \tau_{m_i} \in (t_{h-1}, t_h] \mid \mathbf{X}_{-i,h}\big\} \tag{20}$$

and $\mathscr{A}_{i,h}^c$ is the compliment of $\mathscr{A}_{i,h}$.

The form of the probability in (19) reveals the equivalence between a sequence of price change events and a sequence of H independent Bernoulli trials – independent conditional on the information $\mathbf{X}_{-i,h}$ – with the 'success probability' for an event in the hth subinterval trial given by $\mathbb{P}(\mathscr{A}_{i,h})$. By the definition of the CIF in (15) and (16), the probability of an event in a subinterval $(t_{h-1}, t_h]$ is constant if the payoff relevant state remains $\mathbf{X}_{-i}(t_h)$. Thus, by choosing H sufficiently large that Δ is small relative to the expected transition time of $\mathbf{X}_i(t)$, the success and failure probabilities for the hth trial can be equivalently expressed as:

$$\mathbb{P}(\mathscr{A}_{i,h}) = \lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta + o(\Delta) \tag{21}$$

$$\mathbb{P}(\mathscr{A}_{i,h}^c) = 1 - \lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta + o(\Delta). \tag{22}$$

Substituting these Bernoulli trial probabilities into (19) yields the following joint probability mass function ('PMF') for the sequence $M_{i,1:H}$ conditional on the parameterisation θ_i:

$$\mathbb{P}(M_{i,1:H} \mid \theta_i) = \prod_{h=1}^{H} \big[\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta\big]^{\partial M_{i,h}} \big[1 - \lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta\big]^{1-\partial M_{i,h}} + o(\Delta^{M_i}). \tag{23}$$

This joint PMF is a discrete time analogue of the joint probability density function of the continuous time point process (Truccolo et al., 2005). This joint PMF is also the likelihood $\mathscr{L}_i(\theta_i)$ of observing the sequence of price change events $M_{i,1:H}$ under the probability model characterised by $\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)$ and parameterised by θ_i.

Generalised Linear Model In addition, the joint PMF in (23) is equivalent to the likelihood of independently distributed data $\mathbf{y}_i = \{y_{i,h}\}_{h=1}^{H}$, where $y_{i,h} \in \{0, 1\}$ for $h = 1, ..., H$, under a generalised linear model ('GLM') with Bernoulli distribution and logistic link function (Mc-

Cullagh & Nelder, 1989), also commonly known as a logistic regression model.[19] That is, the joint PMF in (23) is equivalent to the likelihood of data $\mathbf{y}_i$, with $y_{i,h} \equiv \partial M_{i,h}$, under to a model in which:

1. The observations $y_{i,h}$ are assumed to be independently Bernoulli distributed conditional on Q general functions of $X_{-i,h}$:

$$y_{i,h} \mid \left\{g_1(X_{-i,h}), ..., g_Q(X_{-i,h})\right\} \sim \text{Bernoulli}\left(\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta\right) \tag{24}$$

   implying:

$$\mathbb{E}\left[y_{i,h} \mid \left\{g_1(X_{-i,h}), ..., g_Q(X_{-i,h})\right\}\right] = \lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta. \tag{25}$$

2. The expected value of $y_{i,h}$ is conditioned on a linear combination of the general functions $\left\{g_1(X_{-i,h}), ..., g_Q(X_{-i,h})\right\}$:

$$\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta = f^{-1}\left(\sum_{q=1}^{Q} \theta_{i,q} g_q(X_{-i,h})\right) \tag{26}$$

   via a logistic 'link' function:

$$f\left(\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta\right) = \ln\left(\frac{\lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta}{1 - \lambda_i^*(t_h \mid \mathbf{X}_{-i,h}; \theta_i)\Delta}\right). \tag{27}$$

This equivalence implies we can use logistic regression to estimate the relationship between the probability of a price adjustment by station i during the interval $(t, t + \Delta]$ and i's payoff relevant state $X_i(t)$ at time t.

4.3 Consistent estimation of competitors via model selection

The remaining challenge is that we do not know the identities of the markers whose prices enter the payoff relevant state of station i. We address this challenge using the model selection properties of the LASSO. However, we must first define the variables to which we apply LASSO regularisation when estimating the logistic regression model of a station's price adjustments. Therefore, suppose for the time being that we know the identities of the markers $\mathcal{N}_i$ whose prices define the space of payoff relevant states $\mathcal{X}_{-i}$.

State space partition The pricing rule δ_i^* by which station i decides at decision points whether to adjust its price defines a decision boundary partitioning $\mathcal{X}_{-i}$ into two regions. The

[19] For further details on this equivalence see Truccolo et al. (2005), who show there is also an equivalence to the likelihood of a GLM with Poisson distribution and logarithmic link function if H is chosen large enough.

region on one side of this boundary contains all the states in which i will choose to adjust its price:

$$\mathscr{X}_{-i}^{1} \equiv \{\mathbf{k}_{-i} \in \mathscr{X}_{-i} | \delta_i^*(\mathbf{k}_{-i}) \neq k_{-i,i}\}. \tag{28}$$

The region on the other side contains the states in which i will choose to not adjust its price and instead continue in its current state:

$$\mathscr{X}_{-i}^{0} \equiv \{\mathbf{k}_{-i} \in \mathscr{X}_{-i} | \delta_i^*(\mathbf{k}_{-i}) = k_{-i,i}\}. \tag{29}$$

Consequently, the CIF defined in (16) will equal zero for all the states $\mathbf{k}_{-i} \in \mathscr{X}_{-i}^{0}$. Further, the CIF can be decomposed into a linear combination of functions indicating whether the state at time t falls into the region $\mathscr{X}_{-i}^{1}$.

Model selection The problem now becomes identifying which of the Q variables are arguments into the station's actual pricing rule. In principle, we could do this by estimating the logistic regression model with every combination of variables and using a model selection criteria like the Bayesian Information Criterion ('BIC') to select the one with the best fit. In practice this is not possible – there are 2^Q such combinations, which for even moderately sized Q is infeasibly many models to estimate. Instead, we will take the approach of estimating the logistic regression model with the Least Absolute Shrinkage and Selection Operator ('LASSO') to exploit the LASSO's consistent model selection properties.

The LASSO estimator adds a penalty to the log-likelihood of a model that decreases the log-likelihood for every coefficient estimated to have a non-zero value:

$$\mathscr{L}_{LASSO}(\theta, C) \equiv \log \mathscr{L}(\theta) - \frac{1}{C} \sum_{q=1}^{Q} |\theta_q| \tag{30}$$

where C is the 'hyperparameter' parameterising the amount by which the likelihood is penalised for each non-zero coefficient.[20] Consequently, the LASSO estimated coefficients:

$$\widehat{\theta}_{LASSO}(C) \equiv \underset{\theta}{\operatorname{argmax}} \mathscr{L}_{LASSO}(\theta, C) \tag{31}$$

will include fewer non-zero coefficients than the coefficients estimated by maximum likelihood:

$$\widehat{\theta} \equiv \underset{\theta}{\operatorname{argmax}} \log \mathscr{L}(\theta). \tag{32}$$

[20] It is more conventional to write the hyperparameter with the notation $\lambda = 1/C$. However, given I have already used λ_i to denote the Poisson arrival rate of decision points for station i, and $\lambda_i(t|\mathscr{I}_t)$ to denote the CIF of a point process, I follow the `scikit-learn` documentation in using C to denote the inverse of the LASSO hyperparameter. Note that this implies the log-likelihood is penalised *more* by a *smaller* hyperparameter C.

The variables for which the LASSO estimates non-zero coefficients are said to be 'selected' by the LASSO.[21] Denote the set of these selected variables by:

$$\widehat{S}(C) \equiv \{q \mid \widehat{\theta}_{q,LASSO}(C) \neq 0, q = 1, \ldots, Q\}. \tag{33}$$

If we define by S^* the variables entering the true model:

$$S^* \equiv \{q \mid \theta^*_q \neq 0, q = 1, \ldots, Q\}, \tag{34}$$

where θ^*_q are the values of the parameters in the true model, then consistent model selection entails:

$$\mathbb{P}\left(\widehat{S}(C) = S^*\right) \to 1, \text{ as } n \to \infty. \tag{35}$$

Note that this is not the same as – nor necessarily achieved by – consistent parameter estimation, which involves:

$$\widehat{\theta}^n_q - \theta^*_q \xrightarrow{p} 0, \text{ as } n \to \infty \tag{36}$$

where $\widehat{\theta}^n_q$ is the value of the parameter θ_q estimated from an n-observation data set.

Consistent model selection as defined in Equation (35) requires that the parameters in the true model be sufficiently large, and that the Q variables included in the estimated model satisfy a neighbourhood stability condition known as the "irrepresentable conditon" (Meinshausen & Bühlmann, 2006; Zhao & Yu, 2006). How large is sufficiently large depends on n. The parameters θ^*_q for $q \in S^*$ are sufficiently large if they satisfy the "beta-min condition", which given our notation we will refer to as the "theta-min condition":

$$|\theta^*_q| \gg \sqrt{q^* \log(Q)/n} \tag{37}$$

for every $q \in S^*$, where $q^* \equiv |S^*|$. Therefore, the bigger is the data set used in estimation the smaller the true parameters can be and still get consistently selected by the LASSO estimator.

In order to state the irrepresentable condition, let us assume without loss of generality that our Q variables are ordered such that the variables entering the true model come first: $S^* = \{1, \ldots, q^*\}$. Then we can split the $n \times Q$ matrix of our data $\mathbf{G}^n$ into the $n \times q^*$ matrix $\mathbf{G}^n_1$ containing data for the 'true variables' and the $n \times (Q - q^*)$ matrix $\mathbf{G}^n_2$ containing data for the 'false variables'. The irrepresentable condition requires that:

$$\left|\left|\mathbf{G}^{n\top}_2 \mathbf{G}^n_1 \left(\mathbf{G}^{n\top}_1 \mathbf{G}^n_1\right)^{-1} \text{sign}(\theta^*_1, \ldots, \theta^*_{q^*})^\top\right|\right|_\infty < 1 \tag{38}$$

[21] Note that if the optimisation problem in Equation (31) has multiple solutions (for a given C), the variables that have zero coefficients in one solution will have zero coefficients in all solutions.

where $||\mathbf{g}||_\infty \equiv \max_q |g_q|$. This requirement is essentially the requirement that all the coefficients estimated by regressing the one of the false variables on the true variables have magnitude less than one. That is, it requires the false variables to not be closely approximated by any linear combination of the true variables.

The irrepresentable condition is a strong condition to satisfy. It will be easier to satisfy the smaller are the periods Δ into which we discretise time, because in the limit as $\Delta \to 0$, the number of periods in which any two stations' prices differ will grow to infinity. Nonetheless, any empirical specification has a chance of not meeting the condition. In that case, the specification may still satisfy the conditions required for the more conservative goal of variable screening.

Variable screening differs from model selection in that while the latter has the goal of selecting all the true variables, variable screening has the less ambitious goal of selecting all true variables with parameters of a certain magnitude. Let us define the set of 'substantial true variables' by:

$$S^{\text{relevant}(\underline{\theta}^n)} \equiv \{q \mid |\theta_q^*| \geq \underline{\theta}^n,\ q = 1, ..., Q\}, \tag{39}$$

for a threshold magnitude $\underline{\theta}^n$ that depends on the number of observations n. Then, for certain technical 'compatibility conditions' on the smallest eigenvalue of the data matrix $\mathbf{G}^n$ (Bühlmann & van de Geer, 2011), if the regularisation hyperparameter C_n is in the order of $\sqrt{\log(Q)/n}$, for any $\underline{\theta}^n > O\left(q^* \sqrt{\log(Q)/n}\right)$ it will the case that $\widehat{S}(C_n)$ contains $S^{\text{relevant}(\underline{\theta}^n)}$ with high probability for all n. And if in addition the theta-min condition is satisfied, $S^{\text{relevant}(\underline{\theta}^n)} = S^*$.

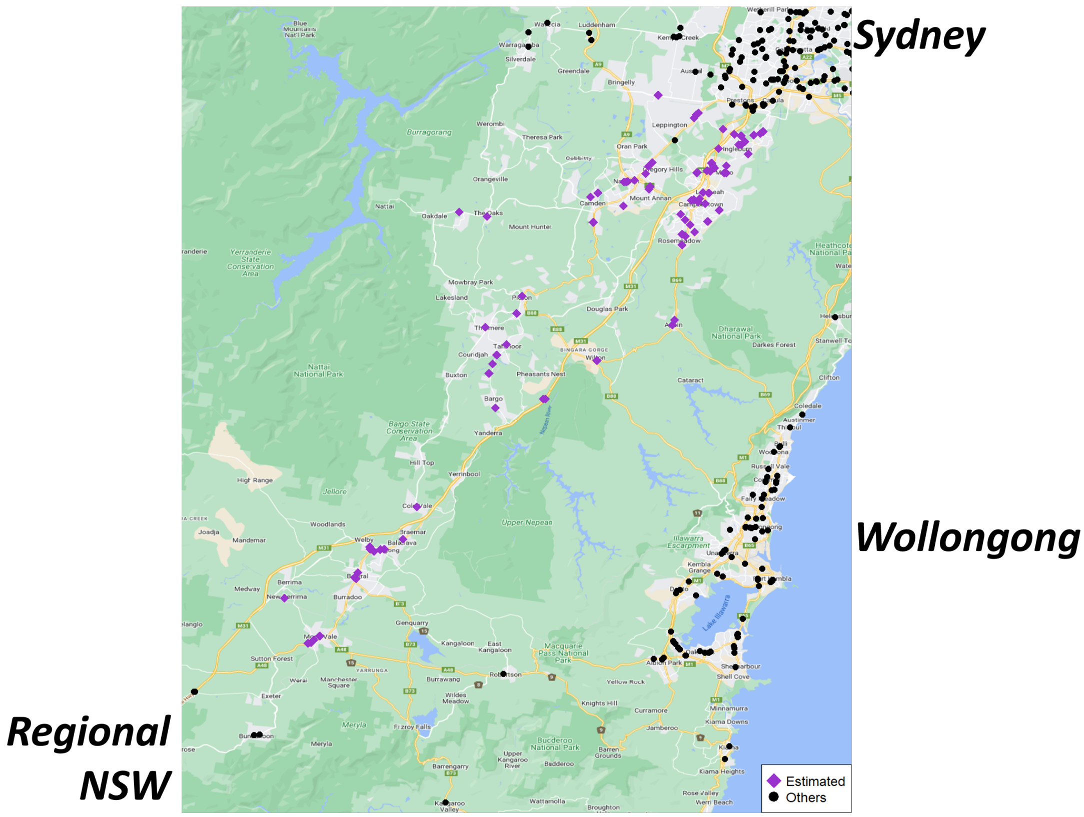
Variable screening essentially says that the LASSO will be a conservative estimator of which variables enter the true model, in the sense that the set of variables selected by LASSO will contain all the substantial variables with a high probability, but possibly also some false positives.

5 Identifying the competitors of NSW gasoline stations

In this section I demonstrate the method proposed in Section 4 by applying it to the FuelCheck data to estimate the competitors of gasoline stations in Sydney. Even though the computational efficiency of the method is greatly enhanced by the logistic regression implementation derived in Section 4.2, estimation is still computationally costly. Therefore I apply the method to a sample of 86 stations in the corridor leading south out of Sydney towards Melbourne. Figure 5 shows the locations of these stations, which includes the stations whose cycling prices are shown in Figure 4 in Section 3.3. I choose this corridor because it spans and connects urban stations in Sydney's southern suburbs and regional stations in the Southern Highlands area of NSW, and is geographically bounded on the east and west by uninhabited wilderness.

Consequently, this subset provides an ideal setting for assessing whether urban and regional stations inhabit distinct geographic markets, or together form a single, interconnected market.

Figure 5: Stations entering marker estimation Map of the corridor leading south out of Sydney's southern suburbs into the Southern Highlands of NSW. Locations of 86 stations for which competitors were estimated are marked by purple diamonds, while black circles mark the locations of additional stations in the depicted area.



5.1 Empirical specification

To implement the method I must specify both the duration Δ of intervals into which we discretise continuous time, and the possible state-space partitioning pricing rules $g_q(\cdot)$ being used to set prices. These specification choices are guided by my understanding of the institutional context, summarised in Section 3.1, and the station pricing behaviour I observe in the data. I describe these choices below.

Discretising continuous time The duration Δ of the intervals into which continuous time is discretised must strike a balance between two contrasting effects on the estimation. On the one hand, the Poisson assumption underpinning our estimation method requires that there be no more than one price adjustment event in any interval. Satisfying this condition requires a sufficiently short interval.[22] On the other hand, shorter intervals result in more rows of data, which increases the computational cost of estimation. Thus interval lengths should be no shorter than necessary.

I discretise my sample period into 5-minute intervals. I am satisfied this interval length strikes the correct balance between minimising the computational cost of estimation and reducing the incidence of multiple price adjustments by a station within any interval. The average time between price adjustments by a station is in no case shorter than 10 hours. This average is dragged up in instances by a long right tail. However, for the majority of station whose minimum inter-adjustment interval is less than 5 minutes, it is also less than 30 seconds. Therefore, given the vast majority of stations change prices less than once a day, I am are not concerned that there is systematic competitive behaviour happening at a temporal resolution finer than five minutes.

Having chosen the interval length, I construct the binary variable indicating the event of a price adjustment by a station within the interval:

$$y_{ih} = \mathbf{1}_{\{p_{i,h+1} \neq p_{ih}\}} = \mathbf{1}_{\{p_{it_h+\Delta} \neq p_{it_h}\}}$$

for every station i and every interval $h = 1, ..., H$.

Partitioning the state space The core of the method is model selection via LASSO regularisation to identify the markers whose prices are arguments to a station's pricing rule, which partitions the station's state space into a region of states in which the station decides to adjust its price and a region of states in which it decides to maintain its current price. Therefore, implementing the method requires a set of variables defining possible boundaries between these two decision regions. I am guided in constructing these variables by the institutional knowledge summarised in Section 3.1 – particularly the knowledge that there will be roughly a dozen markers (at most) for each station – and extensive visualisations of the price timeseries of nearby stations, like those in Figures 2 and 4, which show stations adjusting their prices to maintain a brand-specific differential with respect to other nearby stations during the undercutting phase of cycles.

Therefore, I construct the following variables to define potential partitions of a station's payoff relevant state spaces:

[22]A smaller interval also increases the probability that the variables entering the estimation satisfy the irrepresentable condition necessary for consistent model selection by the LASSO. That is, shorter intervals reduce the chance of the price adjustment times of any two stations always falling in the same interval of time, and thus that their discretised price timeseries are highly correlated.

Dummies for the event that the station's price differential with respect to another station j is greater than a threshold m, for a range of thresholds and a set of other stations:

$$\mathbf{1}_{\{p_{ih}-p_{jh}\geq m\}}$$

for $m=-10,9,...,10$ and $j\in S_i$.

Interactions of these dummies for every pair of other stations:

$$\mathbf{1}_{\{p_{ih}-p_{jh}\geq m\}}\times\mathbf{1}_{\{p_{ih}-p_{j'h}\geq m'\}}$$

for $m,m'=-10,9,...,10$ and $j,j'\in S_i$ such that $j\neq j'$.

Dummies for the event that the station's price margin is greater than a threshold m, for a range of thresholds:

$$\mathbf{1}_{\{p_{ih}-TGP_h\geq m\}}$$

for $m\in\{-2,0,2,5,10,15\}$.

Interactions for these price margin dummies with both the price differential dummies and the interactions between price differential dummies:

$$\mathbf{1}_{\{p_{ih}-TGP_h\geq m\}}\times\mathbf{1}_{\{p_{ih}-p_{j'h}\geq m'\}}$$

and

$$\mathbf{1}_{\{p_{ih}-TGP_h\geq m\}}\times\mathbf{1}_{\{p_{ih}-p_{j'h}\geq m'\}}\times\mathbf{1}_{\{p_{ih}-p_{j''h}\geq m''\}}$$

for $m\in\{-2,0,2,5,10,15\}$, $m',m''=-10,9,...,10$ and $j',j''\in S_i$ such that $j'\neq j''$.

Dummies for the hour-of-day and day-of-week to proxy for systematic fluctuations in demand:

$$\mathbf{1}_{\{hour(t_h)=l\}}$$

and

$$\mathbf{1}_{\{day(t_h)=d\}}$$

for $l=1,...,24$ and $d=1,...,7$, where $day(t)$ and $hour(t)$ are functions extracting the day and hour of t.

The first of these four categories of variables captures the essential feature of station pricing rules – visible in their resulting price dynamics – that they define bands of price differentials with respect to the prices of markers, such that stations do not adjust prices until their markers' prices fall below the lower bound of this band. The range of thresholds allows for the possibility that the lower bound of this band may be above the station's price, as is likely in the case of a station being mass-market brand and its marker being a premium brand with less price-sensitive customers. The second category of variables allows for the possibility that a single marker lowering its price outside of a station's price differential band is not a sufficiently large competitive threat to induce the station to respond, and that rather stations adjust their price only once a combination of their markers have undercut them. The third category of variables allows for cost-based pricing rules, and allows for the observed behaviour leading up to a market-wide restoration when stations' prices have been discounted down to their own price margin floor. The final category of variables accounts for the possibility of pricing rules that respond to systematic fluctuations in demand, which in this market are likely created by hour- and day-based phenomena like rush hour and weekend travel.

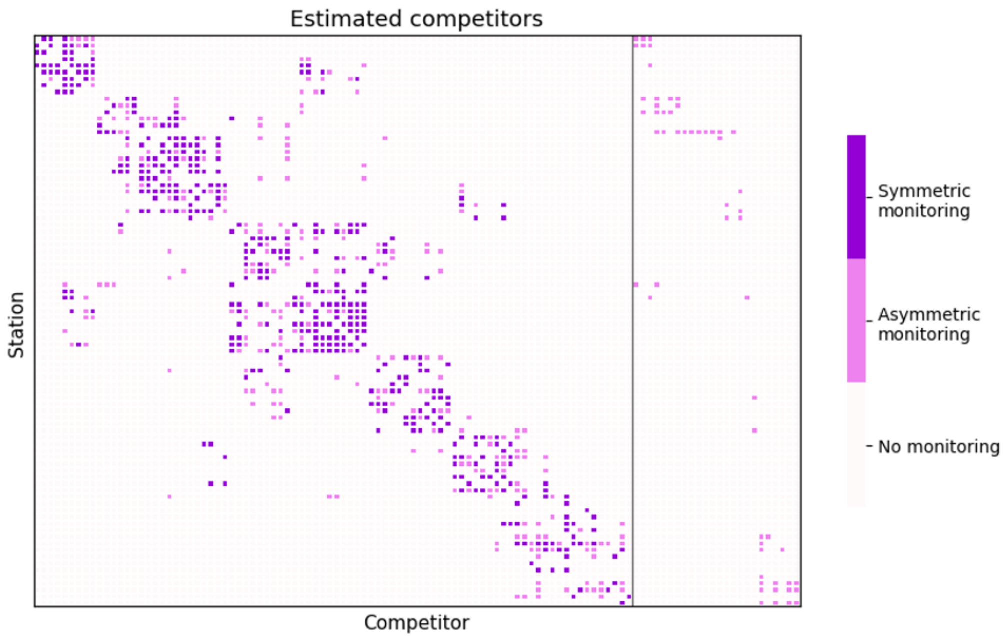
5.2 Estimated competitor sets & competition structure

The stations are monitoring and responding to the prices of very heterogeneous sets of competitors, in terms of both size and composition. I estimate the stations are marking between two and 16 competitors, with markedly different compositions in the stations they mark. I visualise these relationships with an adjacency matrix in Figure 6a. For each row station, the coloured cells mark the column stations they are treating as their competitors. The darker purple colouring indicates the treatment is reciprocated, while the lighter pink colouring indicates asymmetry in the relationship – the row station is monitoring and responding to the prices of the column station, but not vice versa. The stations are roughly ordered in the indices from north at the top / on the left, to south at the bottom / to the right. The column stations to the right of the vertical line are those being treated as a competitor by some station in my sample, but are not themselves a part of my sample. These are stations at the northern and southern of the corridor, depicted by black dots in Figure 5.

Even though stations' competitor sets are very heterogeneous, they overlap to connect all stations in the sample into a single networked competitive structure. I visualise this network with a standard spring layout in Figure 6b. Unfortunately the network cannot be coherently visualised on top of a map like that in Figure 5, but the layout of nodes in Figure 6b roughly corresponds to a 90-degree clockwise rotation of the layout of stations in Figure 5. Nodes with black edges are stations for whom I have estimated competitors, and the nodes without edges are those stations who are a marker of one of the stations in the sample, but are not themselves in the sample.

This networked competitive structure differs substantially from the concept of distinct geographic markets that is commonly used to model and analyse competition, particularly

Figure 6: Estimated competitive relationships Estimated price monitoring and response relationships of 86 stations in sample.



(a) Adjacency matrix

(b) Network

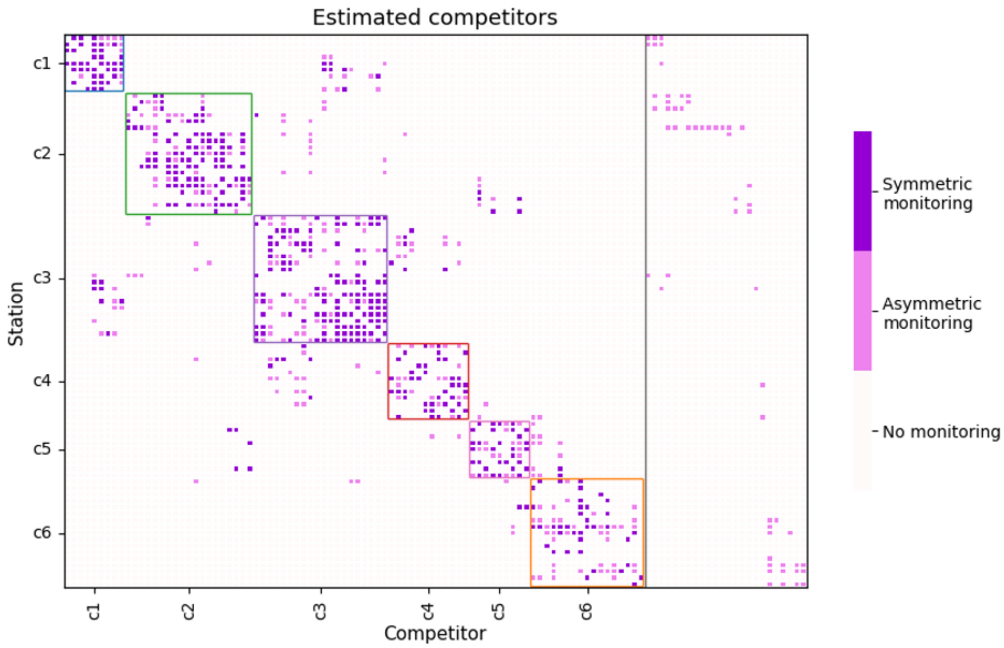
in this industry. If we demarcated distinct markets in this corridor – say using the 3km and 10km radii used by the ACCC in its 2017 merger review (ACCC, 2018b) – not only would we be ignoring the heterogeneity and asymmetry in the competitor sets of neighbouring stations, but we would be ignoring competitive influences on the stations stemming from the broader competitive network.

Market-like clusters However, there are dense groupings of competing stations visible in Figures 6a and 6b, might approximate distinct markets closely enough to support an accurate analysis of competition if these clusters of stations were treated as disjoint markets. I therefore detect clusters of stations within the network using the label propagation algorithm of Raghavan et al. (2007), which resembles competitive price matching by stations. The algorithm is initialised by giving every station a unique label. The algorithm then cycles through the stations, updating their labels as it goes to match the most common label amongst their monitored competitors. The algorithm iterates this process until convergence, resulting in densely-connected clusters of stations who all share a common label. Therefore, the algorithm resembles the process of stations updating their prices to match their monitored competitors, and the clusters it finds are those for whom this process of competitive price matching would result in uniform prices.

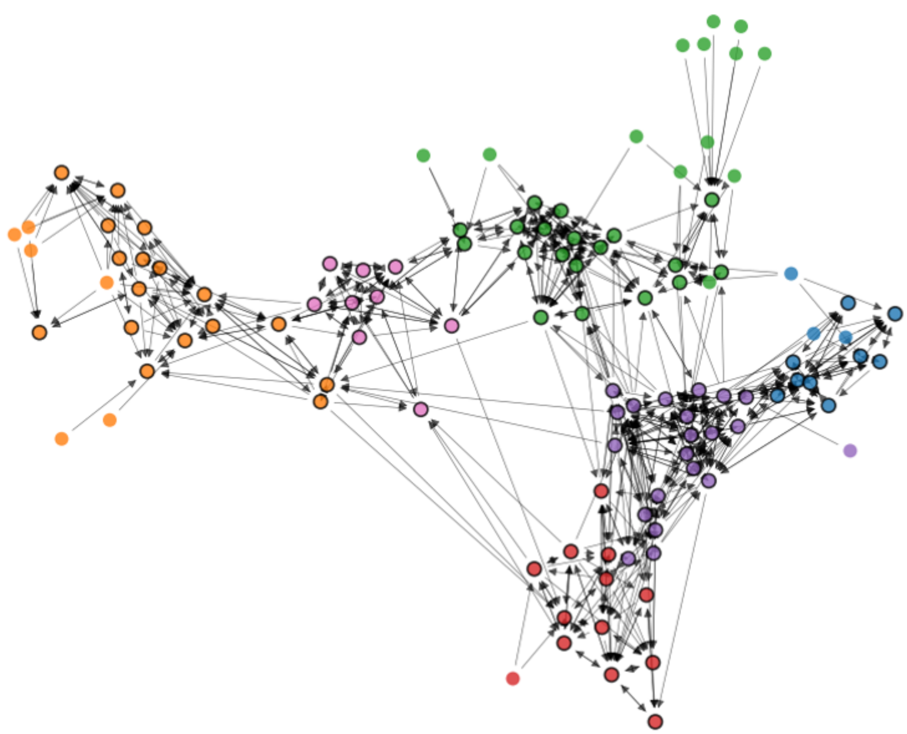
I find six distinct clusters of stations in the network formed from their competitive monitoring. These clusters are depicted by node colourings in the network in Figure 7b and blocks in the adjacency matrix in Figure 7a. The latter visualisation highlights the degree to which the networked competitive structure connecting stations resembles the outcome of a stochastic block model (Holland et al., 1983). That is, the competitive structure of this market resembles that which would be generated if stations chose to mark stations in their own cluster with a probability p^H, and stations in other clusters with a lower probability $p^L < p^H$.

The six clusters are easier to interpret when compared to the geographic location of their respective stations. In Figure 8 I colour the points marking station locations on the map of the corridor south of Sydney to show the stations' cluster membership. This visualisation reveals the six clusters are geographically contiguous. It also suggests some of the partitions between clusters can be explained by features of the road network that would create barriers for motorists wanting to substitute between stations in different clusters. For example, stations in the second cluster are separated from those to their south in the fifth cluster by a drive of at least 20 minutes, and they are separated from those to their east in the first, third and fourth clusters by both the Hume Motorway and Campbelltown railway. However, inspection of the road network does not as easily explain why the stations in the first, third and fourth clusters form three separate clusters, rather than being joined together into one common cluster. Therefore, while the cluster results suggest there could be some uniformity of prices

Figure 7: Estimated clusters of stations Six station clusters detected by the (asynchronous) label propagation algorithm.



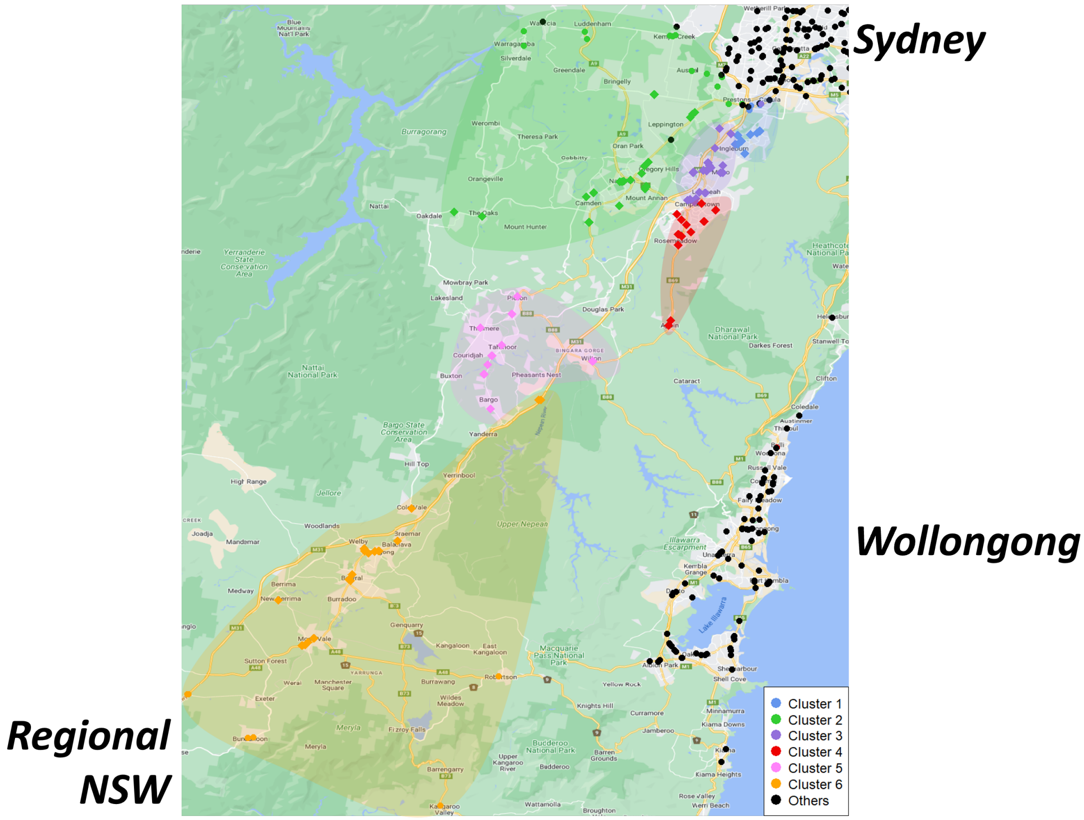
(a) Adjacency matrix



(b) Network

with geographically-distinct, market-like clusters of stations, it is not clear ex ante where those clusters should lie.

Figure 8: Locations of station clusters Map of stations in the corridor leading south out of Sydney, with station's membership in six clusters depicted by colour.



6 Conclusion

In this paper I introduced a method for identifying who retail firm managers are considering and treating as their competitors when they adjust prices. The method exploits the idea that managers choose to adjust their prices when they notice their price position vis-à-vis their competitors has become unfavourable, and thus that there is information about the identities of those competitors contained in the relation between the timing of price adjustments and relative prices at the times of price adjustments. The method I propose exploiting this idea is derived from a structural model of price competition, but consists of estimating a reduced-form GLM expression with model selection via LASSO regularisation. Thus while the method has both a theoretical foundation and econometric consistency, its implementation involves no more than LASSO-regularised GLM regression on feature-engineered data, meaning the method can be implemented by industry analysts and data scientists using modern, highly-optimised machine learning toolkits.[23]

I also demonstrated the method with an application to the retail gasoline industry in New South Wales, Australia. Using 30 months of high-frequency, publicly, available price data from the FuelCheck price-monitoring program, I estimated the competitor sets of 86 stations located in a corridor south of Sydney. Collectively these competitor sets revealed three novel insights about the competitive structure of a retail market: (1) the size and composition of stations' competitor sets were heterogeneous, even for closely neighbouring stations; (2) pairwise competition between stations was asymmetric, in the sense that one station could be monitoring and responding to the prices of another without the reverse being true; and, (3) overlaps in stations' competitor sets connected them all together into a networked competition structure.

These insights have deep implications for the way we think about and characterise competition structure, which can be seen by comparing these insights to the competition structure the ACCC used to analyse gasoline station competition in its 2017 merger review (ACCC, 2018b). In that review, the ACCC assumed that the market in which each acquiree station competes constituted the stations within a 3km or 10km radius. And further, by using the concentration of that market as a sufficient measure of the competitive constraint on the acquiree station's pricing power, the ACCC implicitly assumed all the stations in that market competed with each other and only each other. This assumed competition structure is markedly different from the one revealed by my application, which not only suggests competition between the stations in this defined market would be sparse, and would include stations beyond the border of this market, but that concentration in this market would say little about the constraints being imposed on the acquiree station's pricing by competition. Therefore, the insights from my application imply a need to broaden our characterisations of competition structure to capture the networked nature of retail competition and study the impact of networked competition structures on the propagation of price adjustments and the ability for information about movements in costs and demand to be efficiently impounded into market prices.

[23]For example, the method can be efficiently implemented using the `scikit-learn` module in `python`, or the `GLMnet` package in `R`.

References

ACCC. (2015). Undertaking to the australian competition & consumer commission given for the purposes of section 87b by informed sources (australia) pty ltd.

ACCC. (2017). Statement of issues: Bp — proposed acquisition of woolworths' network of retail service station sites.

ACCC. (2018a). Petrol price cycles in australia.

ACCC. (2018b). Public competition assessment: Bp australia pty ltd — proposed acquisition of woolworths limited's network of retail service station sites.

ACCC. (2019a). Report on the australian petroleum market – june quarter 2019.

ACCC. (2019b). Report on the australian petroleum market – march quarter 2019.

ACIL-Tasman. (2009). Petroleum import infrastructure in australia: Main report.

AIP. (2019). Facts about the australian wholesales fuels market & prices.

Arcidiacono, P., Bayer, P., Blevins, J. R., & Ellickson, P. R. (2016). Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *Review of Economic Studies*, *83*, 889–931.

Atkinson, B. (2009). Retail gasoline price cycles: Evidence from guelph, ontario using bi-hourly, station-specific retail price data. *The Energy Journal*, *30*(1), 85–109.

Atkinson, B., Eckert, A., & West, D. S. (2009). Price matching and the domino effect in a retail gasoline market. *Economic Inquiry*, *47*(3), 568–588.

Bimpikis, K., Ehsani, S., & İlkılıç, R. (2019). Cournot competition in networked markets. *Management Science*, *65*(6), 2467–2481.

Bühlmann, P., & van de Geer, S. (2011). *Statistics for high-dimensional data: Methods, theory and applications*. Springer.

Byrne, D. P. (2019). Gasoline pricing in the country and the city. *Review of Industrial Organization*, *55*, 209–235.

Byrne, D. P., & de Roos, N. (2019). Learning to coordinate: A study in retail gasoline. *American Economic Review*, *109*(2), 591–619.

Byrne, D. P., Nah, J. S., & Xue, P. (2018). Australia has the world's best petrol price data: Fuelwatch and fuelcheck. *The Australian Economic Review*, *51*(4), 564–577.

Campbell, A. (2013). Word-of-mouth communication and percolation in social networks. *American Economic Review*, *103*(6), 2466–2498.

Castanias, R., & Johnson, H. (1993). Gas wars: Retail gasoline price fluctuations. *The Review of Economics and Statistics*, *75*(1), 171–174.

Chan, T. Y., Padmanabhan, V., & Seetharaman, P. B. (2007). An econometric model of location and pricing in the gasoline market. *Journal of Marketing Research*, *44*(4), 622–635.

Chintagunta, P. K., & Rossi, F. (2016). Price transparency and retail signs: Evidence from fuel price signs in the italian highway system. *Journal of Marketing Research*, *53*(3), 407–423.

Chintagunta, P. K., & Rossi, F. (2018). Price uncertainty and market power in retail gasoline: The case of an italian highway. *Marketing Science*, *37*(5), 753–770.

Daley, D. D., & Vere-Jones, D. (2003). *An introduction to the theory of point processes: Elementary theory and methods* (2nd ed., Vol. 1). Springer.

Daley, D. D., & Vere-Jones, D. (2008). *An introduction to the theory of point processes: General theory and structure* (2nd ed., Vol. 2). Springer.

Doraszelski, U., & Judd, K. L. (2012). Avoiding the curse of dimensionality in dynamic stochastic games. *Quantitative Economics*, *3*, 53–93.

Doyle, J., Muehlegger, E., & Samphantharak, K. (2010). Edgeworth cycles revisited. *Energy Economics*, *32*, 651–660.

Eckert, A., & West, D. S. (2004). Retail gasoline price cycles across spatially dispersed gasoline stations. *The Journal of Law & Economics*, *47*(1), 245–273.

Erdem, T. (1996). A dynamic analysis of market structure based on panel data. *Marketing Science*, *15*(4), 359–378.

Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, *62*(1), 53–82.

Fershtman, C., & Pakes, A. (2012). Dynamic games with asymmetric information: A framework for empirical work. *The Quarterly Journal of Economics*, *127*(4), 1611–1661.

Foros, Ø., & Steen, F. (2013). Vertical control and price cyles in gasoline retailing. *The Scandinavian Journal of Economics*, *115*(3), 640–661.

Goyal, S., & Moraga-González, J. L. (2001). R&d networks. *The RAND Journal of Economics*, *32*(4), 686–707.

Grice, R. (2023). Chain-linked markets.

Guadagni, P. M., & Little, J. D. C. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science*, *2*(3), 203–280.

Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic bloxkmodels: First steps. *Social Networks*, *5*(2), 109–137.

Hyoryung, N., Joshi, Y. V., & Kannan, P. K. (2017). Harvesting brand information from social tags. *Journal of Marketing*, *81*(4), 88–108.

Israeli, A., & Anderson, E. (2023). Learning about demand in the long-tail: The role of competitive monitoring.

Kannan, P. K., & Sanchez, S. M. (1994). Competitive market structures: A subset selection analysis. *Management Science*, *40*(11), 1484–1499.

Kim, J. B., Albuquerque, P., & Bronnenberg, B. J. (2011). Mapping online consumer search. *Journal of Marketing Research*, *48*(1), 13–27.

Kim, S., Putrino, D., Ghosh, S., & Brown, E. N. (2011). A granger causality measure for point process models of ensemble neural spiking activity. *PLoS Computational Biology*, *7*(3).

Lee, T. Y., & Bradlow, E. T. (2011). Automated marketing research using online customer reviews. *Journal of Marketing Research*, *48*(5), 881–894.

Lewis, M. S. (2009). Temporary wholesale gasoline price spikes have long-lasting retail effects: The aftermath of hurricane rita. *The Journal of Law & Economics*, *52*(3), 581–605.

Li, J. Q., Rusmavichientong, P., Simester, D., Tsitsiklis, J. N., & Zoumpoulis, S. I. (2015). The value of field experiments. *Management Science*, *61*(7), 1722–1740.

Li, J., Netessine, S., & Koulayev, S. (2018). Price to compete ... with many: How to identify price competition in high-dimensional space. *Management Science*, *64*(9), 4118–4136.

Maskin, E., & Tirole, J. (1988). A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica*, *56*(3), 571–599.

Maskin, E., & Tirole, J. (2001). Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, *100*(2), 191–219.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall.

Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *Annals of Statistics*, *34*(3), 1436–1462.

Moe, W. W. (2006). An empirical two-stage choice model with varying decision rules applied to internet clickstream data. *Journal of Marketing Research*, *43*(4), 680–692.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science*, *31*(3), 521–543.

Nodelman, U., Shelton, C. R., & Koller, D. (2002). Continuous time bayesian networks, 378–387.

Noel, M. D. (2007a). Edgeworth price cycles, cost-based pricing, and sticky pricing in retail gasoline markets. *The Review of Economics and Statistics*, *89*(2), 324–334.

Noel, M. D. (2007b). Edgeworth price cycles: Evidence from the toronto retail gasoline market. *The Journal of Industrial Economics*, *55*(1), 69–92.

Novak, T. P. (1993). Log-linear trees: Models of market structure in brand switching data. *Journal of Marketing Research*, *30*(3), 267–287.

Novak, T. P., & Stangor, C. (1987). Testing competitive market structures: An application of weighted least squares methodology to brand switching data. *Marketing Science*, *6*(1), 82–97.

Pinkse, J., Slade, M. E., & Brett, C. (2002). Spatial price competition: A semiparametric approach. *Econometrica*, *70*(3), 1111–1153.

Raghavan, U. N., Albert, R., & Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, *76*(3), 036106.

Remer, M. (2012). An empirical investigation of the determinants of asymmetric pricing. *International Journal of Industrial Organization*, *42*, 46–56.

Ringel, D. M., & Skiera, B. (2016). Visualizing asymmetric competition among more than 1,000 products using big search data. *Marketing Science*, *35*(3), 511–534.

Sheikhattar, A., Miran, S., Liu, J., Fritz, J. B., Shamma, S. A., Kanold, P. O., & Babadi, B. (2018). Extracting neuronal functional network dynamics via adaptive granger causality analysis. *PNAS*, *115*(17), E3869–E3878.

Shugan, S. M. (1987). Estimating brand positioning maps using supermarket scanning data. *Journal of Marketing Research*, *24*(1), 1–18.

Siegrist, K. (1997). *Random: Probability, mathematical statistics, stochastic processes*. Retrieved April 25, 2023, from https://www.randomservices.org/random/markov/index.html

Thomadsen, R. (2005). The effect of ownership structure on prices in geographically differentiated industries. *The RAND Journal of Economics*, *36*(4), 908–929.

Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., & Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology*, *93*(2), 1074–1089.

Urban, G. L., Johnson, P. L., & Hauser, J. R. (1984). Testing competitive market structures. *Marketing Science*, *3*(2), 83–112.

Ushchev, P., & Zenou, Y. (2018). Price competition in product variety networks. *Games and Economic Behavior*, *110*, 226–247.

Vitorino, M. A. (2012). Empirical entry games with complementarities: An application to the shopping center industry. *Journal of Marketing Research*, *49*(2), 175–191.

Wang, Z. (2009). (mixed) strategy in oligopoly pricing: Evidence from gasoline price cycles before and under a timing regulation. *Journal of Political Economy*, *117*(6), 987–1030.

Yang, Y., Zhang, K., & Kannan, P. K. (2021). Identifying market structure: A deep network representation learning of social engagement. *Journal of Marketing*, *84*(4), 37–56.

Zhao, P., & Yu, B. (2006). On model selection consistency of lasso. *Journal of Machine Learning Research*, *7*, 2541–2563.

Zimmerman, P. R., Yun, J. M., & Taylor, C. T. (2013). Edgeworth price cycles in gasoline: Evidence from the united states. *Review of Industrial Organization*, *42*, 297–320.