DOES INCREASING CONCENTRATION HIT POORER AREAS MORE? A STUDY OF RETAIL PETROLEUM MARKETS*

Peter L. Ormosi† Farasat A.S. Bokhari‡
Sean Ennis§ Franco Mariuzzo¶

A central tenet in the field of industrial organisation is that increasing/decreasing market concentration is associated with increased/reduced markups. But does this variation affect every consumer to the same extent? Previous literature finds price dispersion exists even for homogeneous goods, at least partially as a result of heterogeneity in consumer engagement with the market. We study this question by linking demographic and income heterogeneity across local areas to the impact of changing market concentration on markups. With 15 years of station-level motor fuel price data from Western Australia and information on instances of local market exit and entry, we apply a non-parametric causal forest approach to explore the heterogeneity in the effect of exit/entry. The paper provides evidence of the distributional effect of changing market concentration. Areas with lower income experience a larger increase in petrol stations' price margin as a result of market exit. On the other hand, entry does not benefit the same low-income areas with a larger reduction in the margin than in high-income areas. Policy implications include a need to further focus on increasing engagement by low-income consumers.

I INTRODUCTION

The relationship between market concentration and prices is among the most intensively researched theoretical and empirical topics in industrial organisation. There is substantial

*We are grateful for comments from Steve Davies, Adriaan Dierx, and David Byrne. The paper has also benefited from comments from participants of the CCP Seminar Series (2020), the Competition Commission South Africa - Economic Research Bureau (2021), the University of Amsterdam conference (2021), Should Wealth and Income Inequality Be a Competition Law Concern?, the Competition and Markets Authority seminar series (2021), EARIE (2021), and the Compass Lexecon Economic Conference (2022). All remaining errors are ours. Data and code used in this paper are available at: https://github.com/PeterOrmosi/distributional_effects.

†Corresponding author. Compass Lexecon, and Norwich Business School, University of East Anglia, NR4 7TJ, Norwich, UK, *e-mail: p.ormosi@uea.ac.uk*. The views expressed in this paper are the sole responsibility of the authors and cannot be attributed Compass Lexecon or any other parties.

‡Loughborough Business School, Loughborough University and Centre for Competition Policy, University of East Anglia, United Kingdom*e-mail: f.bokhari@lboro.ac.uk*

§Norwich Business School and Centre for Competition Policy, University of East Anglia, NR4 7TJ, Norwich,UK.*e-mail:s.ennis@uea.ac.uk*

¶School of Economics and Centre for Competition Policy, University of East Anglia, NR4 7TJ, Norwich, UK.*f.mariuzzo@uea.ac.uk*

evidence showing that increased concentration is associated with higher prices and that more competition lowers prices. This evidence includes analyses of the impacts of changing concentration from market exit and entry. These findings have served as the basis for economic policies to liberalise markets and promote competition for efficiency and consumer welfare benefits.

But this literature has focused mainly on average effects. In this paper, our interest shifts away from these average effects to breaking down the effects by income group. We ask the question: does increasing or falling market concentration affect everyone the same way? To find an answer, we look at a homogeneous good, retail petroleum. Walrasian theory would suggest that in a homogeneous goods market, consumers pay the same price, therefore if entry and exit affect prices, the price change will be the same for every consumer. But extensive search literature has proven that this is not the case if consumers differ in how much they engage with the market [Salop and Stiglitz, 1977; Varian, 1980], and in their willingness to pay [Diamond, 1987]. If finding low prices requires consumers to engage in costly search, and consumers differ in their search costs, and if consumers also differ in their willingness to pay, economic theory and evidence show that in equilibrium it can be optimal for different stores to charge different prices for the same good. Much empirical work supports these arguments [Wilson and Price, 2010; Woodward and Hall, 2012; Allen et al., 2014; Stango and Zinman, 2016; Lach and Moraga-González, 2017]. A stylised and somewhat simplistic synopsis of these papers is that a high willingness to search is associated with lower prices. Logically, it would follow from these findings that if changing market concentration changes the equilibrium price, the price change will reflect this heterogeneity in consumers' engagement with the market.

Our main objective is to find out whether this is the case, and more specifically, whether such heterogeneity in consumer characteristics leads to distributional effects, i.e. do lower-income households pay more or less for increasing concentration, and do they benefit more or less from an increase in competition? This would depend on whether low/high-income households have lower or higher willingness to pay and whether they are more or less likely to search. The answer is not intuitively obvious. For example, households on higher income may be expected to search more as they have higher consumption and so stand to make higher absolute savings from search, but equally, as the opportunity costs of time are likely to be higher for richer consumers, they may choose to search less. Earlier works that have linked search and income - reviewed by Byrne and Martin [2021] - offer mixed evidence, although they point more towards the conclusion that lower-income consumers search less.

To bring additional empirical evidence to this debate, we make use of the geographical and temporal variation of market structure and prices in local petrol retail markets in Western Australia to design a natural experiment exploring the impact of increasing and decreasing local market concentration on the retail margin. In the absence of individual-level purchase data, we link local demographic and census data to petrol stations. To ensure that the local characteristics adequately represent the people who shop at the stations, we limit our analysis to petrol stations in less busy areas, where we can be more certain that dominantly local customers shop at the stations. We then utilise the variation in local demographic characteristics to investigate the relationship between the impact of changing market concentration and demand-side heterogeneity (with a particular focus on income). By employing an event study design, we can adjust the event window in a way that ensures that exit and entry can be considered exogenous in our experiments.

Our paper makes contributions to two main strands of work. First, we draw from the recent empirical literature on price dispersion in retail petroleum. These works are motivated by the observation that price dispersion exists, even with homogeneous goods and where price information is readily available. Closest to our work is Lach and Moraga-González [2017], who look at the relationship between the number of firms in a market and retail petrol price dispersion. Using data from the Dutch petrol retail market they find that price distributions in less competitive markets first-order stochastically dominate price distributions in more competitive markets and that consumer gains from the increasing competition are larger for more informed consumers. Lach and Moraga-González [2017] make an important theoretical contribution, by showing that increased competition has an effect on prices only when it changes the level of consumer informedness. Pennerstorfer et al. [2020] set out on a similar task but assume sequential search, and formulate a theoretical model in which they link consumer information to price dispersion and predict an inverted U-shaped relationship. They test this model with retail petrol price data about Austria. To proxy for consumer information, they assume that drivers with longer commuting distances are more informed of petrol prices than those who commute less. Byrne and de Roos [2017] look at the intensity of search in the petrol market, using the same data as ours - but accessing the number of visits to the FuelWatch website and instead of changes in market structure, investigate search intensity as a function of price dispersion.

The merger retrospective literature on retail petroleum mergers looks at how changing market concentration (through mergers) contributes to price dispersion. The emerging evidence is mixed. Simpson and Taylor [2008] find no evidence of higher prices following exit through mergers, Hastings [2004] and Taylor et al. [2010] find price increases of different magnitudes. Regarding entry, Barron et al. [2004] show that adding one fuel station within a local market (i.e., 2.4 km ring) leads to a price reduction that varies across cities from 1.84 to 5.26 US cents per litre. However, Hosken et al. [2008] use a larger dataset and find no relationship between firm density and market price.

To add to this stream of empirical literature, instead of focusing on the magnitude of price (margin) dispersion, we look at the impact of changes in market concentration on the expected price (margin). We do this partly because we do not observe individual shopping decisions, but also because we do not observe the exact location of consumers within each local area. Instead, our interest is in the asymmetric impact of a change in the competitive pressure exerted on firms, conditional on local average consumer characteristics. Unlike most previous works, we use information on the time of exit/entry in each local market (as opposed to exploring different levels of concentrations in a cross-section of local markets) to estimate their impact on the price of each petrol station that remained (or had been incumbent) in the market. Our quasi-experimental design allows us to look at how the retail margin of the petrol stations that remain in the market change, after one petrol station exits from their proximity.

The second main strand of relevant works where we offer new evidence is related to the distributional impact of market power. Baker and Salop [2015] set out the problem at a conceptual level and offered an agenda for further work. Some of these works focus on the link between market power and inequality [Khan and Vaheesan, 2017; Ennis et al., 2019]. Despite the increased focus, much of the currently available analysis is either purely conceptual or is based on empirical work with aggregated macro data [Dierx et al., 2017; Ennis et al., 2019; Zac et al., 2020].

There is much less evidence at the market level, which is hardly surprising; linking changes in market concentration to different demographic groups is not a trivial exercise given its intensive data requirements. To offer more micro-level evidence we approach this problem differently, motivated by the question: do the poor pay more for goods and services, the leading thought behind Caplovitz [1963]. Nevertheless, a handful of papers have taken a micro level look to analyse differences in low- and high-income consumers. Handbury [2021] looked at product offerings in poor and wealthy areas and found that relative to low-income households, high-income households enjoy 40 percent higher utility per dollar expenditure in wealthy cities, relative to poor cities, and much of this variation is driven by the range of product offerings in these local areas. Although it is not their primary focus, Allen et al. [2014] find (using individual-level data from mortgage markets) that the average effect of mergers underestimates the increase in market power, and they show that competition benefits only consumers at the bottom and middle of the transaction price distribution.

Unlike these above works, we focus on a single, homogeneous product, retail petrol, to offer evidence of the difference between low- and high-income consumers. But whereas in Handbury [2021] and Allen et al. [2014] the main drivers of the difference are supply side factors, in our paper we argue that the heterogeneity in how changing market structure affects different consumer groups is driven (at least partially) by demand-side factors (similarly to Lach and Moraga-González [2017]).

On the methodological side, we draw from to the literature on causal forests to estimate heterogeneous treatment effects, as proposed by Athey et al. [2019]. Whereas a large number of location, firm, and time characteristics in our data raise dimensionality issues, which would justify the use of a tree-based approach, at the same time, we have a relatively small cross-section of exits and entries in our sample. To handle this problem, we propose using an ensemble causal forest approach. We demonstrate through simulations, that this performs better than a single causal forest in cases with low number of observations and large number of estimable parameters.

Using a retail petrol price dataset of Western Australian price comparison website, FuelWatch, between 2001 and 2019, we find that, in line with conventional industrial organisation theory, exit leads to an increase (although not significant on the average), and entry causes a drop in the retail margin. When looking at the heterogeneity in these estimates, we find that although low- and high-income areas do not significantly differ in the number of petrol stations (and the composition of these stations), low-income areas experience a larger (and significant) increase in the price margin with exit. At the same time, they do not benefit from a lower drop in the margin with entry. Other factors, such as level of competition, reliance on cars, commuting distance, age, or education also drive some of the heterogeneity but even after controlling for these factors, the difference between low and high-income households remains. This suggests that unobserved differences between low- and high-income households also contribute to differences in market engagement across these groups.

These findings offer evidence that a reconsideration of some of the conventional thinking around competition policies may be warranted. The lack of engagement of lower-income consumers with the market suggests that conventional antitrust policy tools may not be able to attain their objectives of improving consumer welfare. Instead, antitrust should not only focus on restoring the level of competition on the supply-side (for example through enforcement action, or by breaking up monopolies) but assign increased priority to improved demand-side remedies to enhance consumer engagement.

The paper is organised as follows. First, we provide a stylised economic framework, which pulls together some of the canonical theories from previous literature. This is followed by an introduction and description of our data, and a discussion of the methodology. We then present the results of our causal forest estimates before offering results from a linear regression, which also allows us to offer results that account for the potential endogeneity of exit and entry.

II THE ECONOMIC FRAMEWORK

The theoretical motivation of this work links to the vast search literature, which highlights that differences in search and decision costs are likely to influence consumer engagement, and therefore even prices of homogeneous goods display some dispersion.[1] Some of these works assume sequential search, such as Pennerstorfer et al. [2020], who build on Varian [1980] and Stahl [1989] to model search in homogeneous goods, where consumers differ in their degree of informedness, and extend this model where consumers differ also in their willingness to pay for the product. This allows distinction between consumers based on how informed they are about the price of petrol. For some, obtaining an additional price quote is costly; others are aware of all prices charged in the relevant market as they have access to the 'clearinghouse'. This setup leads to a mixed equilibrium due to the tension between charging a high price to exploit uninformed consumers and charging a low price to attract informed consumers. As a result, the authors find an inverse-U shape relationship between price dispersion and the proportion of informed consumers.

Whilst sequential search models could be relevant to the study of petrol price dispersion in certain settings, our view is that it does not fit our purpose for three reasons. Firstly, in our case consumers are unlikely to search sequentially. Instead it is probably more fitting to assume that consumers do not drive around searching for petrol prices, but instead obtain price quotes in a more passive way, for example through their daily commutes. Secondly, in our setting consumers also have access to the FuelWatch price comparison tool, which allows an easy online comparison of prices. Thirdly, as our focus is on consumer heterogeneity, we do not want to assume that uninformed consumers within a market face the same search cost. It is plausible that regional income differences are not only related to the share of informed consumers and differences in willingness to pay, but also to search cost differences (we indeed proxy for this heterogeneity in search costs through commuting distance and access to internet). Finally, unlike Pennerstorfer et al. [2020], our focus is on the impact of

[1] It may seem obvious to link our work to works on the demand elasticity of petrol, such as Wadud et al. [2010], who provide estimates of motor fuel elasticities for different income levels. They look at the heterogeneity in petroleum demand elasticity, and find, among others an inverse relationship between income and demand elasticity. Whilst this bears some relevance to our study, it is tangential to our research question, as we are interested in how consumers choose between different suppliers, i.e. the brand elasticity of petrol demand, rather than product elasticity.

changing market structure, rather than the change in price dispersion.

Our theoretical inspiration is closer to that of Lach and Moraga-González [2017], who proposed using a generalisation of Varian [1980] and Armstrong et al. [2009] in a way that allows for richer heterogeneity in consumer price information. To do this, they rely on a probability generating function for the number of prices observed by consumers. In their interpretation consumers differ in driving and commuting patterns, as well as in attentiveness to posted prices, but the formulation of their model is general enough to capture other sources of information. Because of the conflicting forces of the desire to steal business from its competitors by offering better deals, and the willingness to extracting surplus from consumers who do not compare prices, the market is characterised by a mixed strategy equilibrium. In this setting, if the amount of price information consumers have is heterogeneous, prices are typically dispersed in equilibrium. For this reason, a change in the number of suppliers does not simply affect the average price, but the distribution of prices across this heterogeneous set of consumers. Our paper offers an empirical test of this important contribution.

We do not observe individual level data on fuel purchases and on consumer informedness. Instead, we observe local variation, both in prices and in our proxies of how informed consumers are (such as differences in commuting habits, internet access, or income). This variation would imply that some of our consumers (area average) are more informed than others.[2] Therefore in our study design the unit of observation is not the individual, but the area, and we look at price dispersion and variation in consumer heterogeneity across these areas. Market concentration changes over time in these areas (through exits and entries) and we test the price impact of a change in market concentration, conditional on a given level of consumer informedness.

Although Lach and Moraga-González [2017] do not explicitly incorporate income, we assume that the level of informedness (i.e. the heterogeneity in search costs) is related to income for the following reasons. Motor fuel is a non-discretionary part of household expenditure. Such products (similar to rent or food) typically display significant distributional differences in that they constitute a larger fraction of poorer households' expenditure.[3] Moreover, low-income households may be less likely to be able to switch to more expensive substitutes, which may require replacing the car or changing working habits if the price of petroleum goes up.

II(i) *Income and market engagement*

Although it may seem counter-intuitive at first glance that there is less engagement with the market in lower income areas, the literature has identified various frictions that can counter the conventional neoclassical view that lower income consumers are more sensitive to price changes.

Lower-income consumers often face challenges that limit their ability to benefit from market competition. Certain factors can disproportionately affect lower-income consumers,

[2] Of course, in the absence of individual level data, we do not know which consumer drives past exactly which petrol station. Moreover, we also do not observe the heterogeneity in consumer attentiveness of the prices they drive past.

[3] For the UK, Mattioli et al. [2018] finds that the poorest households often spend around 20% of their income on motor fuel, and frequently more than this.

hindering their ability to reap the benefits from competition. Byrne and Martin [2021] provide a carefully constructed review of the relevant literature and concludes that most evidence points in the direction that low-income households engage less with the market.

More specifically, low-income consumers/households are associated with lower propensity to search [Kiser, 2002; De los Santos, 2018; Nishida and Remer, 2018; Eizenberg et al., 2021]. Low-income consumers search less for example because search may be more challenging or costly for lower-income consumers if there's a longer physical distance to travel [Allen et al., 2019; Chung and Myers Jr, 1999; Eizenberg et al., 2021], if they have less access to online search platforms [Goldfarb and Prince, 2008], or if they are less able to search or negotiate effectively [Stango and Zinman, 2023; Letzler et al., 2017; Calvet et al., 2009; Morton et al., 2003].

In a study on low-income consumers, the UK Competition and Markets authority highlighted that low-income consumers may lack access to information about the products and services available in the market.[4] They might not have the resources or time to research various options, compare prices, or make informed decisions, which can lead to sub-optimal choices.[5] Low-income consumers may also face additional transportation costs and time constraints.

The supply-side of the market may strategically exploit this relative consumer inertia. Reviewing the relevant literature, Byrne and Martin [2021] refer to evidence from Allen et al. [2019] and Byrne et al. [2022] in arguing that suppliers may make more expensive offers if their customers have higher expected observable search costs.

Based on the above literature, in the context of motor fuel we have three reasons to think that prices may be higher in low-income areas. Firstly, motor fuel is a non-discretionary part of household expenditure. Such products (similar to rent or food) typically display significant distributional differences in that it constitutes a larger fraction of poorer households' expenditure. For example in the UK, Mattioli et al. [2018] finds that the poorest households typically spend around 20% of their income on motor fuel, and frequently more than this. At an intuitive level, higher motor fuel prices eventually encourage the average consumer to cut back on driving or switch to more fuel-efficient vehicles. However in the short-run, low-income households may have few options but to continue buying motor fuel and cut back on other expenditures (or get further into debt).

Secondly, low-income consumers could also encounter challenges when seeking out favourable deals if they have limited access to information on prices, which could be if they have more limited access to technology (to access price comparison websites like FuelWatch), or they drive, and therefore search less.

Finally, there could be unobserved differences between low- and high-income consumers, for example if the former are more likely exposed to behavioural biases that hinder their ability to search effectively.

Some of the considerations in the previous literature apply in other industries but not in our specific case. For example, the literature has documented evidence on 'retail deserts' in other markets [Allen et al., 2019; Chung and Myers Jr, 1999; Eizenberg et al., 2021], and also lower competition in rural areas [Alm et al., 2009]. However, we show in our descriptive analysis that we do not find such a difference in the number of rivals in our specific case of

[4] The full report is available online at the following URL: `https://tinyurl.com/32e4ad5p`.

[5] See also Du and Ma, 2022 for an analysis applied to energy consumption in China.

petrol retail in Western Australia.

III PETROL RETAIL MARKETS AND THE DATA

The Australian petrol retail market is characterised by a small number of very large, and many fringe players. These can be divided into three distinct types. Refiner-wholesalers are vertically integrated retailers such as BP, Caltex, Mobil, and Viva Energy/Shell. This includes refiner–wholesaler controlled sites and independently operated but refiner–wholesaler branded sites. Large independent retail chains are independent retailers such as 7-Eleven, United, Puma Energy, and On The Run. Some supermarkets also sell fuel, such as Coles Express and Woolworths. At the country level, the combined retail market share (based on sales volume) of the large vertically integrated firms dropped significantly from over 80% in 2002, to under 40% in 2017. During the same period, the market share of supermarkets and independents increased substantially. Regarding the individual brands, Shell/Viva Energy (trading under Coles Express) and Woolworths (a supermarket) had a respective market share of 20-25% over our study period, followed by BP and Caltex, just under the 20% mark. The remaining sales volume is supplied by independent retailers.[6] Regarding the number of retail units in our sample, BP (260 stores) Caltex (254 stores), and Shell (216 stores including Coles) are the largest.

The main component of our data is daily prices at petroleum retail outlets in Western Australia, for the period 2001-2019, which was downloaded from FuelWatch.[7] FuelWatch is a price comparison website to motorists in Western Australia. At the time of introduction, the website was a response to policy concerns about the levels of price dispersion in the country, implying that some consumers would have been paying largely over the odds. Byrne et al. [2018] offers a detailed description of the FuelWatch data, here we focus on the most important features that are relevant for this paper. FuelWatch is empowered to monitor and report on WA wholesale and retail fuel prices under the Petroleum Products Pricing Act 1983 (amended to accomodate the creation of FuelWatch in 2000/2001). The Act gives the Government authority to enforce price transparency.

FuelWatch was launched in 2001, and its scope was largely extended in 2003. Today it covers approximately 80% of regional and 100% of metropolitan retail outlets in Western Australia. The website collects and shares information about the geographical location of the retail outlet (precise address), the brand of the operator, and prices for a range of motor-fuel products, such as unleaded petrol (ULP), premium unleaded petrol (PULP), 95 RON (octane) petrol, diesel, branded diesel, and liquefied petroleum gas (LPG). Not all outlets sell the whole range of products. The way consumers can access FuelWatch has significantly improved since its launch and has been used to feed into various smartphone apps going back to 2010. Through FuelWatch, consumers have free access to the next day's petrol prices at almost all petrol stations, reducing switching costs for consumers who use the internet to search for the best petrol deals.

We had a total of 15,638,524 observations of daily petrol station level prices across all products at 1299 (1051) petrol stations. For some of these, there were ownership changes in

[6] https://www.accc.gov.au/system/files/Petrol-market-shares-report.pdf.
[7] https://www.fuelwatch.wa.gov.au

our sample period, which is why we have 1299 distinct petrol stations but only 1051 unique forecourts. Most of the price data is available for two products, unleaded petrol (ULP) (3,964,180) and diesel (3,811,105). As we are not directly interested in the daily variation of prices, and also to eliminate issues from rogue missing observations, we averaged the price data at the weekly level and limited our focus to the most popular product, ULP. The main reason we excluded diesel from our analysis is that we do not observe individual purchases, but instead use variation in local demographic factors around petrol stations to answer our research question. Because diesel has a disproportionate number of purchases from commercial vehicles, it makes local demographics less relevant for fuel purchases.

We also collected weekly average wholesale prices for the Western Australian region. We acknowledge that not all retail units pay the same wholesale price. Vertically integrated companies have better distribution systems and lower costs than independents. However, this brand-level wholesale price information is not publicly available. Instead, in our estimates, we will control for the brand to account for this cost variation. The wholesale price data was only available from 01 January 2004 onwards, which further reduced our sample size, leaving us with 489,721 observations of weekly petrol station level prices.

Figure 1 shows the over-time variation of ULP prices. There is significant seasonality in the data. We removed weekly and yearly seasonality for two reasons: (1) our main empirical method of causal forests would not allow us to control for time fixed effects,[8] and (2) our interest is in the immediate shocks as a result of exit/entry net of seasonal trends.

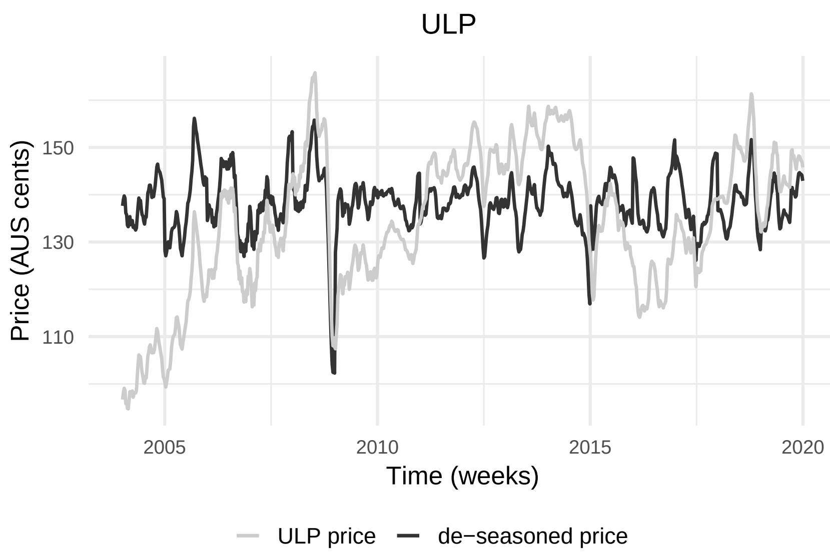

Figure 1
Weekly retail ULP prices with and without de-seasoning

For each petrol station, we acquired the longitude and latitude coordinates (using its address) and applied the Haversine formula,[9] to identify which petrol stations are located

[8]Cohen et al. [2023] also de-seasonalise and standardise their outcome measure for a causal forest analysis.
[9]The distance over the earth's surface. We could have used more sophisticated distance measures (e.g.

within 1, 2, and 5 miles from each other. To give an example, Figure 2 shows the number of competitors a selected independent petrol station had within a 1, 2, and 5-mile radius in our observed study period. Not all of these petrol stations were always available and competing in the entire sample period. Figure 3 illustrates this for the same independent petrol station. It shows that on 1 Jan 2004 it had 4 active rivals (Mobil, BP, Ampol, Puma). Then in early 2004 Ampol exited the market. In 2008 Puma also left the market, and in 2010 BP also left. Later in 2010 BP, and soon after that Ampol re-opened. Please note that there are two exits close together in a short span. To avoid any complications due to confounding effects, we removed exits of this type from our sample (see Section III(i)).

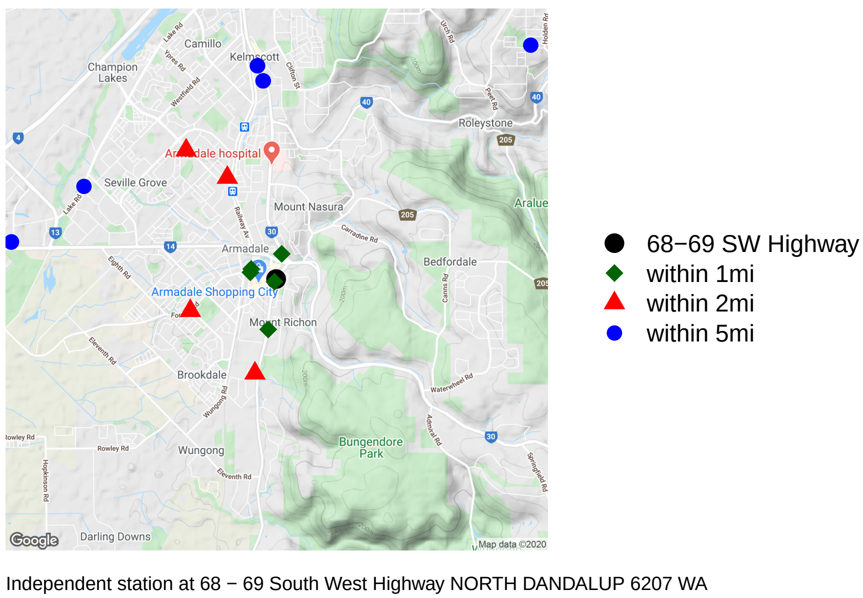

Figure 2
Example petrol station and surrounding competition

We do not observe individual purchases of ULP, neither do we know where exactly each consumer is located. Previous works on price dispersion looked at the distribution of prices within geographical areas around petrol stations. These were following the idea that engaged consumers would find the cheapest prices, therefore what matters for them is not the average, but the lowest price in their proximity. Admittedly, one weakness of this approach, is that because the location of consumers is unobserved (with the exception of Pennerstorfer et al. [2020] who observe consumers commuting routes), these papers look at a radius around petrol stations and assume that the consumer is located exactly where the petrol station is, therefore what falls within a given radius of the petrol station also falls within a similar radius for the consumer and can be thought of as the local market (put differently, it assumes that that the same geographical market applies to all consumers within the boundaries of this radius). This is unlikely to be true in our case (especially in

the driving distance between two places) but our objective was not to precisely estimate the relationship between travelling distance and shopping behaviour, rather look at how a change in the number of petrol stations in proximity of a petrol station affects prices.

spread-out rural areas).

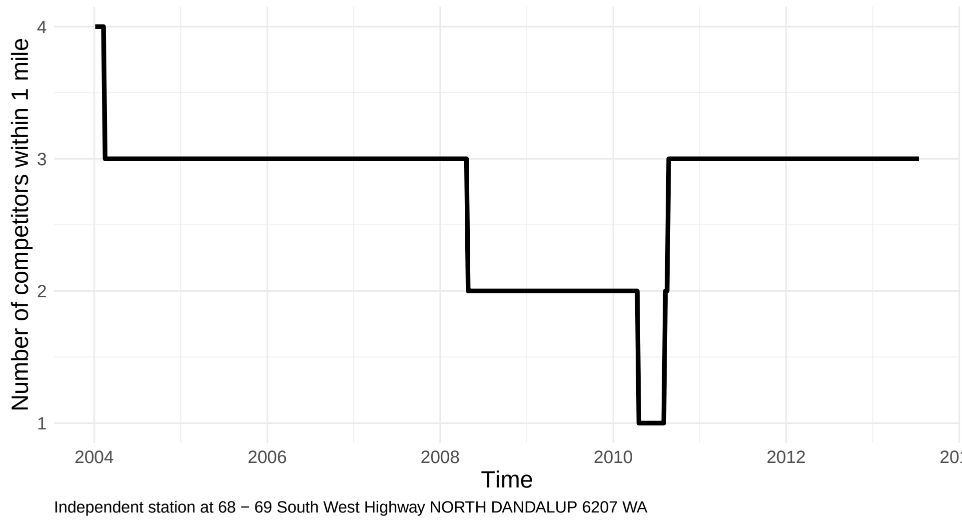

Figure 3
Example of entry and exit around a specific petrol station

To avoid this issue, and also because we employ a different research design, we do not look at price distributions around each firm. Instead we take a different approach, in which we look at each petrol station individually, and assume that their prices are affected (directly) by the characteristics of consumers around them, and (indirectly) by the number of rivals in their vicinity. More precisely, our focus is on how the price of each firm that remain in the market (following the exit of a station), or had already been in the market (before the entry of a new petrol station) changes in response to exit/entry, given this heterogeneity in the composition of nearby consumers.

We have 141 petrol stations for which the first observed reported price data dates after the start of our relevant study period (01 January 2004). But from the data available to us we had no way of telling whether a station entered the market during the study period (and had not existed before) or whether they just joined the FuelWatch scheme later. For this reason we removed these first entries (i.e. the first appearance in our dataset).

Figure 4 helps understand our approach of preparing the data for our study design. For each firm in Western Australia we look at whether there had been an exit/entry within their vicinity (1, 2, and 5 mile radii). For example, on Figure 4 petrol station A witnessed an exit within 1 miles. We record this as a treatment for A, and collect local level information for the area where A is located (which will include for example, average household income, or that the number of rivals A had within 1 mile, changed from 3 to 2). For B however there was no exit within 1 mile (note that an analysis similar to Lach and Moraga-González [2017] would have needed us to look at the price at B when looking at price dispersion around A, even though under our assumption B was not impacted by the exit).

Of course, looking only at the number of competitors masks information about the identity of the competitor. For example, the petrol station used as an example in Figures 2 and 3 had 5 BP stations competing within a 5-mile radius, therefore the finding that

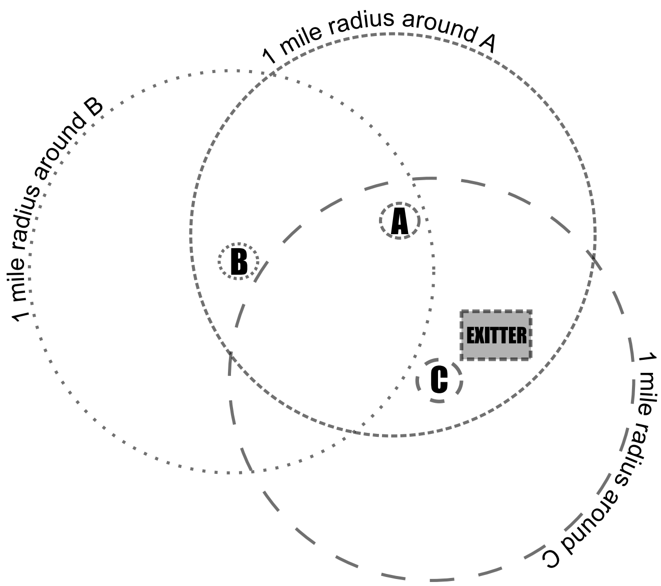

Figure 4
An illustrative example of our local market definition

the BP station within 1 mile from the independent station decided to exit may have to be interpreted in this context. We deal with this by introducing variables such as the number of same brand competitors in the area, and a vertical chain dummy.

III(i) *The meaning of exit/entry*

We created our data on entry/exit using information on prices reported by petrol stations. Because under the Petroleum Products Pricing Act 1983, petrol retailers must notify the Petrol Commissioner (appointed by the Australian Competition and Consumer Commission) of their prices on a daily basis, we assume that if prices are not reported for a longer period, it is a sign of temporary or permanent exit.

We defined exit as the reduction, and entry as the increase in the number of rivals a petrol station has within a 1-mile radius, (in the Appendix we provide our main results for exit/entry within 2 and 5-mile radii). To ensure that the reduction/increase in rival numbers reflects a genuine and lasting change in market structure, and not just missing data or a short-term, temporary shutdown of one of the stations (for example for small restoration or development work), we removed from our sample all instances of exit and entry, where the same petrol station, or any other stations exited or entered the vicinity of another petrol station in our pre-, and post-treatment periods (26 weeks before and after the treatment). This also ensured that there was no confounding effect from another change in local market structure well before and after the treatment. Figure 3 offers a visual representation of these market structure dynamics. It shows instances of exit and entry over a 10 year period around one of the petrol stations, and shows instances of short spells of “market exit”.

The reduced sample included 392 instances of exit and 354 instances of local market entry. Figure 5 shows the annual distribution of these exit and entry events. Over time the number of exits drops, partly because with fewer petrol stations in the market, there are fewer potential stations to exit. Entries happen at more stable rate over time. The figure also displays the quarterly breakdown of exit and entry. This shows an increased number of exits and entry in the second (exit) and third (entry) quarters of the year, which is likely to do with the end of the tax year (end of June), i.e. more exit happens before the start of a new tax year, and more entry happens after the start of a new tax year.

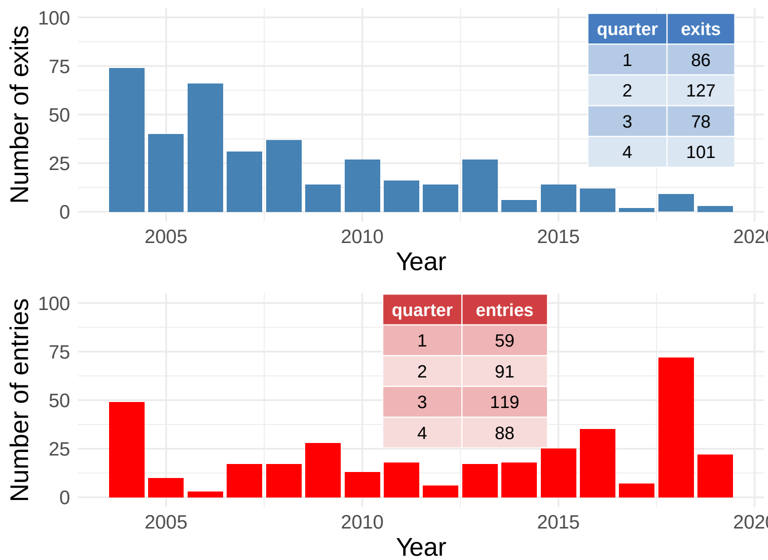

Figure 5
Number of entries and exits by year and quarter

III(ii) *Local areas*

As our next step we link the price and market structure data to local characteristics. These characteristics are available at Statistical Area Level.[10] We do this by connecting the petrol station postcodes to their corresponding SA2 areas.

There are 137 distinct SA2 areas in our sample, which include 195 distinct postcode areas. SA2s generally have a population range of 3,000 to 25,000 persons, with an average population of about 12,000 people in our sample. SA2s in remote and regional areas generally have smaller populations than those in urban areas. Using the SA2 code, we then link each petrol station to local characteristics, using data from the 2016/2011/2006 Australian censuses, and the Personal Income in Australia report of the Australian Bureau of Statistics (ABS). We match this data with the corresponding petrol stations for each census data (prices before 2008 were matched to the 2006 census, prices between 2009 and 2013 were linked to the 2011 census, and prices from 2014 onwards were linked to the 2016 census).[11]

[10] https://www.abs.gov.au/websitedbs/d3310114.nsf/home/australian+statistical+geography+standard+(asgs)

[11] The reason we did not use an exact matching is that demographic features are unlikely to change right

The full list of our variables and their summary statistics are given in Tables B.1 and B.2 in the Appendix.

Income: We record the annual taxable income at the postcode level for the study period 2004-2019. Unlike some of the other area characteristics, this data is time-variant (annual). The data is collected from the taxation statistics of the Australian Taxation Office.[12]. For individual income, the coverage is complete for 2004-2018. Although not at the heart of our analysis, we also control for business income. For this we have net business income for 2004-2018, and net rent for the same period. For the other business income variables, data is available between 2011-2018. For these, for the period 2004-2010, we assumed the same value as in 2011.

Search: We measure search through two main variables, the median commuting distance in an area, and the level of home internet penetration in the area.[13] These two measures account for two dimensions of search: (1) people who drive more to work have a lower opportunity cost of search, as they already survey the prices as they drive past them; (2) people with home internet access are more likely to engage with the FuelWatch price comparison tool. Both can be thought of as different dimensions of search costs. Internet access allows consumers to search online. In households without internet access, search has to be physical, which is associated with higher costs. Several previous works have looked at Internet use as a potential proxy for consumer informedness [Brown and Goolsbee, 2002; Tang et al., 2010; De los Santos et al., 2012; Sengupta and Wiggins, 2014] Commuting distance allows us to incorporate the cost of physical search: the longer someone commutes to work, the more petrol stations they sample without incurring extra search costs. Previous empirical works that draw on information on commuting patterns to study heterogeneity in petrol prices include Cooper and Jones [2007]; Houde [2012]; Pennerstorfer et al. [2020].

ABS indices: The Australian Bureau of Statistics introduced several indices to measure the economic and social conditions in an area. The Index of Relative Socio-economic Disadvantage is a general socio-economic index. A low score indicates a relatively greater disadvantage in general. For example, an area could have a low score if there are many households with low income, or many people with no qualifications, or many people in low-skill occupations. The Index of Education (the level of qualification achieved or whether further education is being undertaken) and Occupation (classifies the workforce into the major groups and skill levels) reflect the educational and occupational level of communities. This index does not include any income variables. The Index of Economic Resources is a proxy for the financial aspects of relative socio-economic advantage and disadvantage (it summarises variables related to income and wealth). This index excludes education and occupation variables.

Other area characteristics: We have data on the age structure in each SA2 area (age, and % of people in various age brackets), the level of education, the average, and the mean commuting distance, and the means of commuting.

Brand: As we have a homogeneous product, there should not be much quality variation

at the time of taking the census. For example the local demographic characteristics in 2015 are likely to be better represented by the 2016 census rather than the 2011 one.

[12] https://data.gov.au/data/dataset/taxation-statistics-postcode-data

[13] We only had data on home internet. We believe this gives us a reasonable proxy for internet access for at least the first half of our study period (2004-2019).

in the actual product, but there might be in the services linked to the product. The same good sold in two different stores could also be differentiated by the retail environment in which it is sold, with ‘high-quality’ stores charging more. Controlling for brands allows one to control for quality variation noise in our price variation data. Implicitly this assumes that heterogeneity in quality might exist across, but not within brands.

Note that we have much more diverse data on demand-side factors than supply-side ones. However, even in the absence of firm-level data, we can allow for supply-side heterogeneity by controlling for area-level supply-related factors, such as the average business expense in an area, the average business tax paid, the average business income, and the average rent paid by businesses.

III(iii) *Linking local characteristics to petrol stations*

To ensure that the above local characteristics are indicative of the shoppers who purchase petrol at each petrol station, we took two further steps.

Firstly, we removed from our sample all petrol stations that were on highways, where a substantial consumption typically comes from non-local shoppers.

Secondly, we collected traffic information in the vicinity of each of our petrol stations. For this we used data from 4461 traffic count sites in Western Australia.[14] This data expresses the average number of vehicles and heavy vehicles at each location on a day. For each of our petrol stations, we took all traffic count sites within a 1, 2, and 5 mile radius, and summed up the traffic count from these nearby count sites.

Table I
Traffic counts and highway address

	number of petrol stations	average traffic count from metres within		
		1 mile	2 miles	5 miles
not highway	861	13467	13552	13637
highway	190	11221	10948	10700

Table I shows that there is more traffic in the vicinity of petrol stations that are not located on a highway. The table confirms our strategy to remove both highway stations (because of the high proportion of pass-through traffic), and the highest traffic intensity petrol stations (because it is likely that people from other geographic areas shop at these stations). The table also reveals that the average traffic-metre level metric gives a stable measure of the level of traffic around petrol stations, irrespective of whether we include traffic-metres from further away from the petrol station.

With the help of this traffic data, we removed the busiest quarter of petrol stations from our sample to leave us with the less busy ones, where it is more reasonable to believe that most of the shoppers are from local areas. This exercise left us with 599 unique petrol stations (299 exits and 191 entries).

[14] https://data.aurin.org.au/dataset/wa-govt-mrwa-mrwa-traffic-digest-2018-na

IV DESCRIPTIVE ANALYSIS AND STUDY DESIGN

IV(i) *Descriptive analysis*

As shown in previous works, there can be substantial dispersion in the price of petrol [Pennerstorfer et al., 2020; Byrne and de Roos, 2017]. To get a better understanding of the source of dispersion in our data, this section presents some descriptive information on the retail margin.

First, Table II confirms conventional IO theoretical and empirical wisdom on how the retail margin varies with the level of competition.[15]

Table II
Margin by number of rivals

number of rivals within 1 mile	0	1	2	3	4	5	6	7	8	9	10
ulp margin	1.139	1.118	1.113	1.124	1.111	1.122	1.120	1.101	1.099	1.074	1.092

Table III shows the margins for different levels of income and competition. Low and high in both cases are defined as split around the respective median values. These averages suggest that lower income areas are associated with higher petrol prices, and the difference seems larger in low-competition areas. We posited above that some of this price dispersion may be due to the heterogeneity in the level of engagement with the market.

Table III
Mean margin by level of competition and income

	low income	high income
low competition	1.154 (0.073)	1.138 (0.09)
high competition	1.114 (0.049)	1.101 (0.058)

Standard deviation in parentheses.

Table IV shows that averaging across the total sample, low-income areas experience higher margins in general. Margins are also higher in low-education areas, areas where people commute less, areas with less home internet penetration, and places with lower educational levels or a higher proportion of people over 65.[16]

The frequency of entry and exit also varies across the different areas (Table V). It appears that areas with more competitors see more shifts in market structure. Moreover, low-income areas are associated with proportionately more exit in high-competition areas. high-income areas on the other hand see proportionately slightly more entries.[17]

[15] For an overview of how margins differ depending on the level of competition within 1, 2, and 5 miles, see B.3 in the Appendix.

[16] Table B.4 in the Appendix adds more detail about these margins.

[17] In the Appendix, B.5 shows the number of exits/entries by brand.

Table IV
ULP margin by low and high levels of key factors

	low	high
competition 1mi	1.123	1.118
	(0.078)	(0.066)
competition 2mi	1.142	1.099
	(0.085)	(0.053)
competition 5mi	1.151	1.09
	(0.086)	(0.043)
income	1.134	1.113
	(0.076)	(0.074)
education	1.137	1.111
	(0.082)	(0.067)
% of people +65 age	1.132	1.114
	(0.09)	(0.058)
commuting distance	1.147	1.101
	(0.088)	(0.052)
% people internet home	1.136	1.111
	(0.086)	(0.062)

Standard deviation in parentheses.

Table V
The ratio of exits and entries to the total number of petrol stations by income, competition and population size

		low income		high income	
		exits	petrol stations	exits	petrol stations
exits	low competition	72	205	57	160
	high competition	60	101	51	115
entries	low competition	64	205	60	160
	high competition	52	101	53	115

There is also significant variation in the price margin across the brands. Small independent stores operate with the highest margins, but the large vertically integrated companies (BP, Caltex, Shell/Coles) are also in the top third. The bottom half of the distribution (lowest margins) constitutes mainly independent chains (See Table B.8 in the Appendix). This would suggest that cost-efficiency is likely to be dominated by other factors when it comes to setting the margin, as vertically integrated companies are likely to have lower retail costs, but still choose to have a high margin. Some of these differences may be explained by local cost conditions. Stores in urban areas may face higher rental prices and labour costs, which would lead them to charge more, though they might also face more intense local competition, leading them to charge less. Independent 'corner' shops are unable to exploit economies of scale in wholesale purchasing or other costs and so charge higher prices than large, national retailers for the same product. If poor households are concentrated in areas with high-retail costs, this may explain any finding that poorest areas pay more.

It is also possible that firms behave strategically when choosing whether to open stations in rich or poor areas, which means that in some areas, particularly in rural or under-served communities, there may be limited competition due to geographical or infrastructural constraints.

Table VI
Margin by brand by competition by income

station	low income	med income	high income	station	low income	med income	high income
BP	1.169	1.099	1.136	Coles Express	1.108	1.09	1.112
n	66	61	70	n	15	15	18
comp 1mi	2.476	1.548	2.215	comp 1mi	2.356	2.203	2.431
comp 5mi	13.206	28.724	28.633	comp 5mi	31.78	43.328	32.121
Caltex	1.14	1.099	1.12	Shell	1.169	1.1	1.133
n	71	59	64	n	51	33	45
comp 1mi	2.254	1.646	2	comp 1mi	2.129	1.434	1.659
comp 5mi	12.741	30.465	26.431	comp 5mi	11.983	19.552	27.305

Table VI shows that this does not seem to be the case in our data.[18] Looking at the largest brands we can see that low-income and high-income areas have similar number of competitors, but at the same time also higher margins. To further confirm this, Table VII shows that when the market is defined as a 1, or a 2-mile radius, the number of rivals is similar in low- and in high-income areas (there is a difference when one looks at the significantly wider 5-mile radius geographical market). Moreover, there seems to be a difference in consumer informedness, with high-income areas associated with better informed (longer commute and more internet) consumers.

Table VII
Main data features by income groups

	N within 1mi	N within 2mi	N within 5mi	internet	commute
low income	2.884	5.487	15.632	0.237	4.225
	(1.771)	(3.808)	(16.935)	(0.062)	(2.622)
high income	2.755	6.197	27.132	0.259	4.978
	(1.584)	(4.037)	(21.372)	(0.065)	(2.993)

Standard deviation in parentheses.

Although these descriptive tables are useful for understanding the data, to test our

[18]Table B.6 in the Appendix shows the same data for all brands.

hypotheses, we need an approach that brings together all possible effects into the same model. This is what we set up with our study design.

IV(ii) *Study design*

We estimate the causal impact of the exit and entry of petrol retail forecourts on the retail margin and investigate the heterogeneity of these estimates across different area characteristics. Because we do not observe the same markets with and without exit/entry at the same time, we rely on observational data and employ an event study design to line up all relevant events (exits and entries).

We created an event window (15 weeks before and 15 weeks after the treatment), which we applied to every petrol station that experienced exit/entry.

In our event-study design, we align each instance of local market exit and entry for the ULP price data. This creates unbalanced panel datasets of treated and units for the years 2004-2019 with varying dates of treatment application. Each instance of treatment (exit and entry) is therefore defined as petrol stations that had another station exit or enter the market within a 1-mile radius. For identification, we draw from the sample of non-treated units (petrol stations that did not experience a change in the number of rivals $\pm$ 26 weeks from the time of the treatment), to design a control group that can be used as a stand-in for the outcome that would have happened in the absence of exit/entry.

Studies with a similar research design, that are only interested in the average treatment effect often average over all these potential control units. In our case, we are interested in the heterogeneous treatment effect. Averaging the control units would mean averaging their characteristics, making them ill-suited for our purpose. Instead, we decided to take the most similar petrol stations (based on the observable features in the 25 weeks period before exit/entry). For this we employed nearest neighbour matching, using the propensity score difference to specify distances from the treatment petrol station, and selecting the petrol station with the lowest distance. In a set of experiments, we looked at the nearest 2, 5, 10 neighbour(s) (each treatment petrol station is matched with 2, 5, 10 control petrol stations) but in our main discussion, we focus on the nearest 5 neighbours. We offer a sensitivity analysis of this choice in the Appendix. This gives us, for each instance of exit and entry, a set of 6 petrol stations (1 treatment and 5 control). Table B.7 in the Appendix compares the average and the standard deviation of the treatment and control groups to demonstrate similarity on observables.

In observational studies, a cursory look at the raw data can often provide a useful indication of whether the testable hypotheses hold. Figure 6 shows how the average retail margin varies for the treatment and control groups for ULP for the lowest and highest income terciles. The vertical lines represent the time of exit. We report a longer (25 week) time horizon after the event to provide more information on the longer term price-path following the event and to demonstrate that the effects are not temporary (acknowledging that the further away we are from the event, the more likely that confounding events also come to play). Following exit, the price margin assumes a higher level over time, but only in the low-income areas.

Figure 7 shows that following entry, the ULP margin continues at a lower level for both low- and high-income areas. In both areas an immediate large drop in prices is followed by

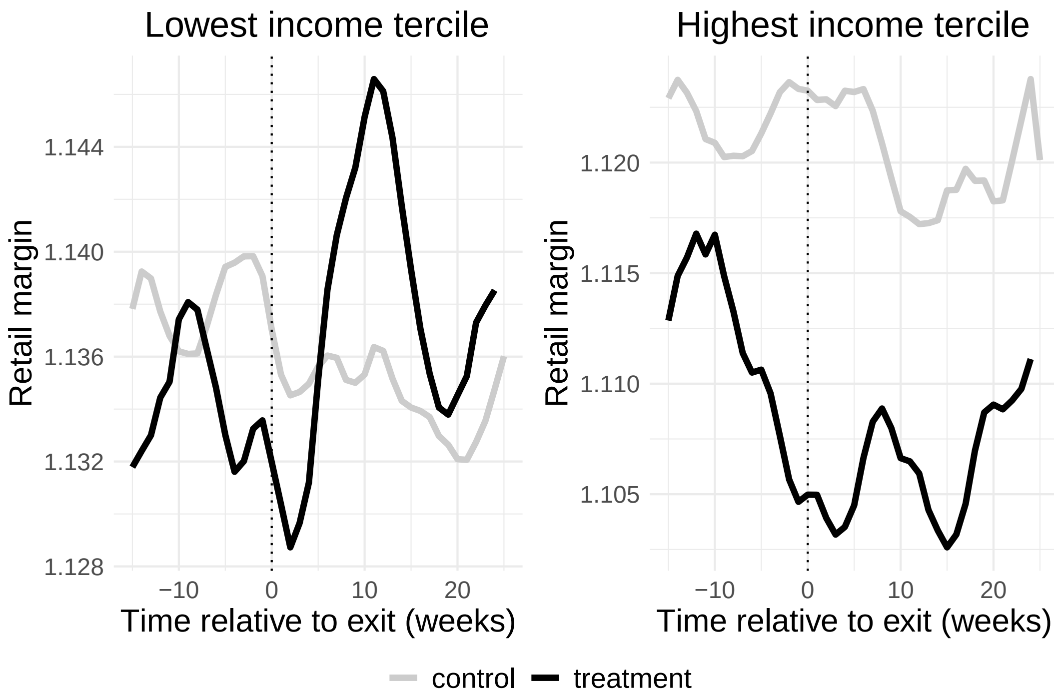

Figure 6
Retail ULP price margin before and after exit at different income levels

an increase (although not back to pre-event levels).

Both of these figures confirm conventional industrial organisation theory and previous empirical findings, i.e. market exit leads to a price increase, and entry is likely to bring down prices, although there seems to be a difference in the level and the change in the level of margins between low- and high-income areas. In the following section, we formally test this difference.

IV(iii) *Exogenous treatment*

Central to our identification is the assumption that market exit/entry are exogeneous. If the decision to exit/enter is done at the same level as the outcome decision (pricing), this assumption would not hold. One form of endogeneity, reverse causality, would mean that changes in the price/margin trigger the treatment, not the other way around. It is also possible that some unobserved factor is behind the variation in both the treatment (exit/entry) and the outcome variable (price/margin). The way this has been handled in previous work depends on their study design. In a study on the impact of an acquisition of a petrol retail brand, Hastings [2004] assume that the disappearance from the market of a rival brand is exogenous in local competitor stations' pricing decisions, conditioned on station-specific fixed effects and city-time effects. Hosken et al. [2011] also assume that mergers are exogenous to the local pricing decisions of rival petrol stations. These studies treat mergers as a natural experiment, which would justify the exogeneity assumption. In our case, there is no similar single natural experiment that drives exit/entry in our

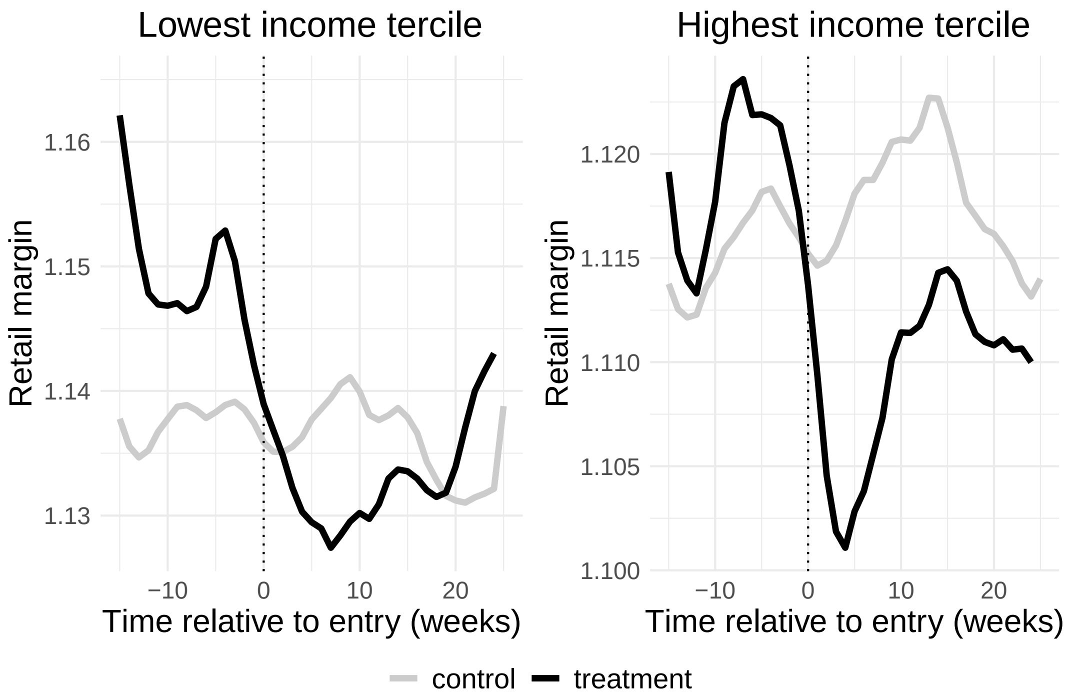

Figure 7
Retail ULP price margin before and after entry at different income levels

geographically and temporaneously disperse set of markets. However, below, we offer two arguments in support of our assumption of exogeneous exit and entry.

IV(iii).1 Timing of the exit/entry

Our study design means that we look at the relationship between the short terms variation in margin/price, and the incidence of exit/entry. In this setup endogeneity should only concern us if the margin in the vicinity of exit/entry is what drives the decision to exit/enter. Given the cost of exit/entry, it is highly unlikely that these decisions are made and implemented on the whim of relatively short-term fluctuations in the margin. To support this argument, Figure 8 shows the weekly distribution of exit and entry. It displays a large spike in the number of exits and entries at week 25, which is the end of the tax year in Australia. For administrative purposes, it makes sense for businesses to close down at the end of the tax year, or open up right at the beginning of the new tax year. This implies that at least a large number of exit and entry instances in our data were not made in response to a short-term change in the retail margin, and provides support for our exogenous treatment assumption.[19]

It may also seem theoretically plausible that changing local characteristics are causing exit/entry. This is not likely to be an issue in our study, because we estimate impacts for

[19]In Section V(v) we present estimates for the sample that only includes exits/entries around the end of the tax year.

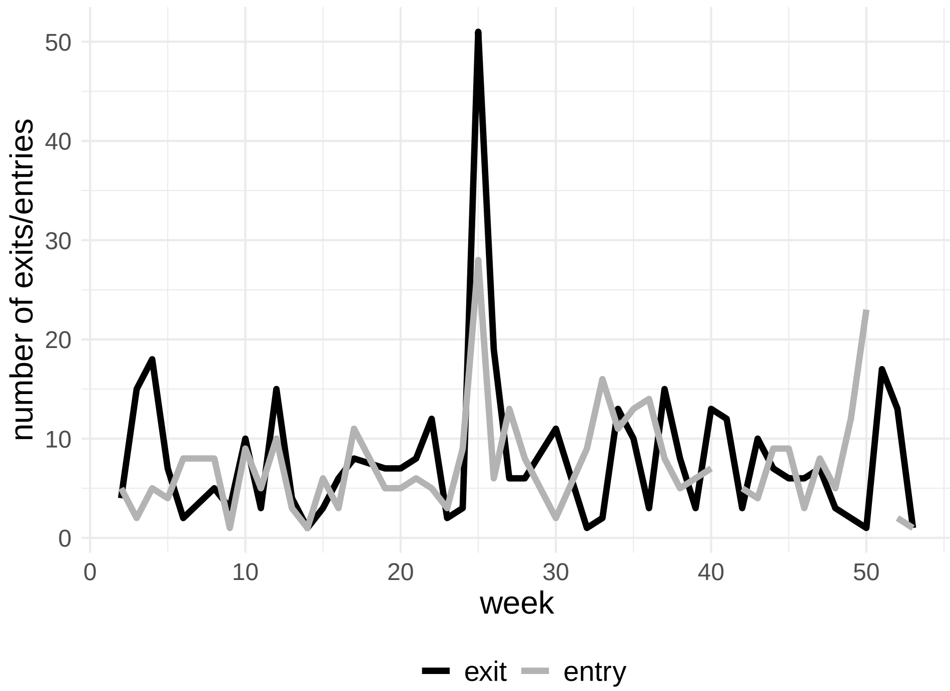

Figure 8
The number of exits and entries by calendar week

±15 weeks around the incidence of exit or entry. Local characteristics do change over time, but they are likely to change at a rate slower than these 30 week event study intervals.

IV(iii).2 Which firms exit/enter?

To decide whether exit/entry are exogenous, it is useful to understand the reason for each instance of exit and entry in our sample. This was not possible for most cases.[20] Nevertheless, our data allows us to verify a few things. One sign of endogeneity would be if the exiting firm was systematically the least or the most expensive option within a given local geographical market (which fact could have triggered the exit). The first two columns of Table VIII show the % of weeks a petrol station spent as the least/most expensive option within a 1mi radius. The first two rows compare the 25 weeks before the exit with the weeks preceding this period for the exiting petrol station. There does not seem to be significant difference: i.e. the relative pricing of the exiting stations (on average) do not differ in the weeks immediately preceding the exit from the other weeks the same firm was in the market.

This finding is consistent with Lach and Moraga-González [2017], who suggest that the prices observed should be consistent with mixed-strategies. The fact that no single firm is the the most/least expensive option for the entire study period, implies that firms do indeed engage in mixed strategies, which sometimes make them more and other times less

[20]For a small number of exits we found evidence from local newspapers or from historical archives of Google Maps StreetView, that they shut down for refurbishing or complete renovation works.

expensive than their rivals.

Table VIII
Characteristics of exiting and entering firms

	% of weeks spent as cheapest station within 1mi	% of weeks spent as most expensive station within 1mi	number of same brand stations within 1mi	number of same brand stations within 2mi	number of same brand stations within 5mi
exiting firm					
<25 weeks pre-exit	0.293	0.433	0.341	0.787	2.500
>25 weeks pre-exit	0.270	0.466	0.339	0.841	2.841
entering firm					
<25 weeks post-entry	0.423	0.430	0.241	0.623	2.766
>25 weeks post-entry	0.476	0.314	0.203	0.532	2.386

V ECONOMETRIC METHOD AND MAIN RESULTS

V(i) *Estimating heterogeneous effects*

The conceptual problem of estimating the treatment effect of exit and entry on the retail margin is similar to that formulated in Rubin [1974]. Denote a vector of covariates for petrol station i by X_i. We let the treatment (exit and entry from and to the market) indicator W_i take on the values 0 (the control group, i.e. no exit/entry) and 1 (the treatment group, i.e. exit/entry). For petrol station i, $i = \{1, ..., N\}$, let Y_i denote the observed outcome and the outcome of interest (the retail margin) in the case of receiving the treatment as $Y_i(1)$ and when not receiving the treatment as $Y_i(0)$. The causal effect of exit/entry for petrol station i is therefore $Y_i(1) - Y_i(0)$. The Conditional Average Treatment Effect (CATE) is given by:

$$\tau(x) = E[Y_i(1) - Y_i(0)|X_i = x]. \tag{1}$$

The problem of causal inference is that we do not observe both $Y_i(1)$ and $Y_i(0)$ at the same time. Instead we estimate CATE (conditional on observable variables X) by a difference in means $\overline{Y}_t - \overline{Y}_c$, where $\overline{Y}_t$ and $\overline{Y}_c$ are the means of the outcome variable for the treated (t) and control (c) groups, respectively. For identification, we assume unconfoundedness, $\{Y_i(1), Y_i(0)\} \perp W_i|X_i$, i.e. the markets that experience exit/entry are selected randomly conditional on the observable covariates.

Our objective is to conduct estimation and inference on the function $\tau(x)$ to gain insight into the heterogeneity of the treatment response, across our observable local area characteristics. One way to do this would be through introducing interaction terms in the estimation of $\tau(x)$. Alternatively, we could estimate $\tau(x)$ for different sub-samples of the data, $\tau_b(x)$, $b = 1, 2, \ldots, B$. The problem with these multiple interaction terms is that they run the risk of using a misspecified model, and even if the correct model was estimated, it can quickly run into dimensionality problems.[21]

[21] For k potential sources of heterogeneity, this would mean adding $2^k - k - 1$ interaction terms to our

We use generalised random (causal) forests as proposed by Wager and Athey [2018] and Athey et al. [2019]. The method fits our problem for multiple reasons. As a non-parametric tree-based method, it does not require us to specify a (potentially complex) relationship between our covariates and the treatment effect. It also allows the efficient handling of large covariate spaces. In our case the number of possible sources of treatment heterogeneity (accounting for all interactions) is much larger than the sample size, therefore methods such as OLS cannot be considered without the research filtering which features to use first. Moreover, Athey et al. [2019] showed that the estimates achieve asymptotic normality and as such, it is suitable for hypothesis testing on the treatment effects. Finally, independent variables are time-invariant in our event-study setup (they change every 1 or 5 years). Estimating a 2-way fixed effects linear model would mean simply losing important covariates as they would be subsumed by the fixed effects dummies.

Below we provide a brief introduction to tree-based methods and causal forests. This should be sufficient for those who are unfamiliar with these methods to understand its intuition, but for details, we refer the reader to Athey and Imbens [2015] and Athey et al. [2019]. Regression trees are a non-parametric machine learning approach, which are frequently used for prediction problems in data science.[22] Assume we have k covariates and N observations, and we want to partition the covariate space $\mathcal{X}$ into M mutually exclusive regions $R_1, ..., R_M$, where the outcome for an individual i with covariate vector X_i in region R_m is estimated as the mean of the outcomes for training observations in R_m. Denote the subset of covariates observations corresponding to R_m as $\mathbb{X}_m$. Let X_j be a splitting variable and s be a split point. For the initial stage, with $M = 2$, define the observations of covariates associated with observations of X_j that exceed the point s as $R_1(j, s) = \mathbb{X}_1 = \{X \mid X_j \leq s\}$ and similarly, $R_2(j, s) = \mathbb{X}_2 = \{X \mid X_j > s\}$. The algorithm selects the pair (j, s) that solves:[23]

$$\min_{j,s} \left[\sum_{i:X_i \in \mathbb{X}_1} (Y_i - \overline{Y}_1(j, s))^2 + \sum_{i:X_i \in \mathbb{X}_2} (Y_i - \overline{Y}_2(j, s))^2 \right] \tag{2}$$

where $\overline{Y}_1(j, s)$, and $\overline{Y}_2(j, s)$ are the mean outcomes in $R_1(j, s)$ and $R_2(j, s)$. Eq.2 splits the data into two regions, then the process is repeated on each of the two resulting regions. Regression forests are ensemble methods, whereby the forest predictions are constructed as the average of the tree-based predictors. (Eq.2) can also be thought of as the 'growing' or 'splitting' part of constructing regression trees.

Causal trees build on the same concept, but for each node, instead of minimising the mean squared error (MSE) for the difference between the average outcomes for each node, it minimises the MSE for the difference in the estimated treatment effects.

Athey et al. [2019] proposes using an honest approach to estimating these causal trees, i.e. they grow the tree on a sample of the data, and they estimate it using a different sample. In the context of causal trees, the idea is that the (regions) are small enough that

model.

[22] For details see: Breiman et al. [1984].

[23] The number of splits is chosen through cross-validation in order the balance the tradeoff between low bias and high variance of regression trees.

the (Y_i, W_i) pairs for each leaf had come from a randomised experiment. In this case, the treatment effect in the small space of each leaf with the corresponding set $\mathbb{X}_m$ is given by:

(3)
$$\hat{\tau}_{\mathbb{X}_m} = \frac{1}{|\{i : W_i = 1, X_i \in \mathbb{X}_m\}|} \sum_{\{i:W_i=1,X_i\in\mathbb{X}_m\}} Y_i - \frac{1}{|\{i : W_i = 0, X_i \in \mathbb{X}_m\}|} \sum_{\{i:W_i=0,X_i\in\mathbb{X}_m\}} Y_i$$

Finally, to construct a causal forest, we draw repeated bootstrap samples of size B from the training data to recursively estimate a number of causal trees. The prediction for an individual with a vector of covariates X_i is then $\hat{\tau} = \frac{1}{B}\sum_{b=1}^{B} \hat{\tau}_b$, where $\hat{\tau}_b$ is the estimate produced by tree b. Athey et al. [2019] show that the estimated treatment effect is asymptotically normal.

Causal forests are useful for finding heterogeneity in the treatment effect in a cross-section setup. The variables in X_i are area petrol station and area characteristics and can be considered constant within the 30-week event window of our analysis. The outcome variable Y_i on the other hand is time-variant. Therefore to use a causal forest, we transform the outcome variable in the following way so that the outcome variable of interest Y_i is the change in the retail margin for petrol station i before and after the exit/entry:

$$Y_i = \frac{1}{|\{T_1\}|} \sum_{t\in\{T_1\}} margin_{it} - \frac{1}{|\{T_0\}|} \sum_{t\in\{T_0\}} margin_{it} \tag{4}$$

T_0 and T_1 represent the pre- and post-treatment periods, respectively. This means that we forego the possibility of estimating time-dependent treatment effects (for example in the sense of traditional event study designs). We believe this trade-off is justified as we are primarily interested in the heterogeneity in the treatment effect, rather than its dynamics.[24]

In X_i we include the features listed in Table B.1, which reveals some overlap. For example, we have many different ways of measuring education, or wealth, and we have no a priori knowledge, which one of these is important in driving the treatment effect heterogeneity. Athey and Wager [2019] proposes removing the least important features from the estimation of causal trees to improve estimates. This is a feasible option, but we are specifically interested in the effect of some variables on the treatment effect, and this solution may eliminate some of our variables of interest. Instead, we add an extra layer to causal forests with a bagging ensemble learning method. The idea is, to randomly draw several features, add our features of interest, and re-estimate the forest in each draw, on this reduced sample of features. This way the estimated ensemble treatment effects are $\tau^G = \frac{1}{G}\sum_{g=1}^{G} \hat{\tau}_g$, where G is the number of causal forests we run to get our ensemble treatment effects. The standard errors are derived from the bootstrapped standard errors of the individual causal forest and the squared deviation of the treatment effects:

$$\sigma_{\hat{\tau}^G} = \sqrt{\frac{\sum_{g=1}^{G}[\hat{\sigma}_g^2(\hat{\tau}_g - \bar{\tau})^2]}{G}} \tag{5}$$

[24]This also explains why we de-seasonalised the data.

We argue that this ensemble method is more fitting in cases where there is a relatively small sample size, and a large number of parameters, and we have specific (theory-driven) interest in a selected set of these features. In Section A of the Appendix, we provide details and simulations to justify our approach.

V(i).1 Unconfoundedness assumption

Up to this point, we have assumed that selection to treatment was random (unconfoundedness assumption), conditional on our observable variables. Non-random selection means that unless all relevant variables are observed, our estimates will be biased. A frequent violation of the random assignment assumption is when unobserved factors are correlated with the treatment and the outcome variable in question, leading to biased estimates (omitted variable bias). Conventionally, researchers try to remedy this problem by employing fixed effects, or instrumental variables in their models. The problem with this approach is that it relies on strong assumptions that may not hold, and it is hugely limited by dimensionality issues in a conventional linear regression setup. For example, in our study, non-linear interactions between the independent variables may affect the treatment, and not controlling for this would lead to biased estimates. But the use of a linear model constrains researchers in how many of these interactions they can include in their models. Our choice of method handles this problem and allows a much richer set of observable factors to control for. Although it is never possible to observe and account for all relevant factors, under our model the conditional independence assumption relies on a much wider range of attributes than would be possible in linear models. We can include a large number of observed variables and their interactions with the way the treatment affects the outcome, reducing the risk of omitted variable bias.

V(ii) *Difference-in-differences results*

As a first step, we estimated the impact of exit and entry on the ULP margin using a simple difference-in-differences model with the event study design explained above. We have a balanced panel dataset of 299 treatment units and 1495 control units[25] for studying exit, and 191 treatment units and 955 control units for studying entry, over a study period of +/- 15 weeks around the time of exit/entry, in which we observe the outcome (ULP margin) Y_{it} for unit i at time t respectively. We estimate the following linear model:

$$Y_{it} = \alpha_i + \beta T_t + \delta D_{it} + \epsilon_{it} \tag{6}$$

where α_i are unit fixed effects, T is a before-after event dummy, D is a binary treatment variable, which is 1 for the treated units after the treatment, and zero otherwise. The treatment effect is the δ coefficient. The inclusion of unit fixed effects accounts for unit-specific (but time invariant) confounders in a flexible manner. As local area characteristics are unlikely to shift significantly in short intervals (within the six months of our study period), this allows us an easily executable framework to look at only the impact of exit.

[25] The same petrol station could have been a treatment/control unit multiple times over our study period of 15 years.

Because our study brings together events over a 15 year period, we did not include time fixed effects.

To understand the heterogeneity in the impact of exit and entry, we split the sample into low-, and high-income and low-, and high-competition areas (we do the splits around the respective medians), and estimate eq.6 for the resulting four sub-samples.[26]

Table IX below shows the treatment effects and the associated standard errors for the four sub-samples. For comparability across the different models, we standardise all variables, therefore the coefficients can be read as (proportions of) standard deviation changes in the price margin as a result of exit/entry. We also report z-scores for the difference in low-, and high-income coefficients.

Table IX
The impact of exit and entry on the ULP margin - linear regression results

		low income	high income
exit	low competition	0.181	-0.118
		(0.06)	(0.061)
	z-score of difference	3.49	
exit	high competition	0.122	-0.036
		(0.04)	(0.033)
	z-score of difference	3.05	
entry	low competition	-0.185	-0.156
		(0.053)	(0.04)
	z-score of difference	0.44	
entry	high competition	-0.284	-0.216
		(0.037)	(0.031)
	z-score of difference	-1.41	

Standard errors in parentheses.

These preliminary estimates suggest that exit leads to the highest increase in the ULP margin in low-income and low-competition areas. In high-income and high-competition areas the impact is not significantly different from zero. The difference between low- and high-income areas is significant. We note the negative effect in low competition - high-income areas (prices fall after exit), which could be due to unobserved factors, some of which we control for in our causal forest models. For entry, these preliminary results suggest that prices fall after entry, and they fall more in low-income areas, but this difference from high-income areas is not significant.

Although in these models we control for unit fixed effects, the estimates mask the individual contribution of each area characteristic. For this reason, and the reasons explained in Section V(i) we now turn to using causal forests to offer a more detailed analysis of our research question.

[26]We chose to do this instead of using interaction terms, for the relative simplicity of interpreting the results.

V(iii) *Causal forest results*

Table X shows the conditional average treatment effects (CATE) and the conditional average treatment effects on the treated (CATT). The estimates from each i iteration of our ensemble method are summarised by the treatment effect estimator, $TE = \sum_{i=1}^{k} TE_i/\sigma_i^2$, and standard error $SE(TE) = \sqrt{1/\sum_{i=1}^{k} \sigma_i^2}$.

Exit, on average, triggered a small (statistically not significant) increase in the margin. Entry, on average led to a larger drop in prices. Our interpretation of the asymmetry between exit and entry is to do with the level of market concentration in markets where we observed exit and markets where we were sampling instances of entry. Indeed, in our data, the average number of rivals in markets where entry happened is lower (2.6 within 1 mile, 5.7 within 2 miles, and 15.3 within 5 miles) than in markets where exit happened (2.9 within 1 mile, 6.6 within 2 miles, and 23.3 within 5 miles). In more concentrated markets the effect of a change in market concentration on the retail margin is more pronounced.

Table X
Conditional average treatment effects

Exit		Entry	
CATE	CATT	CATE	CATT
0.006	0.009	-0.249	-0.279
(0.019)	(0.019)	(0.035)	(0.035)

Standard errors in parentheses.

Table XI presents a more detailed breakdown of the treatment effects related to exit, split across our two main variables of interest. We defined low and high for these variables by taking the values corresponding to their 10th and 90th percentiles respectively. We then used our estimated causal forests to predict the treatment effect at these low- and high-income and competition values, assuming mean values for all other covariates.

Table XI
Predicted treatment effects by income and competition

	exit		entry	
	low income	high income	low income	high income
low competition	0.148 (0.009)	0.030 (0.007)	-0.298 (0.019)	-0.276 (0.019)
z-score of difference	10.349		-0.819	
high competition	0.064 (0.008)	-0.039 (0.005)	-0.286 (0.017)	-0.265 (0.017)
z-score of difference	10.918		-0.873	

Standard errors in parentheses.

Several stylised findings can be deduced from this exercise. Most importantly for our investigation, with exit there is a larger price increase in lower-income areas. The price increasing effect of exit is more pronounced in markets where competition is lower. This suggests that increasing market concentration increases price dispersion (the extent to which businesses choose to price discriminate) with low-income areas seeing a larger price increase.

Entry on the other hand results in a fall in prices of similar magnitude in low- and high-income areas - it is higher in low-income areas, but not significantly higher than in high-income areas. The price drop is marginally higher in low-competition areas, which is intuitive as, on the margin, areas with low levels of competition can gain more from a new rival.

Altogether, these results suggest two main effects that hold for each product and both exit and entry. (1) Low-income areas experience a significantly larger increase in the price margin with exit, but a not-significantly larger drop in margins with entry. (2) More concentrated markets witness a larger rise in the price margin from exit, and a larger (but not significantly larger) fall in prices with entry.

V(iv) *Other sources of heterogeneity*

In search of an explanation to the above results, first we looked at whether heterogeneity in the ability to switch to outside options (not buy petrol) explains some of the variation in the impact of exit. For this we distinguished between areas based on the average household's reliance on cars (measured through the % of people who drive to work, and the average number of cars per household). This could be thought of as an indirect proxy for willingness to pay.

Table XII shows that in areas with a higher reliance on cars the impact of exit is larger.[27]

Table XII
Predicted treatment effects of exit in ULP, by different levels of competition, income, and car reliance

		average number of cars per household		% who drive to work	
		low	high	low	high
low competition	low income	0.136 (0.006)	0.155 (0.006)	0.127 (0.007)	0.143 (0.006)
	high income	0.017 (0.005)	0.041 (0.005)	0.010 (0.005)	0.026 (0.005)
high competition	low income	0.054 (0.005)	0.076 (0.005)	0.053 (0.006)	0.066 (0.006)
	high income	-0.048 (0.003)	-0.020 (0.004)	-0.049 (0.004)	-0.037 (0.004)

Standard errors in parentheses.

[27]Here our focus is on the impact of exit. Results on the impact of entry are provided in the Appendix, in Tables C.1 to ??.

To proxy for search efforts, we considered two metrics of consumer informedness (home internet penetration and commuting distance), as reported in Table XIII, but once we controlled for income and competition neither of them performed well in explaining heterogeneity in treatment effects.

Table XIII
Predicted treatment effects of exit in ULP, by different levels of competition, income, and measures of search

		commuting distance		home internet penetration	
		low	high	low	high
low competition	low income	0.160	0.148	0.134	0.137
		(0.012)	(0.012)	(0.012)	(0.011)
	high income	0.056	0.032	0.027	0.03
		(0.011)	(0.008)	(0.009)	(0.009)
high competition	low income	0.074	0.070	0.058	0.062
		(0.012)	(0.011)	(0.010)	(0.009)
	high income	-0.024	-0.034	-0.036	-0.032
		(0.009)	(0.007)	(0.007)	(0.007)

Standard errors in parentheses.

We looked at a number of other local area characteristics. Table XIV shows that areas with lower educational and occupational levels are associated with a higher increase in the margin in response to increasing market concentration. Areas with a higher proportion of over 65 population are associated with a larger increase in the margin following exit.

Table XIV
Predicted treatment effects of exit in ULP, by different levels of competition, income, and education and age

		level of education		proportion of over 65s	
		low	high	low	high
low competition	low income	0.142	0.125	0.118	0.146
		(0.006)	(0.006)	(0.006)	(0.006)
	high income	0.031	0.018	0.007	0.05
		(0.005)	(0.005)	(0.005)	(0.005)
high competition	low income	0.066	0.047	0.046	0.058
		(0.005)	(0.005)	(0.005)	(0.005)
	high income	-0.034	-0.046	-0.052	-0.024
		(0.004)	(0.003)	(0.004)	(0.004)

Standard errors in parentheses.

We also looked at sources of heterogeneity on the supply-side, and whether it matters who the exiting/remaining petrol stations are. First, we considered differences between the remaining (non exiting) petrol station (the treated firm) being independent, or part of a large chain. Table XV shows that exit is followed by a significantly larger increase in the retail margin of large chain petrol stations. The difference between low- and high-income areas remains even where we fix the remaining station to be independent or part of a large chain. On the other hand, the right half of Table XV suggests that it matters less whether the exiting petrol station is a large chain or an independent.

Table XV
Predicted treatment effects of exit in ULP, by different levels of competition, income, and large chain or independent

		the treated firm		the exiting firm	
		independents	large chain	independents	large chain
low competition	low income	0.059 (0.003)	0.133 (0.006)	0.158 (0.007)	0.161 (0.006)
	high income	-0.029 (0.002)	0.076 (0.003)	0.025 (0.005)	0.028 (0.005)
high competition	low income	0.035 (0.004)	0.077 (0.006)	0.076 (0.006)	0.080 (0.006)
	high income	-0.071 (0.001)	-0.007 (0.003)	-0.042 (0.003)	-0.041 (0.003)

Standard errors in parentheses.

V(v) *Further tests on the exogeneity assumption*

As we discussed above, a large proportion of exits and entries happen around the end/start of the tax year. Although one could argue that endogeneity in exit/entry is less likely to be a problem in our event study that spans over 15 years, a feature of our data allows us to limit our sample to exits and entries where we have a stronger argument that the treatment is exogenous.

We noticed earlier that there was a higher propensity of exit/entry around the end of tax year. For this reason we re-estimated our causal forests for the subgroups of exits happening within 10 weeks before the end of the tax year, and subgroups of entries happening within 10 weeks after the start of the tax year (to keep the sample size large enough for a meaningful estimation of our causal forests). These reduced sub-samples contain 59 instances of market exit and 51 instances of entry. Table XVI compares the mean and standard deviation of our main variables of interest for these sub-samples with the rest of the sample. There appears to be no systematic difference between these exits/entries and the rest of the sample.

Table XVII shows the conditional average treatment effects for the sub-samples of exits/entries around the end of tax year. The CATEs reported in Table XVII are similar to those estimated for the total sample (Table XI). If one accepts the claim that these ex-

Table XVI
Comparing the samples around the end of tax year (ETY)

		income	internet	commuting	comp1mi	comp2mi	comp5mi
exit	reduced sample	47624.919	0.229	4.083	2.645	6.047	20.678
		(9545.2)	(0.073)	(2.863)	(2.171)	(4.640)	(19.851)
	rest of the sample	45535.9	0.238	4.145	2.771	6.321	20.124
		(7908.4)	(0.074)	(2.932)	(2.096)	(4.655)	(19.597)
entry	reduced sample	47269.6	0.230	4.101	2.663	6.099	20.667
		(9303.9)	(0.073)	(2.860)	(2.154)	(4.631)	(19.789)
	rest of the sample	49091.3	0.230	3.898	2.615	5.756	19.585
		(10556.6)	(0.074)	(3.069)	(2.305)	(4.832)	(20.348)

Standard errors in parentheses.

its/entries around the end of the tax year are exogenous, this finding would support our claim that our main results are not biased by potential endogeneity (reverse causality).

Table XVII
Predicted treatment effects for samples around the end of tax year

	exit		entry	
	low income	high income	low income	high income
low competition	0.144	0.105	-0.367	-0.363
	(0.004)	(0.004)	(0.004)	(0.004)
z-score of difference	6.894		-0.707	
high competition	0.143	0.104	-0.368	-0.364
	(0.004)	(0.004)	(0.004)	(0.004)
z-score of difference	6.894		-0.707	

Standard errors in parentheses.

It is also possible that not the treatment, but an event before the treatment consistently confounds our estimates. To test this, we look at pre-treatment parallel trends by focusing on the period preceding the treatment. If pre-treatment the parallel trend assumption is not violated, we would expect to see zero treatment effect. For this exercise we assumed a placebo treatment to happen 6-months before the real treatment. Table C.3 in the Appendix shows our results for the impact of exit on the ULP margin and finds that the effect is not significantly different from zero in any of the tested instances.

V(vi) *Other sensitivity checks*

In the results highlighted above, our focus was on the impact of exit/entry within a 1-mile radius. We looked at what happens to the margins if a rival from a 2-mile and a 5-miles radius exits or enters the market. In getting these estimates we followed the same logic as for the headline results. We first estimated a causal forest and used it to predict the treatment effect for various levels of competition, and income. We find that the results are qualitatively the same, but that the effect size increases as we are looking at the impact of a petrol station exiting/entering at a further distance (see Table XVIII below).

Table XVIII
Predicted treatment effects of exit (within 2 and 5 miles), by different levels of competition and income

	exit within 2 miles		exit within 5 miles	
	low income	high income	low income	high income
low competition	0.204	-0.002	0.248	0.078
	(0.002)	(0.001)	(0.002)	(0.001)
high competition	0.133	-0.053	0.223	0.064
	(0.002)	(0.001)	(0.002)	(0.001)

Standard errors in parentheses.

As with other tree-based methods, causal forests allow the clustering of estimands (see Athey and Wager 2019). We use this as a robustness check to cluster the petrol stations that experience the same exit/entry as one of their competitors. Results with causal forests clustered by postcode are show in TableXIX. Once again, our results remain qualitatively the same.

Table XIX
Predicted treatment effects of exit in ULP, by different levels of competition and income - clustered by postcode

	low income	high income
low competition	0.149	0.029
	(0.010)	(0.008)
z-score	9.370	
high competition	0.064	-0.039
	(0.010)	(0.006)
z-score2	9.522	

Standard errors in parentheses.

Finally, we also looked at how sensitive our results are to choosing a different nearest neighbour matching to select our control group petrol stations. Table XX shows the results for choosing the 2 and the 10 nearest neighbours. Our story remains qualitatively unchanged.

VI DISCUSSION OF THE RESULTS

In this paper we presented evidence supportive of exit leading to a small increase, and entry triggering a larger drop in the price margin. We argue that the asymmetry is because in our

Table XX
Predicted treatment effects of exit (with 2 and 10 nearest neighbours as control), by different levels of competition and income

	2 nearest neighbours		10 nearest neighbours	
	low income	high income	low income	high income
low competition	0.105	0.009	0.088	-0.03
	(0.003)	(0.002)	(0.002)	(0.001)
high competition	0.047	-0.041	0.026	-0.079
	(0.001)	(0.001)	(0.001)	(0.001)

Standard errors in parentheses.

sample entry tends to happen in more concentrated markets. Although the effect of exit is not significant on average, by looking at treatment effect heterogeneity, we identify the cases where it leads to a significant increase in the margin. These findings of heterogeneous treatment effects offer strong support to the Lach and Moraga-González [2017] model.

First of all, the margin-increasing-effect of exit (and the margin-reducing-effect of entry) is larger in less competitive markets. Second, low-income areas experience a larger increase in the retail margin of petroleum products when market concentration increases. At the same time, we do not find that the same low-income households enjoy a significantly larger drop in margins when competition intensifies. These results adds support to earlier findings that heterogeneity in the level of engagement with the market can lead to price dispersion even in homogeneous goods. Our results also that lower income areas experience higher post-exit increase in the margins imply that increasing market concentration can have distributional effects.

We argue, based on previous literature (see Section II(i)) that this could be due to both demand-side (e.g. how low-income households differ in their market engagement) and supply-side factors. More specifically, we looked at the following potential explanations.

- Lack of short-term alternatives: We started on the assumption that the % of people driving to work, and the average number of cars per household are indicative of how quickly a household can respond to a petrol price increase (assuming that a larger reliance on cars reduces the scope of short term alternatives). We found that exit led to a larger increase in the margin in areas with more reliance on cars. Under the assumption that higher reliance on cars is correlated with willingness to pay, this result implies a larger price increase in areas with higher willingness to pay.

- Higher search costs: To introduce two approximations of search, we looked at commuting distance (assuming that for motor fuel, more commuting means lower search costs, as people drive past more petrol stations on the way to work), and home internet penetration (given the role of petrol station price comparison tools in Western Australia). We found that margins are higher in low-commuting areas and in areas with lower internet access. But once we control for income (and other factors), these factors do not drive heterogeneity in treatment effect estimates.

- Other demand-side reasons: We focused on two main demand-side features that are often highlighted in the search literature: educational levels and age. In previous works, it has been shows that lower educational attainment and older age can both reduce market engagement. For example, looking at financial markets, Hastings et al. [2017] found that less-educated workers are less financially literate and are more influenced by sales force concentration. Our findings were in line with this previous literature. Although lower income households are also more likely to be lower on educational attainment and have lower financial literacy the impact of education remains even after controlling for income. Age played a similar role, areas with a larger proportion of older population experience higher increase in the margin after exit. Once again, this is in line with previous literature [Waddams Price and Zhu, 2016], who also found that age plays an important role in willingness to search: older people are more deterred from searching, and by longer switching time, and are less affected by their own experience of switching in other markets. Lusardi and Tufano [2015] links this age-differential to financial literacy.

- Supply-side reasons: Supply-side responses to exit may also explain variation in the impact of changing market concentration. We show that large chains respond with higher price increases following the exit of one of their rivals. However, we do not find that large chains are more/less likely to be found in low/high-income areas. Moreover, differences in income remain even after controlling for the type of the retailers.

- Unobserved factors: Even with the above factors fixed, the difference between low-, and high-income areas remained. This suggests that unobservable factors play an important role in how low-, and high-income households engage with the market. Byrne and Martin [2021] for example argues that differences on a cognitive level, differences in biases, and in how people process information may also be behind low-income people engaging less with the market.

An important implication of our findings for policy is that competition alone cannot reduce prices. Even in markets with more choice, margins increased more after exit in low-income areas. If consumers in these households engage less with the market, the benefits of competition are less likely to be transferred to them. On the other hand, as we show for high-income areas, increasing market concentration is less likely to leave them facing increased margins even in concentrated markets, which is in line with previous literature that links income to some of the drivers of market engagement, such as education.

These translate to two main messages for policymakers. First of all, the harm avoided by blocking a harmful increase in concentration includes some regressive distributional effects (acknowledging of course that not all exits are harmful on average, and some may reflect improved efficiency in the firms that remain in the market). Second, getting the market structure right may only offer a partial solution to a competition problem. Demand-side remedies may also be needed to ensure that consumers engage with the market. Moreover, our findings also give support to arguments that even where blocking or breaking up concentration is not possible, demand-side remedies may help mitigate harmful effects, provided that some choice still exists for consumers.

Finally, these findings should also offer useful lessons to merger retrospectives.[28] Most previous studies focus on the average price effect of exit through mergers, but in mergers with geographically distinct local markets, it would be useful to also look at distributional effects, using an approach similar to ours.

VII CONCLUSION

Motor fuel is a non-trivial part of poorer households' expenditure, which means that poorer households already pay a larger share of their income on transport-related fuel. If they pay a higher price for increased market concentration, the impact is much more pronounced in relative terms. This is important because it implies that antitrust needs to revisit some of its conventional wisdom and account for the possibility that some people benefit less from the elimination of anti-competitive conduct that reduces competition, and this should be reflected in the design of remedies, i.e. remedies should not have the average consumer in mind, but account for the heterogeneity of the impact of remedies across different income groups.

An important implication of our findings is that they offer support for the argument that antitrust could help address inequality while staying true to its mission of promoting competition.[29] We do not argue that income or wealth equality should be incorporated directly into competition policies. But we emphasise that ill-designed and executed competition policy and enforcement can contribute to increased inequality. Moreover, the success of the competition policy should not be evaluated for the average consumer. Instead, competition policy, when possible, should consider the possibility of a differential impact and impose remedies accordingly.

Motor fuel is similar to food in the sense that it is a non-discretionary part of household expenditure, which also displays significant distributional differences. Mattioli et al. [2018] identify a distinct group of households, around 10% of the UK's population who are in car-related economics stress: on low income, experience high-motoring costs, and a low response to fuel price changes. This thinking is seemingly also gaining some consideration in the regulatory review of mergers. The UK Competition and Markets Authority specifically emphasised the difference in local areas regarding food and petrol expenditure (lower-income areas spending a relatively larger proportion of their income on food/petrol) in the Sainsbury's/Asda merger.[30] Whilst we think this is an important and welcome development, we also believe that more micro-level evidence is needed on this topic. To build up the evidentiary toolkit of competition authorities, we hope that this paper will help foster the drive to deliver more merger retrospectives that estimate not only the average but the distributional effects of mergers as well.

[28] Kwoka [2014], and Mariuzzo and Ormosi [2019] provide an overview of these retrospectives.
[29] See for example Baker and Salop [2015], or Shapiro [2018].
[30] Para. 8.283.

REFERENCES

Allen, J., Clark, R. and Houde, J.-F., 2014, 'The effect of mergers in search markets: Evidence from the canadian mortgage industry,' *American Economic Review*, 104(10), pp. 3365–96.

Allen, J., Clark, R. and Houde, J.-F., 2019, 'Search frictions and market power in negotiated-price markets,' *Journal of Political Economy*, 127(4), pp. 1550–1598.

Alm, J., Sennoga, E. and Skidmore, M., 2009, 'Perfect competition, urbanization, and tax incidence in the retail gasoline market,' *Economic Inquiry*, 47(1), pp. 118–134.

Armstrong, M., Vickers, J. and Zhou, J., 2009, 'Consumer protection and the incentive to become informed,' *Journal of the European Economic Association*, 7(2-3), pp. 399–410.

Athey, S. and Imbens, G. W., 2015, 'Machine learning methods for estimating heterogeneous causal effects,' *stat*, 1050(5), pp. 1–26.

Athey, S., Tibshirani, J., Wager, S. et al., 2019, 'Generalized random forests,' *The Annals of Statistics*, 47(2), pp. 1148–1178.

Athey, S. and Wager, S., 2019, 'Estimating treatment effects with causal forests: An application,' *arXiv preprint arXiv:1902.07409*.

Baker, J. B. and Salop, S. C., 2015, 'Antitrust, competition policy, and inequality,' *Geo. LJ Online*, 104, p. 1.

Barron, J. M., Taylor, B. A. and Umbeck, J. R., 2004, 'Will open supply lower retail gasoline prices?' *Contemporary Economic Policy*, 22(1), pp. 63–77.

Breiman, L., Friedman, J., Stone, C. J. and Olshen, R. A., 1984, *Classification and regression trees* (CRC press).

Brown, J. R. and Goolsbee, A., 2002, 'Does the internet make markets more competitive? evidence from the life insurance industry,' *Journal of political economy*, 110(3), pp. 481–507.

Byrne, D. P. and de Roos, N., 2017, 'Consumer search in retail gasoline markets,' *The Journal of Industrial Economics*, 65(1), pp. 183–193.

Byrne, D. P. and Martin, L. A., 2021, 'Consumer search and income inequality,' *International Journal of Industrial Organization*, p. 102716.

Byrne, D. P., Martin, L. A. and Nah, J. S., 2022, 'Price discrimination by negotiation: A field experiment in retail electricity,' *The Quarterly Journal of Economics*, 137(4), pp. 2499–2537.

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

Calvet, L. E., Campbell, J. Y. and Sodini, P., 2009, 'Measuring the financial sophistication of households,' *American Economic Review*, 99(2), pp. 393–398.

Caplovitz, D., 1963, 'The poor pay more: Consumer practices of low-income families,' Technical report.

Chung, C. and Myers Jr, S. L., 1999, 'Do the poor pay more for food? an analysis of grocery store availability and food price disparities,' *Journal of consumer affairs*, 33(2), pp. 276–296.

Cohen, M. c., Jacquillat, A., Serpa, J. C. and Benborhoum, M., 2023, 'Managing airfares under competition: Insights from a field experiment,' *Management Science*.

Cooper, T. E. and Jones, J. T., 2007, 'Asymmetric competition on commuter routes: the case of gasoline pricing,' *Southern Economic Journal*, pp. 483–504.

De los Santos, B., 2018, 'Consumer search on the internet,' *International Journal of Industrial Organization*, 58, pp. 66–105.

De los Santos, B., Hortaçsu, A. and Wildenbeest, M. R., 2012, 'Testing models of consumer search using data on web browsing and purchasing behavior,' *American economic review*, 102(6), pp. 2955–80.

Diamond, P., 1987, 'Consumer differences and prices in a search model,' *The Quarterly Journal of Economics*, 102(2), pp. 429–436.

Dierx, A., Ilzkovitz, F., Pataracchia, B., Ratto, M., Thum-Thysen, A. and Varga, J., 2017, 'Does eu competition policy support inclusive growth?' *Journal of Competition Law & Economics*, 13(2), pp. 225–260.

Du, Y. and Ma, T., 2022, 'Impact of electricity pricing strategy on residential energy consumption decisions in China: Empirical evidence from a household survey,' *Journal of Environmental Economics and Policy*, 11(2), pp. 154–171.

Eizenberg, A., Lach, S. and Oren-Yiftach, M., 2021, 'Retail prices in a city,' *American Economic Journal: Economic Policy*, 13(2), pp. 175–206.

Ennis, S. F., Gonzaga, P. and Pike, C., 2019, 'Inequality: A hidden cost of market power,' *Oxford Review of Economic Policy*, 35(3), pp. 518–549.

Goldfarb, A. and Prince, J., 2008, 'Internet adoption and usage patterns are different: Implications for the digital divide,' *Information Economics and Policy*, 20(1), pp. 2–15.

Handbury, J., 2021, 'Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities,' *Econometrica*, 89(6), pp. 2679–2715.

Hastings, J., Hortaçsu, A. and Syverson, C., 2017, 'Sales force and competition in financial product markets: the case of mexico's social security privatization,' *Econometrica*, 85(6), pp. 1723–1761.

Hastings, J. S., 2004, 'Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in southern california,' *American Economic Review*, 94(1), pp. 317–328.

Hosken, D., Silvia, L. and Taylor, C., 2011, 'Does concentration matter? measurement of petroleum merger price effects,' *American Economic Review*, 101(3), pp. 45–50.

Hosken, D. S., McMillan, R. S. and Taylor, C. T., 2008, 'Retail gasoline pricing: What do we know?' *International Journal of Industrial Organization*, 26(6), pp. 1425–1436.

Houde, J.-F., 2012, 'Spatial differentiation and vertical mergers in retail markets for gasoline,' *American Economic Review*, 102(5), pp. 2147–82.

Khan, L. M. and Vaheesan, S., 2017, 'Market power and inequality: The antitrust counterrevolution and its discontents,' *Harv. L. & Pol'y Rev.*, 11, p. 235.

Kiser, E. K., 2002, 'Predicting household switching behavior and switching costs at depository institutions,' *Review of Industrial Organization*, 20, pp. 349–365.

Kwoka, J., 2014, *Mergers, merger control, and remedies: A retrospective analysis of US policy* (MIT Press).

Lach, S. and Moraga-González, J. L., 2017, 'Asymmetric price effects of competition,' *The Journal of Industrial Economics*, 65(4), pp. 767–803.

Letzler, R., Sandler, R., Jaroszewicz, A., Knowles, I. and Olson, L. M., 2017, 'Knowing when to quit: Default choices, demographics and fraud,' *The Economic Journal*, 127(607), pp. 2617–2640.

Lusardi, A. and Tufano, P., 2015, 'Debt literacy, financial experiences, and overindebtedness,' *Journal of Pension Economics & Finance*, 14(4), pp. 332–368.

Mariuzzo, F. and Ormosi, P. L., 2019, 'Post-merger price dynamics matters, so why do merger retrospectives ignore it?' *Review of Industrial Organization*, 55(3), pp. 403–429.

Mattioli, G., Wadud, Z. and Lucas, K., 2018, 'Vulnerability to fuel price increases in the uk: A household level analysis,' *Transportation Research Part A: Policy and Practice*, 113, pp. 227–242.

Morton, F. S., Zettelmeyer, F. and Silva-Risso, J., 2003, 'Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities?' *Quantitative marketing and Economics*, 1, pp. 65–92.

Nishida, M. and Remer, M., 2018, 'The determinants and consequences of search cost heterogeneity: Evidence from local gasoline markets,' *Journal of Marketing Research*, 55(3), pp. 305–320.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C. and Yontcheva, B., 2020, 'Information and price dispersion: Theory and evidence,' *International Economic Review*, 61(2), pp. 871–899.

Rubin, D. B., 1974, 'Estimating causal effects of treatments in randomized and nonrandomized studies.' *Journal of educational Psychology*, 66(5), p. 688.

Salop, S. and Stiglitz, J., 1977, 'Bargains and ripoffs: A model of monopolistically competitive price dispersion,' *The Review of Economic Studies*, 44(3), pp. 493–510.

Sengupta, A. and Wiggins, S. N., 2014, 'Airline pricing, price dispersion, and ticket characteristics on and off the internet,' *American Economic Journal: Economic Policy*, 6(1), pp. 272–307.

Shapiro, C., 2018, 'Antitrust in a time of populism,' *International Journal of Industrial Organization*, 61, pp. 714–748.

Simpson, J. and Taylor, C., 2008, 'Do gasoline mergers affect consumer prices? the marathon ashland petroleum and ultramar diamond shamrock transaction,' *The Journal of Law and Economics*, 51(1), pp. 135–152.

Stahl, D. O., 1989, 'Oligopolistic pricing with sequential consumer search,' *The American Economic Review*, pp. 700–712.

Stango, V. and Zinman, J., 2016, 'Borrowing high versus borrowing higher: price dispersion and shopping behavior in the us credit card market,' *The Review of Financial Studies*, 29(4), pp. 979–1006.

Stango, V. and Zinman, J., 2023, 'We are all behavioural, more, or less: A taxonomy of consumer decision-making,' *The Review of Economic Studies*, 90(3), pp. 1470–1498.

Tang, Z., Smith, M. D. and Montgomery, A., 2010, 'The impact of shopbot use on prices and price dispersion: Evidence from online book retailing,' *International Journal of Industrial Organization*, 28(6), pp. 579–590.

Taylor, C. T., Kreisle, N. M. and Zimmerman, P. R., 2010, 'Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in southern california: Comment,' *American Economic Review*, 100(3), pp. 1269–76.

Varian, H. R., 1980, 'A model of sales,' *The American economic review*, 70(4), pp. 651–659.

Waddams Price, C. and Zhu, M., 2016, 'Empirical evidence of consumer response in regulated markets,' *Journal of Competition Law & Economics*, 12(1), pp. 113–149.

Wadud, Z., Graham, D. J. and Noland, R. B., 2010, 'Gasoline demand with heterogeneity in household responses,' *The Energy Journal*, 31(1).

Wager, S. and Athey, S., 2018, 'Estimation and inference of heterogeneous treatment effects using random forests,' *Journal of the American Statistical Association*, 113(523), pp. 1228–1242.

Wilson, C. M. and Price, C. W., 2010, 'Do consumers switch to the best supplier?' *Oxford Economic Papers*, 62(4), pp. 647–668.

Woodward, S. E. and Hall, R. E., 2012, 'Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence,' *American Economic Review*, 102(7), pp. 3249–76.

Zac, A., Casti, C., Decker, C. and Ezrachi, A., 2020, 'Competition law and income inequality: A panel data econometric approach,' *Available at SSRN*.

APPENDIX

A SIMULATIONS TO DEMONSTRATE OUR ENSEMBLE BAGGING MODEL OF CAUSAL FORESTS

We conducted some experiments to justify using an ensemble causal forest method in cases where the data has a relatively small number of individuals and many potential sources of treatment heterogeneity. We simulate a dataset with $n = 1000$ individuals, where the number of factors increases from $k = 20$ to $k = 120$. In each loop, the treatment effect is a linear function of $J = k/5$ of these factors.

Our data generating process (DGP) is as follows:

$$Y_i = \alpha X_i + \beta_W W_i + \beta W_i X_i + U_i \tag{7}$$

Where W_i is the treatment variable, following a binomial distribution $W_i \sim B(n, 0.5)$, and $U_i \sim N(0, 1)$. The variable X_i represents a vector of J covariates, generated from a multivariate normal distribution. β is a vector of J parameters with $\beta_j = 1$ (for $j = 1, \ldots, J$) (i.e. as we increase J we add covariates. The starting DGP is defined as $J = 2$; $k = 10$, $n = 1000$.

Assume that we are interested in the effect of a set of factors X_1, X_2, X_3, where theory supports some relationship between the factors and the treatment effect. For this reason we then record estimates of $\beta_{1,2,3}$, as we systematically vary the k (20,30,40,50,60,70,80,90,100), and correspondingly the J (the number of variables causing heterogeneity in treatment) parameter (4,6,8,10,12,14,16,18,20), while holding everything else constant.

We compare the following two processes:

- Single causal forest: We follow Athey and Wager [2019] and start by training two random forests for Y and W and use its parameters as parameter choices to run our causal forest. Similarly to Athey and Wager [2019] we first train a pilot causal forest including all features, and then train a second forest only on those features that had most splits in the first forest (features that had at least the average share of splits). This helps in focusing efforts on the most important features. Our change in comparison to this formula is that we force our feature of interest ($X_{1,2,3}$) to be in the second, smaller pool of features as we are specifically interested in their role in treatment heterogeneity.

- Ensemble causal forests: We estimate $C = 1000$ causal forests. In each iteration, we repeatedly draw a random sample of $J/10$ features, plus we add our features of interest and estimate the causal forest on this small sample of features. The idea is that through our iterations, each feature has interacted with $X_{1,2,3}$. We then average over the estimates to give our ensemble estimate. For example for X_1 we get $\hat{\beta}_1 = 1/C \sum_{c=1}^{C} \hat{\beta}_{c,1}$.

Figure A.1 shows the estimates (and standard errors) for $\hat{\beta}_1$. The horizontal red line marks the true effect β_1. Using the single forest method, the estimates drop as we have an increasing number of features and a small sample size. Using the ensemble method, the estimates are not affected by the increase in the number of features.

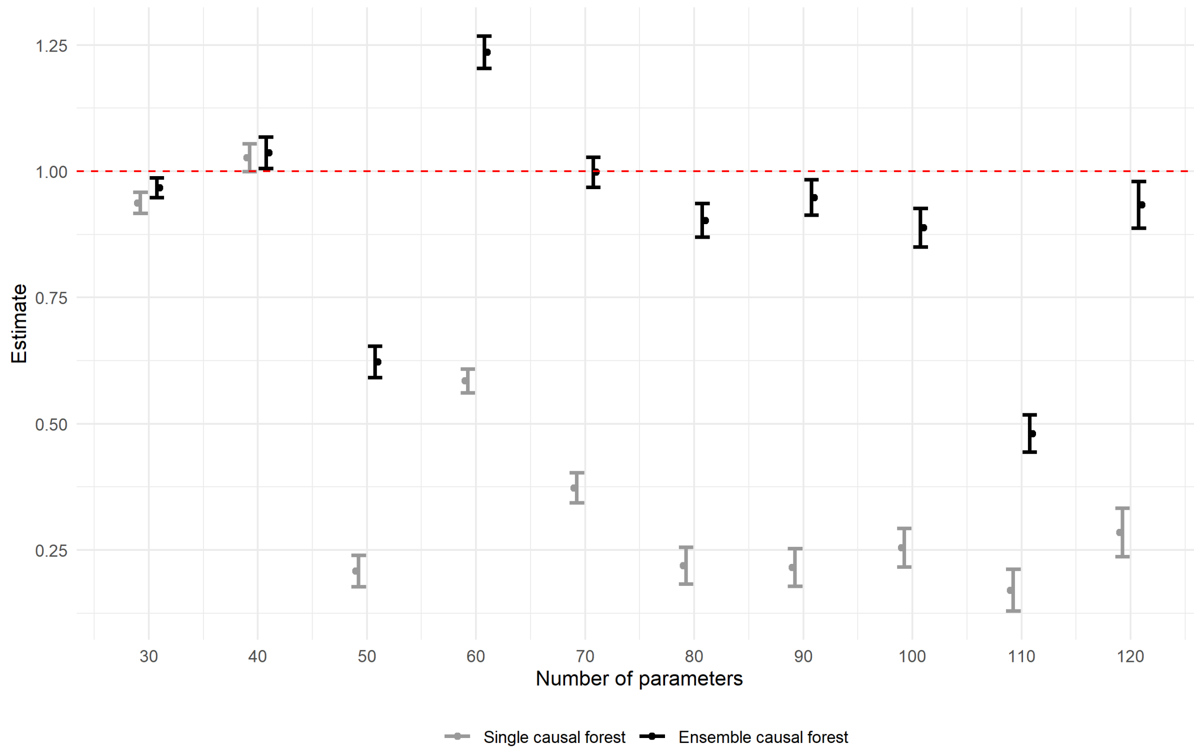

Figure A.1
Ensemble v single causal forest - simulation results

B FIGURES AND TABLES FOR DESCRIPTIVE PART

Table B.1
Main features of the data

topic	variable	topic	variable
time	year	housing	average rent
	quarter		weekly rent $1-74
wealth	gini coefficient		weekly rent $75-99
	income share of top 1%		weekly rent $100-124
	income share of top 5%		weekly rent $125-149
	income share of top 10%		weekly rent $150-174
	% of people in lowest quartile (relative to AUS)		weekly rent $175-199
	% of people in second quartile (relative to AUS)		weekly rent $200-224
	% of people in third quartile (relative to AUS)		weekly rent $225-249
	% of people in highest quartile (relative to AUS)		weekly rent $250-274
employment	number of employed people living in region no		weekly rent $275-299
	number of earners		weekly rent $300-324
	age of earners		weekly rent $325-349
	sum income		weekly rent $350-374
	median income		weekly rent $375-399
	mean income		weekly rent $400-424
business income/cost	mean ind income before tax		weekly rent $425-449
	mean total business income		weekly rent $450-549
	mean total business expense		weekly rent $550-649
	mean net business income		weekly rent $650-749
	mean estimated business tax		weekly rent $750-849
	mean gross rent		weekly rent $850-949
	mean net rent		weekly rent $950 and over
age	age 0-4		nil rent payments
	age 4-10	internet access	internet accessed from dwelling (%)
	age 10-15		internet not accessed from dwelling (%)
	age 15-20		internet access from home / population
	age 20-25	number of cars	no cars
	age 25-30		one motor vehicle
	age 30-35		two motor vehicles
	age 35-40		three motor vehicles
	age 40-45		four or more
	age 45-50		average no cars
	age 50-55	commuting	average commuting distance (mi)
	age 55-60		median commuting distance (mi)
	age 60-65		interquartile range of commuting (mi)
	age 65-70		standard deviation of commuting (mi)
	age 70-75		train
	age 75-80		bus
	age 80-85		ferry
	age 85-99		tram
	age 65+		taxi
	age 35-65		car as driver
	age 15-35		car as passenger
	age 0-15		truck
	population		motorbike scooter
education	index of education and occupation		bicycle
	advanced diploma and diploma level		walked only
	bachelor degree level		worked at home
	certificate I II level		did not go to work
	certificate III IV level	competition	number of rivals within 1 mile
	certificate level		number of rivals within 2 mile
	certificate level nfd		number of rivals within 5 mile
	graduate diploma and graduate certificate level	brand	brand size
	level of education not stated		top brand (bp, shell, caltex)
	postgraduate degree level		number of same brand stations within 1 mile
	index of economic resources		number of same brand stations within 2 mile
	index of relative socio-economic advantage and disadvantage		number of same brand stations within 5 mile
	index of relative socio-economic disadvantage		

Table B.2
Summary statistics of the main variables

	mean	sd	10%	25%	50%	75%	90%
ulp retail price	138.28	10.30	127.29	131.92	137.16	143.72	151.05
diesel retail price	146.19	8.80	136.49	140.77	145.46	150.56	156.83
ulp wholesale price	123.54	6.85	115.94	119.65	123.61	127.78	131.78
diesel wholesale price	129.62	6.77	122.01	125.62	129.46	133.69	137.08
cushing price	25.81	2.75	22.85	24.55	25.84	27.07	28.53
ulp margin	1.12	0.07	1.06	1.08	1.10	1.15	1.21
diesel margin	1.13	0.06	1.07	1.09	1.12	1.15	1.19
number of rivals (within 1mi)	1.85	1.70	0	1	2	3	4
number of rivals (within 2mi)	5.19	4.24	0	2	5	8	10
number of rivals (within 5mi)	21.82	20.94	1	4	14	37	56
median income	50607.2	9084.0	41979	45225	50049	54605	60254
mean income	64962.1	18375.5	51568	55371	61552	68390	79770
usual resident population	12249.7	7150.1	4297	5870	11790	16517	23065
people aged 0-14 years	18.90	4.61	14.3	16.7	19.2	21.7	24.4
people aged 15-64 years	66.35	9.08	60.7	63.2	67	70.4	73.6
people aged 65 years and over	14.75	8.85	6.2	10.1	14.1	18.4	21.1
median age	38.46	6.57	32.2	33.5	37.6	41.5	44.7
sex ratio	109.50	41.00	92	96.9	99.7	105.1	120
earners age	43.07	4.63	37	39	44	47	48
number of earners	7123.57	4754.28	2040	3303	6589	9942	13621
no educational attainment	53.46	72.83	3	10	26	67	136
average commuting distance (mi)	11.27	8.34	4.33	6.18	9.36	13.95	18.33
median commuting distance (mi)	6.74	4.65	1.79	2.79	6.55	8.96	12.99
car as driver	0.28	0.08	0.22	0.25	0.29	0.32	0.34
one motor vehicle	0.25	0.07	0.17	0.20	0.26	0.30	0.32
index of relative socio-economic disadvantage	993.10	80.00	917	975	997	1040	1071
index of relative socio-economic advantage and disadvantage	991.43	74.62	901	956	988	1041	1084
index of economic resources	1003.85	79.78	925	973	1016	1050	1089
index of education and occupation	980.20	73.46	886	928	977	1016	1101

The price statistics are reported for 489,721 weekly observations, the area characteristics are reported for the 1051 distinct petrol station locations.

Table B.3
Margin by competition

level of competition within 1 mile	2 miles	5 miles	ulp margin
low	low	low	1.153
high	low	low	1.189
low	high	low	1.124
high	high	low	1.148
low	low	high	1.094
high	low	high	1.092
low	high	high	1.085
high	high	high	1.094

We split the number of competitors within each radius around their median values. For within 1 mile: 0-2 (low) versus 3 or more (high) competitors, for within 2 miles: 0-5 (low) versus 6 or more (high) competitors, and for within 5 miles: 0-15 (low) versus 16 or more (high) competitors.

Table B.4
Mean margin by levels of competition, income, and search

		low internet		high internet	
		low commute	high commute	low commute	high commute
low competition	low income	1.198 (0.093)	1.124 (0.06)	1.161 (0.066)	1.124 (0.065)
	high income	1.227 (0.095)	1.115 (0.061)	1.207 (0.104)	1.094 (0.048)
high competition	low income	1.142 (0.083)	1.077 (0.027)	1.109 (0.043)	1.087 (0.033)
	high income	1.085 (0.04)	1.076 (0.028)	1.088 (0.032)	1.087 (0.032)

Standard deviation in parentheses.

Table B.5
Number of exits and entries by brand

brand	exitter	entrant	total_frequency
Amgas	1	NA	6
Ampol	15	1	36
BP	17	21	260
Caltex	36	12	254
Caltex Woolworths	1	11	57
Eagle	1	NA	6
Gull	14	NA	61
Independent	7	7	75
Kleenheat	3	NA	13
Kwikfuel	3	NA	5
Liberty	15	4	31
Mobil	10	6	35
Peak	6	5	30
Puma	2	12	89
Shell	24	15	163
Swan Taxis	1	NA	2
United	4	1	26
Wesco	1	NA	12
7-Eleven	NA	3	46
Better Choice	NA	1	8
Coles Express	NA	7	53
Metro Petroleum	NA	2	1
Vibe	NA	5	20
Black and White	NA	NA	1
BOC	NA	NA	5
FastFuel 24/7	NA	NA	1
Oasis	NA	NA	2
United Fuels West	NA	NA	1

Table B.6
Margin by brand by competition by income

station	low income	med income	high income	station	low income	med income	high income
7-Eleven	1.088	1.089	1.084	Liberty	1.111	1.101	1.124
n	6	20	9	n	9	11	8
comp 1mi	1.761	1.899	2.764	comp 1mi	1.022	2.389	2.261
comp 5mi	46.821	33.118	31.138	comp 5mi	13.911	22.951	46.469
Ampol	1.158	1.076	1.091	Mobil	1.081	1.072	1.068
n	15	6	13	n	5	15	7
comp 1mi	1.683	1.866	2.515	comp 1mi	2.158	2.584	3.149
comp 5mi	7.092	53.11	42.823	comp 5mi	47.131	29.382	47.531
BP	1.169	1.099	1.136	Peak	1.118	1.079	1.072
n	66	61	70	n	4	11	13
comp 1mi	2.476	1.548	2.215	comp 1mi	1.976	1.069	1.252
comp 5mi	13.206	28.724	28.633	comp 5mi	30.032	18.292	24.896
Caltex	1.14	1.099	1.12	Puma	1.11	1.078	1.091
n	71	59	64	n	15	26	24
comp 1mi	2.254	1.646	2	comp 1mi	1.269	1.854	2.687
comp 5mi	12.741	30.465	26.431	comp 5mi	19.933	28.445	29.227
Caltex Woolworths	1.102	1.081	1.123	Shell	1.169	1.100	1.133
n	12	13	13	n	51	33	45
comp 1mi	2.491	0.92	2.195	comp 1mi	2.129	1.434	1.659
comp 5mi	21.877	27.859	18.925	comp 5mi	11.983	19.552	27.305
Coles Express	1.108	1.090	1.112	United	1.118	1.064	1.094
n	15	15	18	n	7	7	6
comp 1mi	2.356	2.203	2.431	comp 1mi	4.01	0.805	1.663
comp 5mi	31.78	43.328	32.121	comp 5mi	11.671	28.136	18.224
Gull	1.130	1.108	1.106	Vibe	1.109	1.084	1.054
n	23	14	14	n	6	8	5
comp 1mi	2.653	0.946	1.882	comp 1mi	1.756	1.012	3.53
comp 5mi	11.36	17.385	41.297	comp 5mi	8.382	8.85	43.091
Independent	1.184	1.144	1.206				
n	31	17	12				
comp 1mi	1.199	0.752	2.365				
comp 5mi	3.491	3.343	18.521				

Table B.7
Comparison of treatment and control groups

	exit experiments		entry experiments	
	control	treatment	control	treatment
ULP margin	138.304	137.634	138.304	137.634
	(9.845)	(8.785)	(9.845)	(8.785)
median income	1.121	1.116	1.121	1.116
	(0.072)	(0.064)	(0.072)	(0.064)
Median age	47293.46	47646.6	47293.46	47646.6
	(9683.722)	(8309.644)	(9683.722)	(8309.644)
Number of earners	52454.3	52963.63	52454.3	52963.63
	(18037.731)	(18392.807)	(18037.731)	(18392.807)
People aged 0-14 years	2.053	2.004	2.053	2.004
	(0.607)	(0.654)	(0.607)	(0.654)
People aged 15-64 years	5148.438	4864.065	5148.438	4864.065
	(3376.557)	(2818.705)	(3376.557)	(2818.705)
People aged 65 years and over	0.137	0.135	0.137	0.135
	(0.033)	(0.03)	(0.033)	(0.03)
Median commuting distance kms	0.124	0.132	0.124	0.132
	(0.077)	(0.048)	(0.077)	(0.048)
Gini coefficient	0.601	0.605	0.601	0.605
	(0.138)	(0.114)	(0.138)	(0.114)
Car as driver	0.633	0.638	0.633	0.638
	(0.095)	(0.079)	(0.095)	(0.079)
One motor vehicle	2443.159	2739.144	2443.159	2739.144
	(2343.207)	(2501.03)	(2343.207)	(2501.03)
Usual Resident Population	9639.856	10179.63	9639.856	10179.63
	(6563.955)	(5703.378)	(6563.955)	(5703.378)
Index of Education and Occupation	969.171	985.021	969.171	985.021
	(70.149)	(81.323)	(70.149)	(81.323)
number of rivals (1mi)	2.63	3.413	2.63	3.413
	(2.192)	(1.759)	(2.192)	(1.759)
number of rivals (2mi)	6.142	7.647	6.142	7.647
	(4.869)	(4.43)	(4.869)	(4.43)
number of rivals (5mi)	21.031	26.851	21.031	26.851
	(20.537)	(21.9)	(20.537)	(21.9)

Standard deviation in parentheses.

Table B.8
Margin by brand

brand	frequency	petrol	petrol_margin
BP	260	140.4476	1.139175
Puma	89	134.9497	1.093329
Independent	75	142.4966	1.156507
Shell	163	140.9678	1.142926
Caltex	254	138.6466	1.123753
Coles Express	53	136.3636	1.104578
Caltex Woolworths	57	134.7086	1.090946
Eagle	6	145.0823	1.172944
Liberty	31	136.4341	1.106022
United	26	136.5321	1.102224
Ampol	36	138.5711	1.122965
Gull	61	139.3256	1.130417
Peak	30	132.9341	1.07529
Mobil	35	132.531	1.069455
Vibe	20	133.5344	1.08093
Better Choice	8	131.8751	1.065882
7-Eleven	46	134.6931	1.089123
Wesco	12	134.7267	1.07488
Amgas	6	133.128	
Kwikfuel	5	134.8179	
Oasis	2	134.3975	
United Fuels West	1	141.8159	1.149817
Swan Taxis	2	135.0044	1.078013
Metro Petroleum	1	130.416	1.062722

Table B.9
Number of rivals in the exit and entry samples used for the causal forest estimation

	exit	entry
number of rivals within 1 mile	2.862	2.579
	(2.216)	(1.789)
number of rivals within 2 mile	6.605	5.690
	(4.066)	(5.517)
number of rivals within 5 miles	23.268	15.241
	(19.893)	(21.687)

Standard errors in parentheses.

C TABLES FOR THE RESULTS SECTION

Table C.1
Predicted treatment effects of entry in ULP, by different levels of competition, income, and car reliance

		average no.cars/household		% who drive to work	
		low	high	low	high
low competition	low income	-0.296	-0.287	-0.338	-0.267
		(0.014)	(0.015)	(0.019)	(0.013)
	high income	-0.269	-0.262	-0.317	-0.247
		(0.014)	(0.014)	(0.018)	(0.013)
high competition	low income	-0.284	-0.275	-0.323	-0.255
		(0.013)	(0.014)	(0.017)	(0.012)
	high income	-0.256	-0.25	-0.303	-0.238
		(0.013)	(0.013)	(0.017)	

Standard errors in parentheses.

Table C.2
Predicted treatment effects of exit in ULP, by different levels of competition, income, and measures of search

		commuting distance		internet penetration	
		low	high	low	high
low competition	low income	-0.324	-0.264	-0.252	-0.344
		(0.021)	(0.019)	(0.02)	(0.024)
	high income	-0.298	-0.243	-0.226	-0.317
		(0.02)	(0.018)	(0.019)	(0.024)
high competition	low income	-0.308	-0.252	-0.236	-0.335
		(0.02)	(0.018)	(0.018)	(0.022)
	high income	-0.286	-0.236	-0.209	-0.309
		(0.019)	(0.017)	(0.017)	(0.022)

Standard errors in parentheses.

Table C.3
Predicted treatment effects of exit in ULP, by different levels of competition, income, and search - placebo treatment

	low income	high income
low competition	0.020 (0.014)	0.033 (0.023)
z-score	0.680	
high competition	-0.0073 (0.013)	0.019 (0.012)
z-score	0.678	

Standard errors in parentheses. Study period selected as [-35,-10] weeks before the true exit. Placebo exit at -25.