Startup Search Costs*

David P. Byrne[†] Nicolas de Roos[‡]

April 30, 2018

Abstract

Search frictions determine the impacts of policies or technologies in many markets. Workhorse economic models used to study these impacts assume constant marginal search costs: individuals pay the same search cost to become informed about an additional price (sequential search models), or the entire price distribution (non-sequential search models). This paper provides evidence from a natural experiment in retail gasoline on a new form of search costs: startup costs. We empirically document how a temporary, large exogenous shock to consumers' search incentives leads to a substantial permanent increase in search intensity for gasoline prices. This result is difficult to explain with a standard search model but follows directly in a model with a one-time up-front cost to start searching.

JEL Classification: D83, L81

Keywords: Search, Price Clearinghouse, Retail Gasoline

*We recommend viewing this article in color. We thank Simon Loertscher, Jose Moraga-Gonzalez, Matthew Shum, Michelle Sovinsky, Tom Wilkening, Nan Yang and seminar participants from the University of Melbourne, the University of Groningen, the University of Vienna, the 2016 Asia-Pacific IO Conference, and the 2018 Tinbergen Institute Empirical IO Workshop for helpful comments.

[†]Department of Economics, University of Melbourne, 111 Barry Street, VIC 3010, Australia, e-mail: byrned@unimelb.edu.au.

[‡]School of Economics, Merewether Building H04, University of Sydney, NSW 2006, Australia, e-mail: nicolas.deroos@sydney.edu.au.

1 Introduction

To purchase a flight, fill up one's gas tank, find a new book to read, a nice restaurant, a new pair of shoes, or even a romantic partner, consumers must set aside time and effort to search for the most suitable or affordable option. According to the theory of search, the trade off between these costs of search and the expected benefits of search determines the end of the search process and the set of products ultimately considered by the consumer.[1] In applications, parsimony is important, and a typical specification presumes that the marginal cost of search is independent of the consumer's search experience in prior purchases.

In this paper, we provide evidence on a new form of search costs that we call startup search costs. These costs represent one-time, up-front costs individuals incur the first time they search. After the startup search cost is incurred, the marginal cost of search for any future purchases is reduced.

We provide evidence on the existence of startup search costs using a unique dataset and natural experiment in retail gasoline. Section 2 describes the dataset. The data comes from an urban gasoline market that has an online price clearinghouse which consumers can use to become informed about the price distribution day-to-day. We have access to the universe of station-level gasoline prices from 2014-2016, which we match to daily website usage from the price clearinghouse. Our ability to perfectly measure price levels and dispersion, and directly measure search behavior from a price clearinghouse in a homogeneous product market makes this context particularly well-suited for studying the dynamics of search and price dispersion.[2]

In Section 3, we describe the natural experiment and present our main empirical results. The natural experiment stems from a price war that disrupts a stable coordinated pricing equilibrium. In a separate study, Byrne and de Roos (2017b), we document *five years* of stable and highly coordinated pricing in the market from March 2010 to April 2015, which is consistent with tacit collusion. However, in May 2015, a price war occurs and price coordination breaks down. During a three-week conflict period, retail price dispersion spikes and gains from search grow by 100% relative to their pre-war levels. Firms eventually resolve the war and return to stable coordinated pricing, and price dispersion and search incentives return to their pre-war levels after five months. Based on pre-war retail price and search dynamics in our dataset, and anecdotal evidence of a retail ownership change that precipitated the war, we argue in Section 3 that the price war is exogenous to search behavior on the price clearinghouse.

[1] Baye, Morgan and Scholten (2006) overview an extensive and influential literature on search models, which date back to Stigler (1961).

[2] See also De los Santos, Hortascu and Wildenbeest (2012) and Koulayev (2014) for analyses of search models using web search data. See Eckert (2013) for an overview of an extensive literature on retail search and price dispersion in retail gasoline.

We study search intensity on the price clearinghouse one year before and after the temporary shock to price dispersion and search incentives. Our main empirical result is that daily search intensity permanently rises by 70% following the shock. We show that this rise in search intensity comes from an increase in the number of unique searchers using the clearinghouse, and not an increase in average search intensity among searchers. This pattern of history dependent search points to new consumers experimenting with the clearinghouse for a first time in response to the substantial price uncertainty created by the price war. Having tried it once, consumers continue to use the clearinghouse thereafter, even after search incentives return to pre-war levels.

Section 4 formalizes our notion of startup search costs, estimates their magnitude, and shows how ignoring them leads to model misspecification. Specifically, we develop and estimate a standard non-sequential search model that incorporates startup search costs.[3] Exploiting the natural experiment, we estimate the relative magnitudes of startup search costs and marginal search costs. Our simple search model rationalizes the permanent rise in search intensity from a temporary shock to search incentives. By contrast, a standard non-sequential search model is unable to rationalize the persistent increase in search.

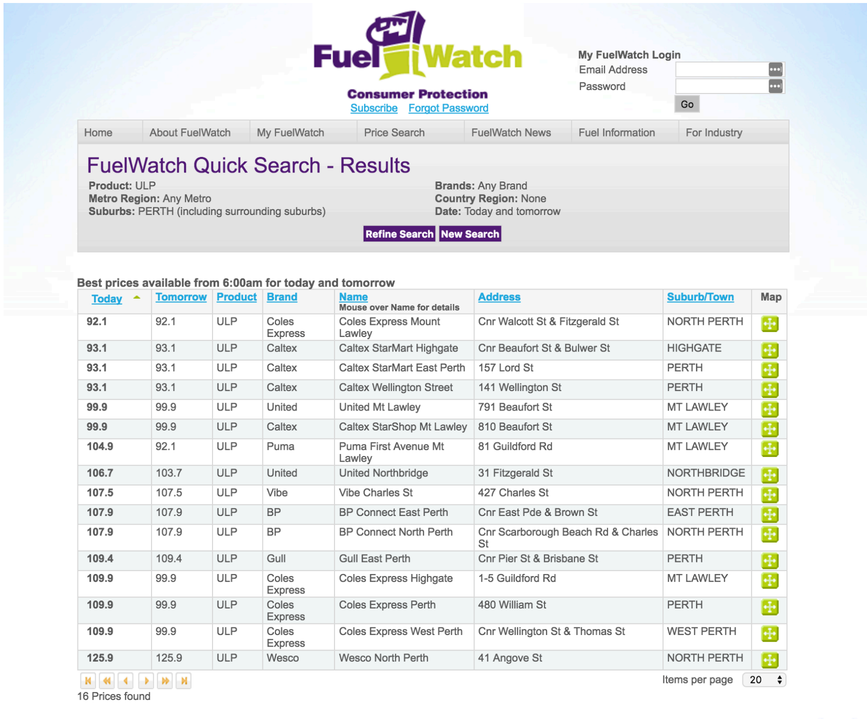
We conclude in Section 5. Our findings suggest the omission of startup search costs leads to qualitatively different inferences regarding search behavior. Moreover, our study illustrates the challenges faced by policymakers when designing tools to enhance price transparency when consumers face startup search costs: the price clearinghouse we study had been in existence for 15 years by May 2015, more than 90% of the market were aware of its existence, yet the price war shock led to a 70% rise in the volume of search. Finally, we highlight avenues for future research. While we provide evidence on the existence of startup search costs, further work is required to elicit their microeconomic foundations. We discuss deeper possible mechanisms for the history dependence in search that we find.

2 Context and data

Our research context is Perth, Australia, a city with approximately 2 million people. Perth, like many urban gasoline markets worldwide, has a concentrated retail gasoline market. Four major firms dominate the market: BP, Caltex, Coles and Woolworths. The former two firms are vertically integrated oil majors, while the latter two firms are major supermarket chains that also sell gasoline. All four firms either directly or indirectly control retail prices day-to-day at their large station networks. All other stations in the market are operated by independent retailers.

[3] Under "non-sequential" search, search reveals the entire vector of prices. This matches the market environment in which a comprehensive price clearinghouse operates.

Figure 1: Fuelwatch Price Clearinghouse



Today	Tomorrow	Product	Brand	Name (Mouse over Name for details)	Address	Suburb/Town	Map
92.1	92.1	ULP	Coles Express	Coles Express Mount Lawley	Cnr Walcott St & Fitzgerald St	NORTH PERTH	
93.1	93.1	ULP	Caltex	Caltex StarMart Highgate	Cnr Beaufort St & Bulwer St	HIGHGATE	
93.1	93.1	ULP	Caltex	Caltex StarMart East Perth	157 Lord St	PERTH	
93.1	93.1	ULP	Caltex	Caltex Wellington Street	141 Wellington St	PERTH	
99.9	99.9	ULP	United	United Mt Lawley	791 Beaufort St	MT LAWLEY	
99.9	99.9	ULP	Caltex	Caltex StarShop Mt Lawley	810 Beaufort St	MT LAWLEY	
104.9	92.1	ULP	Puma	Puma First Avenue Mt Lawley	81 Guildford Rd	MT LAWLEY	
106.7	103.7	ULP	United	United Northbridge	31 Fitzgerald St	NORTHBRIDGE	
107.5	107.5	ULP	Vibe	Vibe Charles St	427 Charles St	NORTH PERTH	
107.9	107.9	ULP	BP	BP Connect East Perth	Cnr East Pde & Brown St	EAST PERTH	
107.9	107.9	ULP	BP	BP Connect North Perth	Cnr Scarborough Beach Rd & Charles St	NORTH PERTH	
109.4	109.4	ULP	Gull	Gull East Perth	Cnr Pier St & Brisbane St	PERTH	
109.9	99.9	ULP	Coles Express	Coles Express Highgate	1-5 Guildford Rd	MT LAWLEY	
109.9	99.9	ULP	Coles Express	Coles Express Perth	480 William St	PERTH	
109.9	99.9	ULP	Coles Express	Coles Express West Perth	Cnr Wellington St & Thomas St	WEST PERTH	
125.9	125.9	ULP	Wesco	Wesco North Perth	41 Angove St	NORTH PERTH	

A key aspect of the market is a price transparency program called Fuelwatch. The program was introduced on January 3, 2001 by Western Australia's state government. By law, before 2pm each day firms must submit CSV files to the state government that contain tomorrow's station-level retail prices. The next day at 6am when stations open, they are required to post the prices that were submitted at 2pm the previous day. Prices are then fixed for 24 hours. From our conversations with the government, compliance with the program is near perfect.

With the Fuelwach price data, before 2pm each day the government posts online today's prices for all stations in the market. This helps customers engage in cross-sectional price search. After a data verification check, at 2:30pm the government further posts tomorrow's price for every station in the market. This helps customers engage in inter-temporal price search. Figure 1 depicts the Fuelwatch price clearinghouse at www.fuelwatch.gov.au. Survey evidence from the Western Australian Government indicates that more than 90% of Perth households are aware of the Fuelwatch price clearinghouse.

2.1 Data

The price clearinghouse generates a uniquely rich dataset for studying retail search and price dispersion. From the clearinghouse, we have access to the universe of retail prices from 2001 to present day. This allows us to perfectly measure daily price levels and dispersion. We match

Table 1: Summary Statistics

	Obs.	Mean	Std. Dev	Min	Max
Price data					
Station Retail Price	146931	128.64	13.67	92.9	169
Market Terminal Gate Price	731	118.3	12.6	95.8	147.0
Station Retail Margin	146931	10.3	6.8	-17.0	35.7
Search data					
Daily website visits	731	19699	6691	8243	53517
Monthly unique website visitors	24	208184	47767	129630	276737

Notes: Sample period is July 1, 2014 to June 30, 2016.

these price data to the daily terminal gate price (TGP) for gasoline, which is the local spot price for wholesale gasoline. These spot prices include a margin for upstream suppliers of gasoline. The difference between the retail price and TGP provides an estimate of the retail margin. It is an appropriate estimate for studying the evolution of margins over time because the TGP is the main time-varying component of stations' wholesale fuel costs. Other time-invariant parts of marginal costs include quantity discounts, shipping costs, wharfage fees, and insurance costs.

Moreover, the state government provided us access to daily website usage data from the Fuelwatch website. More specifically, the government provided daily data on the total number of website hits on the Fuelwatch website, and the total number of unique visitors to the Fuelwatch website month-to-month. These search data, combined with the universe of station-level price data, permit a direct examination of how retail search behavior responds to changes in price levels and dispersion at the market level over time.[4]

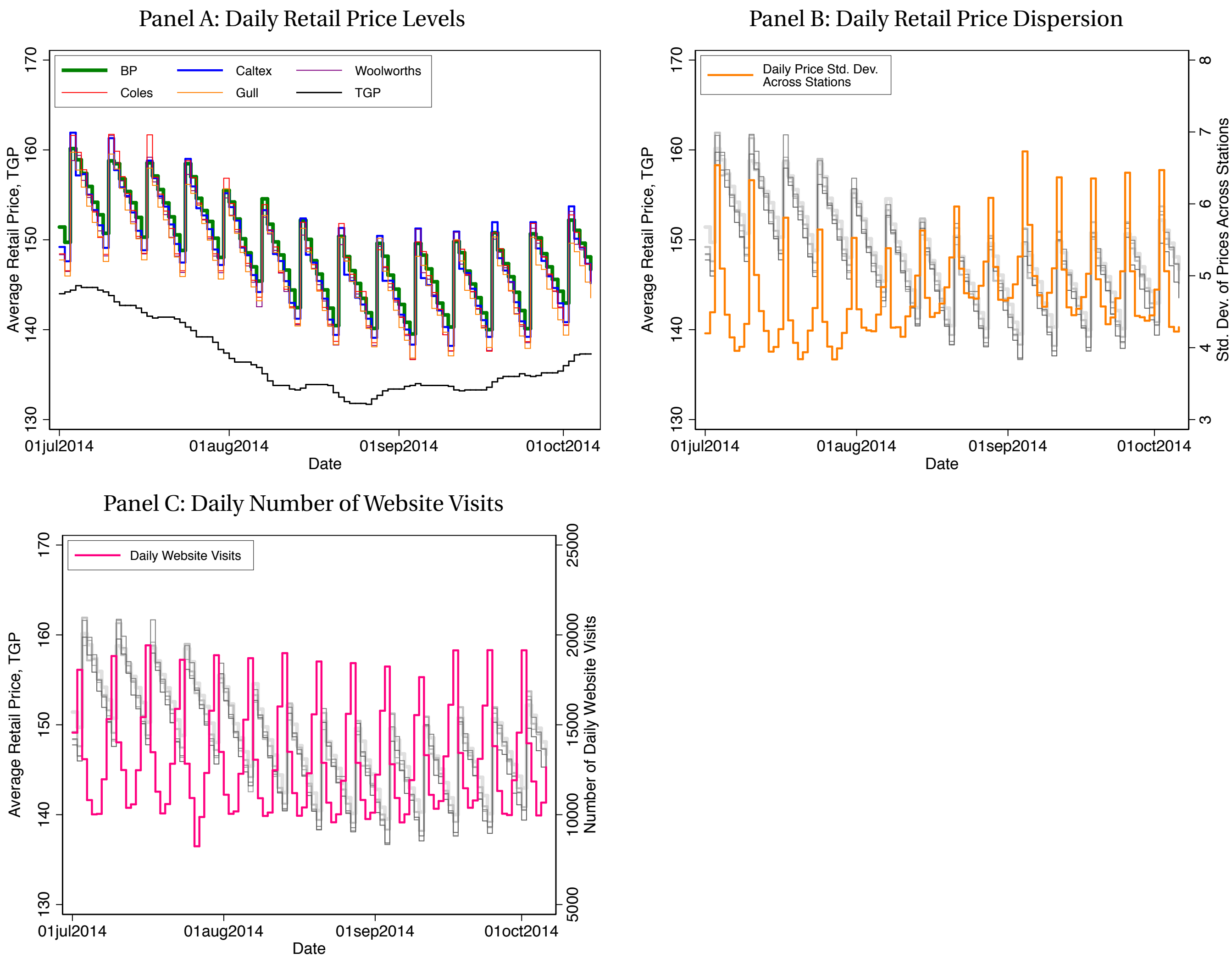
Our sample period spans two-years, from July 1, 2014 to June 30, 2016. Over this period, average station-level retail prices and margins in terms of cents per litre (cpl) are 128.6 cpl and 10.3 cpl, respectively. The average number of Fuelwatch website visits each day is 19699, and there are 208184 unique visitors to the website each month on average. Table 1 presents summary statistics from our dataset.

2.2 Price cycles and search cycles

Figure 2 depicts cyclical patterns in retail price levels, price dispersion, and search that persist over the entire sample period. Panel A, which plots average daily retail prices by retailer and the daily TGP, highlights cycles in price levels. Over the time period depicted in the figure, every Thursday, retail prices jump by approximately 10% of the average market price. Between price

[4]Daily data on the total number of unique visitors is unavailable due to data privacy concerns from the state government. Similarly, daily data on website hits at the individual level or by disaggregated census blocks are not available because of privacy concerns.

Figure 2: Price and Search Cycles Cycles



jumps, most retailers cut prices by 2 cpl each day until the next price jump day occurs. Intertemporal search incentives are therefore heightened on Wednesdays in anticipation of price jumps.

Panel B plots corresponding daily market price dispersion, as measured by the daily standard deviation of retail prices across stations. For reference, we overlay daily price dispersion with the price cycles from panel A in greyscale. The figure highlights spikes in price dispersion on Thursdays, as the average of the standard deviation of retail prices rises from 4.97 at the bottom of the cycle, to 6.77 cpl at the top of the cycle. This large rise in price dispersion reflects a noisy coordination process as market prices transition from the bottom to the top of the cycle. Cross-sectional search incentives thus tend to rise substantially on Thursdays.

Panel C of Figure 2 plots daily search intensity on the Fuelwatch website, as measured by the number of website hits from computers and mobile devices. Like prices, search exhibits cycles, whereby search intensity jumps on Wednesdays and Thursdays. On the day before and day of price jumps, on average the website receives 27049 and 19531 visits, respectively. On all other days, the website on average receives 18263 visits. In sum, search intensity rises with cross-sectional and inter-temporal search incentives just before and during price jumps.

3 Dynamics of price dispersion and search

In this section, we examine the coevolution of retail prices and search over the entire two-year sample period. We document the break out of a price war among the retailers who were previously coordinating on the timing and magnitude of price jumps and cuts. We describe the impact of the price war on price levels, dispersion, and margins. We then show that daily search intensity permanently rises by 70% following the price war, and argue that this response of search to a temporary war-induced shock to search incentives is causal.

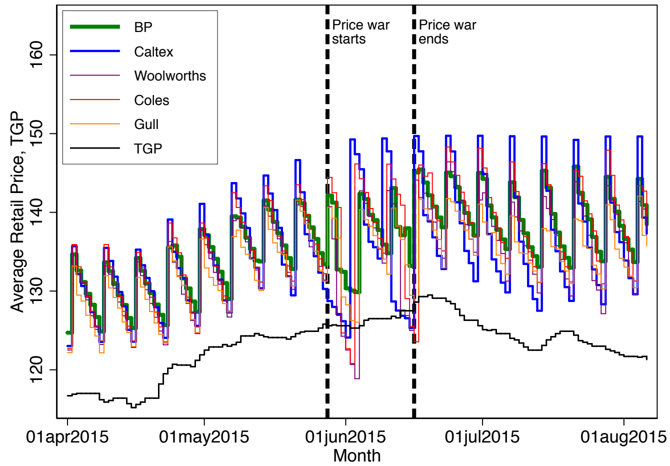
3.1 Price War

As alluded to above, the price cycles in Figure 2 are stable and regular. As we document in Byrne and de Roos (2017b), between March 2010 and April 2015, firms coordinate on price cycles using two simple focal points: Thusday price jumps and 2 cpl price cuts on days of the week between jumps. Over this five year period *every* market price jump occurs on Thursday. The result is a tighty coordinated price cycle that is consistent with tacit collusion.

However, after five years of stability, Caltex breaks from this pricing strategy in May 2015. This is depicted in Figure 3, which plots average daily retail prices for the four major companies and the independent retailer, Gull. The figure highlights the disruption Caltex's defection creates, particularly with the timing of price jumps. In particular, in the last week of May, Caltex defects from Thursday price jumps and instead engages in Tuesday jumps. After three weeks of turmoil and an uncoordinated price cycle, the other major retailers along with Gull transition to Tuesday price jumps, re-establishing a coordinated price cycle. Tuesday price jumps are subsequently stable from June 15 to present day (ACCC, 2017).

Exogeneity of the Price War to Search. We argue that the price war and its corresponding effects on price levels and dispersion is exogenous to search behavior with the price clearinghouse. We believe this for three reasons. First, as mentioned, the coordinated pricing structure is stable for five years prior to the May 2015 price war, and the price war occured without warning. Indeed, media releases from the Western Australian government and local major news

Figure 3: 2015 Price War



outlets highlight a high degree of unpredictability of prices during the price war. For example, a manager at the Western Australian Royal Automobile Club describes the public's surprise in June 2017 over the price war and transition to Tuesday price jumps as follows:[5]

> *"We don't have a full understanding of why the cycle has changed ... and we want to understand why this is happening. It has changed after a period of certainty and we don't know what the future looks like"*

Second, based on our conversations with the Western Australian Government, the only shock in the market that can potentially be linked to the price war is a major change in ownership at Caltex. In March 2015, the Chevron Corporation sold off its 50% share in Caltex Petroleum Australia Pty. Ltd. to Australian shareholders; after this sale, Caltex becomes 100% owned by shareholders.[6] This supply-side shock in ownership may have led to a change in management and pricing tactics, and hence the price war. However, the ownership change is unrelated to local demand or online price search in Perth.

Finally, as we show momentarily in Figure 4 below, search behavior on the Fuelwatch price clearinghouse is stable for the entire year prior to the price war. There is no evidence to suggest that changes in online search precipitates the price war.

Given these empirics and anecdotal evidence, we assume that the price war and related changes in price dispersion are exogenous to search behavior on the price clearinghouse. We

[5]The quote appeared on July 6, 2015 in an Australian Broadcasting Corporation article entitled "Monday cheapest day to buy fuel in Perth in change to long-running petrol cycle", accessed at http://www.abc.net.au/news/2015-07-06/monday-now-cheapest-for-fuel-in-perth/6598290.

[6]See, for example, the Australian Competition and Consumer Commission "Report on the Australian petroleum market", March quarter 2015.

therefore interpret the empirical relationship between search and price dispersion as causal and reflecting demand-side search behavior. We further pursue this interpretation in estimating a search model in Section 4.

3.2 Evolution of search and price dispersion

Figure 4 presents time series for daily average retail price levels (panel A), margins (panel B), price dispersion (panel C) and search (panel D) before and after the price war. In each panel, we plot the raw daily time series in greyscale, and the weekly average of the daily series in color, which more clearly depicts trends. In panel A we see that price levels primarily trend with wholesale costs over time. Panel B highlights cyclical daily margins that arise because of the price cycle. At weekly frequencies, margins hover at around 10 cpl, and average 7 cpl during the price war. There are no other discernable trends in margins before or after the price war.

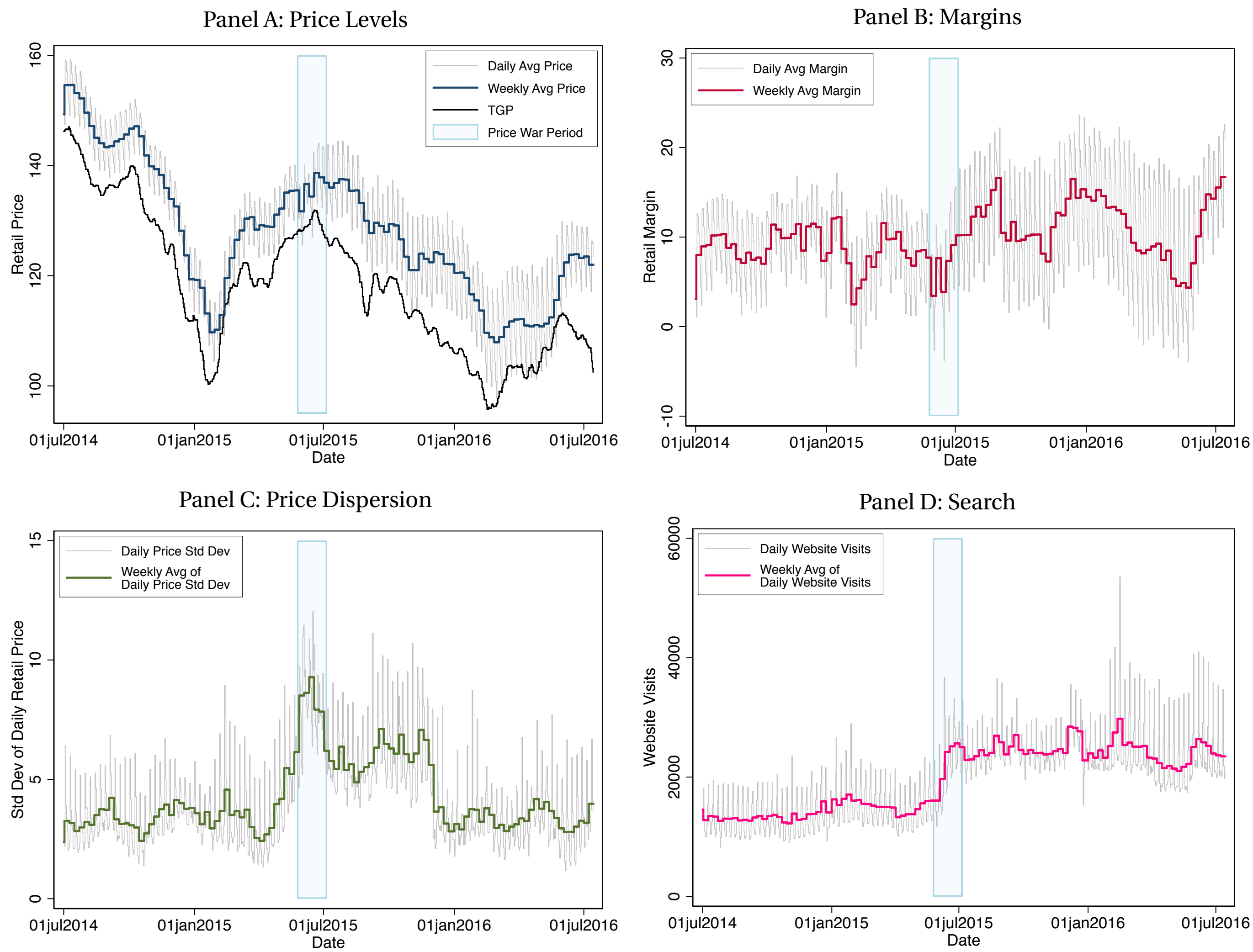
Panel C illustrates how price dispersion, as measured by the daily standard deviation of prices, rises by more than 100% around the price war. During this period, search incentives are higher than they have been in the past five years. Dispersion drops from its peak immediately after the price war is resolved, and gradually returns to pre-war levels six months after the price war in January 2016.[7] That is, the price war creates a large, temporary exogenous shock to search incentives in the market.

How does search respond to these price fluctuations? Panel D of Figure 4, which we view as the paper's central result, provides the answer: despite the temporary shock to search incentives, we find a permanent increase in search intensity with the price clearinghouse. The increase in daily search is large: it rises from an average of 14097 visits before the shock to 24461 visits after the shock, a 70% increase. Importantly, this jump is *not* driven by emails or text messages from Fuelwatch that might cue search behavior; such messages are not sent to consumers. The shift in panel D reflects new and permanent active effort in using the price clearinghouse following the shock.

This substantial and permanent rise in search intensity could reflect a rise in search intensity among existing searchers, or the emergence of new searchers, or both. Figure 5 provides evidence that strongly suggests it is driven by new searchers. With the left axis we plot the number of unique visitors to the Fuelwatch price clearinghouse month-to-month. We find the average number of unique visitors permanently rises from 151,677 visitors pre-war to 239,959 visitors post-war, a 60% increase in the number of unique searchers. With the figure's right axis,

[7] While the transition to Tuesday price jumps is completed after the three week price war, cross-sectional price dispersion remains elevated on price jump days for several months after the price war. This dispersion reflects larger price jumps by Caltex relative to its rivals. By November 2015, firms are again able to coordinate on the size of price jumps, and price dispersion returns to baseline.

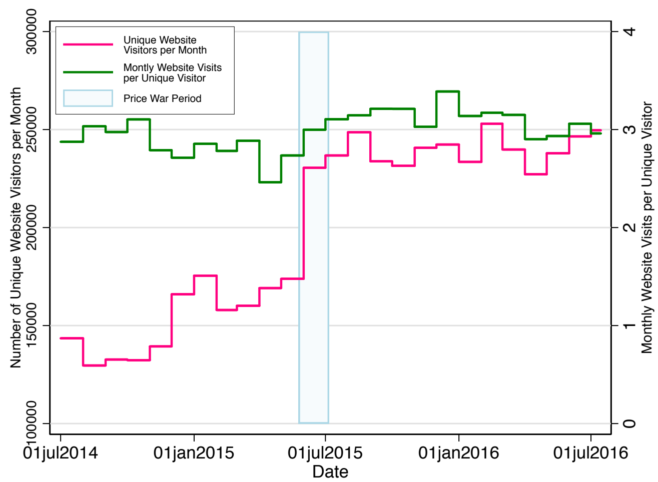
Figure 4: Prices, Margins and Search Before and After the Price War



we plot the monthly number of visits to the Fuelwatch price clearinghouse per unique visitor. This remains stable around three visits per month before and after the price war. That is, search intensity per searcher does not appear to dramatically rise after the price war.

These patterns present a challenge for conventional search models. If searchers are myopic and face search costs that are unchanged irrespective of past search behavior, then search levels should return to their baseline levels as search incentives return to their baseline levels over time. This is clearly not the case in Figures 4 and 5. Past search experience appears to be important for future search behavior. That is, there is history dependence in search. These results suggest that the first time a consumer engages in search, "startup" search costs are potentially high. However, conditional on sinking these costs, the persistence of search levels among new searchers following the shock indicates that marginal search costs from using the price clear-

Figure 5: 2015 Price War



inghouse are small.

4 Estimating startup search costs

In this section, we estimate a simple search model that incorporates startup search costs in an otherwise standard model of non-sequential price search. We introduce the model in Section 4.1, discuss estimation and identification in Section 4.2, and present results in Section 4.3.

Our goal is to provide indicative estimates of the relative magnitudes of startup and marginal search costs in a parsimonious model. We specify non-sequential search on the price clearinghouse because households obtain information on the entire distribution of prices when they search on it. We also presume that consumers are unsophisticated in the sense that they consider only the current-period benefits of a decision to use the search platform. A sophisticated consumer, aware of her startup search costs, would weigh the expected net present value of future benefits of learning to use the search technology, leading to higher estimates of startup search costs.[8] Our specification also abstracts from the incentive for intertemporal search. In Byrne and de Roos (2017a), we find evidence for intertemporal search in this market, which contributes to within-week variation in search. By focusing on cross-sectional search, our model is better suited to identifying longer term trends in search rather than within-week fluctuations.

Our model setup reflects the data we work with. We have market-level search data, and therefore we are unable to identify individual-level heterogeneity in the model's parameters.[9]

[8] In Appendix B, we show that, when consumers are sophisticated, inferred startup costs are higher for more patient consumers, while inferred marginal search costs are unaffected.

[9] See, for example, Koulayev (2014) for an application that studies online price search using individual-level search behavior from an online hotel price search website. As mentioned in Section 2, we requested such data

Moreover, as with virtually all research on gasoline demand, we do not have access to high-frequency data on quantities of gasoline purchased and consumers' fuel tank inventories.[10]

We focus strictly on the demand-side of the market for two reasons. First, as argued in Section 3, daily price changes are plausibly exogenous to online search. Given that we have a direct measure of search intensity, we can use the demand side of a search model alone to identify the search costs from using the price clearinghouse. Second, in Byrne and de Roos (2017b) we argue that firm behavior is consistent with tacit collusion over our sample period. Therefore, we cannot use standard static first-order conditions to model pricing and build supply-side moments, as in Hong and Shum (2006) or Wildenbeest (2011), to help identify search costs. A dynamic model of the supply side of the market is beyond the scope of the current paper.

4.1 Model

Consumer i's indirect utility on date t if she purchases τ_i liters of petrol is given by

$$U(s_{it}, \mathbf{p}_{it}) = \begin{cases} \overline{u} - \tau_i \min\{\mathbf{p}_{it}\} - s_{it} & \text{if she searches} \\ \overline{u} - \tau_i \overline{p}_{it} & \text{otherwise,} \end{cases}$$

where $\mathbf{p}_{it}$ is the vector of prices available to consumer i at date t and s_{it} is the current cost of search for consumer i. As we discuss below, the set of stations available to consumer i is based on geographic proximity to her home address. This formulation assumes that if consumer i on date t engages in price search on the clearinghouse, she becomes fully informed about the price distribution and pays the minimum price in her local choice set, $\min\{\mathbf{p}_{it}\}$. If she does not search, she pays the average price in her local choice set, $\overline{p}_{it}$. Her gains from searching in period t are

$$g_{it} = (\overline{p}_{it} - \min\{\mathbf{p}_{it}\}) \times \tau_i.$$

Given a search cost parameter vector θ, consumer i's search costs in period t are given by

$$\begin{aligned} s_{it}(\theta) &= f_i \times (1 - w_{it}(\theta)) + c_i, \\ w_{it}(\theta) &= 1\{\text{consumer } i \text{ has searched before date } t\}, \end{aligned}$$

where $1\{\cdot\}$ is an indicator function, and f_i and c_i are consumer $i's$ startup and marginal search

for Fuelwatch from the Western Australian Government. Unfortunately, they are prohibited from providing such search data at the individual level or disaggregated into local areas.

[10] Levin, Lewis and Wolak (2016) provide the first ever published research on daily market-level gasoline demand behavior. With their unique data, they are able to distinguish between the binary decision to purchase gasoline and how much gasoline to purchase conditional on deciding to purchase. Their reduced-form study abstracts from search frictions and inventories.

costs. The costs f_i and c_i are assumed to be independently distributed across consumers according to gamma distributions (Hong and Shum, 2006) with shape parameters μ_f and μ_c and scale parameters σ_f and σ_c.[11] We collect these search cost parameters with $\theta = [\mu_f, \sigma_f, \mu_c, \sigma_f]'$.

Recall that the Fuelwatch price clearinghouse allows consumers searching after 2:30pm to discover prices for today and tomorrow. To account for this, consumer i considers the gains from search in periods t and $t+1$ when making her search decision on date t:

$$y_{it}(\theta) = 1\{\max\{g_{it}, g_{it+1}\} > s_{it}(\theta)\}.$$

The share of consumers engaging in online price search on date t is

$$q_t(\theta) = \frac{\sum_{i=1}^{Q} y_{it}(\theta)}{Q},$$

where the market size Q represents the number of consumers considering a gasoline purchase each day.[12]

4.2 Estimation and identification

We estimate the model using a Simulated Minimum Distance estimator that compares the share of searchers predicted by the model, $q_t(\theta)$, to its empirical analogue, $\hat{q}_t$, computed as

$$\hat{q}_t = \frac{n_t}{Q}$$

where n_t is the number of Fuelwatch website hits on date t.

Computing $q_t(\theta)$ and $\hat{q}_t$ requires us to make an assumption regarding the market size, Q. We compute this as $Q = (0.80 \times 1,576,000)/7$, which assumes that 80% of the population in Perth aged between 15 and 79 years plans to fill up their car once every week, and does so uniformly by day of the week. We calibrate the size of a gasoline purchase to $\tau_i = 50$ liters for all consumers. The most popular car in Australia is the Toyota Corolla, which has a 55 liter tank. According to this calibration, each representative consumer fills up their Toyota Corolla when it is almost

[11] We have estimated the model under alternative functional form assumptions, with no qualitative differences in reported results. In particular, we have estimated the model under the assumption that f_i and c_i are drawn from the Log Normal distribution, both under the assumption that startup and marginal search costs are independent, and allowing for correlation between them.

[12] By fixing the market size over time, we implicitly introduce an outside good into the model. Because some consumers are aware of the cyclical nature of prices in this market, there is a cycle in sales volumes over the week. See, for example, Australian Competition and Consumer Commission (2014) on the existence of a demand cycle in the Perth market. Fixing the market size amounts to assuming that, each day, the same volume of consumers considers both whether to purchase gasoline and whether to use the search platform.

empty.[13]

To measure the gains from search, g_{it}, we partition the region of Greater Perth into local districts as classified by the Australian Bureau of Statistics.[14] For each district, we obtain driving age (18-79 years) population and the location of the centroid of the district from the 2011 Census. By matching this to the location of each station, we obtain the set of stations within a 5km radius.[15] Average and minimum prices for each district are defined with reference to this local set of stations. To construct our simulated minimum distance estimator, we then assign each consumer randomly to a district according to weights based on driving age population. By calculating local search gains in this manner, we abstract from commuting patterns.[16]

Let $u_t(\theta) = \hat{q}_t(\theta) - q_t$ be the difference between the model's prediction and the fraction of searchers in the data at date t, and let $u(\theta)$ be the $T \times 1$ vector of prediction errors, where T is the number of dates in the sample. We estimate θ by minimizing the objective function

$$\hat{\theta} = \arg\min_{\theta} G(\theta) = u(\theta)' u(\theta).$$

For a given value of θ, we compute $G(\theta)$ by forward simulating the search shares, $q_t(\theta)$, for each sample date $t = 1, \ldots, T$. In simulating search shares, we must keep track of each simulated consumer's history of search, embodied in $w_{it}(\theta)$. For this forward simulation we need to initialize the set of "active" customers in the market who have already sunk their startup search costs at the start of the sample period, $w_{i0} = 1$. We calibrate the fraction of active consumers using the maximum value of n_t/Q in the two weeks prior to the start of the estimation sample. We then assign active status randomly across consumers such that $\sum_i w_{i0} = \max_{t \in T_0} n_t$, where T_0 is the two week pre-sample. With this method, 10% of consumers purchasing gasoline on a given date have already sunk their startup search costs at the start of the sample.

[13] Data sources for our calculations are as follows. From Australian Bureau of Statistics Table 3235.0, "Population by Age and Sex, Regions of Australia", there were 1,576,479 people aged between 15 and 79 years in the greater Perth area on June 30, 2014. The Australian Competition and Consumer Commission (2007) report into the Australian petrol market commissioned a survey of 775 motorists in Australia, finding that 26% purchase more than once per week, 50% purchase once per week, 20% purchase every 2 weeks, and 4% purchase less than every 2 weeks. According to the Federal Chamber of Automotive Industries (FCAI), the most popular car in Australia in 2014 was the Toyota Corolla (see: https://www.drive.com.au/motor-news/the-10-most-popular-cars-of-2014-20150105-12ihkp). The fuel tank capacity for a Toyota Corolla is 55 liters (see http://www.toyota.com.au/corolla/specifications/ascent-sedan-manual).

[14] We use the finest classification available from the 2011 Census, known as Statistical Areas, Level 1. Of the 3789 Statistical Areas in the Greater Perth region, the average district has 318.7 people (s.d. 128.3).

[15] Our main qualitative findings are unaffected by the choice of search radius. However, as we might expect, increasing the radius of search raises the value of information, and leads us to infer higher search costs. See Appendix A, where we estimate the model using a 2km and a 10km search radius, for details.

[16] See Pennerstorfer et al. (2016) for an analysis of commuting routes and search incentives. Because our goal is to obtain indicative estimates of the relative magnitudes of startup and marginal search costs, rather than engage in counterfactual analysis, we see no reason that the assumption of local search affects our main findings.

Table 2: Search Model Estimation Results

	With startup Search Costs (1)	Without startup Search Costs (2)
Marginal search cost distribution		
μ_c	0.202	0.611
	(0.016)	(0.007)
σ_c	40.769	160.257
	(3.039)	(5.475)
Startup search cost distribution		
μ_f	5.070	
	(0.086)	
σ_f	4.228	
	(0.203)	
Objective function, $G(\hat{\theta})$	0.356	1.061

Notes: Robust standard errors are in parentheses (). The number of observations is $T = 731$ dates. All calculations assume consumers purchase 50 liters of gasoline.

Identification. The distribution of marginal search cost is identified by periods in which the benefits of search are unremarkable. Thus, variation in aggregate search activity associated with variation in search gains, when such gains are moderate, identifies the parameters μ_c and σ_c. By contrast, the startup search cost parameters μ_f and σ_f are identified by the responsiveness of aggregate search to unprecedented gains from search arising from the price war. Finally, over the sample period, aggregate search varies between 5% and 30% of consumers (see Figure 7). Therefore, estimation is well-suited for identifying the search cost distribution corresponding to this range of search intensities, but not the entire search cost distribution.

4.3 Results

Estimation results are presented in Table 2. Column (1) contains estimates for the full model, and column (2) contains estimates for a constrained model in which there are no startup search costs. Panels (A) and (B) of Figure 6 present the cumulative density functions of the startup and marginal search cost distributions based on the point estimates. To illustrate the qualitative difference between the models, consider the 20th percentile consumer. For the model with startup search costs, the 20th percentile startup costs are \$13.30/day. Conditional on having sunk this startup cost, the 20th percentile marginal search cost is \$0.01/day.[17] If startup search costs are removed from the model, estimated marginal search costs are an order of magnitude greater. The 20th percentile consumer has marginal search costs of \$9.94. This conforms with intuition, as these marginal search cost estimates are driven by both startup search costs and marginal search costs.

Turning to model fit, from Table 2 we find the econometric objective function for the full model of $G(\hat{\theta}) = 0.356$ is substantially lower than the model without startup search costs, where $G(\hat{\theta}) = 1.061$. Figure 7 further describes the implications for model fit from accounting for startup search costs. Panel A shows that, despite being simple, our model with startup costs is able to recreate the amplitude and frequency of search cycles, with the amplitude being somewhat underestimated. Importantly, the full model precisely fits the sharp and permanent shift in search intensity after the temporary shock to search incentives caused by the price war.

Panel B of Figure 7 yields a stark contrast for the model without startup costs. While the model is able to capture the amplitude and frequency of search cycles, the model is unable to produce a permanent shift in search behavior after the shock. As the gains from search return to baseline following the shock, in the absence of startup costs, the standard non-sequential search model predicts search intensity will also return to baseline.

Finally, panel C shows the predicted evolution of consumer experience with search. In grey (left scale), we depict the population-weighted average gains from search, and in the foreground (right scale), we show the predicted fraction of active consumers. The fraction of consumers who have incurred startup search costs is approximately a step function over time. When the gains from search are abnormally high in the middle of the sample, the predicted fraction of consumers with search experience grows rapidly from 12% to 18%. This leads to a permanent increase in predicted search activity.

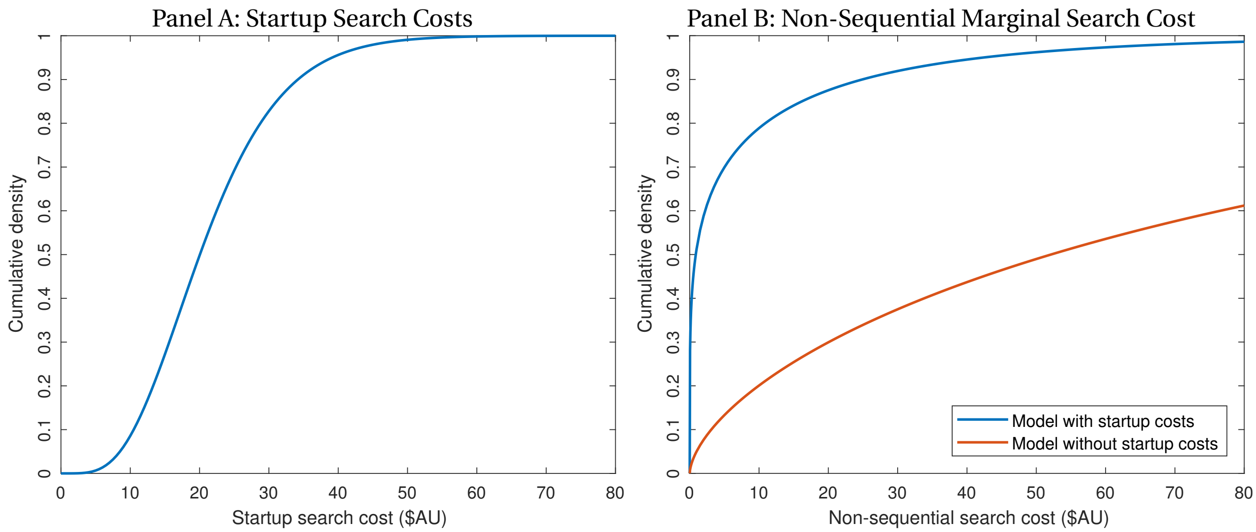
5 Summary and discussion

In studying search frictions, economists have, to date, assumed that search costs are independent of search history. Exploiting a natural experiment in retail gasoline, together with unique data on retail prices and search behavior, we have provided evidence on a new form of search costs that we call startup search costs. The novel evidence of history dependence in search behavior that we find suggests that the first search cost sunk by a consumer is drastically different from subsequent search costs. Search experience matters.

The results from our simple empirical search model highlight the implications of startup

[17]Note that, because startup and nonsequential search costs are independently drawn, one cannot simply add the 20th percentile of each distribution to obtain the 20th percentile aggregate search cost.

Figure 6: Search Cost Distributions



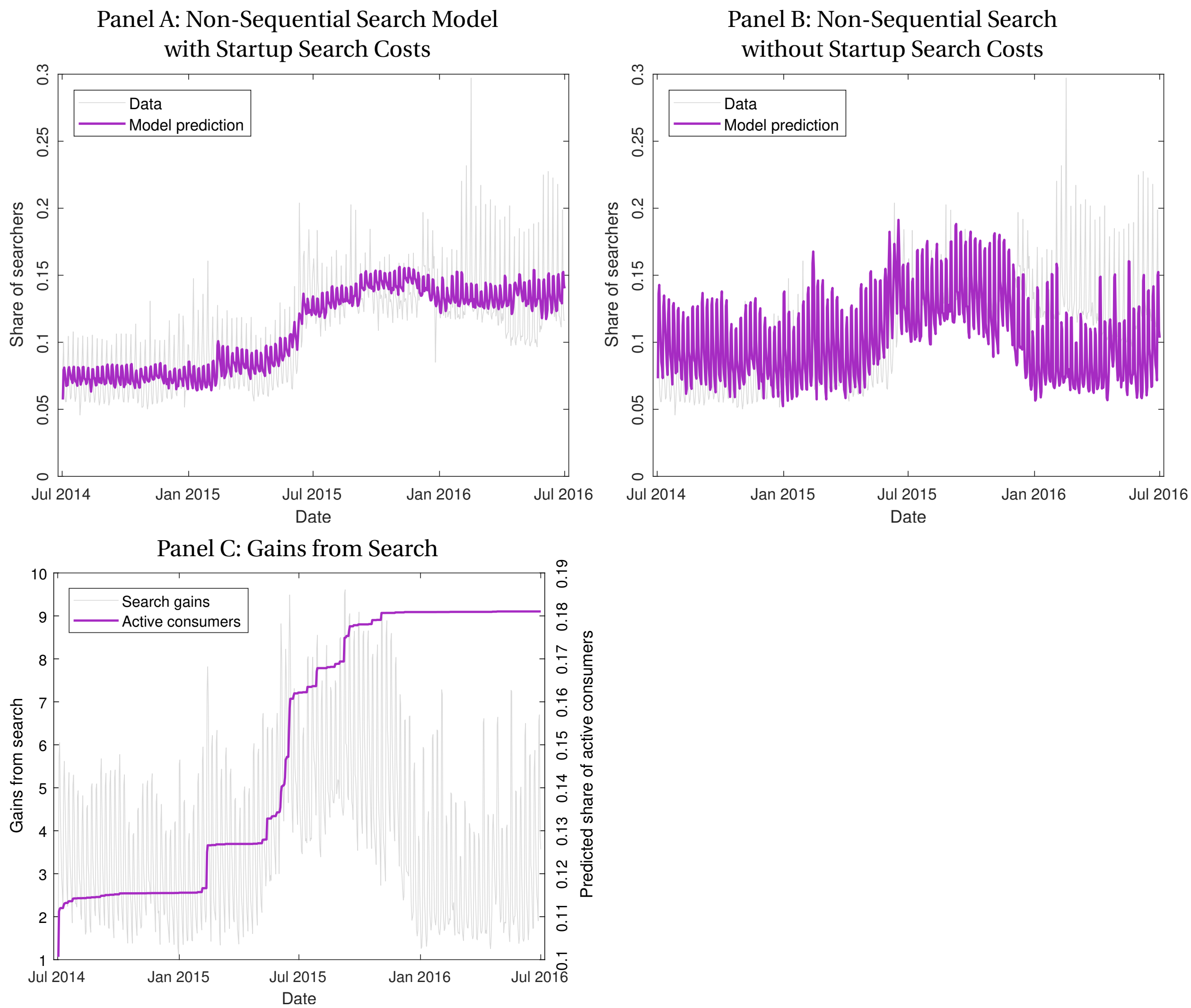
search costs for the measurement of search frictions. Model misspecification that ignores startup search costs yields overestimated marginal search costs. Moreover, we have shown how a standard non-sequential search model is unable to account for a permanent rise in market-level search from a temporary exogenous shock to search incentives.

Because we work with market-level and not individual-level search data, we are unable to identify deeper microfoundations for startup search costs.[18] We can think of four possible mechanisms to explore in future research. First, startup search costs could be driven by technology adoption costs (Foster and Rosenzweig, 2010) with online price clearinghouses. Second, startup search costs could reflect consumers holding biased beliefs about the value of search (Koulayev and Alexandrov, 2017), and updating these beliefs after trialing the clearinghouse during the price war. Third, consumers could rapidly form habits after trialing the clearinghouse (Becker and Murphy, 1988), with minimal rates of habit decay. Finally, the inertia in non-adoption of the clearinghouse for 15 years before the war, followed by a permanent shift in usage after, could reflect procrastination or time-inconsistency (O'Donoghue and Rabin, 1999) in learning to use the clearinghouse.

Understanding the role of startup search costs and their underlying mechanisms is important for policy. We believe that the evidence presented here points to a new and important policy challenge with online search platforms aimed at promoting price transparency and market efficiency: policymakers need to get consumers "over the hump" in starting to use such

[18]Our specification of startup search costs in the simple model we estimated is akin to the reduced-form nature of marginal search costs in standard sequential and non-sequential search models. With regard to the latter models, researchers have similarly begun to explore underlying mechanisms for marginal search costs including spatial frictions (Pennerstorfer et al. 2016; Startz 2017; Buchholz 2017), cognition (Crawford et al., 2013), or consumers' opportunity cost of time.

Figure 7: Search Model Predictions



platforms. We have found that this hump prevented consumers from engaging with a well-established price clearinghouse for 15 years. It took a three-week, temporary price shock to substantially increase online price search. The lesson for policy is that large, temporary shocks to search incentives can help consumers overcome startup search costs and lead to long-run adoption of search platforms. Policy interventions that encourage customers to experiment with such platforms are a potential remedy for overcoming startup search costs.

Within industrial organization, our study raises a separate question for future research: what is the impact of startup search costs for firms' pricing decision? In the retail gasoline market that we study, we obtain an interesting implication for firms' pricing decisions. In Byrne and de Roos (2017b), we find evidence consistent with tacit collusion in this market. In this context, by encouraging consumers to engage with search, the temporary price war may have led to an increase in demand elasticity, and therefore collusive outcomes may have become more difficult to sustain following the war. This suggests a new trade-off – price variation could lead to a sustained increase in demand elasticity – facing cartel members contemplating either an adjustment of pricing policies or defection from the cartel.

References

Australian Competition and Consumer Commission. Petrol prices and Australian consumers. Report of the ACCC Inquiry into the Price of Unleaded Petrol, 2007.

Australian Competition and Consumer Commission. Petrol prices and Australian consumers. Report of the ACCC Inquiry into the Price of Unleaded Petrol, 2014.

Australian Competition and Consumer Commission. Report on the australian petroleum market, september quarter 2017, 2017.

Michael R. Baye, John Morgan, and Patrick Scholten. Information, search, and price dispersion. 2006. Chapter 6 in Handbook in Economics and Information Systems Vol. 1 (T. Hendershott, Ed.), Amsterdam: Elsevier.

Gary S. Becker and Kevin M. Murphy. A theory of rational addiction. *Journal of Political Economy*, 96:675–700, 1988.

Nicholas Buchholz. Spatial equilibrium, search frictions and efficient regulation in the taxi industry. mimeo, Princeton University, 2017.

David P. Byrne and Nicolas de Roos. Consumer search in retail gasoline markets. *Journal of Industrial Economics*, 65:183–193, 2017a.

David P. Byrne and Nicolas de Roos. Learning to coordinate: A study in retail gasoline. mimeo, The University of Melbourne, 2017b.

Vincent P. Crawford, Miguel A. Costa-Comes, and Nagore Iriberri. Structural models of nonequilbrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51:5–62, 2013.

Babur de Los Santos, Ali Hortascu, and Matthijs R. Wildenbeest. Testing models of consumers search using data on web browsing and purchasing behavior. *American Economic Review*, 102:2955–2980, 2012.

Andrew Eckert. Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys*, 27:140–166, 2013.

Andrew D. Foster and Mark R. Rosenzweig. Microeconomics of technology adoption. *Annual Review of Economics*, 2:395–424, 2010.

Han Hong and Matthew Shum. Using price distributions to estimate search costs. *RAND Journal of Economics*, 37, 2006.

Sergei Koulayev. Search for differentiated products: Identification and estimation. *Rand Journal of Economics*, 45:553–575, 2014.

Sergei Koulayev and Alexei Alexandrov. No shopping in the u.s. mortgage market: Direct and strategic effects of providing information. CFPB Working Paper 2017-01, 2017.

Laurence Levin, Matthew S. Lewis, and Frank A. Wolak. High frequency evidence on the demand for gasoline. *American Economic Journal: Economic Policy*, 27, 2016. forthcoming.

Ted O'Donoghue and Matthew Rabin. Doing it now or later. *American Economic Review*, 89: 104–124, 1999.

Dieter Pennerstorfer, Philipp Schmidt-Dengler, Nicolas Schutz, Christoph Weiss, and Biliana Yontcheva. Information and price dispersion: Theory and evidence. 2016.

Meredith Startz. The value of face-to-face: Search and contracting problems in nigerian trade. mimeo, Yale University, 2017.

George J. Stigler. The economics of information. *Journal of Political Economy*, 69:213–225, 1961.

Matthijs R. Wildenbeest. An empirical model of search with vertically differentiated products. *Rand Journal of Economics*, 42:729–757, 2011.

A Specification of search gains: robustness

Recall that the gains to search for consumer i at time t are defined by

$$g_{it} = (\overline{p}_{it} - \min\{\mathbf{p_{it}}\}) \times \tau_i.$$

In the body of the paper, we presumed that average and minimum prices for consumer i were taken with respect stations within a 5km radius of the centroid of her local district. In this section, we also consider search radii of 2km and 10km.

As the search radius varies between 2km and 10km, there is a noticeable impact on the consumer's choice set. With a search radius of 2km, 5km, and 10km, there are on average 1.96 stations (s.d. 1.57), 10.29 stations (s.d. 6.32), and 34.26 stations (s.d. 18.86) within the choice radius for each district. For estimation purposes, we eliminate all districts with less than two stations inside the search radius.

For comparability, we retain the same format for the presentation of results. Table 3 contains estimation results. Columns (1) and (2) ((3) and (4)) contain estimates based on a 2km (10km) search radius. For each specification of the search radius, the left and right columns contain, respectively, esimates based on the full model and a constrained model that does not include startup search costs. Figures 8 and 9 depict, for the 2km search radius specification, the cumulative distribution of estimated startup and marginal search costs, and predictions for the model, respectively. Figures 10 and 11 contain analogous information for the 10km search radius specification.

The main qualitative features we highlighted earlier carry over to alternative specifications of the search radius. In particular, the model with startup costs leads to qualitatively different inferences over marginal search costs, the fit of the model is much improved by incorporating startup search costs, and only the model with startup search costs is able to explain the permanent increase in search activity following the temporary shock to search gains.

Adjusting the search radius does lead to some variation in results. When the search radius is reduced, measured search gains tend to be lower. This can be seen by comparing Panel C of Figures 7, 9, and 11. As a result, to rationalize observed search activity, estimated startup and marginal search costs are lower when the search radius is reduced. This is best seen by comparing Figures 6, 8, and 10. Consider first the model with startup search costs. Estimated 20th percentile startup costs are \$9.54, \$13.30, and \$15.51 when the search radius is 2km, 5km, and 10km, respectively. At the left tail of the distribution, estimated marginal search costs are negligible in each specification of the search radius. The change to our search gain specification makes a bigger difference to estimates for the model without startup search costs. The 20th percentile consumer now has estimated marginal search costs of \$3.98, \$9.94, and \$22.06 when

Table 3: Estimation Results with Alternative Search Radii

	2km Search Radius		10km Search Radius	
	With Startup Costs	Without Startup Costs	With Startup Costs	Without Startup Costs
	(1)	(2)	(3)	(4)
Marginal search cost distribution				
μ_c	0.205	0.455	0.215	0.421
	(0.017)	(0.005)	(0.017)	(0.003)
σ_c	16.907	176.596	36.781	1322.721
	(2.341)	(6.759)	(3.829)	(17.370)
Startup search cost distribution				
μ_f	1.945		6.201	
	(0.081)		(0.144)	
σ_f	12.110		3.811	
	(0.613)		(0.135)	
Objective function, $G(\hat{\theta})$	0.374	0.816	0.375	0.934

Notes: Robust standard errors are in parentheses (). The number of observations is $T = 731$ dates. All calculations assume consumers purchase 50 liters of gasoline.

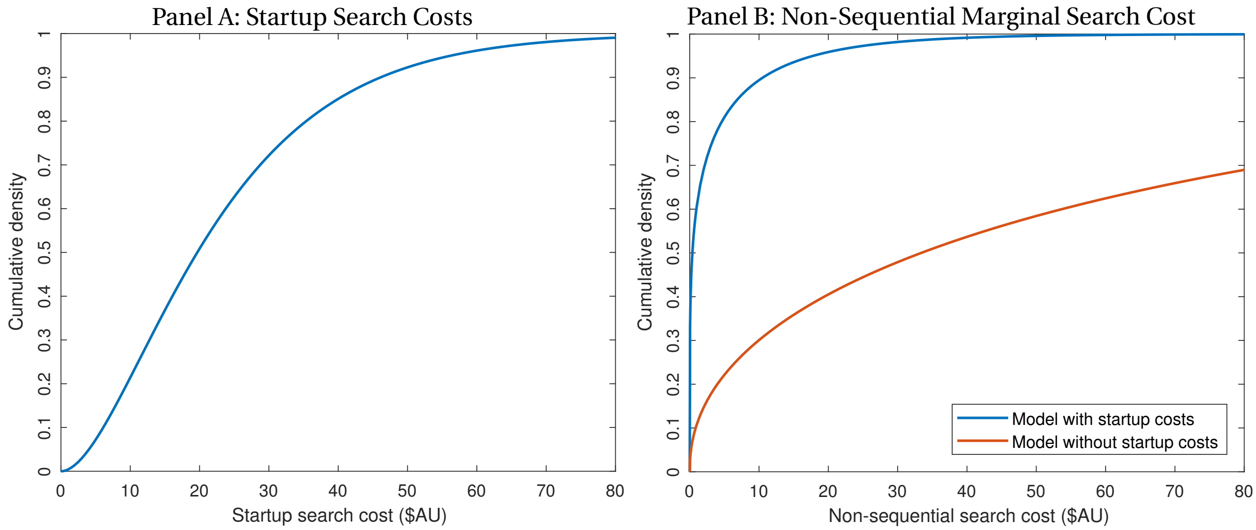
the search radius is 2km, 5km, and 10km, respectively.

Comparing Figure 9 and 11 reveals some subtle differences in model predictions. When search gains are defined more locally, this accentuates the volatility in search gains, and there is an associated increase in the high-frequency volatility in predicted search, both for the full model (Panel A) and the restricted model without startup search costs (Panel B). Finally, Panel C also suggests that, when search gains are defined more locally, to rationalize the volume of search activity, a greater proportion of consumers are predicted to incur their startup search costs.

B Consumers with a dynamic perspective

In the model of Section 4, we presumed consumers take a static perspective when deciding whether to incur startup search costs. In this section, we illustrate the implications of relaxing this assumption. Consider the perspective of consumer i who evaluates the impact of today's search decision on the search environment that she will face in the future. To fix ideas, we begin by laying out the Bellman equation faced by consumer i at time t when she adopts this dynamic perspective.

Figure 8: Search Cost Distributions, 2km Search Radius



Given current search state w_{it}, search cost parameters c_i and f_i, and price vector $\mathbf{p}_t$, her current valuation is given by

$$V(w_{it}) = \max_{\chi_{it}\in\{0,1\}} \chi_{it}\left(\overline{u} - \tau_i \min\{\mathbf{p}_{it}\} - s_{it}\right) + (1-\chi_{it})\left(\overline{u} - \tau_i \overline{p}_{it}\right) + \delta \mathbb{E}_t V(w_{it+1}),$$

$$w_{it+1} = w_{it} + (1-w_{it})\chi_{it},$$

where $\chi_{it} = 1$ indicates a decision to search today and $\chi_{it} = 0$ indicates no search; and $\mathbb{E}_t$ indicates period-t expectations over future price distributions.[19] The parameter δ describes the rate at which consumers discount the next fuel purchase, and could reflect impatience, and concerns about the decay or obsolecence of current knowledge of the search process.

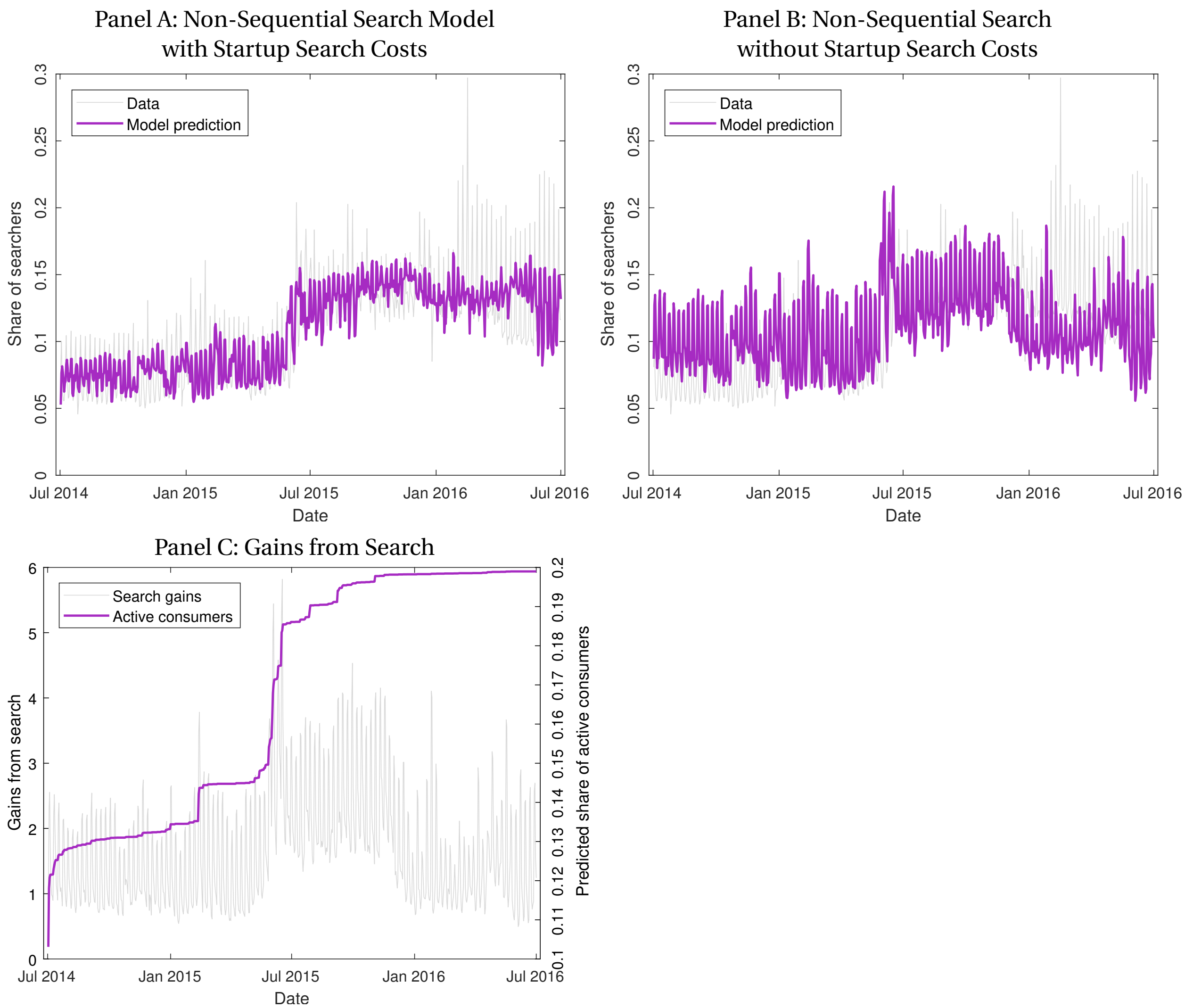
Expectations of future prices play an important role through their influence on the continuation value of the consumer's dynamic problem. For illustration, we consider a simple expectations process. We say that consumer i adopts stationary expectations if she anticipates the current price distribution to be observed in subsequent periods: $\mathbf{p}_{t+k} = \mathbf{p}_t$, for $k > 0$. This leads to the following proposition.

Proposition 1. *Suppose consumer j adopts a static perspective with search costs c_j and f_j, and consumer i adopts a dynamic perspective with stationary expectations and search costs c_i and f_i. Then consumers i and j are observationally equivalent if $c_i = c_j$ and $f_i = f_j/(1-\delta)$.*

Proof. First, consider consumer j. Based on her static perspective, she searches iff $g_{jt} > s_{jt}$. If $w_{jt} = 1$, she searches iff $g_{jt} > c_j$; if $w_{jt} = 0$, she searches iff $g_{jt} > c_j + f_j$.

[19] For simplicity, we ignore the intertemporal searh opportunities presented by the Fuelwatch program in this formulation, and consider search gains in period t based solely on the period t price distribution.

Figure 9: Search Model Predictions, 2km Search Radius

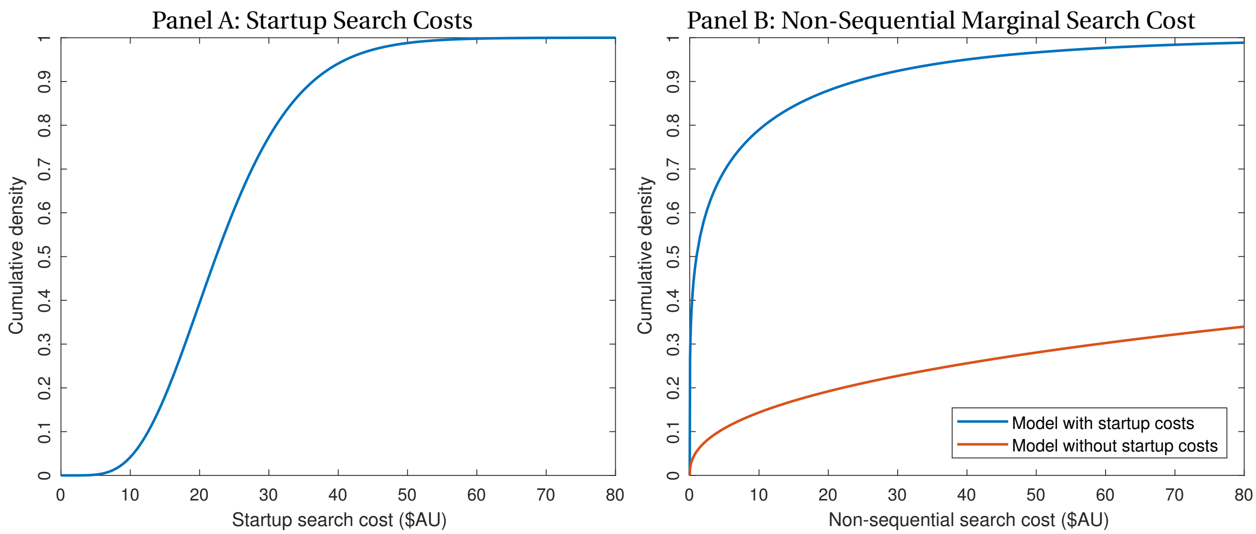


Next, consider consumer i and suppose $w_{it} = 1$. In this case, she has already sunk her startup search costs and, as a result, her current search decision has no dynamic consequences. Thus, she chooses to search iff $g_{it} > c_i$. Because $w_{it} = 1$ is an absorbing state, we can solve for the value $V(1) = U(c_i, \mathbf{p}_{it})/(1-\delta)$, where $U(.)$ is as defined in Section 4.

Suppose instead $w_{it} = 0$ and observe that consumer i has value

$$V(0) = \max_{\chi_{it}\in\{0,1\}} \chi_{it}\left(\overline{u} - \tau_i \min\{\mathbf{p}_{it}\} - c_i - f_i\right) + (1-\chi_{it})\left(\overline{u} - \tau_i \overline{p}_{it}\right)$$
$$+\delta\left(\chi_{it} V(1) + (1-\chi_{it}) V(0)\right).$$

Figure 10: Search Cost Distributions, 10km Search Radius



Consumer i searches in period t iff

$$\overline{u} - \tau_i \min\{\mathbf{p}_{it}\} - c_i - f_i + \delta \frac{U(c_i, \mathbf{p}_{it})}{1-\delta} > \overline{u} - \tau_i \overline{p}_{it} + \delta V(0).$$

Observing that consumer i makes the same decision whenever $w_{it} = 0$, we can deduce that she searches iff

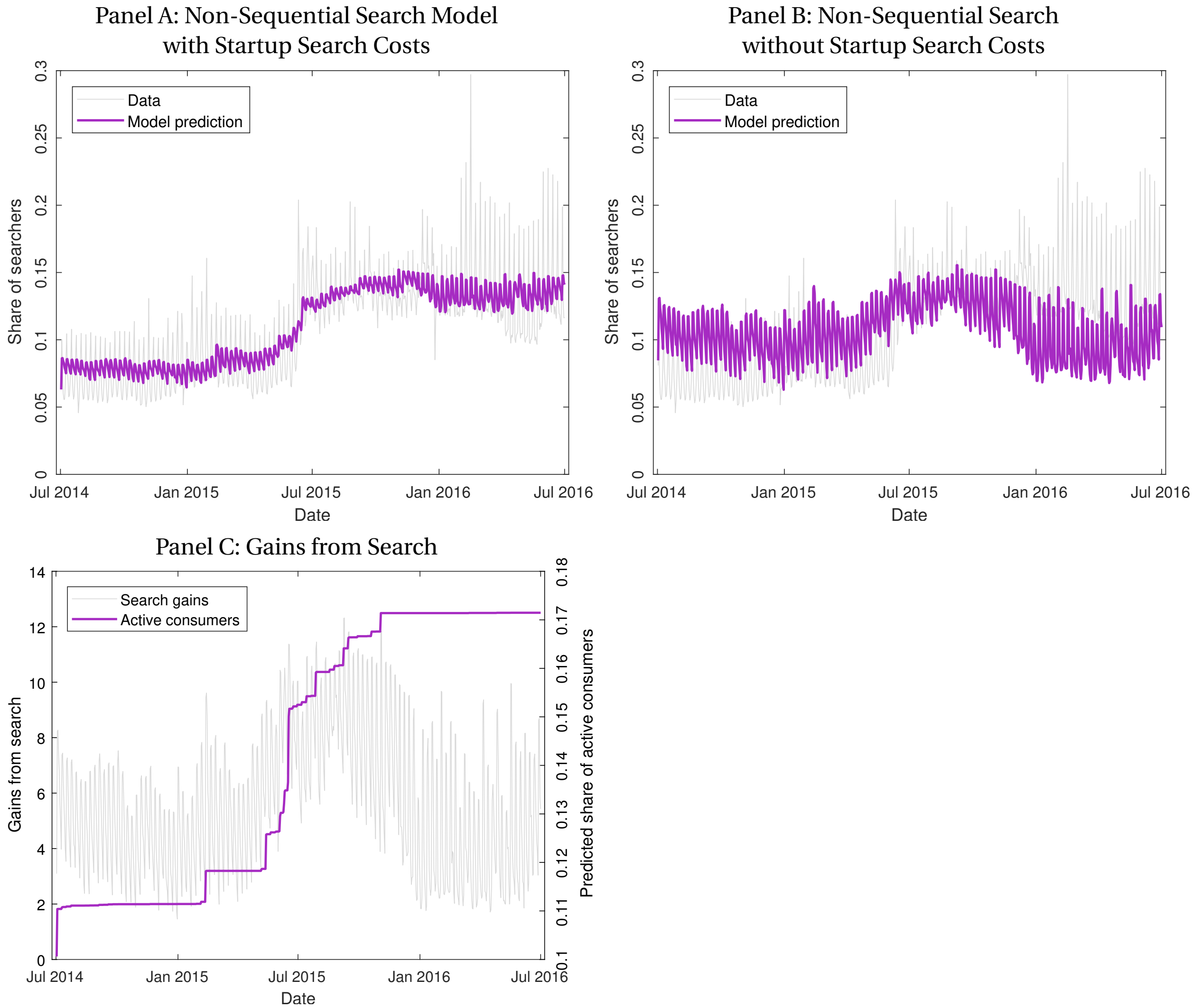
$$(1-\delta)\left(\overline{u} - \tau_i \min\{\mathbf{p}_{it}\} - c_i - f_i\right) + \delta U(c_i, \mathbf{p}_{it}) > \overline{u} - \tau_i \mathbf{p}_{it}. \tag{1}$$

Next, we show that, when $w_{it} = 0$, consumer i searches iff $g_{it} > c_i + (1-\delta) f_i$. We break this into two steps. First, observe that if $g_{it} > c_i$, then a consumer who had already sunk her startup search costs would choose to search. This means that $U(c_i, \mathbf{p}_{it}) = \overline{u} - \tau_i \min\{\mathbf{p}_{it}\} - c_i$. Substituting into (1) leads to the conclusion that i searches iff $g_{it} > c_i + (1-\delta) f_i$. Next, suppose that $g_{it} \leq c_i$. In this case, $U(c_i, \mathbf{p}_t) = \overline{u} - \tau_i \overline{p}_t$. Suppose further that $\chi_{it} = 1$. Substituting into (1) leads to the condition $g_{it} > c_i + f_i$, a contradiction. Therefore $\chi_{it} = 0$ whenever $g_{it} \leq c_i$. Combining the two cases, we have our desired result that consumer i searches iff $g_{it} > c_i + (1-\delta) f_i$.

Finally, comparing consumers i and j leads to the conclusion that their choices are identical if $c_i = c_j$ and $(1-\delta) f_i = f_j$, as required. □

Proposition 1 gives a feeling for the impact of the consumer's perspective on inferences about search costs under the assumption of stationary expectations. The perspective adopted by consumer i has no impact on the inferences we make about her marginal search costs. However, particularly for patient consumers, inferred startup search costs will be substantially

Figure 11: Search Model Predictions, 10km Search Radius



higher if we presume consumers adopt a dynamic perspective.

The logic of the proof of Proposition 1 provides an indication of the impact of the assumption of stationary expectations. Suppose that in period t, consumer i decides to first engage in search. Under the stationarity assumption, she anticipates that she would also have chosen to initiate search in period $t+1$ had she not chosen to search in period t. Thus, she derives a benefit of f_i in every subsequent period. Similarly, if instead she anticipates that price variation and the gains to search will increase over time, then she will also anticipate engaging in search in each period, and the value to her of initiating search will be the same. Alternatively, if she expects the gains to search to fall, she may anticipate that there are future periods in which she would not be willing to initiate search. In this case, by assuming stationarity, startup search costs will be overestimated.