Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market*

Stephanie Assad[a], Robert Clark[b], Daniel Ershov[c], Lei Xu[d]

December 2020

Abstract

We provide the first empirical analysis of the relationship between algorithmic pricing (AP) and competition by studying the impact of adoption in Germany's retail gasoline market, where software became widely available in 2017. Because adoption dates are unknown, we identify adopting stations by testing for structural breaks in AP markers, finding most breaks to be around the time of widespread AP introduction. Because station adoption is endogenous, we instrument using headquarter adoption. Adoption increases margins, but only for non-monopoly stations. In duopoly markets, margins increase only if both stations adopt, suggesting that AP has a significant effect on competition.

*The views in this paper do not reflect those of the Bank of Canada. Helpful comments were provided by John Asker, David Byrne, Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolò, JP Dubé, Avi Goldfarb, Joe Harrington, JF Houde, Alex MacKay, Jeanine Miklós-Thal, Catherine Tucker, and seminar participants at the University of Bologna, TSE, Paris TelecomTech, Tel Aviv Coller, CESifo, FTC, University of Amsterdam conference on algorithmic collusion, Netherlands Authority for Consumers and Markets, Bank of Canada - University of Toronto Joint Conference on Collusion, the 2020 NBER Summer Institute IT and Digitization, and the 2020 NBER Economics of AI. Daniel Ershov would like to acknowledge support received from ANR under grant ANR-17-EUR-0010 (Investissements d'Avenir program) and from ANITI. Correspondence to: [a]Stephanie Assad - Queen's University, Dunning Hall, 94 University Avenue, Kingston, ON; Email: assads@econ.queensu.ca, [b]Robert Clark - Queen's University, Dunning Hall, 94 University Avenue, Kingston, ON; Email: clarkr@econ.queensu.ca, [c]Daniel Ershov - Toulouse School of Economics, Université Toulouse 1 Capitole, 1 Esplanade de l'Université, 31080 Toulouse, France, Email: daniel.ershov@tse-fr.eu., [d]Lei Xu - Bank of Canada, 234 Wellington St Ottawa, ON; Email: lei.xu2@gmail.com

1 Introduction

Pricing-algorithm technology has become increasingly sophisticated in recent years. Although firms have made use of pricing software for decades, technological advancements have created a shift from mechanically-set prices to AI-powered algorithms that can handle large quantities of data and interact, learn, and make decisions with unprecedented speed and sophistication. The evolution of algorithmic-pricing software has raised concerns regarding possible impact on firm behaviour and competition. The potential for algorithms to facilitate collusion, either tacit or explicit, has been a popular discussion-point among antitrust authorities, economic organizations, and competition-law experts in recent years (OECD 2017; Competition Bureau 2018; Autorité de la Concurrence and Bundeskartellamt 2019; UK Digital Competition Expert Panel 2019; Ezrachi and Stucke 2015, 2016, 2017; Varian 2018; Goldfarb et al 2019). Since the goal of algorithms is to converge to an optimal policy, AI agents could learn to play a collusive strategy to achieve a joint-profit maximizing outcome. Algorithmic pricing software can also facilitate collusion through increased ease of monitoring and speed of detection, and through punishment of possible deviations.

The literature on algorithmic collusion is expanding, with contributions from the fields of economics, law, and computer science. At present, there is no theoretical consensus as to whether algorithms facilitate tacit collusion (Kühn and Tadelis 2018; Calvano et al 2020; Miklós-Thal and Tucker 2019; Brown and MacKay 2020). Despite some evidence that collusive algorithmic behaviour can appear in synthetic environments, there are questions about whether it can and will arise in practice. As of yet, there is no empirical evidence linking the adoption and use of pricing algorithms to market outcomes related to competition. The objective of this paper is to supplement existing theoretical literature by conducting the first empirical analysis of the impact of wide-scale adoption of algorithmic pricing software. We focus on the German retail gasoline market, where, according to trade publications, algorithmic pricing software became widely available beginning in 2017, and for which we have access to a high-frequency database of prices and characteristics for every retail gas station in the country from January 2016 to December 2018.[1]

Investigating the impact of the adoption of algorithmic-pricing software on competition requires overcoming three important challenges. First, even with access to detailed pricing data, adoption decisions are typically not publicly observed. Second, adoption is endogenous, since the decision to

[1]**Legal disclaimer:** This paper analyses the impact of adoption of algorithmic pricing on competition strictly from an economic point of view. We base our understanding of the facts on publicly-available data on prices from the German Market Transparency Unit for Fuels. To our knowledge, there is no direct evidence of anticompetitive behavior on the part of any algorithmic-software firms or gasoline brands mentioned in this paper.

adopt is correlated with factors that are unobserved to the researcher. Finally, even if adoption can be causally linked with higher prices or margins, it is not clear whether these can be attributed to changes in competition intensity rather than to other factors, such as an improved ability to detect underlying fluctuations in wholesale prices or predict demand.

To overcome the first challenge we test for structural breaks in pricing behaviours that are thought to be related to the use of sophisticated pricing software: (i) the number of price changes made in a day, (ii) the average size of price changes, and (iii) the response time of a station's price update given a rival's price change. We focus on these measures since they capture the promised impacts of algorithmic software in the retail gasoline market. Leading algorithmic pricing software providers explain that their software performs high frequency analysis to "rapidly, continuously and intelligently" react to market conditions. We use a Quandt-Likelihood Ratio (QLR) test (Quandt 1960) to look for the best-candidate break date. For each of the three pricing behaviour measures, we test for structural breaks at each station for each week in a large window around the time of supposed adoption. For each measure, the best-candidate structural break for a given station is the week with the highest F-statistic. Breaking in one of the three measures could occur for any number of reasons, but breaking in multiple markers in close proximity should provide a strong indication of adoption. Therefore, we classify a station as an algorithmic-pricing adopter if it experiences a best-candidate structural break in at least two out of three pricing behaviours within a short time period, which we take to be four weeks, but is robust to alternative specifications. We find that approximately 30% of stations in our data set experience best-candidate breaks in multiple pricing behaviour measures within a four week window. The majority of these breaks occur in mid-2017, just as algorithmic pricing software supposedly became widely available in Germany.

After having identified adopters, we examine the impact of their adoption on retail prices and margins. For retail gasoline, margins are a clear indictor of profitability and market power: the ability of stations to mark-up retail prices over wholesale prices. Previous studies on coordination and collusion in this market use margins to evaluate competition (Clark and Houde 2013, 2014; Byrne and De Roos 2019), and theory papers on algorithmic competition also make clear predictions related to margins (Calvano et al 2020; Brown and MacKay 2020). Although we control for time and station-specific effects, as well as time-varying market level demographics, individual station adoption decisions may be correlated with station/time specific unobservables (managerial skills, changing local market conditions, etc). OLS estimates are likely attenuated. We address this challenge by instrumenting for a station's adoption decision. Our main IV is the adoption decision by the station's *brand* (i.e., by brand headquarters). As demonstrated by previous technology-adoption episodes in the gasoline retail market, brands can facilitate adoption by their stations. "Adopting" brands provide support/subsidies/training to individual stations, reducing adoption costs.[2] Brand-level decisions should not be correlated with individual station-specific unobservables, making this instrument valid. Since brand adoption decisions are also unobserved, we use a proxy as our instrument: the fraction of a brand's stations that adopt AI pricing. The idea being that if a large fraction of a brand's stations adopts AI, it is likely that the brand itself adopted and facilitated adoption by the stations. As a robustness check, we also use an alternative set of instruments: annual measures of broadband internet availability and quality in the local area around each station. Most algorithmic pricing software are "cloud" based and require constant downloading and uploading of information. Without high-speed internet, adoption will not be particularly useful. Conditional on local demographic characteristics broadband quality should not depend on station-specific unobservables, but stations should be more likely to adopt algorithmic pricing software once their local area has access to reliable high-speed internet.

Using brand-adoption as an IV we find that following adoption mean station-level prices increase by approximately 0.6 cents per litre. Margins increase by 0.8 cents per litre, or roughly 9%.[3] These findings provide evidence of the causal impact of adoption of algorithmic pricing software on prices and margins. However, it is not clear whether these higher margins can be attributed to changes in the degree of competition intensity rather than to other factors, such as an improved ability to detect underlying fluctuations in wholesale prices or better predict demand. To isolate the effects of adoption on competition we focus on the role of market structure, comparing adoption effects in monopoly (one station) markets and non-monopoly markets. If adoption does not influence competition, effects should be similar for monopolists and non-monopolists. We also perform a more direct test of theoretical predictions by focusing on duopoly (two station) markets. We compare market-level margins in markets where no stations adopted, markets where one station adopted and markets where both stations adopted. In the first market type, competition is between human price setters. In the second it is between a human price setter and an algorithm, while in the last it is between two algorithms. By comparing all three market types we are able to identify the effect of algorithmic pricing on competition.

We observe heterogeneity in outcomes based on market structure suggesting that algorithmic pricing software affects margins and prices by changing competition. Adopting stations *with no*

[2]Below we provide examples of other episodes of technology adoption in retail gasoline markets.

[3]Estimates using alternative broadband availability IVs are qualitatively similar to the main estimates, although somewhat larger. See Appendix E.4 for additional discussion of these results.

competitors in their ZIP code (i.e. monopolists) see no statistically significant change in their mean margins or prices. In contrast, adopting stations *with* competitors in their ZIP code see a statistically significant mean margin increase of 0.9 cents per litre and the distribution of their margins and prices generally shifts right. We also find that the 95th percentile of prices falls for monopolist adopters, suggesting that algorithms may have some benefits for consumers by reducing the highest prices charged. Our market-level results suggest that relative to markets where neither station adopts, markets where both do see a mean margin increase of 3.2 cents per litre, or roughly 38%. Mean prices increase by 4 cents per litre. Markets where only one of the two stations adopts see no change in mean margins or prices. These results show that market-wide algorithmic-pricing adoption raises margins, suggesting that algorithms soften competition. The magnitudes of margin increases are consistent with previous estimates of the effects of coordination in the retail gasoline market (Clark and Houde 2013, 2014; Byrne and De Roos 2019).

To provide further evidence of the impact on competition and to better understand the mechanism we examine whether algorithms are unable to learn how to compete effectively, or whether they actively learn how not to compete (i.e., how to tacitly collude) by testing the *timing* of price and margin changes. If there is no learning, margin and price changes should occur immediately after market-wide adoption. If algorithms are learning not to compete, it should take longer for them to train and converge to tacitly-collusive strategies (Calvano et al 2020). We find evidence that margins do not start to increase until about a year after market-wide adoption, suggesting that algorithms in this market learn tacitly-collusive strategies. These findings are in line with simulation results in Calvano et al (2020). We also examine the pricing behaviour that emerges in markets where both duopolists are algorithmic adopters. We show that in a market where both duopolists adopt, a station is more likely to respond to a rival's price decrease with an immediate price decrease of their own. There is no comparable change in the propensity to respond to price increases by a rival. We also find that when both stations adopt algorithmic pricing, the duopolist setting higher prices is less likely to undercut the duopolist setting lower prices. The timing of these effects is consistent with the timing of the price and margin increases. Altogether, these findings provide further evidence that adoption affects competition and they suggest that the algorithms learn that undercutting will not be profitable, since the undercutter will always be followed to the lower price by its rival.

Our results have important policy implications. Antitrust authorities around the world are considering adjustments to their toolkits to *address the challenges of the digital economy* (Autorité de la Concurrence and Bundeskartellamt 2019; UK Digital Competition Expert Panel 2019). Currently, competition authorities expend substantial resources pursuing hard-core cartels on an individual basis. In so doing they may overlook what may be a much broader set of collusion-facilitating devices that do not even require a conspiracy. Algorithmic pricing may be one such strategy. Communication via earnings calls is another (see Aryal et al 2020). We provide some policy recommendations in Section 9.

The remainder of this paper is laid out as follows. The next section discusses relevant literature. Section 3 provides a background discussion and an overview of the relevant players in the German market. Sections 4 and 5, respectively, discuss the data and methodology we use in our analysis. Sections 6 and 7 discuss, respectively, our results regarding (i) identifying adoption and (ii) the impacts of AI adoption on station and market outcomes. We also conduct a number of robustness checks. In Section 8 we provide evidence to support the idea that outcome results are driven by algorithms learning to tacitly collude. We present a brief policy discussion and some conclusions in Section 9.

2 Related Literature

This paper is most closely related to the recent literature concerning the potential link between algorithmic pricing and collusion. Theoretical and experimental results remain ambiguous. Several papers have shown that when algorithmic-pricing competition is modelled in a repeated game framework collusive outcomes are inevitable under certain conditions (Salcedo 2015; Klein 2019; Calvano et al 2020); however, others argue that improved price response to demand fluctuations may provide increased incentives for firm deviation from a collusive price (Miklós-Thal and Tucker 2019; O'Connor and Wilson 2019). Klein (2019) and Calvano et al. (2020) use computational experiments to study the effect of Q-learning algorithms on strategic behaviour of competing firms. Both studies find that these repeated games will converge to collusive outcomes including supra-competitive pricing and profits, as well as punishment of competitor deviation. While Miklós-Thal and Tucker (2019) find that improved demand prediction may lead to the possibility of collusion in markets where it is previously unsustainable, in other markets it may create incentives for deviation that were absent with less prediction capabilities. O'Connor and Wilson (2019) come to similar conclusions. Brown and MacKay (2020) develop a model where firms compete in pricing algorithms (rather than prices) and show that prices may increase even without collusion. Overall, there is little certainty as to whether algorithmic competition will lead to collusive outcomes in reality. There is, as far as we are aware, no empirical research regarding this question in the economics literature.[4]

The question as to whether algorithm usage may result in coordinated behaviour is of widespread interest and has been studied in fields outside of economics such as law and computer science. There are several papers in the computer science literature studying coordination of algorithms in repeated games. A number of these papers, including Kaymak and Waltman (2006, 2008) and Moriyama (2007, 2008) have indicated that reinforcement learning algorithms can result in cooperative outcomes; however, these outcomes are not always the most likely and are dependent on various specifications of the algorithm. Legal scholars generally express more certainty that the use of algorithmic pricing can lead to collusive behaviour. Authors including Ezrachi and Stucke (2015, 2016, 2017) and Mehra (2015) have expressed concern over this issue and its implications for competition policy.

We also relate to an extensive literature on the retail gasoline market. There is a literature on collusion in gasoline markets. Earlier work includes Borenstein and Shepard (1996), as well as Slade (1987, 1992). More recently Wang (2008, 2009), Erutku and Hildebrand (2010), Clark and Houde (2013, 2014), and Byrne and de Roos (2019) have all studied anti-competitive behaviour in the retail gasoline industry. There have been a small number of papers looking specifically at the German retail gasoline market (Dewenter and Schwalbe 2016, Boehnke 2017, Cabral et al 2018, Montag and Winter 2019).

A related area of literature studies the impact of technological advancements on price discrimination. A consequence of the rapid expansion of Big data and AI driven market analysis by firms is that personalized pricing strategies may become increasingly feasible and sophisticated. As technology advances, it can be better used to learn more about consumer tastes as well as to more accurately price products as a function of these tastes. In particular, authors have noted that Big data may facilitate first-degree price discrimination, which has generally been seen as challenging to implement in many markets (Ezrachi and Stucke (2016)). It is possible that more accurate determination of optimal personalized pricing can increase firm revenues (Shiller and Waldfogel 2011; Shiller 2014). Kehoe, Larsen, and Pastorino (2018) find that firm profit, as well as consumer surplus, may increase or decrease under personalized pricing depending on consumers certainty regarding their product tastes. They also find that in every case, total welfare is higher under discriminatory pricing in

[4] Decarolis and Rovigatti (2019) find that common bidding intermediaries in online advertising markets lead to anti-competitive effects, reducing prices for bidders at the expense of the platform. Bidding in this market is done through algorithms, which leads to parallels with the algorithmic pricing literature and regulatory concerns about multiple competitors in a market adopting the same pricing algorithm. Their findings suggest that algorithms could serve as “hubs” in a hub-and-spoke cartel (Harrington 2018b). Unlike this paper, the primary focus of Decarolis and Rovigatti (2019) is on increasing intermediary concentration rather than on algorithmic pricing software behaviour and the mechanism through which bidding decisions are made.

comparison to uniform pricing. Dubé and Misra (2018) show through experiments that personalized pricing improves firm profits and that a majority of consumers benefit.

3 Background

3.1 The German Retail Gasoline Market

Similar to other retail gasoline markets around the world, distinct retail brands play an important role in Germany. Most stations in the market are affiliated with brands.[5] ARAL and Shell are the dominant brands, together making up over 25 percent of stations in Germany.[6] There are a number of other large brands with over 350 stations each: Esso, Total, Avia, Jet, Star, BFT, Agip, Raiffeisen, and Hem. In terms of market shares, ARAL, Shell, Jet, BFT, Total and Esso together account for 84 percent of fuel sales in the German retail gas market.[7]

There are two notable features of competition in the German gasoline market that relate to our analysis: the presence of price transparency and a price-matching policy initiated by Shell in 2015. Price transparency was instituted in August 2013 in response to concerns about tacit collusion and high consumer prices by German regulatory authorities. As part of this initiative, stations that change their price must report their new prices "in real time" to the German Market Transparency Unit for Fuels (www.bundeskartellamt.de). Price changes are shared with consumer-facing information service providers. They are integrated into websites and mobile applications as well as into car GPS systems like TomTom.[8] The stated use of these data is to allow "motorists...to gain information on the current fuel prices and find the cheapest petrol station in their vicinity or along a specific route" and to "increase competition" (www.bundeskartellamt.de). There is conflicting evidence on the effects of this policy on prices and margins in Germany (Dewenter, Heimeshoff and Luth 2017, Montag and Winter 2019).[9]

The second major competition related event is Shell's 2015 price matching guarantee. Under this

[5]Our data set does not specify which stations are vertically integrated and directly owned by the brands and which are owned by independent franchisees who enter into a licensing agreement in exchange for the brand name and some technical support. Both are common in retail gasoline markets (Convenience.org).

[6]Detailed summary statistics of station numbers at the brand level are in Section 4.

[7]2019 fuel sales market shares for each brand are 21 percent for ARAL, 20 percent for Shell, 16 percent for BFT, 10.5 percent for Jet, 9.5 percent for Total, and 7 percent for Esso (bft.de).

[8]A full list of consumer facing data providers is here: `https://www.bundeskartellamt.de/EN/Economicsectors/MineralOil/MTU-Fuels/mtufuels_node.html`. We obtained our data from Tanker-Konig, one such provider.

[9]See Luco (2019) for an analysis of a similar transparency program in Chile.

policy, each Shell station had to match the lowest price of the 10 stations nearest to them within a 30 minute period. This policy did not apply to all consumers but only to those with Shell loyalty cards. Dewenter and Schwalbe (2016) and Cabral et al (2018) study this price matching guarantee and find that it very quickly increased average retail gasoline prices.

Our paper takes this competitive environment as a given. Our data set begins in January 2016, so we study the *additional* effects of algorithmic pricing software in a market with price transparency and with the Shell price matching policy. We perform several robustness checks to confirm that Shell stations (or stations directly competing with Shell stations) are not driving the main results. We find that excluding them from the analysis does not change our main findings (see Appendix E.1).

3.2 Use of Algorithmic Pricing Software in Retail Gasoline Markets

3.2.1 History of Algorithmic Pricing in Retail Gasoline

Fuel retailers are typically secretive about their pricing technology. Algorithmic pricing software providers are mostly privately-owned companies that are similarly secretive about their customer base. The structure of the"upstream" algorithmic pricing software market is unknown and there is no way to gauge the market share of any given software provider. A *Wall Street Journal* article on the subject mentions certain firms, including the Danish company *a2i Systems* and Belgian company *Kantify*, as notable providers (WSJ.com). A few other firms, not listed in the article but prominently featured on the internet as algorithmic software providers, include *Kalibrate* (Kalibrate.com), *Revionics* (Revionics.com) and *PDI* (PDIsoftware.com).

The use of algorithmic pricing software in European fuel retail markets began in the early 2010s. *a2i* sold their software to Danish fuel retail company OK Benzin in 2011 (a2i Systems). However, the main penetration of machine learning and artificial intelligence based pricing software appears to have happened in the mid 2010s, roughly coinciding with the publication of several newspaper articles about the subject in 2017 (WSJ.com,CSPDailyNews.com).[10] Kalibrate began explicitly distinguishing between rule based pricing and algorithmic pricing on its website in mid-2017 (2016 Kalibrate.com, 2017 Kalibrate.com). *a2i*'s software was tested in workshops with gas station owners in the Netherlands and Belgium in 2015 (servicestationmagazine.be) and adopted by a number of

[10]It is possible that providers sold algorithmic pricing software in Germany before 2016 (the start of our sample). We should not be observing any structural breaks for stations that adopted before the start of our sample. This means that we would be labelling some adopters as non-adopters. If adopters have higher average margins than non-adopters, this would bias our station-level estimates towards zero.

Shell stations in the Netherlands by 2017 (WSJ.com).

In Germany, the December 2017 issue of *Tankstop*, a trade publication for Germany's retail gasoline sector, notes that *a2i*'s software had been available to gas station operators within Germany since that summer (see Figure A1).[11] Kalibrate's website explicitly refers to German markets as benefitting from "agile" (i.e., algorithmic) pricing (Kalibrate.com). Kalibrate has had contracts with German brands Orlen and Tamoil/HEM (Kalibrate, businesswire.com).

3.2.2 How does algorithmic pricing software work?

Most software providers reveal few details about their algorithms. Promotional materials generally describe their pricing software as based on "machine learning" or "artificial intelligence," with references to "neural networks" and "deep learning" (Kalibrate.com, PDISoftware, a2i.com). They describe software that can help station owners "master market volatility with AI-powered precision pricing, and respond rapidly to market events and competitor changes" (Kalibrate.com) and take advantage of "superhuman expertise" (a2i.com). Additional promoted benefits include optimizing for long-term revenues and avoiding price wars (Kantify).

All providers stress the ability of their algorithms to incorporate market conditions and variables such as own and competitor prices, sales volumes, costs, and weather and traffic events into their decision-making. Figure A2, reprinted from *The Wall Street Journal*, presents a general summary of the functioning of gasoline pricing algorithms, but it is rather vague (as are other summaries on software providers' websites). *a2i Systems* provides more detail, outlining its algorithm in Derakhshan et al (2016).[12] It is described as a "multi-agent-system" based on the interaction of two agents: a consumer and a gas station. Agent behaviour is described by a "belief-desire-intentions" (BDI) model, a popular approach in computer science and information systems research. An agent's "beliefs," "desires" and "intentions" roughly correspond to information, payoffs and actions/strategy in decision-theory.[13]

a2i's algorithm works in three repeating steps. The first is "observation," where the gas station

[11] In conversations with us, *a2i* claims that, contrary to statements in these advertising materials, they were never active in the German market.

[12] This algorithm is based on the earlier papers Derakhshan et al (2006) and Hammer et al (2006). These papers look at interactions of children at a playground with the goal of encouraging more physical activity.

[13] Individual station owners can set different goals such as market share maintenance or constraints such as minimum price. They can also change the goals over time or adjust them. However, substantial changes by station owners does not happen much in practice. One algorithmic software provider states that approximately 80-90% of station owners *do not customize or interfere* with the default operations of the algorithm (Kalibrate.com).

agent collects data from the environment and forms its "beliefs." As mentioned previously, these data include own prices, sales, traffic and environmental factors. Competitor behaviour is not explicitly modelled but the competitor station prices are included as inputs in this step. In the second step, "learning," the gas station agent uses an Artificial Neural Network (ANN) to map inputs into outcomes.[14] The outcomes are not explicitly outlined in Derakhshan et al (2016), but they likely correspond to sales, revenues and/or profits.[15] These are the "desires"/payoffs in the BDI model. The last step is "adaptation," where the gas station agent sets prices to achieve their "desires"/maximize the objective function.

Many questions remain about how this algorithm or other algorithms of this type operate in practice. Derakhshan et al (2016) does not explicitly state whether the "desires" and "intentions" (or the objective function and strategies) in the model are static or dynamic and whether the algorithm only sets current or both current and future prices. It is also not clear how the algorithm learns. This is important since Milgrom and Roberts (1990) show that agents characterized by "passive learning" (based on past rival responses) and who optimize their static best response cannot reach collusive equilibria. This is not the case for "reinforcement learning" algorithms, such as Q-learning, that can experiment with temporarily sub-optimal strategies to maximize the overall net present value of future payoffs. Reinforcement learning has been the focus of existing simulations-based evidence of the possibility of algorithmic collusion (Calvano et al 2020, Klein 2019).

Derakhshan et al (2016) implies that station agents have dynamic objective functions and set both current and future prices. In an illustrative example they mention that their algorithm can "predict the volume through the day (24 h) at the start of the day." Elsewhere, objectives for algorithmic pricing software are described dynamically (i.e., "maintain market shares"). Derakhshan et al (2016) does not mention using a Q-learning algorithm or any algorithmic "exploration"/"experimentation." The broad description of the algorithm appears to be closer to the "passive learning" approach. However, it also cites reinforcement learning literature (e.g., Shoham et al 2003). More generally, the BDI model provides an attractive setting for reinforcement learning (Guerra-Hernandez et al 2005), and the combination of BDI and reinforcement learning has been an active field of research in computer science in the last 20 years (Albrecht and Stone 2018). We should also mention that there is no detailed information about algorithms available from other providers. They may very well be based on Q-learning or other reinforcement-learning mechanisms.

[14] This step also implicitly models consumer behaviour, but this is not described.

[15] In the earlier papers on children's playgrounds that form the basis of this algorithm, outcomes are categories that capture whether children are playing fast or slow, continuously or discontinuously, etc (Derakhshan et al 2006).

Even with adaptive learning, as long as algorithms set a sequence of prices rather than simply optimize the static best response, the introduction of algorithms in many gasoline retail markets could lead to increased cooperation between stations. This is because of price disclosure initiatives that have been introduced in many countries. In Germany, France, Spain, Chile, Argentina, and other countries, gas stations must report price changes within minutes of changing their prices at the pump.[16] Price information is then immediately and publicly displayed on price comparison websites. This policy creates a market with *perfect monitoring*. Pricing algorithms can process information and react faster than humans to changes in rival behaviour. Derakhshan et al (2016) presents an illustrative example where the algorithm of one station detects changes in the pricing of another station and responds rapidly. Algorithms, therefore, increase the speed of interaction. In a setting with perfect monitoring, increases in the speed of interaction facilitate cooperation, since it is easier to detect and punish deviations from tacitly-collusive equilibria (Abreu et al 1991).

3.3 Algorithmic Pricing Software Adoption

As in other cases of corporate technology adoption (e.g., Tucker 2008), technology adoption in gasoline retail happens at two levels: at the brand HQ (headquarters) level and at the individual station level. Brands make big-picture decisions about the technology they would like their stations to use. They provide stations with employee training, technical support and maintenance and subsidies. Individual station owners make adoption decisions specific to their stations. This involves incurring investment costs such as pump and Point of Sale (PoS) terminal upgrades. The costs can be substantial and are not necessarily fully subsidized by the brand.

An example is the adoption of electronic payment systems in the 1990s. Analogous to algorithmic pricing software, this is a technology that clearly benefits brands and that brands would want their stations to adopt, but that some stations may not want to adopt because of the costs involved. BusinessWeek reports that as part of a brand-wide roll-out of a contactless electronic payment system in 1997 by Exxon Mobil (Esso's US parent company), individual station owners "have to install new pumps costing up to $17,000–minus a $1,000 rebate from Mobil for each pump" (BusinessWeek). Partial investment subsidies by brands help explain staggered or delayed technology adoption in this market. We provide additional evidence for staggered technology adoption in the gasoline retail market in Appendix D. We look at the adoption of electronic payments from 1991 to 2001 by Canadian

[16]In Germany, stations must do it within 1 minute. In France and Chile, stations must report within 10 and 15 minutes of changing prices at the pump, respectively.

gasoline retail stations and document that it takes years after the first appearance of this technology for a substantial fraction of stations belonging to the five biggest brands in the market to adopt. Even after 10 years of availability, fewer than 50% of stations owned by leading brands adopted the technology (Figure B1).

There is no reason to suspect that algorithmic pricing software adoption is different. Anecdotal evidence suggests that gasoline brands have entered into long-term strategic partnerships with AI pricing and analytics providers, either directly or indirectly. For example, in Denmark *a2i* directly entered into a partnership with the large Danish retail fuel company OK Benzin (a2isystems.com). More indirectly, AI-pricing software providers enter into partnerships with IT companies that provide integrated services to brands. *Tankstop*'s December 2017 issue mentions that *a2i*'s services are supported by WEAT Electronic Data Service GmbH, a provider of cash-free payment systems and technical and logistical support for a number of petrol brands within Germany (WEAT.de). *a2i* also has a strategic partnership with Wincor Nixdorf, a retail technology company providing services such as Point of Sale (PoS) terminals and self-checkout solutions (DieboldNixdorf.com).

However, if a brand decides to "adopt," or enter into a partnership with an AI pricing software provider, its stations do not necessarily automatically and instantaneously adopt. There are many reasons why not every one of a brand's stations would adopt this technology. Cloud-based AI-pricing software potentially requires substantial infrastructure investments and not all station owners are in a position to incur these costs right when the technology becomes available, or possibly ever. For example, high-speed internet *and* high-speed internet enabled PoS terminals and pumps are likely required for the software to work. In Germany, many areas do not have access to stable high speed internet connections.[17] Equipment upgrades of this sort are expensive, costing thousands to tens of thousands of Euros (mobiletransaction.org).[18] Station operators also require training with the software to set its parameters and deal with potential errors.

4 Data

This section provides a general description of the datasets we use in our analysis. The Data Appendix includes more details about data construction. The main dataset comes from the German Market

[17] Reports suggest that many areas and regions in Germany receive sub-par services and speeds that are compared to the "old dial-up days" (NPR.org). We use broadband internet availability as an alternative instrument. See discussion in Appendix E.4.

[18] Again, this is analogous to previous cases of technology adoption and upgrading decisions by gas station owners, including allowing for chip cards or automated payment at the pump (Chicago Tribune).

Transparency Unit for Fuels. It contains price data for the most commonly used fuel types, Super E5, Super E10, and Diesel for every German gas station in 1 minute intervals. Our sample covers January 2016 to December 2018.[19] We focus on E5 fuel, which has over 80% market share in Germany (bdbe.de).[20] We calculate an average weekday (non-weekend or holiday) price from 7am to 9pm for each station.[21] We have additional information on each station, including their 5-digit ZIP code, latitude and longitude coordinates, station name, and associated brand. In total, there is information on 16,027 stations. We combine these retail price data with daily regional wholesale prices from Oil Market Report (OMR), a private independent German gasoline information provider. Regional wholesale prices are average daily ex-terminal prices in eight major German refinery and storage areas. We calculate the distance between each gas station and all refinery and storage areas and use wholesale prices from the nearest refinery.[22] Prior to subtracting wholesale price, we also subtract German VAT (19%) from retail prices. We compute station-level daily margins and take the monthly mean and 5th, 25th, 50th, 75th, and 95th percentiles for our station-month level analysis.

We merge in annual regional demographics from Eurostat. We include data on total population, population density, median age, employment (as a share of total population) and regional GDP. These data are at the "Nomenclature of Territorial Units for Statistics 3" (NUTS3) level, which is frequently used by EU surveys. A NUTS3 region is roughly equivalent to a US county and larger than a 5-digit ZIP code. We also incorporate daily weather information from the German Meteorological Service (DWD). These data are collected daily from thousands of local weather stations. We compute the average distance between each gas station and all local weather stations and use weather data from the nearest weather station. We include monthly means and standard deviations of temperature (in degrees Celsius) and precipitation (in mm).

Finally, to construct our second instrument, we collect data on local fixed-line broadband internet from the EU Commission's netBravo initiative (netBravo): whether the local area around the gas-station has widespread availability of 10 Mb/s, 15 Mb/s, 30 Mb/s internet in a given year,[23] and the

[19]Additional data exist for 2014 and 2015, but there are inconsistencies between 2014-2015 and 2016-2018 in station price reporting. Our sample period begins over two years after the start of the transparency initiative and one year after a Shell price matching policy (see Section 3.1 for more details). Results are robust to alternative samples (Appendix E.1).

[20]Super E10 is an ethanol based fuel with 10% ethanol and 90% unleaded petrol. Super E5 is an ethanol based fuel with 5% ethanol and 95% unleaded petrol. In Assad et al (2020), we find similar results using E10 fuel.

[21]Similar to previous studies of this market, in the absence of demand data we focus on the main salient period of time for consumers.

[22]This is a standard approach in the gasoline retail literature. We may be understating retail margins if stations belong to vertically integrated retailers.

[23]We define speed X to be widely available in an area if average speed-tests in that area in that year exceed that speed. As well, we assume that if an area has speed X widely available in a year, it also has the same speed widely

reliability of broadband signals in the area in that year. Reliability is computed by average signal strength (in dBm) and the variance of signal strength.[24]

4.1 Station-Level Descriptive Statistics

Table 1 shows summary statistics, including the number of stations per brand, the number of stations per ZIP code and the average distance between stations. Out of the 16,027 stations in our data set, single-operating stations account for approximately 11 percent. With our IV strategy, these stations are not part of our final estimating sample. The remainder of stations are affiliated with brands. The data set does not specify whether the stations are vertically integrated and directly owned by the brands, or whether they are owned by independent franchisees who have entered into a licensing agreement in exchange for the brand name and some technical support. Both are common in retail gasoline markets (Convenience.org). There are 258 distinct brands in the data, of which 239 have between 2 and 100 stations and 19 have more than 100. The top 5 brands account for 43 percent of stations and the 19 largest brands (those with more than 100 stations) account for 71 percent of total stations (11,752 stations total).

The market definition we use in the main text is a 5-digit ZIP code. In Europe, this is the most detailed ZIP code available. There are 5,781 5-digit ZIP codes in our data of which 2,094 have a single station (are monopoly markets), and 1,307 have two stations (are duopolies). The mean number of stations per ZIP code is around 3 and the median is 2. Only 81 ZIP codes have more than 10 stations.[25] The majority of stations are within 5km of their closest competitors (about 94 percent) and the average distance of a station to its closest competitor is 1.4km.

Table 1: Station Summary Statistics

Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**25%**	**75%**	**Max**
Stations per Brand	258	57.6	227.0	2	2	19	2417
Stations per ZIP Code	5,781	2.77	2.15	1	1	4	17
Distance to nearest station (KM)	16,027	1.40	1.77	0	0.30	1.69	17.19
# of other stations within 1KM	16,027	1.09	1.34	0	0	2	17

available in every subsequent year. More details on the construction of these variables are in the Data Appendix.

[24] Raw data are available at the monthly frequency. We choose to aggregate to the annual level since the number of speed-test/quality observations at the monthly level is small. It is also more likely that stations make decisions based on larger average trends rather than monthly fluctuations.

[25] ZIP codes reflect population patterns, so urban ZIP codes are much smaller in terms of area than rural ZIP codes.

4.2 Station/Month-Level Descriptive Statistics

Table 2 shows summary statistics at the station/month level, including prices, margins and regional demographics and weather. The average price that a station charges is 1.36 Euros per E5 litre, but the mean monthly margin that the average station earns over wholesale regional price (after subtracting VAT) is 8.3 cents per litre. The average station is located in a fairly dense NUTS3 region, with population density of 760 persons per square-km. The median age of the population around a station is 46 years and 53 percent of the population is employed. Over 85 percent of gas stations are located in areas with widely available 10Mb/s internet access, but less than 8 percent of stations are located in areas with widely available 30Mb/s internet access.

Table 2: Station/Month Summary statistics

Variable	Observations	Mean	Std. Dev.
Prices and Margins			
Mean Monthly E5 Price (EUR/litre)	448,221	1.362	.083
Mean Monthly E5 Margin (EUR/litre)	448,221	.083	.032
Regional Demographics and Weather Controls			
ln(Total Regional Population)	448,221	12.419	.816
Regional Population Density (pop/km^2)	448,221	758.238	1022.41
Regional Median Population Age	448,221	46.018	3.125
ln(Regional GDP)	448,221	9.083	.976
Regional Employment Share (employed/pop)	448,221	.527	.134
Mean Temperature (degrees Celsius)	448,221	10.417	6.87
Std. Dev. Temperature (degrees Celsius)	448,221	3.079	.806
Mean Precipitation (mm)	448,221	1.94	1.399
Std. Dev. Precipitation (mm)	448,221	3.603	2.605
Broadband Availability			
10 Mb/s Internet Available Dummy	330,977	.861	.346
15 Mb/s Internet Available Dummy	330,977	.488	.5
30 Mb/s Internet Available Dummy	330,977	.078	.268
Average Internet Signal Strength (dBm)	330,977	-83.497	3.964
Average Internet Signal Variance (dBm)	330,977	3.511	2.057

5 Methodology

Our empirical analysis of the effect of algorithmic pricing on competition in the German retail gas market proceeds as follows. First, we identify AI-adopting stations by performing Quant Likelihood Ratio (QLR) tests to establish structural breaks in the time series of pricing strategies of gas stations in our sample period. Second, we investigate the impact of AI adoption on market outcomes related to competition such as margins over wholesale prices. We compare outcomes for adopting and non-adopting stations. Selection bias coming from the differences between adopters and non-adopters, as well as endogeneity due to the timing of adoption, would attenuate OLS estimates from true effect of adoption. We use a rich set of controls for observables, station and time fixed effects and an instrument based on brand HQ-level adoption decisions to deal with endogeneity concerns. As a robustness check, we also use an alternative set of instruments related to the availability and quality of broadband internet in the local area around a gas station.[26] To test whether effects from the adoption of algorithmic pricing software on market outcomes are driven by competition we examine whether the effects differ across different market structures (monopoly stations vs. non-monopolists). We also perform a more direct test of theoretical predictions by focusing on duopoly markets (geographic markets with two stations). We compare market-level outcomes in markets where no stations adopted, markets where one station adopted and markets where both stations adopted. In the first market type, competition is between human price setters. In the second it is between a human price setter and an algorithm. In the last market type, competition is between two algorithms. This comparison allows us to capture the effect of algorithmic pricing on market competition.

5.1 Identifying Adoption

5.1.1 Identifying Station-Level Adoption

We do not have information on the algorithmic-pricing adoption decisions of individual stations or brands. Our approach is to take advantage of the high-frequency price data to identify these decisions. We focus on three measures of pricing behaviour (aggregated to a weekly level) to identify the adoption of algorithmic pricing at the station level: (i) number of price changes, (ii) average size

[26] Estimates with these IVs are qualitatively similar to our main estimates. We also test a "placebo IV" which uses the brand HQ-level adoption decision by a random brand (not the brand of the station) as an instrument and find null effects. See Appendix E.4 for additional discussion.

of price changes, and (iii) rival response time. We focus on these measures as a means to capture the promised impacts of *a2i*'s pricing software. *a2i*'s website states that their software "rapidly, continuously, and intelligently react[s]" to market conditions; automatically setting optimal prices in reaction to changes in demand or competitor behaviour or to maximize margins without eliciting a change in behaviour by consumers or competitors. We expect that after AI-adoption, stations may make more frequent updates of their prices, due to quicker and more precise detection of demand fluctuations or changes in competitor behaviour. Along these same lines, with faster detection of, and response to, competitor behaviour, we expect to see stations reacting more quickly to changes in competitors' prices. The effects of algorithms on the size of price changes are ambiguous. The algorithm may determine that consumers and competitors are very responsive to price change size and reduce the average size of price changes. Alternatively, the algorithm may determine that consumers and competitors are not responsive to price change sizes, allowing stations to increase their price changes without affecting consumer and competitor behaviour.

These measures of pricing behaviour line up with what is described in the economic and legal literature on algorithms. Ezrachi & Stucke (2015) point out the ability for algorithmic software to increase the capacity to monitor consumer activities and the speed of reaction to market fluctuations. Mehra (2015) points out the ability of AI pricing agents to more accurately detect changes in competitor behaviour and more quickly update prices accordingly. Brown and MacKay (2020) note that two significant features of pricing algorithms are their ability to (i) lower the cost of more frequent price updates and (ii) react quickly to price changes of other firms in the market. Our measure of rival response time follows a similar intuition to the approach taken by Chen et al. (2016) who identify algorithmic pricing users in Amazon Marketplace by measuring the correlation of user pricing with certain target prices, such as the lowest price of that given product in the Marketplace.[27]

In order to identify structural breaks in these measures we use Quandt-Likelihood Ratio (QLR) tests (Quandt 1960, Andrews 1993). This method tests for a structural break in a time-series measure for each period in some interval of time and takes the largest resulting test statistic. It is useful when an exact break date is unknown and has been suggested and used in previous work involving collusive behaviour (Harrington 2008; Clark and Houde 2014; Boswijk et al 2018; Crede 2019; Byrne and de Roos 2019). We conduct a QLR test for each station in our data set and for each variable of interest. Further details on these tests can be found in Appendix B.

[27]By looking at rival responses we may over-state adoption in the market if non-adopting firms are automatically labelled as adopters when they react to the more frequent price changes of adopters. This does not appear to be the case in our data. When examining duopoly markets, we find asymmetric adoption in many markets where one duopolist adopts and the other does not. More details are in Section 6.1.

5.1.2 Identifying Brand-Level Adoption

We use brand HQ-level adoption decisions as an instrument for station level adoption decisions. We do not observe an indicator for whether a brand HQ decided to enter into a strategic partnership with an AI-pricing software provider. However, we can use findings from the station-level classification to infer brand-level adoption. We use a probabilistic definition, computing the probability that a brand adopted by time t as the percentage of a brand's stations that have been classified as adopters by time t. This approach captures underlying brand-level decisions. As mentioned in Section 3.3, brand-level decisions should facilitate the adoption by individual stations. A brand for which a small percentage of stations adopted by time t is unlikely to be an adopter at time t, while a brand for which a large percentage of stations adopted is more likely to be an adopter. Alternative definitions could classify a brand as an adopter as soon as *any* one of its stations is classified as an adopter, or only after *all* of its stations are classified as adopters. These alternative approaches do not reflect technology adoption in this market.[28]

5.2 Evaluating the Impact of Adoption

5.2.1 Evaluating the Impact of Adoption on Station Outcomes

We want to capture the effects of station i's adoption of algorithmic pricing on the distribution of daily margins (above regional wholesale prices) and prices in month t. We evaluate outcomes at six points in each of the margin and price distributions: the mean daily margin or price in month t, as well as each of the 5th, 25th, 50th, 75th, and the 95th percentile daily margin or price in month t. Our OLS specification is as follows:

$$y_{it} = \alpha_i + \alpha_t + \beta D_{it} + \gamma X_{it} + \epsilon_{it}, \tag{1}$$

[28] Brand adoption is *not a necessary condition* for a station to adopt algorithmic pricing software. There are many providers of algorithmic pricing software that cater to small or medium enterprises (e.g., Prisync.com or Comptera.net). *a2i*'s 2017 advertisements target individual station owners and emphasize that all stations, regardless of their brand, can adopt their technology. Defining a brand as an adopter if *any* one of its stations is classified as an adopter would be sensitive to outliers and amplify noise from our station-level adoption measure. The opposite approach, defining a brand as an adopter only if *all* of its stations is inconsistent with the history of technology adoption in gasoline retail markets. As explained in Section 3.3, brand subsidies to stations for technology adoption are often incomplete and technology adoption is highly staggered. In Appendix D we show that it took years for a substantial share of gasoline stations belonging to top brands to adopt electronic payments in the 1990s. Figure 3 in Section 6.2 and Table B5 in Appendix B.6 show similar staggered patterns for brand-level algorithmic-pricing adoption in Germany.

where y_{it} is the outcome variable for station i in time t, α_i and α_t are, respectively, station and time fixed-effects, and D_{it} is a dummy variable equal to 1 if station i has adopted algorithmic pricing in time t and 0 otherwise. X_{it} are time-varying station specific controls. Most importantly, X_{it} includes the number of other gas stations that are in the same postal code as station i. The key coefficient in this regression is β, which captures the effect of AI adoption on y_{it}.

The OLS specification assumes that adoption is exogenous and as-good-as-random (conditional on observables). This is likely not the case. Algorithmic adoption could be correlated with unobservable time-varying station characteristics (ϵ_{it}). Stations with "high" unobservables (for example, better managed stations) could be more likely to adopt algorithmic pricing software and use it effectively.[29] Stations with different ϵs could also have very different market outcomes. This would attenuate the adoption effect towards zero. Stations could also choose to adopt in response to unobservable station-specific shocks - these would also affect both the adoption decision (D_{it}) and outcomes (y_{it}).

To address this issue we include station and time fixed effects and control for a rich set of station-level observable characteristics (i.e., local weather, regional demographics). We also use an instrument for D_{it}. We need a variable that is correlated with an individual station's adoption decision but is not affected by station-specific unobservable shocks. We propose *brand-HQ level adoption* as an instrument.[30] As explained in the previous section, we measure brand-level adoption by computing the share of stations belonging to each brand that have been identified as AI adopters by month t. For station i at time t our IV is the share of stations in brand i's brand that adopted algorithmic pricing by time t. We exclude station i from this share. Brand level decisions likely influence the adoption decisions of individual stations (see Section 3.3 for additional discussion). Brands provide individual stations with employee training, technical support and maintenance (Convenience.org). This happens for both chain-operated stations as well as for more independent lessees. For previous waves of technology adoption (such as electronic payments) brands also directly subsidized some costs associated with required station upgrades. Such support is important for drastic technical

[29]Table B4 in Appendix B.5 shows that adopting and non-adopting stations vary along a wide range of observable characteristics.

[30]As a robustness check, we propose an alternative set of instruments: the availability and quality of broadband internet in the local area around a gas station. As with *brand-HQ level adoption*, the availability of broadband internet should have an effect on a station's decision to adopt algorithmic pricing software. Most algorithmic pricing software are "cloud" based and require constant downloading and uploading of information. Without high speed internet, adoption of such software is not particularly useful for a station. However, the availability of broadband internet in the region should be uncorrelated with station unobservables after conditioning on observable local characteristics. Our estimates with these IVs are qualitatively similar to our main estimates. See Table E7 for results and Appendix E.4 for additional discussion. We also test a "placebo IV" which uses the brand HQ-level adoption decision by a random brand (not the brand of the station) as an instrument and find null effects. Additional discussion is also in Appendix E.4.

changes such as AI adoption. At the same time, brand level decisions should not be influenced by station-level specific conditions.[31] This identification assumption is similar to Hastings (2004) and Allen et al. (2014). In both cases, decisions taken at the national/HQ level are exogenous to local market conditions.

To test whether any observed changes in prices and/or margins come from a reduction in competition and increased market power, or from a better understanding of underlying fluctuations in wholesale prices and consumers' demand elasticity, we look separately at stations that are monopolists in their ZIP code and stations that are not monopolists.[32] If the adoption of algorithmic pricing software does not change competition but benefits station operations in other ways, we should expect to see effects for monopolist adopters. If adoption also affects competition we should expect to see additional non-zero effects for non-monopolist adopters on top of the effects for monopolist adopters. If adoption *only* affects competition, we should expect to see zero effects for monopolist stations and non-zero effects for non-monopolists.

5.2.2 Evaluating the Impact of Adoption on Market-Level Outcomes

In a more direct test of theoretical predictions about the effects of AI on competition, we compare outcomes between adopting and non-adopting *markets*. We focus on duopoly station markets since most theoretical analysis is done for two firms (i.e., Calvano et al 2020, Miklós-Thal and Tucker 2019). A duopoly market is a ZIP code in which there are only two stations.[33] For market m in month t, we use the following OLS specification:

$$y_{mt} = \alpha_m + \alpha_t + \beta_1 T^1_{mt} + \beta_2 T^2_{mt} + \epsilon_{mt}, \tag{2}$$

where y_{mt} is the outcome variable for market m at time t, α_m and α_t are, respectively, market and time fixed-effects. T^1_{mt} is a variable equal to 1 if only one of the two stations in market m is labelled as an adopter.[34] T^2_{mt} is a variable equal to 1 if both stations in market m are labelled as adopters.[35] The two key coefficients in this regression are β_1 and β_2. β_1 captures the effects of AI adoption by

[31] Table B6 shows that conditional on brand size, brand adoption shares are uncorrelated with market characteristics. See additional discussion in Section 6.2.

[32] As a robustness check, we use an alternative market definition based on 1km radius circles drawn around each station. See Appendix E.2 for additional discussion.

[33] As a robustness check, we use an alternative market definition based on 1km radius circles drawn around each station. See Section 7.3 and Appendix E.2 for additional discussion.

[34] $T^1_{mt} = D_{1mt}(1 - D_{2mt}) + D_{2mt}(1 - D_{1mt})$, where 1 and 2 are the stations in market m.

[35] $T^2_{mt} = D_{1mt}D_{2mt}$, where 1 and 2 are the stations in market m.

one of the two firms in a duopoly market and β_2 captures the effects of AI adoption by *both firms* in a duopoly market.

As in the station-level regression, endogenous AI adoption by stations in response to market/time varying unobservables is a concern. Following the logic of our main station-level instruments, we construct market-level IVs using brand-level adoption decisions.[36] The instruments for T^1_{mt} and T^2_{mt} are functions of the brand-level adoption decisions for the brands in market m:

$$\begin{aligned} IV^1_{mt} &= B_{1mt}(1 - B_{2mt}) + B_{2mt}(1 - B_{1mt}) \\ IV^2_{mt} &= B_{1mt}B_{2mt}, \end{aligned} \tag{3}$$

where B_{1mt} is the share of other stations belonging to market m station 1's brand that have been identified as AI adopters in month t. B_{2mt} is the share of other stations belonging to market m station 2's brand that have been identified as AI adopters in month t.

6 Results – AI Adoption

In this section we present results regarding the identification of adopters. In the first subsection (6.1) we discuss station-level adoption before then describing brand-level adoption in the second subsection (6.2). With these results in hand, in the Section 7 we study the effect of adoption on outcomes related to competition.

6.1 Station-level adoption

As outlined in Section 5, we calculate **structural breaks (remove QLR statistics)** QLR statistics for each station and each of the three adoption markers: (i) *number of price changes*, (ii) *average size of price changes*, and (iii) *rival response time*. We find that 13,133 stations experience a significant structural break in at least one of three adoption markers at the 5% confidence level. Almost 50% of **best-candidate** breaks in *number of price changes*, almost 30% of breaks in *rival response time* and over 20% of breaks in *average size of price changes* occur in the spring and summer of 2017 (see

[36]As a robustness check for station-level estimates, we propose an alternative set of instruments: the availability and quality of broadband internet in the local area around a gas station. These instruments would only work for market level data if the two duopolists are in the same ZIP code but also have different broadband access/quality conditions. Our broadband access/quality data is calculated at a coarser geographical level than 5 digit ZIP codes, so we are unable to use these instruments for market level data. See additional discussion in Appendix E.4.

Appendix B and Figures B1, B2 and B3 for the distribution of **best-candidate** break dates in each measure and additional discussion).[37]

The structural breaks capture substantial changes in pricing strategies. For the number of *price changes*, on average, a station that experienced a structural break changes their prices 6 times a day before the break and 9 after the break. *Rival response time* decreases from 64 minutes to 54 minutes on average after a structural break, a drop of about 10%. The *average size of price changes* increases from 2.7 to 2.9 cents, but as expected there is substantial heterogeneity across stations. Additional summary statistics and discussion are in Appendix B and Tables B1, B2 and B3.

Classification: There are many factors that may influence a single pricing behaviour measure on its own, but breaking in multiple markers in close proximity should provide a strong indication of adoption. Therefore, we label a station as an adopter of algorithmic-pricing software if it experiences **best-candidate** structural breaks in at least two measures of pricing behaviour within a short period of time.[38] In our main specification, we define a "short period of time" as 4 weeks, but our results are robust to stricter alternative definitions of adoption.[39]

We classify 3,323 stations as adopters. Figure 1 shows the distribution of the average break date for all adopters, defined as is the average year-week between **best-candidate** break dates of the two or three measures in which a station experiences a significant break.[40] Over 50% of these average break dates occur in the spring or summer of 2017.[41] This is consistent with the supposed increased availability of algorithmic pricing software in the summer of 2017 in Germany (see Section 3.3).

[37]One concern is that other structural breaks may occur at significantly different dates if we considered F-statistics that are not the maximum, but close to it. We find that generally F-statistic distributions are unimodal and stations do not have significantly different dates that may be identified as a structural break. Examples of F-statistics distributions are in Figure B5.

[38]Any combination of two measures will result in a station being classified as an adopter.

[39]In Appendix E.3 we change the definition of "a short period of time," requiring stations to experience **best-candidate** breaks in at least two of the three measures within 2 weeks. We also include an additional definition that only labels stations as adopters if they experience multiple **best-candidate** breaks in *both* E5 and Diesel.

[40]This is a conservative approach. We may be "missing" some adopters, either due to measurement errors in our measures or due to other signals of adoption that we did not consider. In practice, this means that some of the adopters are labelled as non-adopters. This would bias our station-level estimates towards zero and under-state the true effects of adoption.

[41]See Appendix B and Figure B6 for the number of adopters and the distribution of average break dates for each combination of measures.

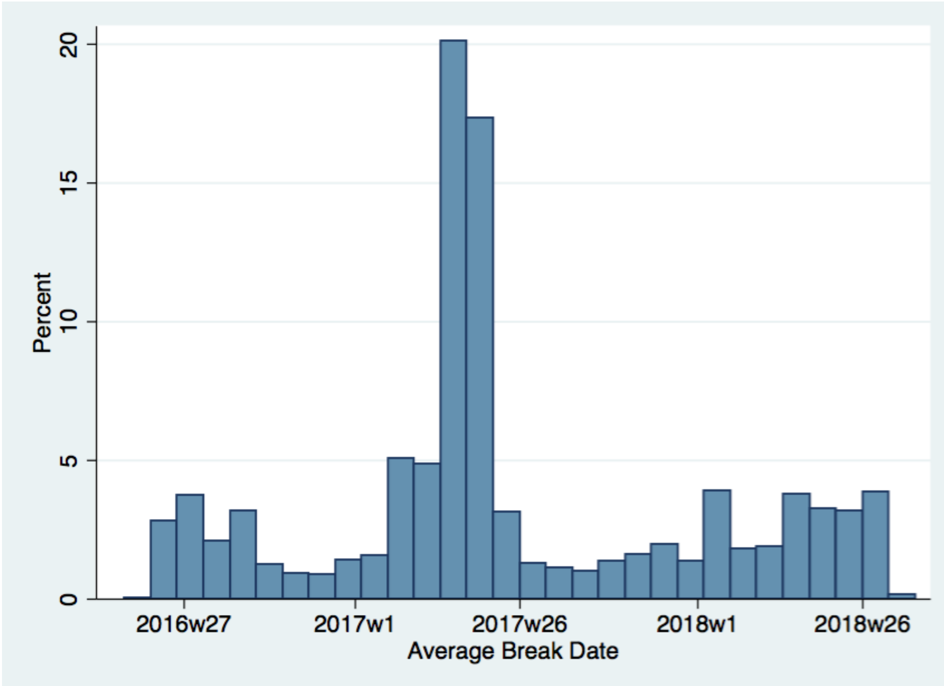


Figure 1: Frequency of Average Break Date for Measures Breaking Within 4 Weeks (3,323 stations)

A concern with our definition of adoption is that stations may be mistakenly labelled as adopters because their responses to an adopting rival's pricing makes them behave as though they also adopted. This does not appear be a regular occurrence. We observe a large number of duopoly markets where one station is classified as an adopter and not its competitor. Out of nearly 1,300 duopoly ZIP markets in our final sample, 780 had no adopters at any point in our sample, 390 had at most one adopter station throughout the sample period, and 69 had one adopter station followed by subsequent adoption by the second duopolist.[42] More generally, Figure 2 shows the geographic distribution of adoption shares in ZIP codes with more than one station in December 2018 (the last month in our data). It can be seen that there are relatively few ZIP codes where adoption shares are higher than 50%.

[42]There are approximately 30 ZIP codes where both duopolists are labelled as adopters in the same month. This could potentially reflect such concerns about mis-labelled adoption. Practically, these markets are not driving our main results. We replicated our analysis without these markets and results remain qualitatively and quantitatively the same.

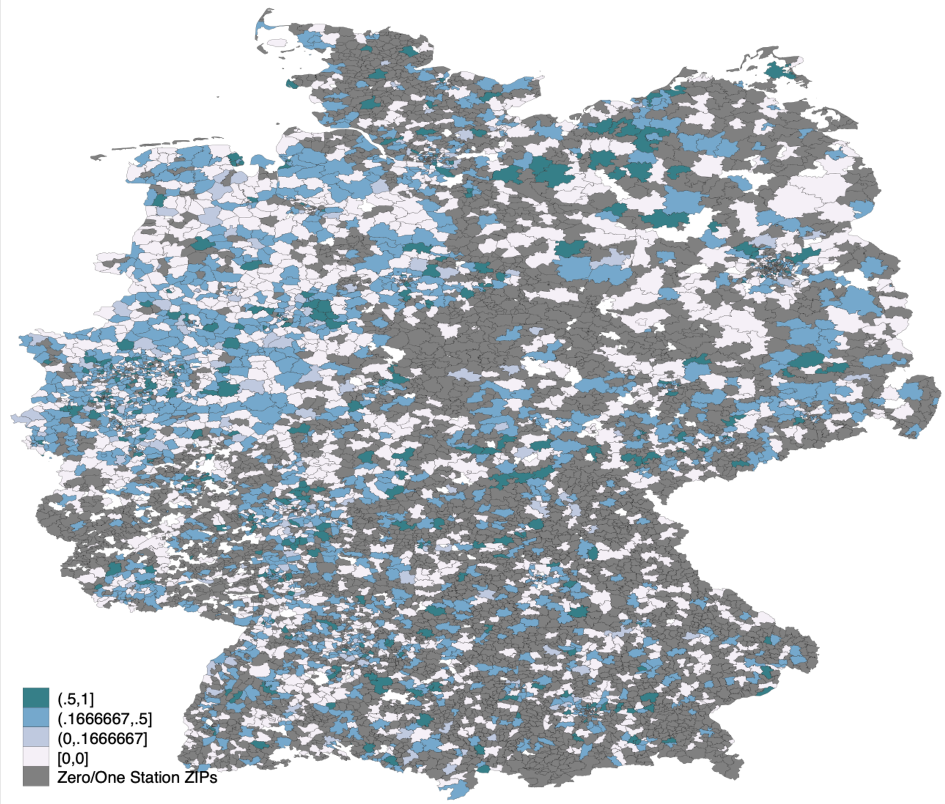


Figure 2: December 2018 ZIP-level Adoption Shares

Adopter and non-adopter stations are different. In Table B4 we find statistically significant differences in market characteristics between adopter and non-adopter stations before any adoption takes place (in 2016). Adopter stations are located in denser areas with different demographic profiles. Adopter stations also face more competition. This suggests that adoption decisions are likely endogeneous, with stations choosing to adopt in response to market conditions. Although we control for observable characteristics and include station and time fixed effects, if adopters and non-adopters are dissimilar in their observables they are also likely dissimilar in time-varying unobservable characteristics (e.g., managerial quality, demand and cost shocks). These findings confirm the need to use an IV strategy to address the endogeneity.

6.2 Brand-Level Adoption

Figure 3 shows the evolution of the share of adopting stations for the Top 5 brands in our data throughout our sample period. Notably, none of these brands have adoption rates over 40% by the end of the sample period.[43] Adoption happens at a staggered rate that varies across brands. All brands experience spikes in adoption patterns that happen around early/mid 2017, likely reflecting the increased availability of the technology. Aral is an early adopter, with 10% of its stations adopting by early 2017. Total catches up and overtakes it by the end of the sample.[44] Esso's and Avia's adoption rates increase at a steadier (albeit slower) pace compared to other brands. The heterogeneity in adoption rates across brands suggests that there is a brand-specific component to AI adoption. As mentioned in Section 3.3, it is likely that some brands were more likely to support the new technology (or adopt at the "HQ" level).

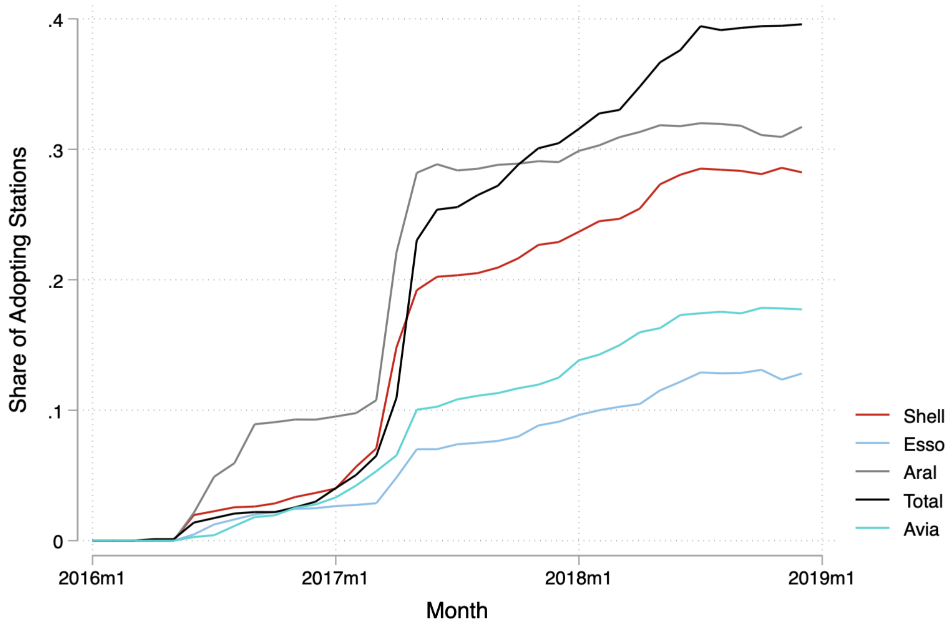


Figure 3: Share of AI Adopters Among Top 5 Brands

The pattern in Figure 3 is similar to the staggered-adoption that was observed for electronic payment adoption by Canadian gasoline retail stations in the 1990s (see Figure B1 in Appendix D). Despite the differences in time, geography and technology, we also find a staggered pattern of

[43]Table B5 present summary statistics regarding the share of a brand's stations that adopt by the end of each calendar year. The share of adopters for smaller brands is lower. The mean adopter share in 2018 for non top-5 brands is 25%. This likely reflects the better support that larger brands can provide to their stations, which would reduce their cost of adoption. We find similar patterns in the distribution of structural breaks for individual adoption measures (i.e., number of price changes) in Figure B7 in Appendix B.

[44]There may be some concerns that Aral's early adoption is a measurement error which is driving the results. We address this in a robustness check by dropping all Aral stations. The main results are quantitatively similar. See Appendix E.1 for more discussion.

technology adoption that appears to be highly brand specific. This suggests that our AI adoption classification captures technology adoption.

We test whether the heterogeneity in brand-level adoption probability is explained by observable brand characteristics. Unlike station-level adoption, brand level adoption is not correlated with brand-level observables after controlling for brand size (the number of stations in the brand). Table B6 shows that conditional on the number of stations in the brand, the share of brand adopters is uncorrelated with average demographic characteristics of a brand's stations. It is also uncorrelated with the average number of competitors that a brand's stations have. This makes intuitive sense. Brands likely spread out their stations across different markets. Local characteristics will inevitably average out. Brands also make broad strategic decisions that should not be influenced by local market conditions. The only statistically significant correlate of adoption probability at the brand level seems to be brand size. Because of this, we control for brand size in the IV estimates below.

7 Results – Effects of AI Adoption

7.1 Impact of Adoption on Station Outcomes

We use 2SLS regressions to measure the impact of algorithmic-pricing adoption on mean daily station margins and prices, along with their distributions. For each station and day in our sample we compute an average station/day price and subtract an average regional daily wholesale price, German gasoline taxes and VAT.[45] This provides us with daily station-level margins. We define mean station-level margins and prices by taking the mean of these daily station-level margins and prices for each station for each month. Also, for each $X \in [5, 25, 50, 75, 95]$ we calculate the Xth percentile of margins and prices by finding the Xth percentile of daily margins and prices for each station within a month.

Station-level IV estimates are presented in Table 3.[46] In each regression we control for the number of competitors in the station's ZIP code, the number of competitors who adopted algorithmic pricing software, region/year demographics and the number of stations in station i's brand. Column (1) shows the first stage of the IV regression. The first stage is strong, with an F-statistic of 35. A 10% increase in the number of other stations affiliated with station i's brand (excepting station i) that

[45] More details about data construction can be found in the Data Appendix.

[46] OLS station-level estimates for margins are in Table C1 in Appendix C. Results from the OLS specification suggest that adoption has a negligible impact on margins. This is unsurprising given the likely endogeneity in adoption decisions. Table B4 shows that adopter and non-adopter stations are very different in their local market demographics and in their competitive environment. They are also likely to be different in their unobservable characteristics.

adopt by period t increases the probability that i adopts by period t by 65%. This is consistent with our intuition that adoption of algorithmic pricing is at least in part a brand-level decision.

Table 3: 2SLS Station-Level Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome:	Adopter	Mean Margin	5th Pctile Margin	25th Pctile Margin	Median Margin	75th Pctile Margin	95th Pctile Margin
Adopter		0.008***	0.015***	0.009***	0.006***	0.005***	0.019**
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.008)
N Competitors in ZIP	-0.003	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	0.000
	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
Share Brand Adopters	0.653***						
	(0.034)						
Non-Adopter Mean Outcome		0.0828	0.0538	0.0681	0.0767	0.0850	0.208
Station FE	YES	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES	YES
Observations	448,221	448,221	448,221	448,221	448,221	448,221	448,221

	(8)	(9)	(10)	(11)	(12)	(13)
Outcome:	Mean Price	5th Pctile Price	25th Pctile Price	Median Price	75th Pctile Price	95th Pctile Price
Adopter	0.006***	0.011***	0.007***	0.007***	0.005**	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N Competitors in ZIP	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non-Adopter Mean Outcome	1.361	1.336	1.349	1.361	1.374	1.387
Station FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
Observations	448,221	448,221	448,221	448,221	448,221	448,221

Notes: Sample is gas station/month observations from January 2016 until December 2018. Margins are computed above wholesale gasoline prices at a regional terminal nearest to station j. Mean Margin/Price is the monthly average pump price for station j in month t. "Xth Pctile" Margin/Price is the Xth percentile of daily pump price or margin for station j in month t. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Share Brand Adopters" is the excluded instrument used in the 2SLS regression. It is equal to the share of stations that belong to the brand of station j that adopted by period t. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the number of stations belonging to station i's brand in month t. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. We also control for the number of other stations in the ZIP code who are adopters at month t. ZIP level clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Columns (2)-(7) of Table 3 show 2SLS estimates with margin outcomes. Column (2) shows that mean margins increase by 0.8 cents per litre on average after AI adoption. This is an increase of about 9% relative to the average non-adopter margin of 8.3 cents.[47] Columns (8)-(13) show 2SLS regressions with price outcomes. Mean prices increase by 0.6 cents per litre on average post-adoption.

[47] 2SLS regressions using alternative instruments based on broadband availability and quality also show that mean margins and mean prices increase after adoption (see Table E7). See Appendix E.4 for additional discussion of these instruments and results.

Each of the 5th, 25th, 50th, and 75th percentile margins and prices increase for adopting stations relative to non-adopters. This indicates that AI-adoption generally leads to higher prices and margins for adopters.

7.2 Impact of Adoption on Competition

Algorithmic pricing can increase station margins and prices through a reduction in competition and increased market power. But there can also be other reasons for such changes. An algorithm could better understand underlying fluctuations in wholesale prices, or identify how price elasticity of demand changes over the day or the week and adjust prices accordingly. We test for these different explanations by allowing for heterogeneous effects across different market structures. We separate our sample into two: one sub-sample of stations that are monopolists in their ZIP code, and one sub-sample of stations that are not.

Results of our 2SLS regression for the two subsamples are presented in Table 4.[48] We find that non-monopolist stations are driving the increase in mean margins, with mean margins increasing for non-monopolist adopters by 0.9 cents post-adoption 11%). By comparison, monopolist adopters have a small and non-statistically significant changes in margins, except at the 5th and 95th percentiles. Average price effects are similar to margins. Mean monthly prices for non-monopolist stations increase by 0.7 cents per litre and the entire price distribution moves to the right. Average prices of monopolist adopter stations do not change, except for the 95th percentile of prices which decrease by 1.2 cents per litre after adoption.

This decrease in the highest prices set by monopolists possibly reflects algorithms' ability to avoid setting prices that are "too high" and captures potential benefits for consumers from algorithmic pricing. Alternatively, the algorithms can help monopolists price discriminate better, which is also not necessarily welfare decreasing. That monopolist margins increase substantially at the 5th and 95th percentiles *without* increasing prices, potentially suggests that monopolist stations were perhaps worse at tracking wholesale prices.

Overall, the mostly statistically null effects on monopolist outcomes and positive effects on non-monopolist outcomes imply that adoption of algorithmic pricing software increases margins as a result of changes in strategic interaction and competition rather than other changes such as better understanding of underlying wholesale price fluctuations and consumers' demand elasticity.

[48] Results using the alternative 1KM radius market definition are in Table E2. See additional discussion of alternative market definitions in Appendix E.2.

Table 4: 2SLS Station-Level Estimates by ZIP Market Structure

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Mean Margin	5th Pctile Margin	25th Pctile Margin	Median Margin	75th Pctile Margin	95th Pctile Margin
			Sample: Monopoly ZIP Stations			
Adopter	0.003	0.013***	0.003	-0.001	-0.004	0.052**
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.026)
Non-Adopter Mean Outcome	0.0826	0.0547	0.0686	0.0770	0.0853	0.197
Observations	67,300	67,300	67,300	67,300	67,300	67,300
			Sample: Non-Monopoly ZIP Stations			
Adopter	0.009***	0.016***	0.010***	0.008***	0.007***	0.014
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.009)
N Competitors in ZIP	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.001
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
Non-Adopter Mean Outcome	0.0829	0.0536	0.0680	0.0766	0.0849	0.211
Observations	380,826	380,826	380,826	380,826	380,826	380,826
Station FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
	(7)	(8)	(9)	(10)	(11)	(12)
Outcome:	Mean Price	5th Pctile Price	25th Pctile Price	Median Price	75th Pctile Price	95th Pctile Price
			Sample: Monopoly ZIP Stations			
Adopter	-0.003	0.003	-0.001	-0.001	-0.005	-0.012**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Non-Adopter Mean Outcome	1.362	1.337	1.350	1.362	1.374	1.387
Observations	67,300	67,300	67,300	67,300	67,300	67,300
			Sample: Non-Monopoly ZIP Stations			
Adopter	0.007***	0.012***	0.008***	0.008***	0.006***	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N Competitors in ZIP	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non-Adopter Mean Outcome	1.361	1.335	1.349	1.361	1.374	1.387
Observations	380,826	380,826	380,826	380,826	380,826	380,826
Station FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES

Notes: Sample includes gas station/month observations from January 2016 until December 2018, split up into two subsamples: one subsample only includes stations that have no competitors in their ZIP code. The other subsample includes only stations that have one or more competitors in their ZIP code. Margins are computed above wholesale gasoline prices at a regional terminal nearest to station j. Mean Margin/Price is the monthly average pump price for station j in month t. "Xth Pctile" Margin/Price is the Xth percentile of daily pump price or margin for station j in month t. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Share Brand Adopters" is the excluded instrument used in the 2SLS regression. It is equal to the share of stations that belong to the brand of station j that adopted by period t. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the number of stations belonging to station i's brand in month t. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. We also control for the number of other stations in the ZIP code who are adopters at month t. Standard errors clustered at ZIP level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Further support for the competition-channel explanation is provided by looking at duopoly market outcomes. Table 5 presents **2SLS** estimates of Equation (4) using the instruments defined in Equation (5) and market-level margins and prices as the outcome variables of interest. We focus on duopoly (two-station) 5 digit ZIP-code markets.[49] First-stage estimates of the 2SLS are in Table C2 in Appendix C. As was the case with the station-level instruments, the partial correlation between market-level instruments and the endogenous variables is strong.

2SLS estimates suggest that AI adoption by only one station in a duopoly market does not affect market-level margins or prices relative to a duopoly market where no stations adopted. However, AI adoption by both stations in a duopoly market does affect market-level margins and prices. Mean market-level margins increase by 3.2 cents per litre after market-wide AI adoption. This is a substantial increase of nearly 38% relative to the baseline. The 5th, 25th, 50th, 75th and 95th percentile market-level margins also shift up after market-wide adoption. Similar effects are observed for market-level prices after market-wide adoption, where mean market prices increasing by 4 cents per litre.

A possible explanation for not seeing changes in mean market-level margins after asymmetric adoption (when one station adopts and the other does not) could be because the adopter's margins increase and the non-adopter's margins fall, cancelling out on average. We test this hypothesis by looking at non-adopter stations in duopoly markets and comparing margins and prices before and after their rival adopts (as before, we instrument for the rival's adoption with the rival brand's adoption decision). Results from these regressions are in Table C3 in Appendix C. We do not see any statistically significant changes in margins and prices following a rival's AI adoption, ruling out this explanation.

These results serve as a direct test of theoretical hypotheses about the effects of AI adoption on market outcomes. Theoretical literature suggests that it is possible for algorithms to facilitate collusion (Calvano et al 2020, Miklós-Thal and Tucker 2019).[50] We cannot be sure what type of algorithms station-owners are using and whether they fully turn over pricing decisions to algorithms. Nonetheless, lack of margin changes from partial/asymmetric adoption and substantial increases in margins and prices after complete adoption is suggestive of algorithms facilitating tacit-collusion. The

[49]Results at the 1km radius market-level are in Table E4. See Appendix E.2 for additional discussion of alternative market definitions.

[50]There is also a possibility that multiple stations in a market turn over their pricing decisions to a common algorithmic software provider. Algorithms in this case serve as the "hubs" of a hub-and-spoke cartel (Harrington 2018b). If multiple stations in a market turn over their pricing decisions to a common algorithmic software provider, our results are in line with the findings of Decarolis and Rovigatti (2019).

magnitude of margin increases in duopoly markets is consistent with previous findings on coordination in retail gasoline markets (Clark and Houde 2013, 2014; Byrne and De Roos 2019). We present additional evidence on the mechanism through which algorithmic competition affects margins in Section 8.

Table 5: 2SLS ZIP Duopoly Market Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Mean Mkt Margin	5th Pctile Mkt Margin	25th Pctile Mkt Margin	Median Mkt Margin	75th Pctile Mkt Margin	95th Pctile Mkt Margin
One Station Adopted	-0.008	-0.003	-0.007	-0.008	-0.008	-0.013
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.037)
Both Stations Adopted	0.032***	0.036***	0.032***	0.028**	0.030**	0.099**
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.044)
Zero-Adopter Mean Outcome	0.0836	0.0551	0.0692	0.0776	0.0857	0.199
ZIP FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Controls	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
Observations	39,148	39,148	39,148	39,148	39,148	39,148
	(7)	(8)	(9)	(10)	(11)	(12)
Outcome:	Mean Mkt Price	5th Pctile Mkt Price	25th Pctile Mkt Price	Median Mkt Price	75th Pctile Mkt Price	95th Pctile Mkt Price
One Station Adopted	-0.016	-0.015	-0.019	-0.017	-0.016	-0.014
	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)
Both Stations Adopted	0.040**	0.045***	0.042**	0.042**	0.039**	0.032**
	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)
Zero-Adopter Mean Outcome	1.350	1.325	1.338	1.350	1.362	1.376
ZIP FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Controls	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
Observations	39,148	39,148	39,148	39,148	39,148	39,148

Notes: The sample includes duopoly market/month observations from January 2016 until December 2018. A duopoly market is defined as a ZIP code with two gas stations. Outcome variable Mean Market Margin is the monthly average of mean market daily differences of pump prices for stations in market m in month t from wholesale price. Outcome variable Mean Market Price is the monthly average of mean market daily pump prices for stations in market m in month t . Xth Percentile Market Margin is the Xth percentile of observed mean market daily margins/prices for stations in market m in month t. "One Station Adopted " is a dummy equal to 1 in month t if one of the two stations in the market experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Both Stations Adopted" is a dummy equal to 1 in month t if *both* stations in the market experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. Instruments for adoption are the "share of brand adopters" of the two stations in the market. 1st stage regression results are in Table C2 in the Appendix. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the sizes of the brands of the two stations at month t. Standard errors clustered at market level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

7.3 Robustness

In this section we briefly outline a series of checks that confirm the robustness of our results to alternative samples, market definitions, adoption classifications and instruments. Results along with further details are in Appendix E. In every case results on the impact of adoption on margins are robust to the proposed check.

1. **Alternative Estimation Samples (Appendix E.1):** We re-estimate the main regressions with alternative samples of stations. We address possible contamination from the Shell price matching promotion from 2015 by (i) dropping observations from markets containing Shell stations, and (ii) dropping all observations from 2016. We also address possible concerns about Aral's stations early adoption of algorithmic prices by removing all observations of Aral stations. We also address concerns about entry/exit of stations from the sample by looking at a balanced sample of stations and a balanced sample of stations and markets, dropping any market where the number of stations changes over time.

2. **Alternative Market Definitions (Appendix E.2):** We use an alternative market definition commonly used in the literature: the direct distance between stations. We define a monopolist as a station with no competitors within a 1km radius. We define a duopoly market as two stations of different brands that are within 1km of one another and that do not have any other competitors within 1km.

3. **Alternative Adoption Definitions (Appendix E.3):** We test the robustness of our "adopter" definition by using alternative classifications. First, we classify adopters based only on measures that do not rely on the presence of a nearby rival (number of price changes, and average price change size). Second, we consider an alternative definition regarding the time between structural breaks. While the baseline model classifies a station as an adopter if they experience a structural break in at least two out of three measures within an 4 week period, in our alternative definition we label stations as adopters if they experience structural breaks in at least two out of three measures within a period of 2 weeks. Last, we consider an alternative definition where a station is classified as an adopter if they experience multiple structural breaks in both E5 and Diesel.

4. **Alternative Instruments (Appendix E.4):** We propose using the availability of broadband access in station j's region as an instrument for adoption. Intuitively, if a station has access to high speed internet and/or reliable internet signals, it should be more likely to adopt algorithmic pricing technology. We use two measures: whether the local area around the gas-station has widespread access to high speed internet in a particular year, and the reliability of broadband signals in that year (measured by the average and variance of signal strength). We also introduce a "placebo" instrument. Rather than using the share of stations of station j's brand that adopted as an IV, we use the share of stations by *another* brand (i.e., the brand of some station k in the market of station j). We expect that there should be no correlation between

the propensity of station j to adopt and average adoption by other brands since they do not directly affect station j's costs.

5. **Alternative Fuel Types:** We use E5 gasoline since it has the highest market share (80%) in Germany. In Assad et al (2020) we use E10 gasoline instead of E5 gasoline. Results are qualitatively and quantitatively similar to the ones in this paper.

8 Mechanism

In this section we use data from duopoly markets to provide evidence of the mechanism through which algorithmic competition increases prices and margins. Our findings suggest that algorithmic learning takes places and that algorithmic competition increases margins and prices only over time. We also provide suggestive evidence of changes in pricing strategies that appear specifically when algorithms compete head to head, involving the softening of competition.

Establishing that learning is taking place is important since there are two possible explanations for why algorithms could reach margins above competitive levels. Pricing algorithms could *fail to learn to compete effectively* (Cooper et al 2015, Hansen, Misra and Pai 2020). For example, algorithms may not fully incorporate rivals' prices or may not best respond to these prices. Alternatively, algorithms could fully incorporate rivals' prices and best respond to them, but they could also *learn how not to compete* (i.e. how totacitly collude).[51] For example, algorithms may learn to punish competitors for reducing prices or other tacitly-collusive strategies. The two explanations have very different implications for competition policy, which should mostly be concerned with algorithms actively learning *not to compete*.

The two explanations have different predictions regarding how quickly prices and margins increase after algorithmic adoption. If the first explanation holds, we would expect to see high margins

[51] Brown and MacKay (2020) present a third possible explanation for why prices and margins increase after algorithmic adoption. In their model, adoption of algorithmic software changes the game that firms play from a standard simultaneous Bertrand pricing game to a stage game. This increases prices and margins relative to a simultaneous Bertrand-Nash equilibrium. We test a key prediction from their model: the bigger the asymmetry in pricing technology, the higher market prices and margins should be. We observe a large number of duopoly markets that feature asymmetric adoption of algorithmic pricing technology. Table C3 shows results from a regression of a non-adopting stations' margins on a dummy variable of whether its rival has adopted algorithmic pricing technology (instrumented by the rival brand's adoption share). We find that there are no statistically significant changes in margins following a rival's adoption. Although the Brown and MacKay (2020) model appears to fit well certain settings (such as cold medicine markets), in our context it does not seem to apply.

throughout the post-adoption period.[52] If the second explanation holds, we would expect to see no initial effects followed by an eventual convergence towards tacitly-collusive price levels and increased margins. Echoing statements made by AI experts, Calvano et al (2020) show in simulations that it takes a long time for algorithms to train and converge to stable strategies. Without "offline" training, their results suggest that training should take several years. Even with offline training, it could take up to a year for their algorithms to converge to stability.

We provide some evidence in favour of the second explanation by examining the timing of adoption effects. Table 6 shows estimates of time-specific effects of one and both stations adopting on mean market margins (i.e. T^1_{mt} and T^2_{mt} from Equation 4), in a regression that includes the controls from Table 5 and market and time FE. Time-specific adoption variables are instrumented by time-specific versions of IV^1_{mt} and IV^2_{mt} from Equation (5). We bin the timing effects into three periods: the first six months after adoption, the second six months after adoption, and a year or longer after adoption. We use these bins since there is only a small number of markets we observe for a very long period of time after adoption.

Table 6 shows 2SLS coefficient estimates of the "Both Stations Adopted" and "One Station Adopted" variables on average monthly market-level margins and prices. Consistent with simulation results in Calvano et al (2020), we find that for roughly the first year after both duopolist stations in a ZIP code market adopt AI there are no statistically significant changes in average market margins or prices at the 95% confidence level.[53] The main effects we find in Table 5 come in only a year after both stations adopt. These results are similar to previous findings on transitions to collusive strategies in other markets. Igami and Sugaya (2019) show that 1990s Vitamin cartels took several years to increase their prices and margins. Clark, Hortsmann and Houde (2020) also show a lengthy adjustment period to high prices for a Canadian bread cartel, as do Byrne and de Roos (2019) in the Australian retail gasoline market.

We provide additional suggestive evidence of how algorithmic competition operates differently from non-algorithmic competition. There are no clear measures of conduct that can be identified in a reduced form setting without an underlying model. In our setting, developing such a model is not straight-forward since we do not know the precise algorithms used by the competitors.[54] Nonetheless,

[52] Or high initial margins followed by lower margins if algorithms learn how to compete more effectively over time.

[53] Figure 10 in Calvano et al (2020) shows that profit margins for algorithms do not substantially change for over 500,000 simulation "periods." Under the assumption that a simulation period lasts for a few minutes, Calvano et al (2020) suggest that this would correspond to at least a year.

[54] Many price setting algorithms including the Q-learning algorithm in Calvano et al (2020) are not designed to play mixed strategies. Other algorithms, as well as humans, are able to play mixed strategies. There are many possible asymmetric equilibria and characterising them without further information is not feasible. We leave this question for

Table 6: 2SLS ZIP Duopoly Price and Margin Timing

VARIABLES	(1) Mean Mkt Margin	(2) Mean Mkt Price
0-6 months since One Station Adopted	-0.000	-0.000
	(0.001)	(0.001)
7-12 months since One Station Adopted	0.001	0.001
	(0.001)	(0.001)
12+ months since One Station Adopted	-0.001	0.000
	(0.003)	(0.003)
0-6 months since Both Stations Adopted	0.006	0.007*
	(0.004)	(0.004)
7-12 months since Both Stations Adopted	0.012*	0.011*
	(0.006)	(0.006)
12+ months since Both Stations Adopted	0.039**	0.043**
	(0.018)	(0.018)
ZIP FE	YES	YES
Year-Month FE	YES	YES
Annual Regional Demographics	YES	YES
N Brand Stations Controls	YES	YES
Weather Controls	YES	YES
Observations	39,148	39,148

Notes: The sample includes duopoly market/month observations from January 2016 until December 2018. A duopoly market is defined as a ZIP code with two gas stations. Outcome variable Mean Market Margin is the monthly average of mean market daily differences of pump prices for stations in market m in month t from wholesale price. Outcome variable Mean Market Price is the monthly average of mean market daily pump prices for stations in market m in month t . "X months since One Station Adopted " is a dummy equal to 1 in month t if one of the two stations in the market has become an adopter in the previous X months and zero otherwise. "X months since Both Stations Adopted" is a dummy equal to 1 in month t if *both* stations in the market become adopters in the previous X months and zero otherwise. Instruments for both "X months since One Station Adopted " and "X months since Both Stations Adopted" include measures of the "share of brand adopters" of the two stations interacted with timing dummies. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the sizes of the brands of the two stations at month t. Standard errors clustered at market level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

we can empirically attempt to evaluate changes in pricing behaviour and the timing of these changes coming directly from duopoly algorithms competing against one another.

We focus on three pricing behaviours. We look at whether a station immediately responds to a price change by its rival: the market-level probability that if one station reduces its price, the other station also reduces its price within 5 minutes, and the market-level probability that if one station increases its price, the other station also increases its price within 5 minutes. We also look at the mean within-market within-day differences in duopoly station prices.[55]

Last, we look at the probability that stations *undercut* one another. In every market and month we label the station that is more likely to set higher average daily prices as the *high-price* station

future research.

[55]In every day and market, we calculate the absolute difference in prices between the two stations and take the average of that measure for each market and month. More details are in the Data Appendix.

Table 7: 2SLS ZIP Duopoly Additional Price Effects

Outcome:	(1) Prob. Response to Price Decrease	(2) Prob. Response to Price Decrease	(3) Prob. Response to Price Increase	(4) Prob. Response to Price Increase	(5) Mean Diff b/w High and Low Price	(6) Mean Diff b/w High and Low Price	(7) Undercutting Prob.	(8) Undercutting Prob.
One Station Adopted	-0.004		0.093		0.008		0.017	
	(0.055)		(0.084)		(0.005)		(0.054)	
Both Stations Adopted	0.181**		-0.108		-0.004		-0.134*	
	(0.073)		(0.076)		(0.006)		(0.071)	
0-6 months since One Station Adopted		0.018***		-0.006		0.000		-0.002
		(0.005)		(0.007)		(0.001)		(0.007)
7-12 months since One Station Adopted		0.021***		0.002		-0.000		0.001
		(0.006)		(0.009)		(0.001)		(0.009)
12+ months since One Station Adopted		0.019		0.027		-0.000		0.010
		(0.015)		(0.019)		(0.002)		(0.018)
0-6 months since Both Stations Adopted		0.062***		-0.038*		-0.000		-0.031
		(0.023)		(0.021)		(0.002)		(0.024)
7-12 months since Both Stations Adopted		0.079***		-0.053*		-0.001		-0.063**
		(0.029)		(0.027)		(0.003)		(0.028)
12+ months since Both Stations Adopted		0.227**		-0.153		0.001		-0.213**
		(0.103)		(0.093)		(0.009)		(0.102)
ZIP FE	YES	YES	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES	YES	YES
N Brand Stations Controls	YES	YES	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	34,370	34,370	33,367	33,367	39,129	39,129	38,664	38,664

Notes: The sample includes duopoly market/month observations from January 2016 until December 2018. A duopoly market is defined as a ZIP code with two gas stations. "X months since One Station Adopted " is a dummy equal to 1 in month t if one of the two stations in the market has become an adopter in the previous X months and zero otherwise. "X months since Both Stations Adopted" is a dummy equal to 1 in month t if *both* stations in the market become adopters in the previous X months and zero otherwise. Instruments for both "X months since One Station Adopted " and "X months since Both Stations Adopted" include measures of the "share of brand adopters" of the two stations interacted with timing dummies. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the sizes of the brands of the two stations at month t. Standard errors clustered at market level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

and the other station as the *low-price* station.[56] Our measure of undercutting is the percentage of days in a month that the station we label as *high-price* sets **lower** prices than the station we label as *low-price*. For markets with no adopters, this probability is approximately 10%. *High-price* stations set lower prices than the *low-price* stations approximately 3 days out of 30.

We estimate the model from Equation (2) using these new outcomes. Coefficient estimates from these regressions are presented in Table 7. Columns (1)-(4) show the effects of algorithmic adoption on immediate responses to rivals' price decreases or increases. Columns (5) and (6) show the effects of one or both duopolists adopting on the mean within-market within-day difference in duopoly station prices. Columns (7) and (8) use market-level undercutting probability as an outcome. Odd columns show aggregated effects and even columns allow for differences in timing for the effects as in Table 6.

Columns (1)-(4) show a striking difference in the effects of algorithmic adoption on the probability that stations respond to rivals' price changes. Column (1) shows that stations are 18% more likely to respond to a rival's price decrease with a price decrease of their own within 5 minutes. Column (2)

[56] If station 1 in a market sets higher average daily prices than station 2 for 21 days out of 30, station 1 is labelled as the *high-price* station and station 2 is labelled as the *low-price* station.

shows that this is propensity is growing over time when both stations are algorithmic adopters but *not* when only one station is an adopter. Notably, this is not the case for price *increases*. Columns (3) and (4) show weak evidence of *decreases* in the propensity of stations to respond to price increases by their rivals after algorithmic adoption, although we find that this propensity does not increase over time and statistical significance for the estimates is low. Columns (5) and (6) show that adoption of algorithmic pricing has no effects on the average differences between station prices in a duopoly market. This is true both immediately after adoption and later. Since Table 5 shows that average market level prices increase after algorithmic adoption, this suggests that both adopting duopolists increase their prices by similar amounts after adoption.

Columns (7) and (8) show that *undercutting*, as defined above, disappears when both stations adopt. In Column (7), we show that undercutting falls by 13% after both stations adopt algorithmic pricing. Relative to a baseline probability of 10%, it becomes a statistical zero in markets where algorithms compete head-to-head. In these markets the *high-price* station never attempts to set lower prices than the *low-price* station, on average. Timing results in Column (8) show that this coincides with the increase in margins and prices. In both Columns (7) and (8), there are no effects when only one station adopts algorithmic pricing.

Together these results are striking and suggest a simple mechanism through which algorithmic competition maintains high prices and margins. Effectively, the algorithms meet any price decrease with a price decrease of their own, teaching each other that undercutting will not be profitable since the undercutter will always be followed to the lower price by the other station.

9 Policy Discussion and Conclusions

We investigate potential links between algorithmic pricing and competition by looking at the widespread introduction of AI-pricing software in the German retail gas market. First, we identify which stations have adopted this pricing software through structural break tests in various measures of behaviour during a sample period of 2016-2018. We then analyze the impact of algorithmic-pricing adoption by comparing competition measures for adopting stations vs. non-adopting stations.

To identify algorithmic-pricing adoption, we focus on stations that experience structural breaks in at least two out of three measures of pricing behaviour within a 4 week period. Comparing breaks in (i) the number of price changes, (ii) the average size of price changes, and (iii) rival response time, we find that the vast majority of breaks occur in mid-2017, the time at which the AI software became widely available.

Having identified adopting stations we investigate the effects of algorithmic adoption on the mean and the distribution of daily margins and prices. Due to the potential endogeneity of station-level adoption decisions, we instrument for station i's adoption using the share of stations in i's brand that have adopted. Results indicate that, overall, AI-adopters with nearby competitors increase mean margins by 9% on average in comparison to pre-adoption levels. Mean prices also increase and the distribution of margins and prices generally shifts right. In contrast, adopters that are a monopolist in their ZIP code do not see changes in their mean margins. Looking at duopoly (two station) markets exclusively, we find that there is no difference in market-level margins between markets in which no stations adopted and markets in which one of the two stations adopted. However, markets in which both stations adopted show a mean margin increase of nearly 38% and the entire distribution of margins shifts to the right (increases). Mean prices increase by approximately 4 cents per litre. These estimates are lower-bounds on the true effects, since measurement errors in the first step of the analysis likely result in labelling some AI adopters as non-adopters.

We investigate the mechanism behind the increases in margins by looking at the timing of effects. If algorithms *fail to learn to compete effectively* we should see immediate increases in margins after both stations in duopoly markets adopt AI. If algorithms *learn how not to compete*, we should see no initial effects followed by eventual convergence to high prices and margins. This is what we find in the data - margins in markets where both duopolists adopt do not change for about a year after adoption and then increase gradually. This is suggestive of algorithms learning tacitly-collusive strategies over time. Overall, the results indicate that the adoption of algorithmic pricing has affected competition and facilitated tacit-collusion in the German retail gas market.

Our findings suggest that regulators should be concerned about the mass-adoption of algorithmic pricing software. Multiple antitrust authorities and economic organizations (OECD 2017; Competition Bureau 2018; Autorité de la Concurrence and Bundeskartellamt 2019; UK Digital Competition Expert Panel 2019) have released reports on algorithmic collusion and competition law. The reports agree that explicit algorithmic collusion would not require any changes to existing competition laws, but would change how competition authorities monitor for and investigate collusive practices. Increased tacit collusion through algorithms could change the legal status of such forms of collusion (in addition to monitoring and investigative practices). Currently, tacitly collusive behaviour is difficult to prove and prosecute as it does not rely on explicit communication. The UK Digital Competition Expert Panel states that with "further evidence...of pricing algorithms tacitly co-ordinating of their own accord, a change in the legal approach may become necessary" (p.110, 2019).

While our evidence is particular to retail gasoline markets in Germany (where high frequency pricing data are available), the same algorithmic pricing software is adopted in gasoline retail markets around the world. At a minimum, our results suggest that competition authorities should undertake a census of retail-gasoline pricing software to understand the market structure of the algorithmic software market and the extent of adoption. Such a census can help separate whether the main effect of algorithmic pricing software on competition comes from multiple stations in a market adopting *the same* or *different* algorithms. We do not directly observe which algorithm competitors adopt and the two possibilities have different implications for regulators and policy-makers.[57]

Our focus in this paper is on the retail gasoline market, but custom-made and "off-the-shelf" algorithmic pricing software is widely available to use for online and offline retailers. Adoption of such algorithms is growing: Brown and MacKay (2020) present evidence of algorithmic pricing by pharmaceutical drug retailers online. *Vendavo*, an AI based retail pricing software reports over 300 global deployments in manufacturing, chemicals, distribution and high tech industries (Vendavo.com). Our results suggest that competition authorities should investigate the relationship between algorithmic pricing software adoption and competition in these and other contexts.

[57] If multiple stations in a market turn over their pricing decisions to a common algorithmic software provider, our results are in line with the findings of Decarolis and Rovigatti (2019). Algorithms in this case serve as the "hubs" of a hub-and-spoke cartel (Harrington 2018b).

References

1. Abreu, D., Milgrom, P., & Pearce, D. (1991). Information and timing in repeated partnerships. *Econometrica* 59, 1713-1733.

2. Agrawal, A., J. Gans and A. Goldfarb (2019). The Economics of Artificial Intelligence: An Agenda. University of Chicago Press.

3. Albrecht, S. V., & Stone, P. (2018). Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence* 258, 66-95.

4. Allen, J., R. Clark and J.F. Houde (2014). The effect of mergers in search markets: Evidence from the Canadian mortgage industry, *American Economic Review* 104, 3365-3396.

5. Andrews, D. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica* 61, 821-56.

6. Aryal, G., Ciliberto, F., and B. Leyden (2020). Coordinated Capacity Reduction and Public Communication in the Airline Industry.

7. Assad, S., R. Clark, D. Ershov and L. Xu (2020), Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market. CESifo Working Paper No. 8521, Available at SSRN: https://ssrn.com/abstract=3682021

8. A2i Systems (2020), https://www.a2isystems.com/a2i-systems/.

9. Autorité de la Concurrence and Bundeskartellamt (2019). Algorithms and Competition.

10. Boehnke, J. (2017), Pricing Strategies, Competition, and Consumer Welfare: Evidence from the German and Austrian Retail Gasoline Market, working paper.

11. Borenstein, S. and A. Shepard (1996). Dynamic Pricing in Retail Gasoline Markets, *The Rand Journal of Economics* 27, 429-451.

12. Boswijk, P., J.H. Bun., and M. Schinkel (2018). Cartel Dating, *Journal of Applied Econometrics* 34, 26-42.

13. Brown, Z. and A. MacKay (2020). Competition in Pricing Algorithms.

14. Bundeskartellamnt (2020). https://www.bundeskartellamt.de/EN/AboutUs/Bundeskartellamt/legislation/legislation_node.html.

15. Byrne, D. and N. de Roos (2019). Learning to Coordinate: A Study in Retail Gasoline, *American Economic Review* 109, 591-619.

16. Cabral, L., N.S. Durr, D. Schober, and O. Woll (2018). Learning Collusion: Theory and Evidence from the Shell Price Matching Guarantee.

17. Calvano, E., G. Calzolari, V. Denicolo, and S. Pastorello (2020). Artificial Intelligence, Algorithmic Pricing and Collusion, *American Economic Review* 110, 3267-3297.

18. Chen, L., A. Mislove, and C. Wilson (2016). An empirical analysis of algorithmic pricing on Amazon marketplace. *In Proceedings of the 25th international conference on world wide web.* International World Wide Web Conference Steering Committee.

19. Clark, R. and J-F. Houde. (2013). Collusion with asymmetric retailers: Evidence from a gasoline price-fixing case, *American Economic Journal: Microeconomics* 5, 97-123.

20. Clark, R. and J-F. Houde. (2014). The Effect of Explicit Communication on pricing: Evidence from the Collapse of a Gasoline Cartel, *The Journal of Industrial Economics* 62, 191-228.

21. Clark, R., I. Horstmann and J-F. Houde (2020). Two-sided hub-and-spoke collusion: Evidence from the grocery supply chain. Mimeo.

22. Competition Bureau (2018). Big Data and Innovation: Implications for Competition Policy in Canada, Discussion paper.

23. Cooper, W.L., T. Homem-de-Mello and A.J. Kleywegt (2015). Learning and pricing with models that do not explicitly incorporate competition. *Operations research* 63, 86-103.

24. Crede, C.J. (2019). A structural break cartel screen for dating and detecting collusion. *Review of Industrial Organization* 54, 543-574.

25. Decarolis, F. and G. Rovigatti. (2019). From Mad Men to Maths Men: Concentration and Buyer Power in Online Advertising. Mimeo.

26. Derakhshan, A., F. Hammer, F., & H.H. Lund (2006). Adapting playgrounds for children's play using ambient playware. In 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 5625-5630). IEEE.

27. Derakhshan, A., F. Hammer and Y. Demazeau (2016). PriceCast Fuel: Agent Based Fuel Pricing. In International Conference on Practical Applications of Agents and Multi-Agent Systems (pp. 247-250). Springer, Cham.

28. Dewenter, R., Heimeshoff, U., and Lüth, H. (2017). The impact of the market transparency unit for fuels on gasoline prices in Germany. Applied Economics Letters, 24(5), 302-305.

29. Dewenter, R. and U. Schwalbe (2016). An Empirical Analysis of Price Guarantees: The Case of the German Petrol Market.

30. Dubé, J.P. and S. Misra (2019). Personalized Pricing and Customer Welfare.

31. Erutku, C. and V. Hildebrand (2010). Conspiracy at the pump. *Journal of Law & Economics* 53, 223-237.

32. Ezrachi, A. and M. Stucke (2015). Artificial Intelligence and Collusion: When Computers Inhibit Competition, University of Tennessee, Legal Studies Research Paper Series #267, 2015.

33. Ezrachi, A. and M. Stucke (2016). How Pricing Bots Could Form Cartels and Make Things More Expensive, *Harvard Business Review*, October 2016.

34. Ezrachi, A. and M. Stucke (2016). *Virtual Competition*, Harvard University Press, Harvard, MA.

35. Ezrachi, A. and M. Stucke (2017). Two Artificial Neural Networks Meet in an Online Hub and Change the Future (of Competition, Market Dynamics and Society), *SSRN Electronic Journal*.

36. Gelman, M., Y. Gorodnichenko, S. Kariv, D. Koustas, M. Shapiro, D. Silverman, and S. Tadelis (2016). The response of consumer spending to changes in gasoline prices (No. w22969). National Bureau of Economic Research.

37. Guerra-Hernández, A., A. El Fallah-Seghrouchni, & H. Soldano (2004). Learning in BDI multi-agent systems. In International Workshop on Computational Logic in Multi-Agent Systems (pp. 218-233). Springer, Berlin, Heidelberg.

38. Hammer, F., A. Derakhshan, Y. Demazeau, and H. H. Lund(2006). A Multi-Agent Approach to Social Human Behaviour in Children's Play. In 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (pp. 403-406). IEEE.

39. Hansen, K., K. Misra, and M. Pai (2020). Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms

40. Harrington, J. (2008). Detecting Cartels. in Handbook of Antitrust Economics (P. Buccirossi Ed.), MIT Press, Cambridge, MA.

41. Harrington, J. (2018a). Developing Competition Law for Collusion y Autonomous Artificial Agents. *Journal of Competition Law & Economics* 14, 331-363.

42. Harrington, J. (2018b). How do hub-and-spoke cartels operate? Lessons from nine case studies. Mimeo.

43. Hastings, J. (2004). Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California. *American Economic Review* 94, 317-328.

44. Igami, M. and T. Sugaya (2019). Measuring the incentive to collude: The vitamin cartels, 1990-1999. Available at SSRN 2889837.

45. Kaymak, L. and U. Waltman (2006). A Theoretical Analysis of Cooperative Behaviour in Multi-Agent Q-Learning, Erasmus Research Institute of Management (ERIM), ERS-2006-006-LIS.

46. Kaymak, L and U. Waltman (2008). Q-learning agents in a Cournot oligopoly model. *Journal of Economic Dynamics and Control* 32, 3275-3293.

47. Kehoe, P., B. Larsen, and E. Pastorino (2018). Dynamic Competition in the Era of Big data

48. Klein, T (2019). Autonomous Algorithmic Collusion: Q-Learning Under Sequential Pricing. Working paper.

49. Kühn K-U. and S. Tadelis (2018). The Economics of Algorithmic Pricing: Is collusion really inevitable?, mimeo.

50. Luco, F. (2019) Who Benefits from Information Disclosure? The Case of Retail Gasoline. *American Economic Journal: Microeconomics* 11, 277-305

51. Mehra, S.K. (2015). Antitrust and the Robo-Seller: Competition in the Time of Algorithms, *Minnesota Law Review*, 100.

52. Miklós-Thal, J. and C. Tucker (2019). Collusion by Algorithm: Does Better Demand Prediction Facilitate Coordination Between Sellers? *Management Science* 65, 1552-1561.

53. Milgrom, P., & Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica* 58, 1255-1277.

54. Monpolkommission (2018), Algorithms and Collusion, *Biennial Report of the Monopolies Commission.*

55. Monpolkommission (2020), https://www.monopolkommission.de/de/monopolkommission/ aufgaben.html.

56. Montag, F and C. Winter (2019) Price Transparency Against Market Power, mimeo.

57. Moriyama, K. (2007). Utility Based Q-learning to Maintain Cooperation in Prisoner's Dilemma Games. In 2007 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'07) (pp. 146-152). IEEE.

58. Moriyama, K. (2008). Learning-rate adjusting q-learning for prisoner's dilemma games. In 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (Vol. 2, pp. 322-325). IEEE

59. O'Connor, J. and N. E. Wilson (2019). Reduced Demand Uncertainty and the Sustainability of Collusion: How AI Could Affect Competition.

60. OECD (2017). Algorithms and Collusion: Competition Policy in the Digital Age.

61. Quandt, R. (1960). Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes, *Journal of the American Statistical Association* 55, 324-330.

62. Salcedo, B (2015). Pricing Algorithms and Tacit Collusion, Working paper.

63. Shiller, B. (2014). First-Degree Price Discrimination Using Big Data

64. Shiller, B. and J. Waldfogel (2011). Music for a song: an empirical look at uniform pricing and its alternatives. *The Journal of Industrial Economics* 59, 630-660.

65. Shoham, Y., Powers, R., & Grenager, T. (2003). Multi-agent reinforcement learning: a critical survey. mimeo.

66. Slade, M. (1987). Interfirm Rivalry in a Repeated Game: An Empirical Test of Tacit Collusion. *Journal of Industrial Economics* 35, 499-516.

67. Slade, M. (1992). Vancouver's Gasoline-Price Wars: An Empirical Exercise in Uncovering Supergame Strategies. *The Review of Economic Studies* 59, 257-276.

68. Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities. *Management Science* 54, 2024-2038.

69. United Kingdom Digital Competition Expert Panel (2019), Unlocking Digital Competition.

70. Varian, H. (2018). Artificial intelligence, economics, and industrial organization (No. w24839). National Bureau of Economic Research.

71. Wang, Z. (2008). Collusive Communication and Pricing Coordination in a Retail Gasoline Market. *Review of Industrial Organization* 32, 35-52.

72. Wang, Z. (2009). (Mixed) Strategy in Oligopoly Pricing: Evidence from Gasoline Price Cycles Before and Under a Timing Regulation. *Journal of Political Economy* 117, 987-1030.

A Background

ZKZ 83988

Der Bezug der Zeitschrift TANKSTOP ist kostenfrei

TANK STOP

Das eft-Fachmagazin für Tankstelle und Mittelstand

Ausgabe 06 | 2017

TITELTHEMA
Allguth und Linde eröffnen Wasserstofftankstelle in München

TANKWORT
Gewinnpotential von über 10 000 Euro p.a. an freien Tankstellen

KLARTEXT
„Waffen gehören in die Hände von Profis"

PRICECAST FUEL©

Benzin

PriceCast Fuel

– damit Sie beim Preis immer richtig liegen!

Dynamic Pricing am Point of Sale

Wann ist der richtige Zeitpunkt für eine Preisjustierung, und was ist dann der optimale Preis?

Das Pricing-System „PriceCast Fuel" gibt für jeden Kraftstoff an jedem Standort und zu jedem Zeitpunkt den Verkaufspreis vor, bei dem Sie am besten Ihre Mengen- und Margenziele erreichen. Sie legen die Strategien für Ihre Tankstellen fest, und „PriceCast Fuel" bestimmt zu jedem Zeitpunkt 24 Stunden im Voraus die optimalen Preise, entsprechend Ihren Budgets und Zielen.

„PriceCast Fuel" ermittelt die Preise durch Anwendung von künstlicher Intelligenz sowohl beruhend auf Vergangenheitsdaten als auch auf aktuelle Transaktionsdaten an Ihren einzelnen Tankstellen. Auch lokale Wettbewerbspreise werden berücksichtigt, sofern sie das Kaufverhalten Ihrer Kunden beeinflussen. „PriceCast Fuel" analysiert mit seinem speziell entwickelten, künstlichen neuronalen Netz die Preisempfindlichkeit Ihrer Kunden zu unterschiedlichen Zeiten und aufgrund variierender Verkehrssituationen und Wettbewerbsbedingungen. So werden Muster und Zusammenhänge erkannt, aus denen fortgeschrittene Algorithmen die Preise errechnen, bei denen Sie ein optimales Ergebnis erreichen hinsichtlich Absatzmenge und Gewinn.

Testen Sie PriceCast Fuel

Je nach Ihrer Markenpositionierung und Ihren Geschäftsziele können Sie mit „PriceCast Fuel" unterschiedliche Strategien einstellen. Am übersichtlichen Kontrollzentrum (Dashboard) des Systems haben Sie jederzeit alle Ihre Tankstellen im Blick: am Ampelsystem – grün, gelb, rot – erkennen Sie sofort, ob Ihre Tankstellen ihre Kennzahlenvorgaben erreichen, und welche Ihre Aufmerksamkeit benötigen.

Buchen Sie unser 14-wöchiges Proof-of-Concept Programm, bei dem Sie an einer Auswahl Ihrer Tankstellen den Effekt von „PriceCast Fuel" überprüfen und nachvollziehen können.

Der Zugewinn an Marge während des Programms übertrifft oft die Kosten für das Proof-of-Concept. Sie werden erleben, wie das tägliche Pricing mit „PriceCast Fuel" transparenter und effizienter wird, denn Sie bekommen mit „PriceCast Fuel" einen besseren Überblick, eine bessere Steuerung und Sie benötigen weniger Zeit. Das Proof-of-Concept Programm bieten wir Ihnen als einfachen Cloud-Service an, d.h. ohne Änderungen an Ihrer bestehenden IT-Infrastruktur.

Für mehr Informationen rufen Sie uns an oder senden Sie uns eine E-Mail.

Copenhagen
a2i systems A/S
International House
Center Boulevard 5
DK-2300 Copenhagen S
Dänemark

Odense
a2i systems A/S
Blangstedgårdsvej 8
DK-5220 Odense
Dänemark

München
a2i systems A/S
Am Schloßberg 14 C
82547 Eurasburg
Deutschland

Ihr Ansprechpartner:
Torben Lauridsen
tbl@a2isystems.com
Tel +49 160 9770 6739

systems

Intelligente Preisstellung leichtgemacht

„Daten sind das neue Öl", sagen Wirtschaftsexperten, und Datenwissenschaftler aus vielen Branchen fördern mit Hilfe von künstlicher Intelligenz (KI) und neuronalen Netzen immer mehr Wissen über Zusammenhänge aus großen Datenmengen (Big Data) zutage. Dieses Wissen nutzen führende Unternehmen für ihren wirtschaftlichen Erfolg, indem es ihnen zu schnelleren und besseren Entscheidungen verhilft – z. B. beim Pricing im Einzelhandel.

„a2i systems A/S" hat den Wert von dynamischem Pricing von Kraftstoffen mit Hilfe von neuronalen Netzen schon 2011 als Erster bei „OK Benzin" in Dänemark durch eine 40 Tankstellen umfassende Untersuchung dokumentiert. Seitdem nutzt „OK Benzin" dieses Verfahren überaus erfolgreich an allen ihren 700 Tankstellen und zwar mithilfe des innovativen Pricing-Systems „PriceCast Fuel®" von „a2i". Weitere Tankstellenbetreiber in Skandinavien und Mitteleuropa sind nach und nach den gleichen Weg gegangen und haben das Pricing in Teilen Europas bereits verändert und neue Maßstäbe für das strategische Pricing mit Wettbewerbsautomatik gesetzt.

„PriceCast Fuel®" unterscheidet sich funktionell deutlich von herkömmlichen Pricing-Systemen, indem es das Kundenverhalten abbildet und danach den optimalen Preis bestimmt, statt nur auf die Preise der Wettbewerber zu schauen und daraufhin zu (re-)agieren.

Das Nachfrageverhalten und die Preisempfindlichkeit der Käufer in Abhängigkeit von Tageszeit, Wochentagen, Feiertagen, Ferienzeiten sowie auch aktuellen Preisen der Wettbewerber werden von „PriceCast Fuel®" stündlich analysiert. Fortgeschrittene Algorithmen berechnen daraus optimale Preisempfehlungen für die jeweils nächsten 24 Stunden. Die Schnelligkeit und die Genauigkeit des Systems, was Timing und Preisgefüge betrifft, ist dabei unübertroffen im Vergleich zu bisherigen regelbasierten Methoden.

Je nach Ihrer Positionierung am Markt und Ihren Geschäftszielen können Sie mit „PriceCast Fuel®" unterschiedliche Strategien einstellen. Am übersichtlichen Kontrollzentrum (Dashboard) des Systems haben Sie jederzeit alle Ihre Tankstellen im Blick: am Ampelsystem – grün, gelb, rot – erkennen Sie sofort, ob Ihre Tankstellen ihre Kennzahlenvorgaben erreichen und welche Ihre Aufmerksamkeit benötigen.

Seit vergangenem Sommer bietet „a2i systems A/S" nun auch unabhängigen mittelständischen Tankstellenbetreibern in Deutschland das hoch innovative Pricing-System als einfach zu nutzende Cloud-Lösung an. Die Beratung, die Anbindung und der Support des Systems werden von „WEAT Electronic Datenservice GmbH" geleistet. Damit ist ein branchenbekanntes und vertrautes Unternehmen als Distributionspartner an Bord, das mit technischer Kompetenz, bundesweiter Reichweite und sehr reaktionsfähigem, deutschsprachigem Support glänzt.

Mit einem 14-wöchigen Proof-of-Concept Programm lässt sich der Effekt und der Nutzen von „PriceCast Fuel®" an einer Auswahl von Tankstellen überprüfen und nachvollziehen. In der Regel ist der Zugewinn an Marge während des Programms sogar höher als der Preis für das Proof-of-Concept Programm. In jedem Fall kann man damit erfahren, wie mit „PriceCast Fuel®" das tägliche Pricing einfacher, transparenter und effizienter wird.

♦ a2i systems A/S

Torben Lauridsen, Business Consultant
Mobil: +49 160 9770 6739
Mail: tbl@a2isystems.com

28 INDUSTRIEPARTNER

Figure A1: December 2017 TANKSTOP Trade Magazine Cover and a2i Advertisements

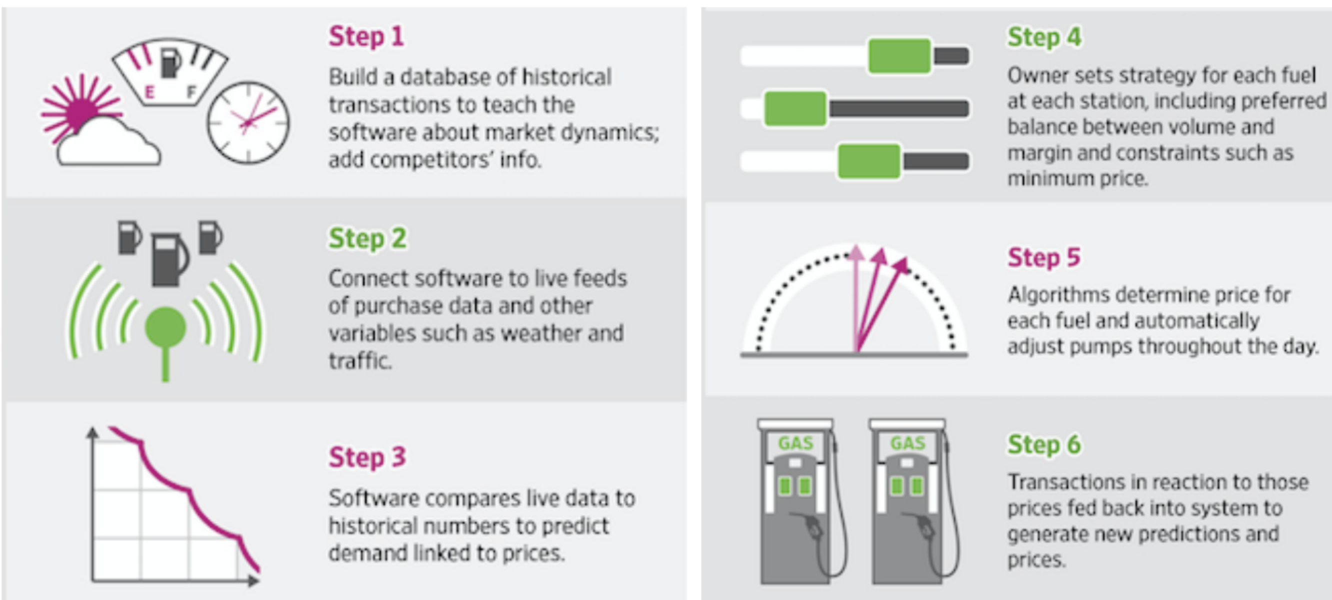


Figure A2: How Algorithms Work (wsj.com)

B QLR Estimation and Results

B.1 QLR Estimation

We estimate the following regression over a range of eligible break periods $\tau_0 \leq \tau \leq \tau_1$ (eligible break periods are measured by week):

$$y_{it} = \alpha_i + \beta_i D_t(\tau) + X_t \gamma_i + \epsilon_{it}, \tag{4}$$

where y_{it} is the variable of interest for station i in time t, $D_t(\tau)$ is a dummy variable equal to 0 if $t < \tau$ and 1 if $t \geq \tau$, and X_t is the crude oil price in time period t, which we use as a control variable. For each regression we test the null hypothesis $H_0 : \beta_i = 0$ and compute the F-statistic $F_i(\tau)$. The QLR statistic is the largest of these F-statistics over the range of eligible break dates:

$$QLR_i = max[F_i(\tau_0), F_i(\tau_0 + 1), ..., F_i(\tau_1)]. \tag{5}$$

The best candidate structural break period for station i is identified as the week τ^* that satisfies $QLR_i = F_i(\tau^*)$.[58] Structural breaks are identified as significant if they exceed a certain critical value.[59] We drop all stations from our data set that do not operate in every week in 2017 (i.e. we keep stations that have 52 observations of a given measure in 2017). We use 30% trimming for our test period, which is standard for QLR testing.[60]

[58] We refer to the QLR statistic as identifying the "best candidate" structural break period because if we look at a test for each time period τ individually, there may be multiple periods in which a structural break would be identified (i.e. has an F-statistic exceeding a certain critical value). The QLR statistic identifies the best candidate break period as it identifies the period with the most significant associated F-statistic.

[59] The distribution of the QLR statistic is non-standard so we cannot use the usual critical values for F-statistics to determine significance. Critical values for QLR statistics are taken from Andrews (2003). Using these values we measure a structural break as significant at the 10% level if $QLR_i \geq 7.12$, at 5% level if $QLR_i \geq 8.68$, and at the 1% level if $QLR_i \geq 12.16$.

[60] We use as our first eligible break date the 15th percentile week in our sample period and as our last eligible break date the 85th percentile week in our sample period.

B.2 Structural Break Test Results

B.2.1 Number of Price Changes

For each station we construct a variable measuring the number of times it changes its price for each date in our sample period. For structural break testing, we aggregate this variable to the weekly level.[61] Out of **13,022 candidate stations, 12,919** experience a significant structural break in the number of price changes at the 5% confidence level. Out of the stations that experience significant breaks, **almost** 50% of the best-candidate breaks occur in the spring and summer of 2017. Figure B1 shows the overall distribution of best-candidate breaks.

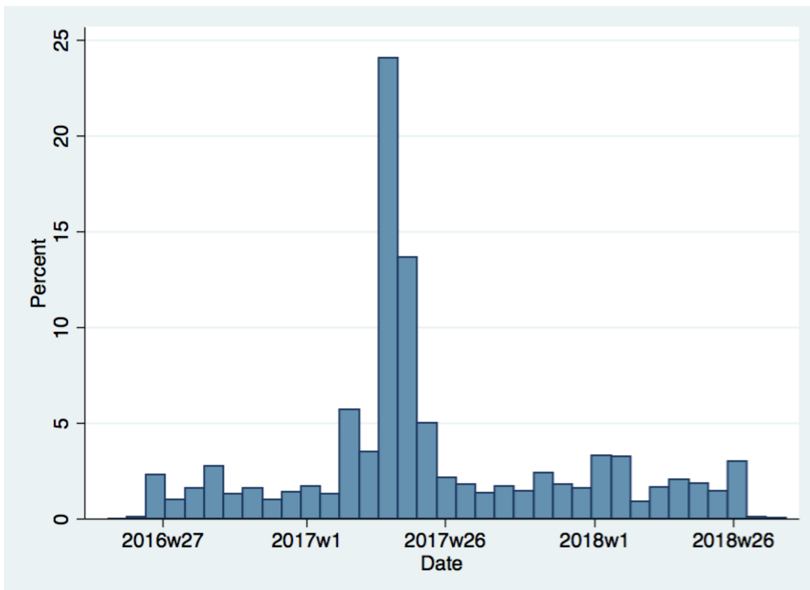


Figure B1: Frequency of Best-Candidate Structural Breaks in Number of Price Changes (12,919 stations included)

We compare the number of price changes before and after best-candidate structural breaks for stations that experienced structural breaks in Table B1. We also include stations that did not experience structural breaks. Adopting stations change their prices more frequently than non-adopting stations, suggesting that our breaks do manage to pick out large changes in pricing strategy. On average, a station that experienced a structural break changes their prices **6.2 times a day before the break and 9.1** after the best-candidate break. There is also a general rightward shift in the

[61] Any stations that do not have a weekly observation for average number of price changes in every week of 2017 are dropped. See more details in the Data Appendix.

distribution of the number of daily price changes after the break.[62] Stations that do not experience breaks experienced **4.3** changes per day.

Table B1: Daily Number of Price Changes

	Mean	Std. Dev.
Post Structural Break Stations	9.1	2.5
Pre Structural Break Stations	6.2	1.6
No Structural Break Stations	4.3	2.2

B.2.2 Rival Response Time

We define a rival for station i as the closest station j that is within a 1KM radius of station i but that belongs to a different brand.[63] Rival response time for station i is calculated as the number of minutes between the time of a price change by rival j and the subsequent price change by station i. If station j changes its price more than once before station i makes a price change, rival response time is taken as the average of the time gaps between each of station j's price changes and station i's subsequent change. When testing for structural breaks in rival response time, we take into account the fact that changes in response time will be mechanically impacted by changes in number of price changes. To identify structural changes separately from this mechanical effect, we control for the number of price changes when running Equation (1) for this measure. Out of **5,646 candidate stations, 5,227 experience significant structural breaks. Out of stations with significant breaks (at at least the 5% level), almost 29%** have best-candidate breaks in the summer of 2017. Figure B2 shows the overall distribution of best-candidate breaks.

[62] At the 25th percentile of number of price changes, a station only changes their prices **5.3 times per day before the break but 7.9** times a day after the break.

[63] This reflects the average distance of stations in the data.

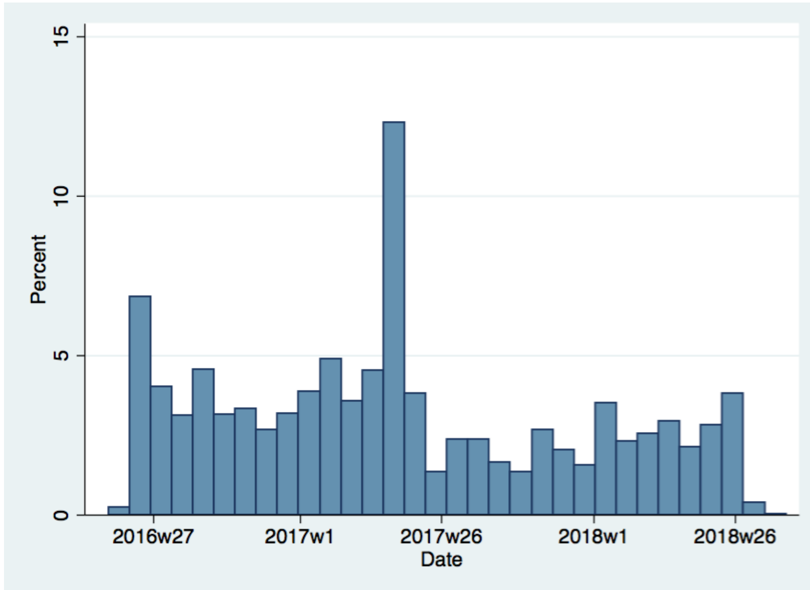


Figure B2: Frequency of Best-Candidate Structural Breaks in Rival Response Time (5,227 stations included)

We compare average rival response times (in minutes) for stations that adopted before and after adoption in Table B2. We find that the structural break captures substantial changes in the measure. On average rival response time decreases from **63 minutes to 54 minutes after the best-candidate break, a drop of about 9%**. There are also decreases at other points in the response time distribution, especially in the right tail.[64] Stations that did not experience a structural break in this measure look more like the stations in the pre-break period, having average response times of over one hour.

Table B2: Rival Response Time (Minutes)

	Mean	Std. Dev.
Post Structural Break Stations	54	24
Pre Structural Break Stations	63	30
No Structural Break Stations	63	31

[64] **At the 75th percentile, response time falls from 74 minutes to 64 minutes. At the 95th percentile, response time falls by 20 minutes.**

B.2.3 Average Size of Price Change

For the average size of price changes, we calculate the average size of price changes made in a day for each station and then average this measure to a weekly level.[65] We look at the distribution of weekly break periods for stations with a QLR statistic that is significant at the 5% level. Out of **12,974 candidate stations, 11,603 experience a statistically significant structural break. Over 20%** of the best-candidate breaks occur in mid-2017. Results are presented in Figure B3. Although there is a spike of stations experiencing best-candidate breaks in average price change size in Spring 2017, there is a large number of breaks in mid-2016 and a number of stations experiencing breaks throughout 2018. The occurrence of breaks in 2016 may be due to prevailing effects of Shell's 2015 price-matching policy, which induced changes in pricing behaviour for some German retail gas stations. In particular, during this time, Shell and ARAL began to interrupt the previously observed price cycles in the market with upward price jumps around midday. Medium and small retail gas brands would follow these increases, although the extent of the midday price jumps for these stations was not as large of those of Shell and ARAL (Cabral et al. 2018).

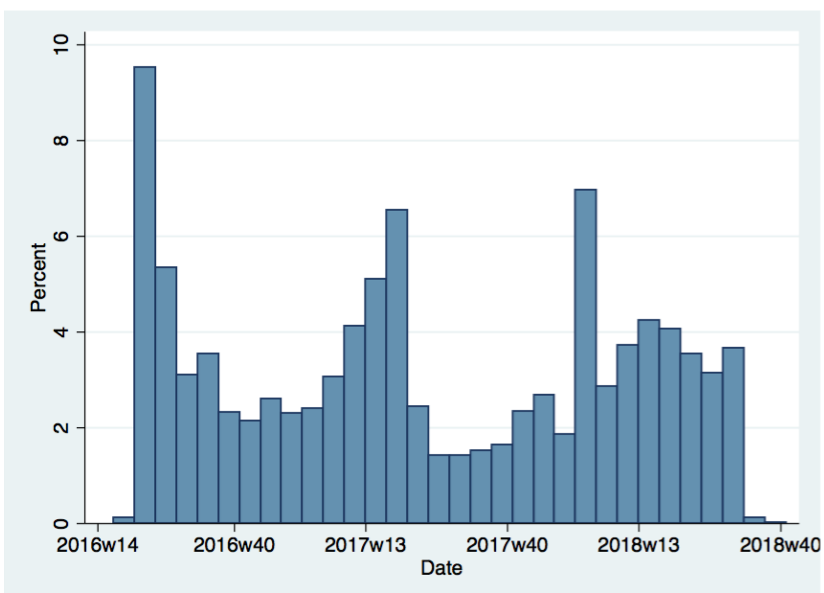


Figure B3: Frequency of Best-Candidate Structural Breaks in Average Size of Price Change (11,603 stations included)

Looking at stations that experience best-candidate breaks in 2016, ARAL and Shell stations make up over 40% of these occurrences. Figure B4 shows the distribution of break periods for Shell and

[65] See more details in the Data Appendix.

ARAL stations in comparisons to all other stations. The figure makes it clear that Shell and ARAL stations drive the observed spike in break frequency in 2016. For all brands, over 20% of stations still experience structural breaks in the spring/summer of 2017. Importantly, we do not use this measure alone to define adoption of algorithmic pricing. As discussed in the main text, we focus on stations that experience best-candidate breaks in at least two measures within a relatively short window of time. We also directly address concerns about Shell in a robustness check, dropping all markets where Shell is present and where the price matching policies would have any effect (see Section 3.1 for additional discussion). Results are qualitatively and quantitatively similar to our main estimates (see Appendix E.1).

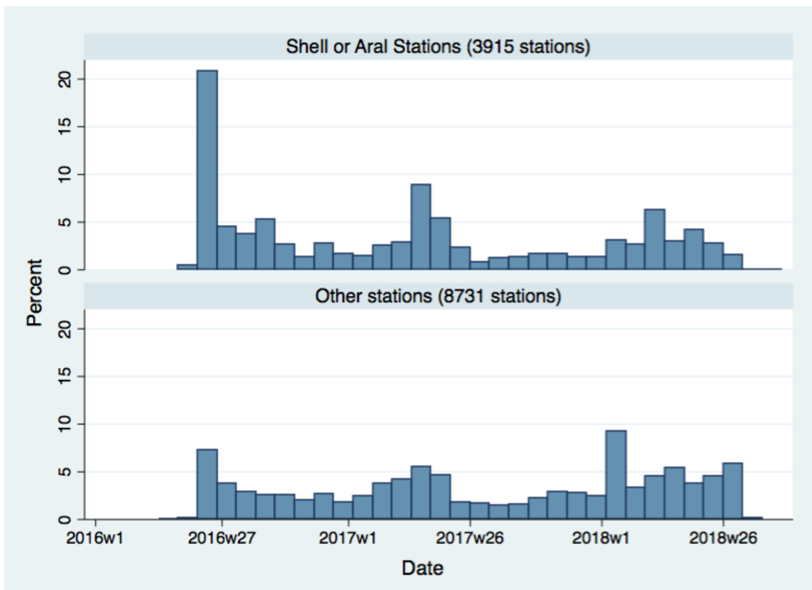


Figure B4: Frequency of Significant Structural Breaks in Average Size of Price Change for Shell and Aral stations vs. other stations)

Table B3 shows differences in average sizes of daily price changes (in cents) before and after the best-candidate breaks. For stations that experience a structural break, the average size of price changes**increases from 2.7 to 2.9 cents. The standard deviation in the average size of price changes falls from 1.0 to 0.7 cents**, suggesting that price fluctuations become more targeted after the structural break.[66]

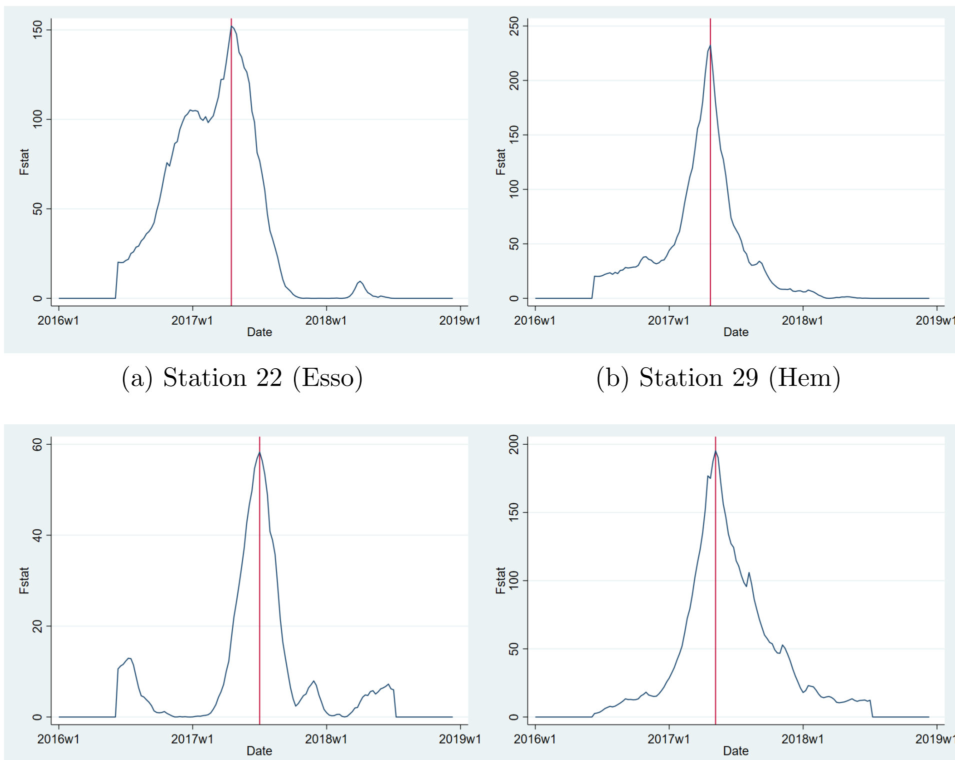
[66]**The average size of price changes also increases at each the 25th, 50th, and 75th percentile. At the 95th percentile, average price change sizes fall from 4.6 cents to 4.1 cents.**

Table B3: Average Daily Price Change Size (cents)

	Mean	Std. Dev.
Post Structural Break Stations	2.9	0.7
Pre Structural Break Stations	2.7	1.0
No Structural Break Stations	2.7	1.0

B.3 Alternative Structural Breaks

We look at the distribution of F-statistics for structural break tests in the number of price changes for stations over the sample period for a few representative stations. We find that generally, stations display a uni-modal distribution in their F-stastistics, meaning we are unlikely to find best-candidate breaks at a significantly different date if we were to, for example, take the second highest F-statistic rather than the maximum. A few examples are shown in B5 of what a typical distribution would look like for a station's F-statistics for structural break tests in the number of price changes.



(a) Station 22 (Esso)

(b) Station 29 (Hem)

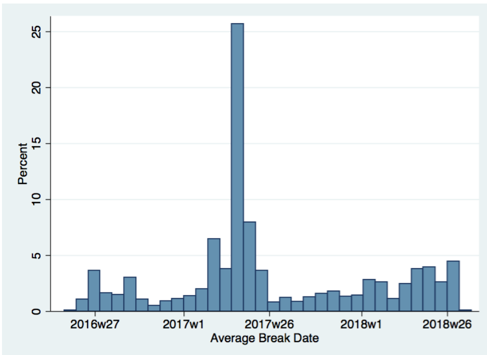
(c) Station 31 (Shell)

(d) Station 163 (ARAL)

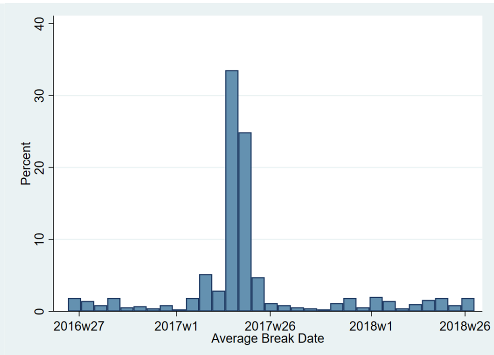
Figure B5: Distribution of F-statistics for Structural Break Tests in Number of Price Changes

To take a more systematic approach to test whether there may be significantly different alternative break dates, for each station, we look at the dates associated with the 2nd highest F-statistics for break tests in the number of price changes. We find that for 75% of stations, these dates are 1 week apart meaning that the next alternative break date would be occur either 1 week before or after the break associated with the highest F-statistic. We find only 10% of stations have difference of 3 or more weeks between the dates associated with the highest and 2nd highest F-statistic.

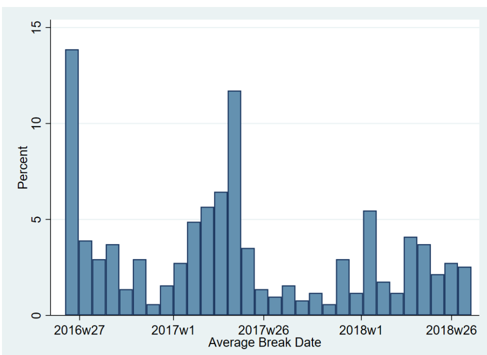
B.4 Distribution of Average Break Dates by Measure Combination



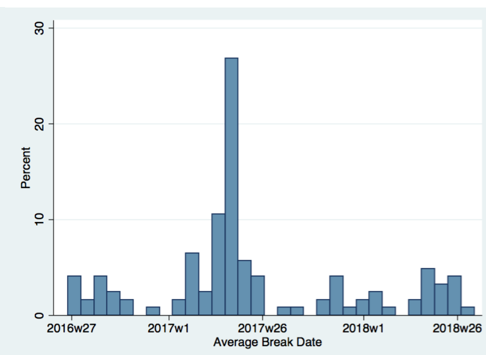
(a) Number of Price Changes and Average Size of Price Change (2,465 stations)



(b) Number of Price Changes and Rival Response Time (695 stations)



(c) Average Price Change Size and Rival Response Time (512 stations)



(d) All 3 Measures (123 stations)

Figure B6: Frequency of Average Break Date for Measures Breaking Within 4 Weeks

Figure B6 shows the distribution of the average break date for each combination of measures, where the average break date is the average year-week between each measure's best-candidate break date. For each measure pair, the largest frequency of average break dates occur in mid-2017. Overall, we see the largest frequency of multiple measure breaks in mid-2017, the suspected period of large scale adoption, suggesting these measures accurately represent changes related to adoption of algorithmic pricing.[67]

[67] We do see some stations that break in both average price change size and rival response time in mid-2016. About **67%** of these stations belong to ARAL and Shell, so it is possible that these breaks may be related to prevailing effects of Shell's 2015 price matching policy (see Section 3.1). We test our sample for robustness by removing all markets where Shell is present in Appendix E.1 and find that it does not affect our main findings.

B.5 Adopter/Non-Adopter Heterogeneity

Table B4: Adopter and Non-Adopter Station Characteristics in 2016

	(1)
Outcome:	Will Station j Adopt AI?
Population Density	0.00003***
	(0.00001)
ln(Population)	0.00443
	(0.03513)
Median Population Age	0.00707***
	(0.00211)
Employment Share	0.09257
	(0.07782)
ln(region GDP)	0.00056
	(0.03241)
N Competitors in ZIP	0.00297*
	(0.00165)
Observations	165,810

Notes: The sample for this regression includes gas station/month observations from January 2016 until December 2016 that are not labelled as AI adopters during this period. The outcome is a dummy variable equal to 1 if the station will eventually be labelled as an adopter in 2017 or 2018, and zero otherwise. Population Density, ln(Population), Median Population Age, Employment Share and ln(regional GDP) are all computed at the NUTS3-year level. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j in month t. We include month fixed effects. Standard errors clustered at the ZIP level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

B.6 Heterogeneity in Structural Breaks/Adoption by Brand Size

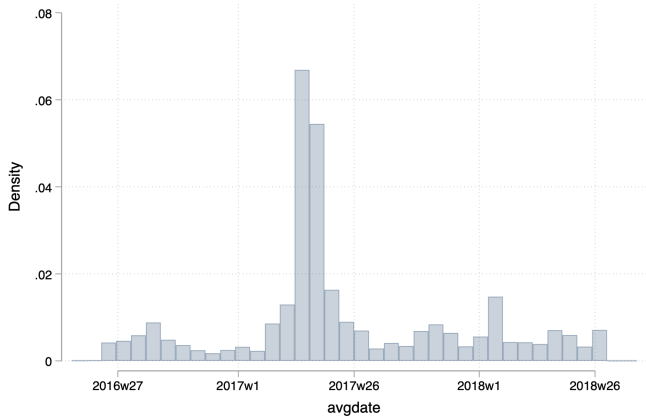
Table B5: Share of Stations Classified as Adopters

	Mean	Std. Dev.	Median
	December 2016		
Top 5 Brands	0.043	0.029	0.030
Other Brands	0.026	0.083	0
	December 2017		
Top 5 Brands	0.208	0.096	0.229
Other Brands	0.135	0.184	0.083
	December 2018		
Top 5 Brands	0.260	0.108	0.282
Other Brands	0.177	0.210	0.125

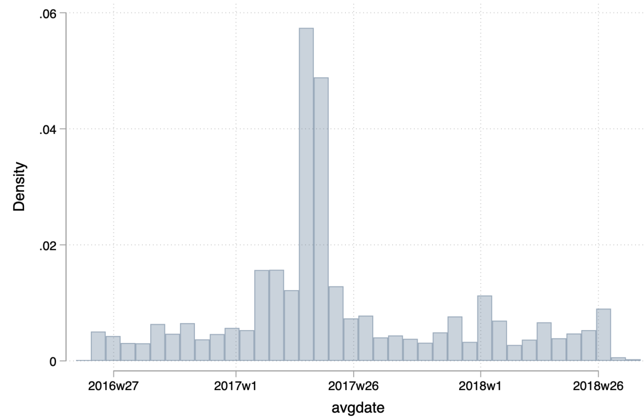
Table B6: Correlates to Brand-Level Adoption Probability

	(1)
Outcome:	Share Brand Adopters
Mean Population Density	0.00003
	(0.00003)
Mean ln(Population)	-0.13789
	(0.11453)
Mean Median Age	0.00667
	(0.00763)
Mean Employment Share	-0.51591
	(0.37037)
Mean ln(region GDP)	0.13839
	(0.10112)
Mean N Competitors in ZIP	0.00147
	(0.00557)
N Brand Stations	0.00003**
	(0.00001)
Observations	6,853

Notes: The sample for this regression includes brand/month observations from January 2016 until December 2018 for brands with two stations or more. The outcome is the share of a brand's stations that are labelled as adopters by month t. Variable "Mean X" is a simple average of variable X across all brand b stations in month t. We include year-month fixed effects. Standard errors clustered at the brand level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

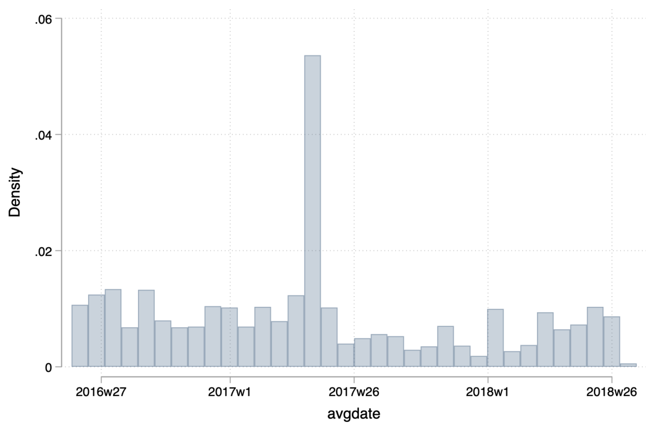


(a) Top 5 Brands

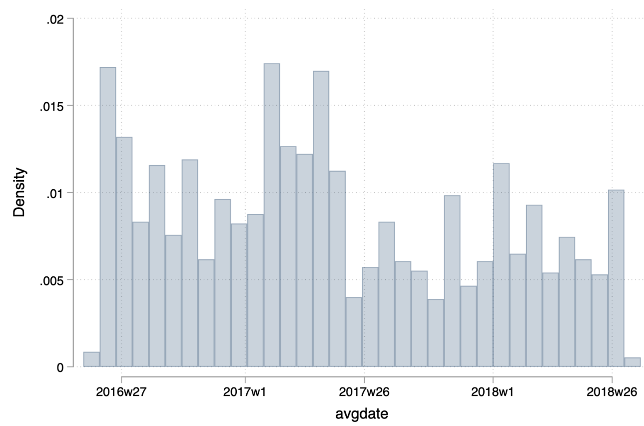


(b) Non-Top 5 Brands

Figure B7: Frequency of Significant Structural Breaks in Number of Daily Price Changes by Brand Size

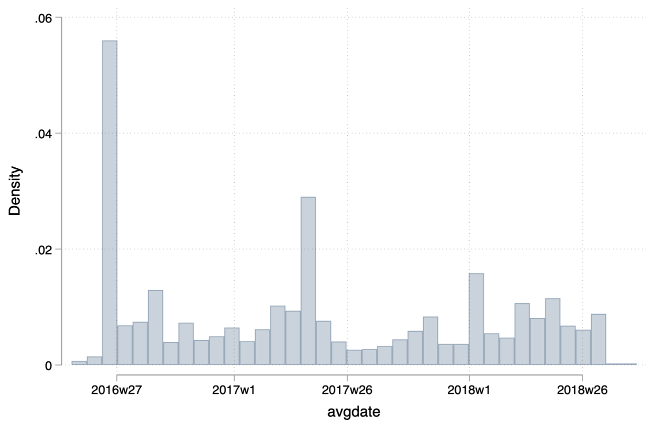


(a) Top 5 Brands

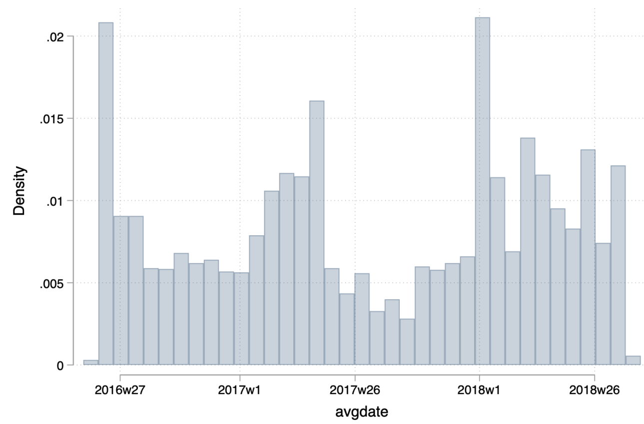


(b) Non-Top 5 Brands

Figure B8: Frequency of Significant Structural Breaks in Rival Response Time by Brand Size



(a) Top 5 Brands



(b) Non-Top 5 Brands

Figure B9: Frequency of Significant Structural Breaks in Average Price Change Size by Brand Size

B.7 Diesel Gas Structural Breaks

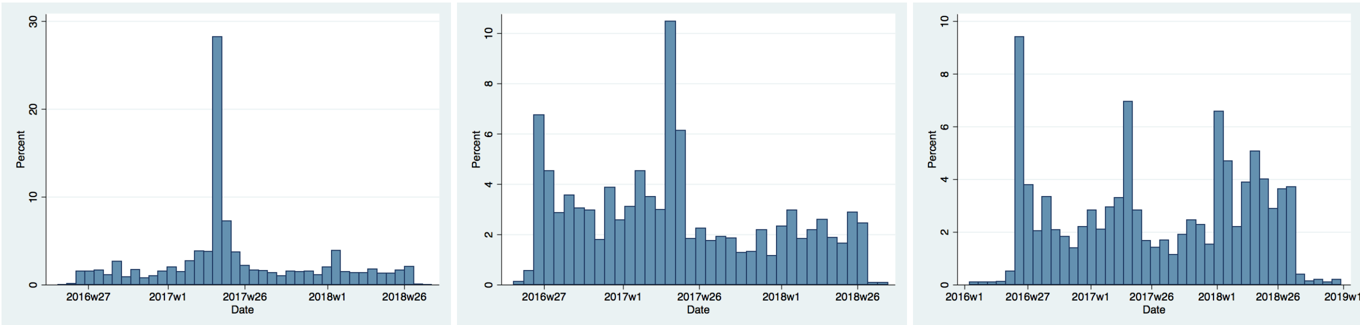


Figure B10: Frequency of Significant Structural Breaks in Number of Price Changes, Rival Response Time, and Average Size of Price Change (Diesel Gas)

C Additional Estimates

Table C1: OLS Station-Level Margin Estimates

Outcome:	(1) Mean Margin	(2) 5th Pctile Margin	(3) 25th Pctile Margin	(4) Median Margin	(5) 75th Pctile Margin	(6) 95th Pctile Margin
Adopter	0.000	0.000**	0.000	0.000	0.000	0.002*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
N Competitors in ZIP	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
Non-Adopter Mean Outcome	0.0824	0.0535	0.0677	0.0763	0.0846	0.206
Station FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
Observations	478,172	478,172	478,172	478,172	478,172	478,172

Notes: Sample is gas station/month observations from January 2016 until December 2018. Margins are computed above wholesale gasoline prices at a regional terminal nearest to station j. Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale gasoline prices. "Xth Percentile Margin" is the Xth percentile of the daily difference of pump price and wholesale gasoline prices for station j in month t. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the number of stations belonging to station i's brand in month t. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors clustered at ZIP level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C2: 1st Stage Results for ZIP Duopoly Markets

Outcome:	(1) One Station Adopted	(2) Both Stations Adopted
IV1	0.627***	0.046
	(0.165)	(0.087)
IV2	-0.173	1.499***
	(0.466)	(0.311)
ZIP FE	YES	YES
Year-Month FE	YES	YES
Annual Regional Demographics	YES	YES
N Brand Stations Controls	YES	YES
Weather Controls	YES	YES
Observations	39,148	39,148

Notes: The sample includes duopoly market/month observations from January 2016 until December 2018. A duopoly market is a ZIP code with two gas stations. "One Station Adopted " is a dummy equal to 1 in month t if one of the two stations in the market experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Both Stations Adopted" is a dummy equal to 1 in month t if *both* stations in the market experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. $IV1$ and $IV2$ use the "share of brand adopters" of the two stations in the market as follows: for market m at time t, $IV1_{mt} = B_{1t}(1 - B_{2t}) + B_{2t}(1 - B_{1t})$, where B_{jt} is the share of brand adopters for station j in this market. Similarly, $IV2_{mt} = B_{1t}B_{2t}$. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors clustered at market level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C3: Rival Adoption Effects

Outcome:	(1) Mean Margin	(2) 5th Pctile Margin	(3) 25th Pctile Margin	(4) Median Margin	(5) 75th Pctile Margin	(6) 95th Pctile Margin
Rival Adopted	0.002	0.005	0.003	0.001	0.002	-0.008
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.018)
IVs	YES	YES	YES	YES	YES	YES
Station FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
Observations	72,647	72,647	72,647	72,647	72,647	72,647

Outcome:	(7) Mean Price	(8) 5th Pctile Price	(9) 25th Pctile Price	(10) Median Price	(11) 75th Pctile Price	(12) 95th Pctile Price
Rival Adopted	0.002	0.002	0.001	0.002	0.001	0.002
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
IVs	YES	YES	YES	YES	YES	YES
Station FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES	YES
Observations	72,652	72,652	72,652	72,652	72,652	72,652

Notes: The sample includes all station/month observations belonging to duopoly markets from January 2016 until December 2018 where zero or one of the duopolists adopted AI. Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale price. The "Xth Percentile Margin" is the "Xh" percentile of the daily difference of pump price and wholesale gasoline prices for station j in month t. "Rival Adopted" is a dummy equal to 1 in month t if the duopoly rival of station j in market m experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. Instruments for a rival's adoption are the "share of brand adopters" of the rival in the market. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors clustered at station level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

D Adoption of Electronic Payments Technology in 1990s

We use annual data from Kent Marketing, a leading survey company in the Canadian gasoline market.[68] It captures annual data from 1991 to 2001 for all retail gasoline stations in seven medium-sized markets in Ontario: Brantford, Cornwall, Guelph, Hamilton, Kingston, St. Catharines and Windsor. The 5 brands with most stations in this data are PetroCanada (98 stations), Esso (84 stations), Shell (61 stations), Sunoco (56 stations) and Pioneer (36 stations). The data includes station characteristics including whether the station accepts "electronic payments."

This is a good benchmark technology for AI adoption. Both could improve station performance as electronic payments allow for a wider set of consumers to purchase gasoline (and larger quantities of gasoline). As for AI, electronic payment companies also have HQ-level deals with retail gasoline brands, but individual station owners had to bear some of the costs of upgrading their equipment. For example, in 1997, Exxon Mobil (Esso's parent company) rolled out the Mobil Speedpass, a contactless electronic payment system. BusinessWeek reported that after the brand-wide rollout, individual Mobil station owners "have to install new pumps costing up to $17,000–minus a $1,000 rebate from Mobil for each pump" (BusinessWeek).

The first appearance of electronic payments at any gas station in the data is in 1993 (the third year of the dataset). Among the five largest brands, no one reached 50% adoption rates of this technology by 2001. The largest share of adopting stations is for Pioneer, where 46% of stations adopted by 2001. Figure B1 shows adoption rates by the top 5 brands (by the number of stations) in this data. It suggests that electronic payment adoption follows a highly staggered pattern. Of the 5 biggest brands, by 1998 (5 years after the technology became available) only two of the brands had *any* adoption. It is also brand specific. Some brands, such as Esso, appear to be continuously upgrading (or supporting the upgrade) of their stations. Stations by other brands, like Pioneer, adopt faster but later. This likely reflects brand-specific strategies.

[68] This is a subset of data used in Clark, Houde and Carranza (2015).

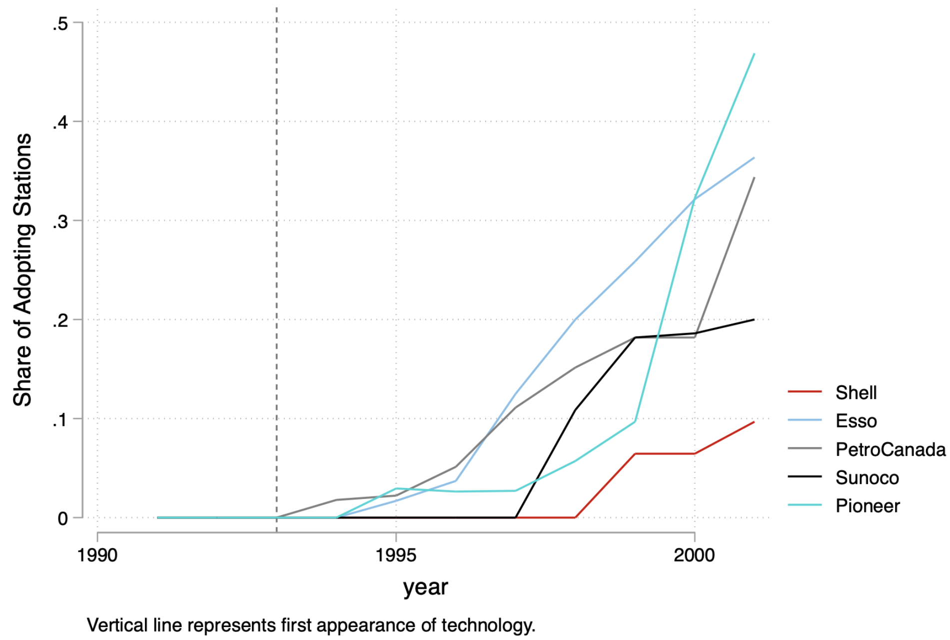


Figure B1: Share of Electronic Payment Adopters Among Top 5 Brands in Canada

E Robustness Checks

E.1 Alternative Estimation Samples

We perform a number of estimation-sample based robustness checks in Table E1. The first two robustness checks deal with concerns about the impact of Shell's 2015 price matching policy (see Section 3.1). The introduction of price matching in 2015 appears to have changed pricing strategies (Cabral et al 2018). These changes in strategies may still be ongoing in 2016. This would confound our results. Shell stations may be mistakenly labelled as algorithmic pricing software adopters. Appendix B.2 shows some evidence that Shell stations have a different distribution of structural breaks for one of our measures (average size of price changes) than other brands, with a large number of stations experiencing breaks in early 2016.[69] Shell stations and their competitors may also set higher prices due to the the price matching guarantee rather than due to the adoption of algorithmic pricing software.

[69]See Figure B4 and Appendix B.2 for additional discussion about structural breaks in the average size of price changes. Our other two structural break measures (average number of price changes and rival response time) do not have the same differences between Shell and other brands.

Columns (1) and (3) in Table E1 deals with this concern by dropping all observations belonging to ZIP codes where the price matching guarantees would be relevant. This includes all Shell stations and stations that are in the same ZIP codes as Shell stations. Results from this sample are quantitatively and qualitatively similar to the main estimates. Even without including any markets where Shell price matching guarantees would have an effect, we find that adoption of algorithmic pricing software increases average margins above wholesale prices by 0.9 cents. Column (3) drops all observations from 2016 (where the Shell effects would be most prominent). Results here are qualitatively similar to the main results but larger in magnitude. This is because it takes time for stations to increase their margins after adoption (see discussion in Section 8).

Column (2) deals with potential concerns about Aral stations. In Figure 3 we show that Aral seems to be a very early adopter of algorithmic pricing software, with a substantial number of stations adopting by 2017. There may be concerns that this is spurious adoption or that this finding (and Aral's adoption) is driven by some brand-specific measurement errors. We address such concerns by removing all observations belonging to Aral stations. Results from this alternative sample are similar to baseline results.

We perform additional robustness checks to address concerns that our main results are driven by entry and exit of stations from the sample - either through the entry of high-quality and high-margin adopters, or through the exit of weak non-adopting stations. In Column (4) we look at a balanced sample of stations. We only include stations that are present in every month of the three year sample period. Results are qualitatively and quantitatively similar to our main estimates. But even these stations can be affected by entry and exit of other stations in their market. In Column (5), we look at a balanced sample of stations and markets that do not change over time. We only include stations that are present in every month of the three year sample period *and* we drop every market where the number of stations changes over time. Results from this subsample are also qualitatively and quantitatively similar to our main results.

E.2 Alternative Market Definitions

There are many possible geographic definitions of "markets." A commonly used definition takes advantage of existing geographic designations such as Census tracts, DMAs, or ZIP codes. In our main results, we define markets based on ZIP code (Tables 4 and 5). Another commonly used definition in the literature looks at the direct distance between stations. Table E2 provides estimates of regressions similar to Table 4 but using the following definition of a monopoly: a station that has

Table E1: Sample Robustness Checks

	(1)	(2)	(3)	(4)	(5)
Sample:	No Shell Markets	No Aral Stations	Dropping 2016 Data	Balanced Sample	Market-Level Balanced Sample
Outcome:	Mean Margin	Mean Margin	Mean Margin	Mean Margin	Mean Margin
	2SLS	2SLS	2SLS	2SLS	2SLS
Adopter	0.009***	0.006***	0.019***	0.008***	0.008***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
N Competitors in ZIP	-0.002***	-0.002***	-0.003***	-0.002***	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Station FE	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES	YES
Observations	200,105	162,199	255,567	138,243	89,972

Notes: All samples include only stations that are in ZIP codes with more than one competitor. Sample in Column (1) includes gas station/month observations from January 2016 until December 2018 that do not belong to a market where a station by a Shell brand is present. Sample in Column (2) includes gas station/month observations from January 2016 until December 2018 that do not belong Aral. Sample in Column (3) includes gas station/month observations from January 2017 until December 2018 (dropping 2016 data). Column (4) includes all gas station/month observations belonging to gas stations that are present in every month of the sample. Column (5) includes all gas station/month observations belonging to gas stations that are present in every month of the sample and are in markets where the number of stations does not vary across the sample period. Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale gasoline prices. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Share Brand Adopters" is the excluded instrument used in the 2SLS regression. It is equal to the share of stations that belong to the brand of station j that adopted by period t. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors are clustered at ZIP-code level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

no competitors within a 1km radius. Non-monopoly stations then are those that have one ore more competitors within a 1km radius. Using this alternative definition yields results that are qualitatively similar to the results in Table 4 but there are differences. In our baseline results we find that only non-monopoly stations increase their prices and margins after adoption. Here, we find that that the mean margins and prices of 1km monopoly stations increase after adoption. However, the magnitudes of the increases are substantially smaller than for 1km non-monopolists. 1km non-monopolists increase their margins by 30% more than 1km monopolists, and increase their prices by over 40% more. The monopolists' price increases are also only statistically significant at the 90% confidence level.

Differences in estimates between Tables 4 and E2 occur because the two market definitions label different stations as "monopolists." While a 1km definition does not vary across different regions, rural area ZIPs are larger than urban area ZIPs. The ZIP-code definition is more conservative. In rural areas, there are many stations that do not have a competitor within 1km (on the same intersection, or an intersection away), but that do have a competitor somewhere nearby (in the ZIP

Table E2: 2SLS Station-Level Results by 1KM Market Structure

	(1)	(2)	(3)	(4)
Outcome:	Mean Margin	Mean Price	Mean Margin	Mean Price
Sample:	Monopoly Stations		Non-Monopoly Stations	
Adopter	0.007**	0.005*	0.009***	0.007***
	(0.003)	(0.003)	(0.002)	(0.002)
N Competitors in ZIP	-0.001*	-0.002***	-0.002***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Station FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES
N Brand Stations Control	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES
Observations	184,278	184,278	246,667	246,667

Notes: Sample is gas station/month observations from January 2016 until December 2018, split up into two subsamples: one subsample only includes stations that have no competitors within a 1KM radius. The other subsample includes only stations that have one or more competitors within a 1KM radius. Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale gasoline prices. Mean Price is the average pump price for station j in month t. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Share Brand Adopters" is the excluded instrument used in the 2SLS regression. It is equal to the share of stations that belong to the brand of station j that adopted by period t. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors are clustered at ZIP level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

code). Table E3 shows a comparison of the number of stations that are labelled monopolists. Using the 1km definition over 6,000 stations are classified as monopolists, whereas only 2,300 stations are when using the ZIP definition. Only 1,800 of those stations overlap, meaning that many of the stations the 1km definition classifies as "monopolists" have some competitors nearby (perhaps 1.5 or 2km away). If there is some competition at ranges beyond 1km, this definition is too lax and would over-state effects for 1KM monopolist stations.

Table E3: Monopoly and Duopoly Market Definition

N ZIP Monopoly Stations	2,323	N ZIP Duopoly Stations	3,093
N 1km Monopoly Stations	6,072	N 1km Duopoly Stations	3,800
N Overlap	1,857	N Overlap	1,126

To a lesser extent, this is also the case for duopoly markets. Approximately 3,000 stations are classified as belonging to a duopoly market based on their ZIP codes. The alternative definition based on a 1KM radius around each station defines a duopoly market as two stations that are within 1KM of one another and that have no other stations within 1KM. 3,800 stations are labelled as belonging to a 1KM-radius duopoly market. Only 1,100 stations belong to a duopoly market according to both definitions. Table E4 replicates some of the regressions in Table 5 using this definition of duopoly markets. Mean margin and price effects are qualitatively and quantitatively similar to the ZIP code definition.

Table E4: 1km Duopoly Market Results

Outcome:	(1) Mean Mkt Margin OLS	(2) Mean Mkt Margin 2SLS	(3) Mean Mkt Price 2SLS
One Station Adopted	-0.000	-0.010	-0.015*
	(0.001)	(0.007)	(0.008)
Both Stations Adopted	0.000	0.036**	0.043**
	(0.001)	(0.015)	(0.018)
IVs	NO	YES	YES
Market FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Annual Regional Demographics	YES	YES	YES
N Brand Stations Controls	YES	YES	YES
Weather Controls	YES	YES	YES
Observations	37,812	37,812	37,812

Notes: The sample includes duopoly market/month observations from January 2016 until December 2018. A duopoly market is defined as two stations that are within 1km of each other and have no other stations within 1km. Outcome variable Mean Market Margin is the average of mean market daily differences of pump prices for stations in market m in month t from wholesale price. "One Station Adopted " is a dummy equal to 1 in month t if one of the two stations in the market experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Both Stations Adopted" is a dummy equal to 1 in month t if *both* stations in the market experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. Regressions in Columns (2) and (3) instrument for adoption using the "share of brand adopters" of the two stations in the market. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors clustered at market level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

This is likely because gasoline stations compete closely with their nearest rivals. Our results in Section 8 suggest that algorithmic adoption by both stations in a duopoly market increases margins by reducing competition. This would also be the case in markets where there are two stations within 1KM of one another and other stations further away.[70] The two nearby stations compete more with one another than with stations that are further away. If both nearby competitors adopt, they will be able to compete less aggressively and increase margins. In that sense, even though the 1km duopoly market definition may be too lax for many (particularly rural) markets, it confirms our baseline

[70]For example, consider a ZIP code with two stations within 500m of each other, and two other stations 3-5KM away. This is not a duopoly ZIP market, but the two stations within 500m of each other would constitute a 1KM duopoly market.

results suggesting that what we find is an effect of AI adoption on competition. In this specification, we also find that the adoption of one competitor reduces mean market level prices and margins. This may be a function of the composition of duopoly markets under this definition.

E.3 Alternative Adoption Definitions

Tables E5 and E6 replicate Column (1) from Table 4 and Column (1) from Table 5 using alternative definitions of AI adoption. We consider three alternative definitions. Our baseline definition classifies a station as an adopter if, within a period of 4 weeks, it experienced a structural break in at least two out of three measures - number of price changes per day, average size of price changes per day and the speed of response to a rival's price change. A potential concern with this definition is that not all stations have a rival within 1km. Such stations could then be less likely be defined as adopters. We address this concern by using only two of the three measures to define an adopter and an adoption date: the number of price changes per day and the average size of price changes. Under this definition, a station is labelled as an adopter if it experiences a structural break in both of these measures within a period of 4 weeks. This is the definition we use in Column (1) of Tables E5 and E6. In Column (2), we label a station as an adopter if they experienced a structural break in any two out of three measures but within a period of 2 weeks. This is a stricter requirement for being labelled as an adopter. Results for these definitions are qualitatively and quantitatively similar to baseline results.

We also consider a stricter adoption measure that involves a station experiencing multiple structural breaks in different fuel types. Under this definition, a station has to experience structural breaks in at least two out of our three adoption measures in **both** E5 and Diesel within a period of 4 weeks. As market structure and demand for E5 and Diesel are fundamentally different, if a station experiences changes in pricing strategy in both fuel types at the same time, it is highly likely to be driven by the adoption of new pricing software. We take the adoption date to be the average between the adoption date of E5 and the Diesel adoption date. Column (3) in each of Tables E5 and E6 present results using this definition of adoption. We find that the results are qualitatively the same as the baseline results at the station level. At the market level, the results are qualitatively similar, but standard errors are larger. We have a much smaller sample at the duopoly-market level than at the station level and with stricter definitions of adopters we lose power. This is particularly the case for the last alternative definition. At the duopoly market level, only approximately 30 of 1,300 markets in our sample have both stations adopting algorithmic pricing under the "E5 and

Table E5: Station Level Results with Alternative "Adopter" Definitions

	(1)	(2)	(3)
Adopter Measure:	N Ch./Ch. Size	2 out of 3 (2 weeks)	E5 + Diesel
Outcome: Mean Margin			
	Sample: Monopoly ZIP Stations		
Adopter	-0.003	-0.002	-0.002
	(0.004)	(0.005)	(0.007)
Station FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Annual Regional Demographics	YES	YES	YES
N Brand Stations Control	YES	YES	YES
Weather Controls	YES	YES	YES
Observations	67,300	67,300	67,300
	Sample: Non-Monopoly ZIP Stations		
Adopter	0.005**	0.009***	0.009***
	(0.002)	(0.003)	(0.003)
N Competitors in ZIP	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)
Station FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Annual Regional Demographics	YES	YES	YES
N Brand Stations Control	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES
Weather Controls	YES	YES	YES
Observations	380,826	380,826	380,826

Notes: Sample is gas station/month observations from January 2016 until December 2018 that have one competitor or more in their ZIP code. Outcome variable Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale prices. In Column (1) "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in both the number of price changes and the average size of price changes within 4 weeks in any previous period. In Column (2) "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures within *2* weeks in any previous period. In Column (3) "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures for both *E5* and *Diesel* gasoline within *4* weeks in any previous period. "Share Brand Adopters" is the excluded instrument used in the 2SLS regression. It is equal to the share of stations that belong to the brand of station j that adopted by period t. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the number of stations belonging to station i's brand in month t. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors are clustered at ZIP code level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Diesel" definition, as compared to approximately 100 markets under our baseline definition.

Table E6: Duopoly ZIP Market Level Results with Alternative "Adopter" Definitions

	(1)	(2)	(3)
Adopter Measure:	N Ch./Ch. Size	2 out of 3 (2 weeks)	E5 + Diesel
Outcome: Mean Market Margin			
One Station Adopted	-0.014	-0.002	0.003
	(0.010)	(0.005)	(0.010)
Both Stations Adopted	0.079**	0.057**	0.016
	(0.040)	(0.027)	(0.029)
IVs	YES	YES	YES
Market FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Annual Regional Demographics	YES	YES	YES
N Brand Stations Controls	YES	YES	YES
Weather Controls	YES	YES	YES
Observations	39,148	39,148	39,148

Notes: The sample includes duopoly market/month observations from January 2016 until December 2018. A duopoly market is defined as a ZIP code with two gas stations. Outcome variable Mean Market Margin is the monthly average of mean market daily differences of pump prices for stations in market m in month t from wholesale price. In Column (1) "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in both the number of price changes and the average size of price changes within 4 weeks in any previous period. In Column (2) a station is labelled as an adopter if it experienced a structural break in any 2 of 3 relevant measures within *2* weeks in any previous period. In Column (3) a station is labelled as an adopter if it experienced a structural break in any 2 of 3 relevant measures for both *E5* and *Diesel* gasoline within *4* weeks in any previous period. "One Station Adopted " is a dummy equal to 1 in month t if one of the two stations in the market adopted in any previous period. "Both Stations Adopted" is a dummy equal to 1 in month t if *both* stations in the market adopted in any previous period. We use the "share of brand adopters" of the two stations in the market as instruments for adoption. 1st stage regression results are in Table C2 in the Appendix. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. We also control for the sizes of the brands of the two stations at time t. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors clustered at ZIP code level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

E.4 Alternative Instruments

E.4.1 Broadband Availability

We propose an alternative set of instruments that correct for endogeneity in station adoption decisions without relying on unobservable brand HQ decisions. The instruments capture the quality of broadband access in station j's region. There is well documented heterogeneity in broadband access and quality in Germany, with some areas and regions receiving sub-par services and speeds that are compared to the "old dial-up days" (NPR.org). In 2017, the second year of our sample, 29% of German users reported internet speeds less than half of those promised by providers (dw.com). A fast and reliable internet connection is a key requirement for the effective use of algorithmic pricing software. Computation is done in "the cloud," so gas stations need fast internet connections to access necessary price information in a timely manner. They also need reliable internet connections to upload their own data and feed and update the software.

Based on data obtained from the EU Commission's netBravo initiative, we have two measures of broadband performance in the local area around each gas station: whether the local area around the gas-station has widespread access to high speed internet in a particular year, and the reliability of broadband signals in that year. We use three indicator variables for high speed internet availability, capturing whether a 10 Mb/s, a 15 Mb/s or a 30 Mb/s connection is widely available in the local area.[71] Reliability is based on average signal strength (in dB) and the variance of signal strength. The intuition behind these instruments is that a gas station should be more likely to adopt algorithmic pricing software once its local area has access to high speed internet. It should also be more likely to adopt algorithmic pricing software if internet signals in its local region are reliable. The availability of internet in the area should not be correlated with station specific unobservables conditional on all other local demographics (income, population density, etc).

There are two downsides to this identification strategy relative to our main approach. First, variation at the region-year level is relatively limited as compared to variation at the brand-month level. Second, because an important source of the variation comes from regional geographic conditions, it is difficult to extend these instruments from station-level analysis to duopoly market level analysis. Duopoly markets, by definition, consist of stations that are close together in geographic space. There are no stations that we consider to be in the same market but that have different

[71]We define speed X to be widely available in an area if average speed-tests in that area in that year exceed that speed. As well, we assume that if an area has speed X widely available in a year, it also has the same speed widely available in every subsequent year. More details on the construction of these variables are in the Data Appendix.

broadband conditions.

Table E7 presents results from regression using these instruments. Qualitatively, the results are similar to those derived using our primary identification strategy. IV estimates show that the adoption of algorithmic pricing software increases mean station margins above wholesale prices. Mean station prices also go up. We once again find that adoption by monopolist stations has no effect on mean margins.

Table E7: Station Level Results with Alternative Instruments

	(1)	(2)	(3)	(4)
Sample:	All Stations	All Stations	ZIP Monopolists	ZIP Non-Monopolists
Outcome:	Adopter	Mean Margin	Mean Margin	Mean Margin
Adopter		0.051***	-0.002	0.045**
		(0.017)	(0.076)	(0.022)
N Competitors in ZIP	-0.008*	-0.002**		-0.002**
	(0.005)	(0.001)		(0.001)
10 Mb/s Internet Available Dummy	0.021**			
	(0.010)			
15 Mb/s Internet Available Dummy	0.018*			
	(0.010)			
30 Mb/s Internet Available Dummy	0.010			
	(0.020)			
Average Internet Signal Strength (dBm)	0.001**			
	(0.001)			
Average Internet Signal Variance (dBm)	-0.002			
	(0.002)			
Station FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Annual Regional Demographics	YES	YES	YES	YES
N Brand Stations Controls	YES	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES	YES
Weather Controls	YES	YES	YES	YES
Observations	330,977	330,977	46,932	283,980

Notes: Samples in Columns (1) and (2) include gas station/month observations from January 2016 until December 2018. Column (3) only includes stations that have no competitors within their ZIP code. Column (4) includes only stations that have one or more competitors within their ZIP code. Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale gasoline prices. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. Excluded instruments used in the 2SLS regressions in Columns (2)-(4) include annual internet speed and signal quality measures: a dummy for whether 10/15/30 Mbps internet was available in that year in that region, and two measures of average broadband signal strength. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. Standard errors are clustered at ZIP level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Quantitatively, point estimates of the effects of adoption are substantially larger than our main estimates. Results from the first stage suggest why this is the case. The instruments shift the adoption variable in expected directions - increases in signal quality increase the probability that a station adopts. The availability of 10 and 15 Mb/s broadband increases adoption. But compared to the brand-level instruments the instruments do not shift adoption probabilities by as much as the brand level instruments. As mentioned previously, there is also less variation in these instruments than in our brand-month instruments.

E.4.2 Placebo IV - Other Brands' Adoption

The main assumption of our baseline instruments is that brand level adoption recovers something about the incentives that the brand provides for their stations to adopt - for example, subsidies for replacing equipment or training. Effectively, we should be capturing brand-specific time varying cost shocks. To test whether this is the case, or whether we are capturing some other set of brand-specific time varying changes, we propose a "placebo" instrument.

This "placebo" instrument for station j in month t is the share of adopting stations at time t by *a different* brand than station j's brand.[72] This instrument has some similar time variation to our baseline instrument (i.e., brand adoption in general is going up over time) but the cost correlation should not exist. Results from this regression are in Table E8. They show that (i) there is no correlation between the propensity of other brands to adopt algorithmic pricing technology and the adoption of station j, and (ii) 2SLS regressions using this instrument do not generate any statistically significant effects of adoption on mean margins or prices.

[72] In practice, we pick a random station in station j's market and use their adoption shares.

Table E8: Station Level Results with "Placebo" Instrument

	(1)	(2)	(3)
Outcome:	Adopter	Mean Margin	Mean Price
Adopter		-0.550	-0.605
		(3.309)	(3.632)
Share Other Brand Adopters	0.005		
	(0.032)		
Station FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Annual Regional Demographics	YES	YES	YES
N Competitors in ZIP Controls	YES	YES	YES
N Brand Stations Controls	YES	YES	YES
N Adopting Competitors Control	YES	YES	YES
Weather Controls	YES	YES	YES
Observations	351,230	351,230	351,230

Notes: Sample included gas station/month observations from January 2016 until December 2018 with at least one competitor in their ZIP code. Mean Margin is the monthly average of daily differences of pump price for station j in month t and wholesale gasoline prices. Mean Price is the average retail price for station j in month t. "Adopter" is a dummy equal to 1 in month t if the gas station experienced a structural break in any 2 of 3 relevant measures in any previous month $\{1, ..., t-1\}$. "Share Other Brand Adopters" is the excluded instrument used in the 2SLS regression. It is equal to the share of stations that belong to a brand present in the same market as station j that adopted by period t. "N Competitors in ZIP" is equal to the number of other stations present in postal code of station j. Regional demographics include GDP, total population, population density, share of population employed and median age a the NUTS3/year level. Weather controls include the mean and standard deviation of monthly temperature and precipitation near station j in month t. We also control for the number of other stations in the ZIP code who are adopters at month t. Standard errors are clustered at ZIP level in parentheses. *** p<0.01, ** p<0.05, * p<0.1