Essays in Applied Microeconomics

Yuan Lyu

St John's College

Judge Business School

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

March 2024

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Chapter 1 is my own work. Chapter 2 has been written in collaboration with Prof. Christos Genakos, Prof. Mario Pagliero and Dr. Fotis Kokkoras and my contribution accounts for 50% of the work. Chapter 3 has been written in collaboration with Prof. Christos Genakos and Prof. Mario Pagliero, and my contribution accounts for 50% of the work.

Yuan Lyu
March, 2024

ABSTRACT

Essays in Applied Microeconomics

Yuan Lyu

This thesis comprises three independent papers on applied microeconomics. The first chapter studies the impact of primary care provider mergers on quality in England. The second chapter investigates the effect of price dispersion on consumer search behavior, drawing evidence from the retail gasoline market in Greece. The final chapter builds on the second, studying the asymmetric price adjustment and the impact of market competition on the asymmetric price adjustment. The details of the three papers are summarized below:

The Effects of General Practice Mergers on Quality in England

The primary care market has witnessed a growing trend of provider consolidation through mergers and acquisitions, yet the implications of this concentration remain uncertain. This study addresses this gap by providing the first empirical evidence on the effects of provider mergers on quality using evidence from the English primary care market. By analyzing all provider mergers from 2014 to 2018, I find predominantly negative effects of mergers on quality. Clinical quality does not change at best, and patient satisfaction decreases dramatically. Notably, the impact on quality varies based on the size of the general practices involved. Mergers between large general practices show a detrimental impact on quality, while mergers between small general practices may yield quality benefits. Additionally, there is no difference in the quality impact between mergers involving parties in the same geographical market and those in different markets. An exploration of the mechanism reveals that mismanagement, rather than changes in market concentration, drives the observed decline in quality following mergers.

The Effect of Competition and Price Dispersion on Search Behavior

We investigate the impact of price dispersion on consumer search behavior, while credibly controlling for market structure. Using the retail gasoline market on isolated, oligopolistic markets, as defined by small Greek islands, we exploit an excise duty tax increase policy as a plausibly exogenous shock to price dispersion. We directly measure consumer search using the number of user visits to a price information platform and mobile application. We find that the tax shock increases price dispersion and that in turn causes a short term increase in consumer search. The effect of price dispersion on consumer search remains regardless of market competition level.

Asymmetric Pass-Through and Competition

We study the pass-through to retail prices of four major changes in taxes for petroleum products (three increases and one subsequent decrease). We use daily pricing data from gas stations on small Greek islands, which define isolated markets with different number of competitors. First, we find that, on average, the pass-through of the tax hikes is five times higher than for the tax decrease. Second, the pass-through of the tax hikes increases with the number of competitors, but that of the tax decrease does not vary with competition (asymmetric competition effect). Third, there is significant asymmetry in the speed of price adjustments. Fourth, the asymmetric adjustment of retail gasoline prices cannot be explained by tacit collusion and the evidence points to search as the most plausible explanation.

ACKNOWLEDGEMENTS

Writing this PhD thesis has been a big journey, and I couldn't have done it without a lot of special people by my side. I would first like to thank my supervisor Christos Genakos. Christos is much more than a supervisor to me; he has become a mentor, a source of inspiration, and a friend. His work ethic and guidance have deeply influenced me, for which I am truly thankful. I am grateful to my advisor, Michael Pollitt for always been there to support me, with patience and great advice whenever I needed it. Huge thanks to Stefan Scholtes. Learning and working with Stefan has been an extraordinary part of my PhD adventure. I'm incredibly thankful for his unwavering support and the wisdom he has shared. A special thanks to Paul Kattuman. Paul always made time to attend my presentations, and his feedback to my work are invaluable.

I'm really grateful for the Judge Business School and the Economics and Policy Research Group. They've given me a place to grow and learn. To Jo - thank you for always being there to help, making everything so much easier. A huge thanks to my amazing friends in Cambridge, Mingmei, Rhys, and Zheying for always being there for me. And to Yayi, Flora, and Jiahe, who, though working in London, made the effort to come visit me once in a while, giving me so much support. To my PhD buddies at CJBS - Luna, Bell, Tom, Pearl, Tirza, Yuxia, and everyone else - you've made this experience so much more fun. St John's College, my home at Cambridge, has been incredibly welcoming and supportive, a place I'm thankful to be part of.

Most importantly, I thank my family and my husband for their endless support and love. And to my best friends back in China - Ziyi, Zihui and Xinyi - thank you for always being there to listen. Your support means the world to me.

Thanks to everyone who has been a part of this journey. This thesis is not just a reflection of my efforts but a testament to the support, guidance, and encouragement you've all generously provided.

CONTENTS

LIST OF FIGURES

List of Tables

CHAPTER 1

THE EFFECTS OF GENERAL PRACTICE MERGERS ON QUALITY IN ENGLAND

1.1 Introduction

Healthcare markets have become increasingly concentrated through mergers and acquisitions (Gaynor et al., 2015). While such consolidation may bring potential benefits, such as improved efficiency through economies of scale and coordinated care, which can lead to better outcomes (Asker and Nocke, 2021; Eliason et al., 2020), it can also lead to a reduction in quality or higher prices (Gaynor, 2004). The existing literature on mergers and acquisitions in healthcare markets has primarily focused on their impact on prices, while research evaluating their effects on quality remains limited. This is concerning as understanding the impact of mergers on non-price strategies is crucial for assessing the overall welfare implications of these mergers (Stiebale and Szücs, 2022; Asker and Nocke, 2021). In this paper, I contribute new evidence by examining the effects of mergers on quality within the primary care market.

Studying the impact of provider mergers in the primary care market is significant for several reasons. First, the primary care market plays a fundamental role in a well-functioning healthcare system. It not only contributes to improvements in population health, longer lives, and greater health equity (McCauley et al., 2021), but also helps prevent the need for more expensive secondary care (Santos et al., 2017). Despite its critical importance, existing empirical studies have predominantly focused on hospital mergers in the secondary care setting, with no research conducted on the primary care market. Unlike hospitals, which mainly offer specialized care, primary care plays a unique role in delivering holistic, long-term care and managing a broad range of patient's healthcare needs. Consequently, the quality of care becomes a significant aspect of the primary care market. Moreover, there has been a persistent trend of provider concentration in the primary care market, observed in various countries such as the United States, EU countries, and the UK (Fulton, 2017; Gravelle et al., 2019; Pál et al., 2021).[1] This trend underscores the relevance of studying merger effects in the primary care market.

The English primary care market is a suitable setting for my study for several reasons. First, primary care is provided free at the point of use in England. This characteristic allows me to exclude the influence of mergers on prices and concentrate solely on examining their effects on quality. Second, the English primary care market offers rich and publicly available data on various quality measures. I use official government measures to assess clinical quality and patient survey data to evaluate patient experiences. This comprehensive evaluation enables me to contribute new evidence on the multiple dimensions of quality affected by mergers. Third, there has been a significant number of practice mergers in the English primary care market in recent years, providing me with a sufficient sample size for my empirical analysis.

Two main incentives drive general practice mergers in England. Some practices merge to ensure their survival and sustainability, particularly in the face of challenges such as doctor shortages. Others merge to achieve efficiency gains through collaborative efforts by forming a larger practice. Yet, from a consumer perspective, whether these mergers offer advantages remains unexplored. As primary care is free at the point of use, quality becomes a salient feature when assessing the merger effects.

The impact of general practice mergers on quality is theoretically ambiguous. On one hand, mergers may yield economies of scale and scope that enhance quality. Firstly, merged practices could realize economies of scale by employing both physician and non-physician staff to implement quality improvement processes and using information

[1]For instance, in the U.S., the percentage of physicians working in large practices with at least 50 physicians grows from 14.7% in 2018 to approximately 17.2% in 2020 (American Medical Association, 2022).

technology to support these initiatives (Mehrotra et al., 2006). Also, physicians often possess distinct specializations, and merged practices can leverage doctors' skills to achieve economies of scope and improve clinical quality (Casalino, 2006). Both arguments suggest potential quality improvements after the merger. However, on the other hand, mergers might reduce the incentive to provide high-quality care if they lead to increased market concentration and undermine competition. Theoretical models show that in systems where providers do not determine prices, such as the primary care context examined in this paper, a positive relationship exists between competition and quality (Gaynor, 2006). Mergers could reduce competition, leading to a quality decrease. Moreover, there could be mismanagement issues that lead to quality decline after the merger. For instance, challenges exist in integrating different cultures and values and establishing effective communication, particularly when the merging parties are geographically distant (Renneboog and Vansteenkiste, 2019). Thus, the direction and magnitude of the merger effect on quality remain an empirical question.

Despite the importance of evaluating provider concentration in the primary care market, to my knowledge, no research has explored this due to the lack of data (Gaynor and Town, 2011). To address this gap, I assemble the first comprehensive dataset documenting the universe of practice mergers in England between 2014 and 2018. I choose these five years because 2014 is the earliest year for which pre-merger outcomes are available, and I want to focus on the period without the disruptions of the coronavirus pandemic.

I use a Difference-in-Differences (DiD) strategy to compare the outcomes of merged practices with those that have never merged before and after the merger. Two main empirical challenges need to be addressed in the analysis. Firstly, I must account for potential endogeneity concerns. To do so, I analyze whether practice mergers appear random by predicting merger likelihood based on practice and local-level characteristics. While I find no consistent observable determinants of practice mergers, I include a comprehensive set of practice-level and local-level controls and practice fixed effects in the regression. I use the propensity score matching (PSM) method to select a valid comparison group from non-merged practices. As a robustness check, I also consider practices that have not yet merged as an alternative comparison group. Secondly, there are variations in treatment timing as mergers occur in different years. I address this staggered roll-out design by adopting a stacked DiD regression approach following Deshpande and

Li (2019) and Cengiz et al. (2019). I conduct numerous robustness checks to validate the effectiveness of the regression approach.

I find that mergers have an adverse impact on quality. The objective clinical quality, assessed across various chronic illnesses, tends to deteriorate over the long run. Also, the subjective patient satisfaction declines dramatically. The overall patient satisfaction experiences a decrease of approximately 3 percentage points. Moreover, patient satisfaction rates for continuity of care with primary doctors and access to care drop by around 3 percentage points and 4 percentage points, respectively. This corresponds to a sharp reduction of 10% and 5%, respectively, for the average practice. Importantly, these merger effects do not appear to be explained by changes in patient mix, as the prevalence of chronic illnesses in the patient pool remains unchanged after the merger. Furthermore, there is evidence suggesting that practices may achieve financial gains following the mergers. These findings imply that mergers may lead to financial benefits but at the expense of a decline in quality.

To further understand these findings, I explore whether merger effects differ based on the types of general practices involved or the nature of the mergers. Firstly, I examine if merger effects differ by the size of the merging parties. I find that mergers involving small general practices yield potential quality benefits, whereas mergers between large general practices result in a significant decline in quality. This finding suggests that mismanagement may be the main driver behind the quality decline, as larger practices might be more challenging to manage, potentially leading to greater mismanagement. Secondly, I compare the effects of within-market mergers, where practices in the same geographical market merge, and cross-market mergers, where practices across different markets merge. The results show no significant difference between them. This finding implies that changes in market concentration might not be the dominating explanation of the merger effects. If it were, we would observe a more negative impact for within-market mergers, which are affected by changes in market concentration, compared to cross-market mergers, which are irrelevant to such changes. Finally, I study mergers of different claimed motivations and find that mergers aimed at achieving efficiency are the most detrimental to quality, while mergers motivated by survival are potentially beneficial. This finding aligns with the results regarding the size of merging parties, as survival-driven practices are typically small, while efficiency-driven mergers usually involve larger practices.

I formally test two potential explanations to understand the mechanisms behind these findings: the market concentration effect and the mismanagement issue. Both explanations could lead to a decline in quality, and it is challenging to determine which mechanism dominates. I address this issue by leveraging within-market and cross-market mergers in my sample. Market concentration changes should only impact within-market mergers, while mismanagement issues should be more salient for cross-market mergers. Therefore, I examine within-market and cross-market mergers separately to test market concentration and mismanagement mechanisms, respectively. If the market concentration effect was the main driver, we would observe a more negative effect on quality in already concentrated markets due to larger changes in market concentration following mergers. However, analyzing the within-market mergers, I find no significant difference between mergers in already concentrated markets and mergers in markets with substantial competition, suggesting that changes in market concentration are not the primary explanation. On the other hand, if mismanagement is the primary factor, we should observe a more adverse effect with larger merging entities. Analyzing cross-market mergers, I find that mergers between larger general practices lead to a more detrimental impact on quality after the merger, which supports the mismanagement mechanism. Identifying the mechanism behind merger effects on quality is another key contribution of this paper. The findings imply that mergers can harm patients regardless of changes in market concentration, primarily due to mismanagement.

These findings have important policy implications, particularly as much of the concentration in physician markets remains unnoticed by regulatory authorities (Gravelle et al., 2019). Based on the findings, I recommend that the government carefully consider the potential harm to patients before approving further mergers. Additionally, special attention should be given to mergers between large practices, given their greater detrimental impact on quality. Although these mergers may intend to bring benefits through achieving efficiency, it is challenging to realize due to mismanagement. Conversely, mergers between small practices hold the potential for positive outcomes.

This paper adds to several areas of literature. First, it is relevant to the research that evaluates the effects of mergers on non-price outcomes. Existing empirical studies have examined the impact of mergers on non-price outcomes in various industries, such as product quality (Fan, 2013) and variety (Sweeting, 2010; Berry and Waldfogel, 2001; George, 2002; Fan and Yang, 2022; Jeziorski, 2014) covering industries including the newspaper industry (George, 2002; Fan, 2013), radio broadcasting industry (Sweeting, 2010; Berry and Waldfogel, 2001; Jeziorski, 2014), and brewery industry (Fan and Yang, 2022). These studies typically employ structural modelling approaches to simulate the effect of hypothetical mergers and quantify the welfare effects of mergers.[2] In the healthcare sector, to the best of my knowledge, all existing studies have examined providers that offer specialized care, such as hospitals.[3] No previous attempts have been made to examine mergers specifically in the primary care market, mainly because identifying merger events in this particular market is challenging. This paper fills the research gap by compiling a comprehensive dataset of merger events in the English primary care market and presenting the first empirical evidence on the impact of provider mergers in this context. I take a retrospective approach and evaluate the outcomes of actual mergers using a reduced-form methodology.

Furthermore, this paper adds to the literature that examines the determinants of quality in the primary care market. Several empirical papers have explored the effect of market competition (Gravelle et al., 2019; Santos et al., 2017; Scott et al., 2022; Dietrichson et al., 2020), financial incentives (Minchin et al., 2018; Kontopantelis et al., 2014), practice list size (Gravelle et al., 2022; Kelly and Stoye, 2014), and behaviour change in physicians (Chauhan et al., 2017) on quality.[4] However, the findings are inconclusive, and further research is needed to identify effective quality improvement tools. This paper contributes to this issue by addressing whether market concentration is a solution for promoting quality. This is particularly relevant considering the current common market concentration trend across different countries in the primary care market.

The rest of this paper proceeds as follows. Section 1.2 provides the institutional background. Section 1.3 describes the data. Section 1.4 outlines the research design. Section 1.5 presents and discusses the results. Section 1.6 tests the sensitivity of the findings. Section 1.7 concludes.

[2]For instance, Beckert et al. (2012) take a structural approach to simulate the effect of mergers between hospitals in England.

[3]For instance, Ho and Hamilton (2000) and Capps (2005) find no significant effect of hospital mergers on most quality indicators in the US. Gaynor et al. (2012) find limited evidence of improvements in quality following hospital mergers in the English NHS. On the other hand, Eliason et al. (2020) examines the effect of mergers on quality in the US dialysis industry and finds a negative effect on patient outcomes.

[4]For a systematic literature review, refer to Ahmed et al. (2021).

1.2 Institutional Background

1.2.1 Primary Care in England

In England, primary care is provided by the general practices market. All residents in England are entitled to choose freely and register with one general practice. For most people, this registered general practice is the first and most commonly used point of contact when they have any physical or mental health concerns.

Once registered with a particular practice, patients are assigned a general practitioner (GP) who becomes their preferred GP. The continuity of care between patients and their preferred GP is highly valued in the primary care market as it is associated with better health outcomes and improved patient experiences (Dossa et al., 2017; Freeman et al., 2010; Kajaria-Montag and Freeman, 2020). However, it is not guaranteed that patients will always be seen by their preferred GP during appointments. In Section 1.5, I investigate the impact of practice mergers on this crucial aspect of continuity of care. Healthcare services are free at the point of use in the primary care market; therefore, quality becomes an essential factor when assessing the primary care system in England.

Each general practice is a small business typically owned by a partnership of several GPs, who have both medical and managerial responsibilities. The English National Health Service (NHS) holds contracts with the practice as a whole, referred to as a "contractor", rather than with individual GP partners. The contractor receives the vast majority of its income from the NHS through these contracts.[5] There are four primary sources of income payments: a global sum payment, which is an annual amount calculated for each contractor based mainly on the number of contractor's registered patients; quality incentive rewards from the Quality and Outcomes Framework (QOF) (explained in more detail in Section 1.3); payment schemes for providing a range of enhanced services, such as several vaccination programs; and payments for specific purposes, such as those related to dispensing services for dispensing practices.[6] General practices are reimbursed for the costs of their premises but have to fund all other expenses, such as hiring practice nurses and clerical staff, from their revenue (Santos et al., 2017).

[5]In addition, practices may generate supplementary revenue by providing certain services, such as issuing private prescriptions or medical certificates.

[6]Contractors who are authorized to provide dispensing services to specific patients will receive payments to cover the costs of providing drugs and appliances, as well as a dispensing fee per item dispensed.

1.2.2 Why GP Practices Merge?

There are three main reasons behind practice mergers.[7] The first type occurs when a GP partner retires, typically in a small practice, and this practice merges with another practice(s) in the same geographical district. The second type of merger involves practices merging to achieve operational efficiencies by consolidating back-office functions. The final category of mergers involves failing practices. I notice that these failing practices are often identified as such by the Care Quality Commission (CQC), an independent healthcare regulator in England. The CQC inspects and rates GP surgeries based on their quality, with ratings ranging from outstanding, good, requires improvement to inadequate. Practices rated as inadequate are placed into special measures, and if they fail to improve within a year, their registration is cancelled by the CQC.[8] Failing practices with poor quality ratings may be motivated to merge to improve their ratings and ensure their survival.

In my analysis, I try to differentiate between these different motivations and explore the heterogeneous effects of mergers across these types in Section 1.5.6.3. I combine the owner-retirement type and failing type of mergers because they both represent cases where merged practices are facing difficulties and are at a low point when the merger occurs. For instance, in the case of practices whose owner retires without another practice to take over, closure becomes a possibility. Therefore, I categorize these mergers as the merge-to-survive type. This category differs from the merge-for-efficiency type, where the merger occurs between normally functioning practices without operational difficulties. My analysis seeks to answer whether the effects of mergers differ depending on whether they are motivated by the need to survive or achieve efficiency.

1.3 Data

1.3.1 GP Practice Merger Data

To my knowledge, no existing dataset documents all instances of general practice mergers in England. Hence, I manually assemble this information with the following procedure.

[7] This follows discussions with insider experts.
[8] For more information, see the guidelines published by CQC.

The first step involves compiling a comprehensive list of general practices that might have engaged in mergers during the sample period. Then, I perform a manual check to determine if each practice on the list participated in a merger, and if so, its merging partner(s). Moreover, whenever possible, I try to identify the motivation behind each practice merger when collecting the data. A detailed explanation of how I collect the practice merger data is outlined in Appendix A.

I collect the universe of practice mergers between 2014 and 2018 in England. I choose these five years because 2014 is the earliest year for which pre-merger outcomes are available, and I want to focus on the period without the disruptions of the coronavirus pandemic.

In my analysis, I define a merger event as the amalgamation of two practices: the target practice and the acquirer practice. The acquirer practice is the primary practice that serves as the main site following the merger, while the target practice becomes a branch site. In some cases, an acquirer takes over more than one target practice within a single year. In such cases, I consider each pairwise combination of the acquirer and target as separate merger events. This approach has the limitation that the acquirer may take over multiple targets simultaneously and, therefore, should be treated as a single merger event. However, due to the nature of my data collection, accurately determining whether the acquirer takes over multiple targets simultaneously is challenging. To simplify the analysis, I classify each combination of the acquirer and the target as distinct merger events. I believe this issue is of less concern for my analysis since I observe only 89 cases (less than 15% of the total sample) where multiple surgeries merge into one in the same year.[9]

In total, I document 787 instances of practice mergers occurring over these five years. Fig. 1.1 plots the number of mergers by year. There is a noticeable increasing trend in the frequency of practice mergers over the study period, suggesting that mergers have become more prevalent among practices over time.

[9] I perform a robustness check by dropping cases where an acquirer takes over multiple targets within the same year, and the findings remain robust. Results are available upon request.

1.3.2 Outcome Variables

My primary focus lies in assessing the impact of general practice mergers on quality. Quality is multi-dimensional. I aim to capture its different aspects by examining both objective clinical quality measures and subjective patient experience measures. While my primary emphasis remains on the quality effect, I also present supplementary evidence exploring the effect of mergers on other outcomes, specifically on financial performance and workload changes. Below, I describe each of my four categories of outcome variables separately. A summary of all my outcome variables is presented in Table 1.1.

1.3.2.1 Clinical Quality Measures

I measure clinical quality using the official QOF data.[10] The QOF is a voluntary scheme introduced in 2004 that financially incentivizes practices to meet quality targets for their registered patients. When it was first introduced, four principal domains were included: the clinical domain; the organizational domain; the patient experience domain; and the additional services domain,[11] where for each domain, there are indicators set out to assess the performance of the general practice. Some indicators simply require certain tasks to be accomplished (e.g., establishing and maintaining a register of patients with coronary heart disease), and points are awarded if the tasks are performed. Other indicators contain designated thresholds against which the practice's performance is assessed using a percentage.[12] Practices earn points on these indicators, which translate into financial rewards. The price per point was £75 when QOF was first introduced but had increased to £179.26 by 2018/19. Though voluntary, participation rates in the QOF are very high over 95%.[13]

[10] The QOF data has been used by previous literature to measure clinical quality. See for instance, Gravelle et al. (2019) and Gravelle et al. (2022).

[11] From 2014 onwards, the QOF has been revised to incorporate three domains only: clinical, public health and public health for additional services. Although the domains have changed, most of the indicators remain, with some indicators being reorganized. For instance, there is an indicator asking practices to support smokers in stopping smoking. This indicator was used to be classified under the organizational domain but has been moved to be included in the clinical domain.

[12] For example, an indicator for blood pressure control in patients with hypertension sets a minimum threshold of a last-time blood pressure reading of less than 150/90 mmHg for at least 45% of the contractor's patients with hypertension. The contractor's performance is assessed by calculating the percentage of patients who meet this target.

[13] See NHS Digital website for more detail.

The clinical conditions covered by QOF are selected specifically targeting those high-priority disease areas where primary care has principal responsibility for ongoing care, and there is evidence that improved primary care will have health benefits.[14] I therefore use QOF data to measure the clinical quality of each practice. Following Gravelle et al. (2019) and Gravelle et al. (2022), I construct three measures. *qofOutcome* is the percentage of the total maximum available points that the practice achieves. These achievement points are what the final payment is calculated and based on. However, using QOF performance as a measure of clinical quality has limitations. First, only around two-thirds of the points are related to clinical quality indicators for specific health conditions, while the rest are for general tasks like keeping records or giving information to patients (Gravelle et al., 2019). Second, there might be potential gaming of exception reporting to get the points. To illustrate, the achievement point for each indicator is calculated as $100\times$A/(T-E), where A is the number of patients for whom the indicator is achieved, T is the total number of patients with the relevant condition, and E is the number of patients reported as exceptions. Doctors are allowed to exclude certain patients from individual clinical indicators (exception reporting) for reasons such as clinical inappropriateness or patient dissent. To address this concern, I construct a second clinical measure, *ClinPA*, using QOF data, focusing only on clinical indicators and practice population achievement. Specifically, I measure performance on each indicator using $100\times$A/T and calculate *ClinPA* as the weighted average of these performance measures, taken over the clinical indicators and disease areas, which are consistently defined between 2013 and 2019. The weights are the maximum points available for the indicators. There are in total 20 consistent clinical indicators over these 6 years. Quite a few indicators get dropped calculating *ClinPA* because the QOF went through a major change in 2019, with several indicators redesigned.[15] Therefore, I also investigate the subsample between 2013 and 2018, which provides a larger sample of consistent clinical indicators (39 indicators), and calculate the corresponding measure, *ClinPA2*.

[14] For a more detailed discussion of QOF data, please see Roland and Olesen (2016) and Roland and Guthrie (2016).

[15] For more information, see the summary by NHS Digital of QOF data in 2019-20.

1.3.2.2 Patient Experiences

I use data from the General Practice Patient Survey (GPPS) to assess patients' experiences with their registered practices. The GPPS is a nationally representative survey that has been sent out to a random 5% of registered patients in each practice across the UK since 2006 (Gravelle et al. 2019).[16]

I first construct two measures to capture patients' overall experience: *OverallSat* is the proportion of patients satisfied with their practice on an overall level (available for 2013-2019), and *Recommend* is the proportion of patients who would definitely or probably recommend their surgery to someone who has just moved to their local area (available for 2013-2018).

Apart from these two general measures, I select several survey questions to capture the different aspects of patient experiences. *Continuity* is the proportion of patients who have a preferred GP and could always or almost always see their preferred GP (available for 2013-2019). This measure reflects the frequency with which patients are able to meet their preferred GP when needed, thereby indicating the level of relational continuity between the patient and their preferred GP. *AppointSat* is the proportion of patients whose overall experience of making an appointment is satisfying (available for 2013-2019). *WaitSat* is the proportion of patients who report that their waiting time at surgery is normally not too long (available for 2013-2017). *OpenHrsSat* is the proportion of patients satisfied with their practice opening hours (available for 2013-2017).

1.3.2.3 Financial Performance

An ideal measure to assess the practice's financial performance would be to use the profits data. However, such information is confidential to the practice and, therefore, unavailable. I thus use the amount of NHS payments to each general practice as a proxy. These payments represent the bulk of the general practice income and thus provide information on revenue.

Using revenue data obtained from NHS Digital, I construct two measures, *RevPerPatient*, which denotes the revenue per patient in logarithm form, and *RevPerGP*, which

[16] During the years of my sample, the survey was administered twice a year for 2014-2016, with the results published in January and July, and was conducted once a year for 2017-2019, with the results published in July. For my main analysis, I follow Gravelle et al. (2022) and use the survey results published in July for 2014-2016.

represents revenue per full-time equivalent (FTE) of GPs in logarithm form.

1.3.2.4 Workload Changes

To assess staff workload, I construct three measures. *PatientsPerGP* represents the number of registered patients per FTE GP, reflecting doctors' workload. Similarly, I construct *PatientsPerNurse* and *PatientsPerAdmin* to capture the workload on nurses and administrative staff, respectively.

1.3.3 Additional Practice Level and Local Level Data

At the practice level, I obtain the number of registered patients and practice prevalences from QOF data. Following Gravelle et al. (2022), I select nine illness conditions: coronary heart disease (CHD), stroke, hypertension, diabetes, epilepsy, chronic obstructive pulmonary disease (COPD), cancer, serious mental illness, and asthma. I collect information on the workforce, including the FTE of doctors, nurses, and administrative staff. Additionally, I include information on whether the practice is a dispensing one. Furthermore, for each practice, I calculate the number of competing GP surgeries located within a 2km radius.

For local area characteristics, I trace each practice's location to the Lower Layer Super Output Area (LSOA) level and assign the corresponding LSOA level characteristics to it. I use the index of Multiple Deprivation (IMD) to account for the socio-economic status of each area. A higher IMD rank implies a less deprived status. Additionally, I account for the rural or urban classification of each LSOA. For a comprehensive list of practice and local characteristics used in my analysis, please refer to Table 1.1.

1.3.4 Summary Statistics

Table 1.2 presents the mean values of the variables used in my analysis, both for the full sample (Column (1)) and the distinction between merged and never-merged practices (Column (2) and Column (3)). For merged practices, I only consider observations from the year before the merger takes place, considering both the acquirer and target practices. Column (4) provides the t-statistic comparing the difference in means between the two subsamples in Column (2) and Column (3). I find that prior to the merger, merged practices exhibit slightly lower QOF performance and patient recommendation rates in comparison to never-merged practices, although the difference is of minor magnitude. Merged practices are smaller in size and tend to have fewer FTE doctors and admin staff. Also, mergers seem to be more likely to occur in more deprived areas. These observations motivate me to use PSM to select a comparable group of practices for comparison with the merged practices. I will explain my methodology next.

1.4 Empirical Methodology

The aim of this paper is to investigate the effect of general practice mergers on outcomes. To do so, I use a difference-in-difference method, comparing practices that experience mergers with a selected group of practices that do not throughout the sample years. The control group is selected using the PSM method, which will be explained in detail later. As there is only one merged entity after the merger, I construct pseudo-merged values for the pre-merger period by combining data from both merging parties. These pseudo-merged values, including both outcome variables and additional control variables, are calculated as weighted averages, with weights determined by the number of patients from each merging party. This approach allows for a consistent comparison of merged practices before and after the mergers.[17] In my main analysis, each treated unit refers to a single merged entity corresponding to each merger event. I use the constructed pseudo-merged values for the pre-merger period and the value of the merged entity for the post-merger period.

The main identification challenges arise from the fact that practice mergers may not occur randomly, and practices that undergo mergers may differ from those that do not in ways that are not observed. These challenges present two main problems for my analysis. Firstly, if there are unobservable factors that affect both a practice's treatment status (i.e., whether it merged) and the outcome variable, there will be endogeneity concerns, which

[17] Gaynor et al. (2012) also construct pseudo-merged values prior to the merger to study the effect of hospital mergers in England. The key difference between our methods is that they only construct values for the outcome variables of the pseudo-mergers and perform PSM separately for each merging party based on their own matching variables prior to the merger. In contrast, I construct pseudo-merged values for both outcomes and control variables prior to the mergers, and perform PSM to select appropriate controls for the single merged entity only. My results remain robust when I conduct separate matching for the target and acquirer practices. Results are available upon request.

make it difficult to make causal claims on the estimates. In addition, if selection bias is present, it can be challenging to identify an appropriate comparison group for the treated practices.

To address these challenges, I first test if practice mergers appear to be random by predicting the likelihood of mergers based on practice and local level characteristics. I perform a logit regression regressing the merger status on these characteristics. For the merged practices, I only use their pseudo-merged values from the year preceding the merger. I run separate regressions for each year to account for the changing pool of potential mergers over time.[18] The results are given in Table 1.3. I find that the variables affecting the likelihood of mergers differ across years. Notably, there seem to be no variables that consistently determine practice mergers. This suggests that mergers may be potentially random. Nonetheless, to enhance the credibility of the causal estimates, I control for these practice-level and local-level characteristics in the regression. Moreover, I include practice fixed effects to adjust for all unobservable differences across practices that are consistent over time.

To mitigate potential selection bias, I use the PSM method to select an appropriate comparison group for the treated practices. The observation that the pool of merged practices changes over the years in Table 1.3 motivates my strategy of conducting separate matching for each year of merged practices. I will explain this approach later. In addition, to further ensure the robustness of the causal estimates, I perform a robustness check using practices that have not yet undergone mergers as the control group. I conduct logit regressions to predict the timing of mergers and find no variables that consistently predict the timing of mergers, suggesting that the timing of mergers may also be potentially random (results are presented in Table B2). Therefore, I use mergers that occur at least two years later as an alternative control group as an additional robustness check.

Apart from the endogeneity concern, I also need to address the issue of staggered roll-out design in my study. Recent research in econometric theory has shown that the standard two-way fixed effects (TWFE) DiD estimators fail to produce valid estimation when applied to settings with variations in treatment timing and treatment heterogeneity (see, e.g., Borusyak et al. 2021; Callaway and Sant'Anna 2021; Goodman-Bacon 2021;

[18] I also conduct an alternative analysis that includes both acquirer and target practices in the regression, rather than using the pseudo-merged values. The results are presented in Table B1, and my main findings remain consistent.

Sun and Abraham 2021). To address this issue, alternative methods have been proposed.[19] I will adopt a *stacked regression* approach in the main analysis and use an alternative method developed by Callaway and Sant'Anna (2021) as a robustness check.

1.4.1 Stacked Regression Method

Performing the standard TWFE DiD regression on a sample with staggered treatment timing introduces *bad comparison* problems that already-treated units are used as comparisons. This is especially problematic when treatment effects vary over time or across groups, even if the parallel-trend assumption is satisfied. To circumvent this issue, I follow Cengiz et al. (2019) and Deshpande and Li (2019) and adopt a *stacked DiD regression* approach.

Specifically, I proceed as follows. First, I categorize all merged units into five cohorts based on the year of the merger, ranging from cohort 2014 to cohort 2018. For each cohort, I construct a comparison group for the treated practices using PSM. This results in five separate datasets, each containing a treatment group (merged entities formed in a specific year) and a corresponding comparison group (never-merged practices selected through matching). Finally, I stack these five datasets together and perform a standard TWFE DiD regression on the stacked dataset, with the dataset-specific unit- and time-fixed effects. This approach ensures that I only use "clean" controls specific to each cohort of treated units for comparison.[20]

The main regression equation is as follows:

$$y_{igt} = \gamma_{ig} + \gamma_{tg} + \beta(Treat_{ig} \times Post_{gt}) + X_{igt}\delta + \varepsilon_{igt} \tag{1.1}$$

where y_{igt} denotes the outcome for merged unit i from dataset g in year t, including outcome measures of quality, financial performance, and workload changes. $Treat_{ig}$ is an indicator equal to 1 for merged units of dataset g. $Post_{gt}$ is an indicator equal to 1 for post-merger years, specified separately for each of the five different datasets. γ_{ig} represent dataset-specific unit fixed effect and γ_{tg} are dataset-specific year fixed effect. X_{igt} are a set of practice-related time-varying covariates that I control for.[21] To avoid

[19] See Baker et al. (2022) for a comprehensive review and comparison of these alternative methods.

[20] This approach can be applied using either a static or a dynamic specification. The estimates of dynamic treatment effect in Section 1.5.5 follow this approach.

[21] At the practice level, I control for the competitive environment, the total number of registered patients,

measurement errors in the timing of the merger, I drop the entire year of the merger for each merged entity. Standard errors are clustered at the cohort-specific unit level. β is the key coefficient of interest that represents the effect of mergers on practice outcomes. Note that the treated units in my analysis are the merged entities resulting from the mergers. Therefore, the estimated coefficients capture the overall effect of the merger on the merged entity.

1.4.2 Matching

The key assumption in a DiD research design is that in the absence of treatment, the treated practices would have evolved in a similar way as the control groups. I use PSM to construct a suitable comparison group. The idea of PSM is that practices are similar if they are equally likely to be treated, i.e., they have the same treatment propensity score (Caliendo and Kopeinig, 2008). Propensity scores are usually estimated by logit or probit regression on a set of matching variables, which are determinants of the treatment assignment. The validity of PSM depends on the assumption that, given these matching variables, the observations are randomly assigned to the treatment and control groups. If this assumption holds, any observed difference in the outcome variable can be attributed entirely to the treatment effect.

For the matching procedure, I use both practice and local-level characteristics as matching variables. At the practice level, I consider the prevalence of nine disease conditions and the number of registered patients. These factors account for the demand for each practice. I also include the FTE of practice staff, including doctors, nurses, and administrative staff, to capture the workforce of each practice. The dispensing status of each practice is also taken into account. At the local level, I use the local-level IMD rank and an indicator variable for urban areas as proxies for the socioeconomic status of the local market. Additionally, the number of competing GP surgeries within a 2 km radius is added to account for the competitive environment.

As shown in Table 1.3, the pool of merged practices changes over the years, and variables that impact the likelihood of mergers differ across years. Therefore, I perform matching separately for each cohort of treated practices. I match with replacement, that

a set of prevalence rates, the workforce composition (including GPs, nurses, and admins FTEs), and dispensing status. At the local level, I incorporate the local IMD rank and a dummy variable indicating whether the area is urban or rural.

is, never-merged practices are re-sampled every year for the matching. I estimate the propensity score using the pre-treatment values of the matching covariates, and I select the three closest matches as the comparison group for each treated practice.

A suitable matching procedure should balance the distribution of the matching variables in the treatment and control groups. I assess the matching quality by evaluating the standardized bias between the matched and unmatched samples as well as carrying out a two-sample t-test. Standardized bias is a commonly used indicator proposed by Rosenbaum and Rubin (1985).[22] The lower the standardized bias, the better the matching quality. The standard practice is to consider a value greater than 20 as large (see, for instance, Gaynor et al., 2012). Moreover, for each matching variable, I perform a two-sample t-test on the sample means between the merged and matched never-merged practices. If the matching is suitable, there should be no significant differences between the covariate means for the two groups.

Results of the balance tests are presented in Table 1.4. For each covariate, I perform the tests twice: once using the raw, unmatched sample and another with the matched sample. Comparing these two shows the extent of the balance problem solved through matching. I find that without matching, significant differences exist between the treated and unmatched controls in variables measuring patient numbers, the staff's FTE, the practice's dispensing status, and the local market condition. The t-tests show that these differences are statistically significant at a 5% level. However, after matching, the t-values indicate that the differences become insignificant, and the standardized bias substantially decreases to below 20. This indicates that the matching procedure effectively identifies a comparable control group.

In the robustness checks (see Section 1.6.2), I vary the number of matches, selecting one, five, and seven closest matches for the analysis. Moreover, contamination of spillover effects may occur if matched practices are in the same local market as the treated practice. However, I find that only two matches are in the same market, so this concern is minor.[23]

[22]For each matching covariate, standardized bias is defined as the difference of means in the merged and matched never-merged subsamples as a percentage of the square root of the average variances in the two groups.

[23]Nevertheless, as an additional robustness check, I also perform the analysis using matches from only outside-market practices. Results are available upon request.

1.5 Results

This paper concentrates on evaluating the effect of mergers on quality, specifically focusing on two key aspects: subjective measures using official clinical quality indicators and objective measures related to patient experiences. Furthermore, I expand my analysis to investigate a range of additional outcomes associated with financial performance and staff workload.

1.5.1 Effect of Practice Mergers on Clinical Quality

The results on the impact of practice mergers on clinical quality are presented in Table 1.5. Columns (1) and (2) suggest that merged practices exhibit a marginal improvement in QOF achievements. On average, merged practices experience a less than 1 percentage point rise in QOF points post-merger. The increase in QOF performance may be due to GP practices selecting and treating healthier patients after the merger. To investigate this issue, I use the prevalence rates of long-term illnesses, including coronary heart disease, stroke, hypertension, diabetes, epilepsy, chronic obstructive pulmonary disease, cancer, serious mental illness, and asthma, as the dependent variables in the regression. The results, shown in Table B3, indicate that the patient pool, in terms of the prevalence of these long-term illnesses, remains stable after the merger. This implies that the positive impact on QOF performance does not result from merged practices selecting healthier patients, which might be a positive sign that mergers bring benefits. However, considering that practices in my sample already perform well in achieving QOF points, with an average practice achievement rate over 95%, the less than 1 percentage point increase represents a minor advancement. Furthermore, this slight increase in QOF performance diminishes when assessing the population achievement of practices in clinical quality indicators, as indicated in Column (3) to Column (6). Based on these results, I conclude that practice mergers do not appear to improve clinical quality significantly.

1.5.2 Effect of Practice Mergers on Patients Experiences

Next, I examine subjective quality measures using patient survey responses.[24] I assess both patients' overall satisfaction levels and their experiences from different aspects.

I begin by examining two indicators that represent patients' overall experience with the practices. Column (1) and Column (2) in Table 1.6 show the effect on patient overall satisfaction. The results indicate that regardless of whether additional covariates are included or not, after the merger, the percentage of patients who feel satisfied with the practice decreases by around 3 percentage points. This decline equates to about a 4% decrease on an average satisfaction rate of around 85%. Similarly, Column (3) and Column (4) reveal a decline of about 3 percentage points in the proportion of patients who would recommend the merged practices to others, translating to a 4% decrease on an average recommendation rate of approximately 78%. Taken together, these results imply that merged practices perform worse from a patient perspective.

Given the observed decrease in patients' satisfaction levels, it is natural to question why this might occur. To explore possible explanations for patients' disappointment, I further select and analyze a group of survey questions that provide more detailed insights into their experiences. I present the results in Table 1.7.

Columns (1) and (2) focus on the merger effect on continuity of care. The results indicate that patients are less likely to see their preferred GPs after the merger, and the decline is substantial in magnitude. Given a mean satisfaction rate of 35% for continuity of care, the additional 3.4 percentage points decrease for merged practices corresponds to a substantial 10% drop from this mean. This sharp decline is particularly concerning as continuity of care is a crucial factor in healthcare outcomes in the primary care market (Dossa et al., 2017; Freeman et al., 2010; Kajaria-Montag and Freeman, 2020). Furthermore, Column (3) and Column (4) reveal that mergers adversely impact access to care. The observed 4 percentage points decline in satisfaction with the appointment process translates to a 5% decrease based on an average rate of 75%. Results presented in Column (5) and Column (6) imply longer waiting times for patients at merged practices. The percentage of patients who report that waiting times are not too long declines by approximately 2.5 percentage points, equating to a 4% decrease based on an average

[24]For the main analysis, I use survey data as the year the survey was taken. However, I acknowledge that merged practices may require time to adapt to changes. Therefore, I lag the survey outcomes by one year and re-estimate the regressions. The results remain robust, and they can be provided upon request.

rate of 58%. This is also noteworthy because long waiting times have been a persistent issue for the NHS, potentially leading to delayed treatment, which is unfavourable for patients.[25] Finally, Columns (7) and (8) reveal that patients also report reduced satisfaction with the practice's opening hours.

In summary, the findings indicate a substantial decline in patient experiences following practice mergers, particularly in continuity of care, access to care, and waiting times. Coupled with the finding of limited improvements in clinical quality, these results collectively point towards a negative impact of practice mergers on quality.

1.5.3 Effect of Practice Mergers on Financial Performance

I provide additional evidence on whether merged practices experience improved financial performance in Table 1.8.

Columns (1) and (2) show that revenue per patient remains unchanged after the merger, regardless of whether additional covariates are included. This finding is consistent with the earlier observation that QOF performance does not improve significantly following the merger. Since quality incentive rewards from QOF can positively influence practice revenue per patient, the absence of significant improvements in QOF achievement explains the lack of substantial changes in revenue per patient.

Furthermore, Columns (3) and (4) demonstrate a significant increase in revenue per FTE GP after the practice merger. The baseline result in Column (3) suggests a rise of approximately 20% in payment per FTE doctor post-merger, with the magnitude further increasing to 24% when additional covariates are incorporated in Column (4).[26] Given that general practices are typically small businesses owned by partnerships of several GPs, these GPs serve both as medical practitioners and as owners who can benefit from increased profits when their practices operate successfully. The increase in revenue per FTE GP after the merger suggests the potential for some financial gains arising from the merger.[27]

[25]See for instance the reports published by the Health Foundation.

[26]Given that RevPerGP and RevPerPatient are in logarithmic terms, we need to exponentiate the estimated coefficients, subtract 1, and multiply by 100 for interpretation.

[27]It is possible that GP partners hire salaried GPs who focus only on medical roles in practices. The FTE GP data I use covers both GP partners and salaried GPs. However, it's worth noting that salaried GPs usually constitute a smaller portion of the FTE GP workforce (as evidenced in reports such as GPOnline). Therefore, I believe using this measurement is a good way to assess the potential financial gains of the

In the sample, the average revenue per FTE GP is approximately £312,951. A straightforward calculation of the roughly 24% growth rate translates to an additional revenue of around £75,108 for the merged practices. These findings suggest that practice mergers can potentially lead to increased earnings compared to the pre-merger levels.

1.5.4 Effect of Practice Mergers on Workload

Finally, I present results on the impact of practice mergers on staff workload in Table 1.9. The upper panel shows the baseline results without additional covariates, while the bottom panel incorporates additional time-varying covariates.

The estimated coefficients indicate that doctors experience an increase in workload after the merger, while the workload of nurses and administrative staff remains unchanged. On average, each FTE doctor has to manage approximately 600 more patients after the merger. This observation aligns with the earlier finding of an increase in revenue per FTE GP post-merger. It is possible that merged practices achieve higher revenue per FTE GP by stretching resources and allocating a greater patient load to doctors. However, this finding is concerning as a high GP workload can lead to poorer practice performance and a negative perception of care from patients (Van den Hombergh et al., 2009; Shanafelt et al., 2010). The observed increase in doctor workload could contribute to the decline in patient outcomes that we find earlier.

Overall, the results highlight that mergers might lead to financial gains at the expense of a deterioration in quality. With these results in mind, I next investigate the persistence of these effects.

1.5.5 Dynamic Effects

It remains unclear whether the merger effects are short-lived or persist over time. To answer this question, I estimate the dynamic treatment effects using the following specification:

$$y_{it} = \gamma_{ic} + \gamma_{tc} + \sum_{\tau=-4}^{\tau=5} \delta_\tau (Treat_{ic} \times D_{ic}^{\tau}) + X_{it}\alpha + \varepsilon_{ict} \tag{1.2}$$

general practice.

where D_{ic}^{τ} are indicators equal to 1 for practice i of cohort c that are τ years after or before the merger year of the cohorts. $\tau = -5$ is left out as the reference year.[28] As in the main specification, the year of merger is dropped altogether for merged practices to prevent measurement errors. Standard errors are clustered at the dataset-specific practice level. Our coefficients of interest are the δ_τ, which capture the difference in outcomes Y between the merged practices and non-merged controls, for τ years after (or before) the merger versus five years before the merger. They shed light on how the effects of practice mergers evolve over time.

I present the estimated coefficients in Table B4 and the corresponding plots in Fig. 1.2 to Fig. 1.4. The pre-merger estimates support the parallel trend assumption, as the confidence intervals include zero for the pre-merger years. Fig. 1.2 reveals no significant change in revenue per patient following the merger. However, there is an immediate increase in revenue per FTE GP after the merger, and this positive effect persists over time. Fig. 1.3 shows no significant change in QOF achievement and some evidence of a decline in the population performance in the clinical indicators in the long run. Fig. 1.4 illustrates a substantial and long-lasting decrease in patient satisfaction following practice mergers.

One drawback of using the event study approach to explore the long-run effect is that the regression is estimated on a reduced sample. For example, when estimating the effect five years after the merger, only mergers occurring in 2014 are included in the analysis. To ensure credible estimates of the short-run and long-run effects, I perform another regression:

$$y_{it} = \gamma_{ic} + \gamma_{tc} + \beta_2(Treat_{ic} \times SR_{ict}) + \beta_3(Treat_{ic} \times LR_{ict}) + X_{it}\alpha + \varepsilon_{ict} \quad (1.3)$$

SR_{ict} and LR_{ict} are two dummy variables representing short-run and long-run respectively. SR_{ict} takes a value of 1 for one year and two years post-merger (corresponding to $\tau = 1$ and $\tau = 2$ in the previous specification), while LR_{ict} takes a value of 1 for $\tau = 3$ to $\tau = 5$. These dummies interact with the $Treat_{ic}$ variable. The estimated coefficients

[28] I choose $\tau = -5$ as the reference year to allow for the possible anticipation effects. The main findings remain consistent when $\tau = -1$ or $\tau = -2$ is used as the reference year. Also, time-varying covariates X_{it}, the same as those used in the static DiD regression, are included here. Results are however robust if no additional covariates are incorporated.

of these interaction variables provide the short-run and long-run effects.

I present the results in Table 1.10. Column (2) confirms an immediate and lasting increase in revenue per FTE GP. For clinical quality, a minor increase in QOF performance in the short term is observed, but this effect disappears in the long run (Column (3)). Moreover, there is evidence of a decline in the population performance in the clinical indicators in the long run (Column (5)). This suggests that clinical quality tends to decline in the long run after the merger. Turning to patient experiences, Columns (6) and (7) indicate that the negative effect of practice mergers on patient satisfaction becomes more pronounced in the long run. This suggests a worsening of patient satisfaction over time following the mergers.

In summary, these findings emphasize the enduring financial gains from practice mergers while also highlighting the long-term decline in quality.

1.5.6 Heterogeneity in the Impact of Mergers

The effect of mergers may differ depending on the types of general practices involved or the nature of the mergers. I explore potential heterogeneous effects along three dimensions: the pre-merger size of merging parties, whether the two merging parties are within the same geographical market or not, and the stated merger motivations. By investigating these aspects, I hope to shed light on the potential impact of mergers in different scenarios.

1.5.6.1 Heterogeneity of Mergers Effect based on the Size of Merging Parties

Mergers between small practices may result in more benefits than mergers between already large ones if there are economies of scale or scope (Gaynor et al., 2012). I therefore explore the potential heterogeneity in the impact of mergers based on the size of merging parties. I categorize the merging parties as small or large based on the number of registered patients one year before the merger, compared to the mean size across all practices in England. The mergers are then differentiated into three types: mergers between small practices, mergers between large practices, and mergers involving one large and one small practice. To examine the heterogeneous effects between these three types, I generate three dummy variables, each representing one of the three types. I then re-estimate Eq. (1.1), substituting $Treat_{ic} \times Post_{ct}$ with three interaction variables, each interacting one of the three dummies with $Post_{ct}$. The omitted group consists of the matched never-merged practices. The regression is estimated with time-varying covariates. The results are presented in Table 1.11.

Column (2) indicates that potential financial gains are achieved across all merger types. In terms of quality, Column (3) shows that mergers between small practices experience improvements in clinical quality, with a significant increase of approximately 1.2 percentage points in QOF performance. This effect is statistically similar to the effect of mergers between one small and one large practice but distinguishes itself from the effect of mergers between large practices. This finding suggests that mergers involving small practices may benefit from economies of scale or scope, resulting in better clinical outcomes. Regarding patient satisfaction, a consistent negative impact is observed across all types of mergers, as shown in Column (6) and Column (7). Notably, mergers between large practices exhibit the most significant drop in patient satisfaction levels.

I further perform a robustness check using more extreme values of the distribution to determine the size of merging parties, distinguishing between small and large practices based on the 25th and 75th percentiles. The results, presented in Table B5, confirm that mergers between small practices appear relatively more beneficial, while mergers between large practices seem most harmful. To better understand the underlying reasons for these heterogeneous effects, I use variables on staff workload as the dependent variables. The results in Table B6 indicate that only mergers between small practices exhibit an increase in patients per GP, although the effect is statistically the same as the merger effect of the other two types. When combined with the observation that these mergers also experience improvements in clinical quality, it suggests that mergers between small practices achieve a more efficient utilization of labour resources. Smaller general practices may find it easier to coordinate their activities post-merger, and the small group setting facilitates closer interactions between staff and patients. In contrast, larger practices may experience mismanagement issues, resulting in quality declines (Robinson, 1999 and Weyrauch et al., 1995).[29]

In summary, the analysis reveals that mergers between large practices exhibit the most substantial decline in quality, emphasizing the need for scrutiny in such cases. Conversely, mergers between small practices show potential benefits in terms of both financial

[29]This finding resonates with the conclusions of Gravelle et al. (2022), who observe that larger practices often exhibit poorer quality performance.

performance and patient care. Although the negative effect on patient satisfaction remains regardless of the merger types, the magnitude of this decline is much smaller for mergers involving small practices.

1.5.6.2 Heterogeneity of Mergers Effect: Within-market vs. Cross-market Mergers

In the English primary care market, two distinct types of mergers are observed: within-market mergers, where general practices within the same geographical market merge, and cross-market mergers, where practices from different geographical markets combine. While antitrust authorities frequently challenge the former, known as horizontal mergers, they rarely do so for the integration of providers across markets (Dafny, 2021).[30] In theory, it is unclear whether the effect on quality differs between within-market and cross-market mergers. Within-market mergers might increase market concentration and undermine competition, potentially leading to a decline in quality. On the other hand, cross-market mergers might face challenges in managing practices located far apart, which could also result in lower quality. An empirical analysis comparing the effects of these two types of mergers is required to address these ambiguities. This analysis will help contribute to the recent debate on whether antitrust authorities should scrutinize mergers involving parties in different markets (Dafny, 2021).

The results in Table 1.12 show no difference between these two types of mergers in both quality outcomes and financial performance. Importantly, clinical quality does not improve for both types of mergers, and patient satisfaction declines dramatically. This finding suggests that changes in market concentration may not be the dominating mechanism behind merger effects on quality. If it were, we would observe a more negative impact on quality for within-market mergers, which are influenced by market concentration changes, compared to cross-market mergers unaffected by such changes. Furthermore, the evidence that cross-market mergers lead to quality declines, especially in terms of patient experiences, adds support to recent calls for scrutiny of mergers involving parties in different markets (Dafny, 2021).

[30] Research on healthcare provider mergers has predominantly focused on within-market mergers, with limited exploration of cross-market mergers. Several studies examine the effects of cross-market mergers in the hospital sector and find that they can be detrimental, leading to significant price increases without commensurate benefits (see, for instance, Dafny et al., 2019; Lewis and Pflum, 2017; Schmitt, 2017).

1.5.6.3 Heterogeneity of Mergers Effect based on Different Motivations to Merge

It is of interest to explore the heterogeneous effects based on different claimed merger motivations. This analysis will help provide insights into whether mergers that aim to achieve efficiency can effectively improve quality through cooperation and coordination, and whether it is a beneficial strategy to save practices that are facing difficulties through mergers.

My manual collection identifies a total of 453 mergers that were proposed to achieve efficiency and 143 mergers with the aim to survive during the sample years. The remaining 191 mergers lack identifiable motivations and are classified as neutral mergers. I analyze if the merger effect differs across these three types: merge-to-survive, merge-to-achieve-efficiency, and neutral mergers.[31] My focus is on comparing the effect of mergers with the motivation to survive and mergers intended to achieve efficiency. It is important to interpret the results for the neutral merger group with caution due to the lack of clear motivations behind those mergers.

The results in Table 1.13 show that the positive effect on financial performance is persistent regardless of the claimed merger motivations (Column 2). However, the impact on quality varies. Mergers aimed at achieving efficiency seem to have the most detrimental effect on quality. This particular type of merger is the only one that does not improve QOF performance (Column 3) and actually leads to a decline in the population achievement of practices in treating their patients (Column 5). Also, they experience the largest drop in patient satisfaction levels. On the other hand, mergers with the motivation to survive experience an increase in clinical quality, and although they also experience a decline in patient satisfaction, the magnitude of the decrease is small.

Based on these findings, I recommend that it might be beneficial to save general practices in difficulties through mergers. For mergers that aim to achieve efficiency, the finding suggests that although they intend to do so, it is hard to realize.

1.5.7 Mechanism: Market Concentration or Mismanagement

The previous analysis demonstrates a predominantly negative effect of mergers on quality. This finding is consistent with multiple explanations. One possibility is that

[31]As a robustness check, I conduct an additional analysis where I examine mergers caused by owner retirement and failure separately, rather than combining them. The main findings hold.

mergers increase market concentration, which can lead to a decline in quality without the discipline of competition. This is especially relevant for within-market mergers. Another possible explanation is mismanagement after mergers, particularly relevant for cross-market mergers, where it might be challenging to manage multiple general practices that are far apart. I explore which explanation is the dominant one in this section.

1.5.7.1 Test the Market Concentration Mechanism

Theoretical models show that in systems where prices are centrally set by an outside body (e.g., government), encouraging competition will lead to quality improvement (Gaynor, 2006). The intuition is that if the price is exogenously set, the dimension of competition for suppliers to attract healthcare users comes down to quality, which will rise as a result. In my context, mergers may decrease the incentive to maintain high quality in the absence of competitive pressures.

To investigate this hypothesis, I focus on the subsample of within-market mergers, which are more susceptible to market concentration effects. I compare merged entities located in highly competitive markets with those in less competitive ones. Mergers in already concentrated markets would experience a larger change in market concentration after the merger compared to mergers in competitive markets. If we observe a persistent drop in quality only in already concentrated markets, it could suggest that changes in market concentration drive the decline in quality after mergers.

To determine the competitive environment, I consider the pre-merger competition level of the acquirer practices.[32] I define a relevant geographic market as a 2km radius around a practice.[33] I classify the merged entity as operating in a highly competitive market if the number of competitors of the acquirer practice before the merger is above the national average; otherwise, it is categorized as operating in a low competitive market.

The results presented in Table 1.14 show no difference in the merger effect on quality based on the pre-merger market concentration level. This implies that changes in market concentration may not be the driving force behind the merger effect on quality. I

[32]The treated unit in my analysis is defined as the combined merged entity. After the merger, the merged entity will use the acquirer practice as the main site. Hence, I focus on the acquirer's market to define the market's competitive level for the merged entity.

[33]Examining general practices situated in the West Midlands area, Santos et al. (2017) document that the mean distance to one's registered practice is 1.88km (median=1.48km).

reinforce this finding by redefining market competition levels based on the 25th and 75th percentiles of the distribution. Results provided in Table B7 provide further support for this finding.

As a robustness check, I use the Herfindahl-Hirschman Index (HHI) to measure market concentration. I first define each merged entity's market as a 2km radius around it. I then calculate the HHI index for this market in the years before and after the merger. Based on the calculated HHI index and the change in the HHI index, I categorize the merged entities into mergers that pose no competitive concerns, mergers that are unlikely to be challenged, and mergers that potentially raise significant competitive concerns.[34] Results presented in Table B8 show that the impact on quality remains similar across these three categories. Among the five quality measures, only the impact on patient recommendation rates differs between mergers with potential concerns and those without concerns. This further supports the notion that changes in market concentration are not the primary driver of changes in quality.

A final robustness check is provided in Table B9 and Table B10, where I examine the subsample of large and small practice mergers separately. By doing so, I essentially control for the potential influence of mismanagement, allowing me to determine if market concentration indeed plays any significant role. Within each subsample, I compare the effect of within-market and cross-market mergers. The results show no significant difference between these types of mergers when accounting for the potential impact of mismanagement. This further supports that market power is not the main explanation behind quality change after the mergers. This finding resonates with the conclusions of Eliason et al. (2020) and Cutler et al. (2017), who, in their studies of the U.S. dialysis industry, similarly find that market concentration alone cannot fully explain the decrease in patient outcomes following mergers.

1.5.7.2 Test the Mismanagement Mechanism

I examine the subsample of cross-market mergers to test the mismanagement mechanism. Specifically, I compare mergers of different sizes. If mismanagement is the primary mechanism, we should observe a more pronounced negative effect on quality in mergers involving large practices, as they may face greater challenges in managing large branch

[34] This classification follows the merger guidelines by the Federal Trade Commission.

surgeries.

The results presented in Table 1.15 provide evidence supporting the mismanagement hypothesis. Mergers involving large practices exhibit lower QOF performance (Column (1)) compared to those involving small practices. A robustness check determining the size of merging parties based on more extreme values of the distribution is given in Table B11. The results confirm that mergers between large practices experience lower clinical quality compared to mergers involving small practices (Column (3) to Column (5)). These findings collectively indicate that mismanagement plays a central role in explaining the observed quality decline after mergers.

1.6 Robustness Checks

1.6.1 Alternative Staggered Roll-out Estimates

I use the estimator developed by Callaway and Sant'Anna (2021) as an alternative to the main specification. Their estimator is designed specifically for staggered roll-out designs. The results, provided in Table B12, confirm the robustness of the findings. Financial performance improves, as indicated by a significant increase in revenue per doctor (Column (2)). However, quality declines, with evidence of a decrease in clinical quality (Column (5)) and a dramatic decline in patient experience (Columns (6) and (7)).

1.6.2 Alternative Control groups

In the main analysis, I construct the control group using PSM by selecting the three closest neighbours that closely resemble the merging practices from a large set of never-merged practices. In this subsection, I examine the sensitivity of the results to the choice of control group.

I re-construct the control group by (i) selecting the closest 1/5/7 never-merged practices by PSM, (ii) using all never-merged practices, and (iii) using future mergers occurring at least two years later as the control group. I replicate the main analysis with the treatment group and these alternative control groups.

Results with varying numbers of matches selected by PSM are presented in Table B13; results with all never-merged practices are presented in Table B14; and the results using mergers occurring at least two years later as the control group are shown in Table B15. The rationale behind using future mergers as the control group is that the timing of mergers is potentially random (as shown in Table B2). Therefore, using not-yet-merged practices as the control group helps ensure a plausible causal estimate. The main findings remain robust across these alternative controls. After the merger, revenue per doctor increases significantly. However, true clinical quality does not improve, and patient experience deteriorates significantly.

1.6.3 Reverse Causality

A potential issue of reverse causality may arise for practices where poor quality is the primary motivation to merge. As I have explained, practices that receive poor quality ratings from the CQC may seek to improve their ratings by merging with other practices. To address this potential issue, I exclude mergers involving merging parties rated as inadequate and analyze the resulting subsample. I present the results in Table B16. The main conclusion holds.

1.7 Conclusion

Over the past few years, there has been an increasing trend of general practice mergers in England, aligning with the current national policy promoting collaboration among practices. However, the impact of these mergers remains unknown. This study provides the first empirical evidence of the impact of general practice mergers in England by assembling a unique dataset of all such mergers between 2014 and 2018.

I find that while there is some indication of financial gains for practices following mergers, the overall impact on quality is negative. On average, merged practices experience an increase in revenue by approximately £75,108, a potentially desirable outcome for addressing the prevalent doctor shortage issue in England through increased financial incentives. However, this financial improvement is not matched by maintaining the same level of patient outcomes. Clinical quality tends to decline in the long run. In addition, patient satisfaction shows a significant decrease. Specifically, there is a substantial 10% decline in the mean satisfaction rate for continuity of care. Additionally, access to care and waiting times also suffer, with mean satisfaction rates experiencing a 5% and 4%

decrease, respectively.

The study explores two potential mechanisms, changes in market concentration and mismanagement, to explain these declines in quality. The analysis reveals that the impact of mergers on quality differs primarily based on the size of the merging parties rather than their geographical proximity. Notably, larger merged entities are associated with worse quality outcomes. This finding suggests that mismanagement plays a role in driving the quality declines, as larger merged entities are more susceptible to mismanagement challenges. On the other hand, there is no difference between the effect of within-market and cross-market mergers, indicating that changes in market concentration do not appear to be the dominating factor. Formal tests differentiating between these two mechanisms provide strong evidence that mismanagement, rather than changes in market concentration, is the principal explanation for the observed quality declines.

These findings have important policy implications, particularly as much of the concentration in physician markets remains unnoticed by regulatory authorities (Gravelle et al., 2019). In light of these findings, I recommend that the government take into account the negative effects of general practice mergers on patients before approving further mergers. While merger guidelines typically focus on assessing potential anti-competitive implications, this research demonstrates that mergers can have negative consequences for patients, regardless of changes in market concentration. Additionally, special attention should be given to mergers involving large practices, given their greater detrimental impact on quality. Although these mergers may intend to bring benefits through achieving efficiency, it is hard to realize due to mismanagement. On the other hand, mergers between small practices may hold the potential for positive outcomes.

Tables

Table 1.1: Variable Definitions

Variables	Definition	Duration
Outcome Variables		
qofOutcome	the percentage of total maximum available points that the practice achieve	2013-19
ClinPA	the weighted average of the performance measures, taken over the clinical indicators and disease areas which were consistently defined between 2013 and 2019.	2013-19
ClinPA2	the weighted average of the performance measures, taken over the clinical indicators and disease areas which were consistently defined between 2013 and 2018.	2013-18
OverallSat	the proportion of patients satisfied with their practice on an overall level	2013-19
Recommend	the proportion of patients who would definitely or probably recommend their surgery to someone who has just moved to their local area	2013-18
OpenHrsSat	the proportion of patients satisfied with their practice opening hours	2013-18
Continuity	the proportion of patients who have a preferred doctor and could always or almost always see their preferred doctor	2013-19
AppointSat	the proportion of patients whose overall experience of making an appointment is satisfied	2013-19
WaitSat	the proportion of patients who report that their waiting time at surgery is normally not too long	2013-18
RevPerPatient	the revenue per patient in logarithm form	2013-19
RevPerGP	the revenue per Full-time Equivalent (FTE) of doctors in logarithm form	2013-19
PatientsPerGP	number of patients per FTE doctor	2013-19
PatientsPerNurse	number of patients per FTE nurse	2013-19
PatientsPerAdmin	number of patients per FTE admin	2013-19
Practice Characteristics		
PreHYP	prevalence of hypertension at practice level	2013-19
PreSTIA	prevalence of stroke and transient ischaemic attack at practice level	2013-19
PreCHD	prevalence of coronary heart disease at practice level	2013-19
PreAST	prevalence of asthma at practice level	2013-19
PreCOPD	prevalence of chronic obstructive pulmonary disease at practice level	2013-19
PreCAN	prevalence of cancer at practice level	2013-19
PreDM	prevalence of diabetes mellitus at practice level	2013-19
PreMH	prevalence of mental health at practice level	2013-19
PreEP	prevalence of epilepsy at practice level	2013-19
GpFTE	total number of FTE doctors at practice level	2013-19
NurseFTE	total number of FTE nurses at practice level	2013-19
AdminFTE	total number of FTE admin staff at practice level	2013-19
Dispensing	an indicator variable of practice's dispensing status, equals to one if it is a dispensing practice	2013-19
NumPatient	total number of registered patients at each practice	2013-19
NumComp	number of competitors within 2km of each practice	2013-19
Area Characteristics		
Urban	an indicator variable equals to one if the area is classified as urban	2013-19
IMD	the composite index of Multiple Deprivation rank for each area	2013-19

Table 1.2: Summary Statistics of the Variables

Variables	(1) Full Sample	(2) Merged	(3) Never-merged	(4) T-statistic
Outcomes				
qofOutcome	95.55	94.51	95.57	3.89
ClinPA	49.99	49.54	50.00	3.40
ClinPA2	79.44	78.55	79.47	4.45
OverallSat	85.25	85.03	85.25	0.77
Recommend	77.80	76.31	77.84	4.04
Continuity	35.18	36.73	35.15	2.93
AppointSat	74.88	75.51	74.86	1.64
WaitSat	58.87	59.27	58.86	0.91
OpenHrsSat	77.29	77.34	77.29	0.20
RevPerPatient	162.84	176.30	163.81	0.47
RevPerGP	312951.43	339548.34	312286.62	1.59
PatientsPerGP	2124.04	2271.34	2120.36	1.37
PatientsPerNurse	5060.89	4613.52	5071.76	3.84
PatientsPerAdmin	1024.22	962.77	1025.72	2.12
Practice characteristics				
NumPatient	7706.65	6779.77	7729.84	6.71
NumComp	8.19	8.26	8.19	0.31
PreHYP	15.03	14.48	15.04	0.57
PreSTIA	1.81	1.76	1.81	0.66
PreCHD	3.37	3.32	3.37	0.26
PreAST	6.32	6.01	6.33	0.96
PreCOPD	2.05	2.07	2.05	0.19
PreCAN	2.69	2.41	2.70	1.79
PreMH	1.01	0.99	1.01	0.38
PreDM	7.23	7.04	7.23	0.61
PreEP	0.82	0.85	0.82	1.06
GpFTE	4.36	3.79	4.38	5.98
NurseFTE	2.02	1.94	2.02	1.30
AdminFTE	8.23	7.56	8.25	3.74
Dispensing	0.13	0.06	0.13	8.82
Local characteristics				
Urban	0.85	0.89	0.85	4.55
IMD	13809.15	12531.29	13841.13	4.69

Notes: The displayed values represent the mean, both for the full sample (Column (1)) and the distinction between merged and never-merged practices (Column (2) and Column (3)). For merged practices, I only consider observations from the year before the merger takes place, taking into account both the acquirer practice and the target practice. Column (4) provides the t-statistic comparing the difference in means between the two subsamples in Column (2) and Column (3). RevPerPatient and RevPerGP are presented in their non-logarithmic format to facilitate comparison.

Table 1.3: Factors that Predict GP Practice Mergers

Variables	Cohort2014	Cohort2015	Cohort2016	Cohort2017	Cohort2018
NumComp	0.012	0.008	-0.020	0.004	0.004
	(0.021)	(0.017)	(0.017)	(0.014)	(0.015)
NumPatient	8.31e-06	-0.000155***	2.31e-05	0.000112***	4.51e-05
	(5.98e-05)	(5.61e-05)	(3.32e-05)	(2.49e-05)	(3.25e-05)
IMD	1.03e-05	5.02e-06	-2.60e-05*	-1.45e-05	-1.53e-06
	(1.59e-05)	(1.42e-05)	(1.35e-05)	(1.13e-05)	(1.14e-05)
PreHYP	0.056	0.033	0.067*	0.069*	0.0629
	(0.055)	(0.049)	(0.039)	(0.041)	(0.041)
PreSTIA	-0.145	-0.0738	0.109	-0.399*	-0.039
	(0.338)	(0.288)	(0.179)	(0.237)	(0.232)
PreCHD	-0.144	-0.141	-0.100	0.112	-0.174
	(0.213)	(0.186)	(0.136)	(0.148)	(0.160)
PreAST	0.214**	-0.015	0.027	0.117*	-0.024
	(0.100)	(0.086)	(0.072)	(0.070)	(0.072)
PreCOPD	-0.043	0.085	0.295**	0.054	-0.040
	(0.186)	(0.150)	(0.125)	(0.127)	(0.127)
PreCAN	0.181	0.066	-0.398**	-0.099	-0.098
	(0.247)	(0.209)	(0.166)	(0.151)	(0.157)
PreMH	-0.126	0.003	0.111	0.053	0.035
	(0.317)	(0.160)	(0.166)	(0.124)	(0.128)
PreDM	0.030	-0.028	-0.156**	0.020	0.062
	(0.080)	(0.069)	(0.064)	(0.051)	(0.053)
PreEP	0.201	0.162	0.030	0.004	0.938***
	(0.506)	(0.406)	(0.067)	(0.365)	(0.337)
GpFTE	0.114**	0.039	0.081*	-0.046	-0.086*
	(0.056)	(0.054)	(0.044)	(0.043)	(0.044)
NurseFTE	0.016	0.083	0.070	0.042	0.105**
	(0.100)	(0.073)	(0.067)	(0.061)	(0.052)
AdminFTE	-0.026	0.118***	-0.018	-0.041*	0.023
	(0.041)	(0.037)	(0.024)	(0.024)	(0.024)
Urban	0.525	-0.145	0.373	0.108	-0.044
	(0.450)	(0.348)	(0.388)	(0.309)	(0.274)
Dispensing	-1.051**	-1.159**	-0.442	-0.276	0.170
	(0.514)	(0.454)	(0.399)	(0.320)	(0.267)
Observations	6,390	6,483	6,500	6,525	6,285

Notes: I run separate logit regressions for each year of mergers, i.e. each cohort. In each regression, I include all practices that have not yet merged in the given year. For example, the regression indicated by the Column heading Cohort2015 excludes practices that merged in 2014. For merged practices, I use the constructed pseudo-merged values from the previous year before the merger in the regression. The dependent variable is whether the practice merged. Robust standard errors are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.4: Balancing Test from PSM

Variables	Sample	Mean		N		T-statistic	%SB
		Treated	Controls	Treated	Controls		
NumComp	Unmatched	8.156	8.192	787	46591	0.150	-0.503
	Matched	8.156	8.008	787	2277	0.523	2.125
NumPatient	Unmatched	8835.389	7649.890	786	45120	5.214	21.593
	Matched	8835.389	9054.014	786	2277	0.835	-3.463
IMD	Unmatched	12377.188	13841.128	783	45909	4.622	-16.145
	Matched	12377.188	12159.543	783	2277	0.595	2.449
PreHYP	Unmatched	14.247	15.043	780	44778	0.803	-0.540
	Matched	14.247	14.246	780	2277	0.003	0.012
PreSTIA	Unmatched	1.747	1.813	780	44778	0.803	-0.553
	Matched	1.747	1.750	780	2277	0.156	-0.623
PreCHD	Unmatched	3.292	3.370	780	44778	0.463	-0.316
	Matched	3.292	3.301	780	2277	0.224	-0.907
PreAST	Unmatched	6.065	6.326	780	44778	0.806	-0.543
	Matched	6.065	6.041	780	2277	0.510	2.040
PreCOPD	Unmatched	2.034	2.049	780	44778	0.116	-0.079
	Matched	2.034	2.031	780	2277	0.069	0.280
PreCAN	Unmatched	2.383	2.698	780	44778	1.926	-1.309
	Matched	2.383	2.384	780	2277	0.020	-0.083
PreMH	Unmatched	0.983	1.010	780	44778	0.439	-0.300
	Matched	0.983	0.981	780	2277	0.077	0.272
PreDM	Unmatched	6.996	7.231	780	44778	0.742	-0.507
	Matched	6.996	7.027	780	2277	0.403	-1.647
PreEP	Unmatched	0.849	0.822	780	44778	1.257	0.906
	Matched	0.849	0.844	780	2277	0.496	1.956
GpFTE	Unmatched	4.903	4.377	784	44476	3.798	15.168
	Matched	4.903	4.963	784	2277	0.382	-1.597
NurseFTE	Unmatched	2.484	2.021	772	43407	5.840	23.719
	Matched	2.484	2.580	772	2277	1.022	-4.141
AdminFTE	Unmatched	9.686	8.251	773	43901	5.271	21.843
	Matched	9.686	9.876	773	2277	0.610	-2.558
Urban	Unmatched	0.898	0.851	787	46473	4.650	14.715
	Matched	0.898	0.898	787	2277	0.024	0.098
Dispensing	Unmatched	0.086	0.128	787	46655	5.071	-14.568
	Matched	0.086	0.091	787	2277	0.469	-1.822

Notes: This table shows the extent of the balance problem solved through PSM. For each covariate, I compare the difference between treated units and untreated units twice: once using the raw, unmatched sample and another with the matched sample.

Table 1.5: Effect of GP Practice Mergers on Clinical Quality

Variable	(1) qofOutcome	(2) qofOutcome	(3) ClinPA	(4) ClinPA	(5) ClinPA2	(6) ClinPA2
Treat × Post	0.690***	0.865***	-0.232*	0.054	-0.577***	-0.240
	(0.231)	(0.310)	(0.118)	(0.161)	(0.204)	(0.288)
Additional Controls	No	Yes	No	Yes	No	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	15,536	15,536

Notes: The coefficient of *Treat* × *Post* represents the impact of the merger on practice clinical quality. Dependent variables are presented in percentage points. For each dependent variable, results are presented twice: once displaying baseline outcomes without additional time-varying covariates and once incorporating additional covariates (not shown to save space). The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (5) and (6) have fewer observations because ClinPA2 is constructed for the subsample covering the years 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.6: Effect of GP Practice Mergers on Patient Experience: Overall

Variable	(1) OverallSat	(2) OverallSat	(3) Recommend	(4) Recommend
Treat $\times$ Post	-2.771***	-2.725***	-2.379***	-2.730***
	(0.336)	(0.414)	(0.511)	(0.631)
Additional Controls	No	Yes	No	Yes
Practice FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	18,144	18,144	15,536	15,536

Notes: The coefficient of *Treat* $\times$ *Post* represents the merger effect on overall patient experiences. Dependent variables represent patients' overall satisfaction levels and are expressed in percentage points. For each dependent variable, results are presented twice: once displaying baseline outcomes without additional time-varying covariates and once incorporating additional covariates (not shown to save space). The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Column (3) and (4) have fewer observations because measures of recommendation rate are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.7: Effect of GP Practice Mergers on Patient Experience: More

Variable	(1) Continuity	(2) Continuity	(3) AppointSat	(4) AppointSat	(5) WaitSat	(6) WaitSat	(7) OpenHrsSat	(8) OpenHrsSat
Treat $\times$ Post	-3.889***	-3.425***	-4.633***	-4.120***	-2.200***	-2.547***	-1.292***	-1.961***
	(0.486)	(0.656)	(0.474)	(0.579)	(0.515)	(0.692)	(0.395)	(0.515)
Additional Controls	No	Yes	No	Yes	No	Yes	No	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	15,536	15,536	15,536	15,536

Notes: The coefficient of *Treat* $\times$ *Post* represents the merger effect on patient experiences. Dependent variables reflect various aspects of patient experiences and are measured in percentage points. For each dependent variable, results are presented twice: once displaying baseline outcomes without additional time-varying covariates and once incorporating additional covariates (not shown to save space). The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Column (5) to (8) have fewer observations because measures of WaitSat and OpenHrsSat are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.8: Effect of GP Practice Mergers on Financial Performance

Variable	(1) RevPerPatient	(2) RevPerPatient	(3) RevPerGP	(4) RevPerGP
Treat × Post	0.009	0.026	0.179***	0.215***
	(0.012)	(0.016)	(0.021)	(0.028)
Additional Controls	No	Yes	No	Yes
Practice FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144

Notes: The coefficient of *Treat* × *Post* indicates the effect of the merger on practice financial performance. The dependent variables for the financial performance outcomes are in logarithmic form. For each dependent variable, results are presented twice: once displaying baseline outcomes without additional time-varying covariates and once incorporating additional covariates (not shown to save space). The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors, clustered at the cohort-practice level, are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.9: Effect of GP Practice Mergers on Workload

Variable	(1) PatientsPerGP	(2) PatientsPerNurse	(3) PatientsPerAdmin
Treat × Post	589.5***	-2.803	33.72
	(146.9)	(287.1)	(77.12)
Additional Controls	No	No	No
Practice FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	18,144	18,144	18,144
Variable	(4) PatientsPerGP	(5) PatientsPerNurse	(6) PatientsPerAdmin
Treat × Post	619.9***	-7.866	35.74
	(162.7)	(308.2)	(73.15)
Additional Controls	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	18,144	18,144	18,144

Notes: The coefficient of *Treat × Post* indicates the impact of the merger on the practice workforce. The upper panel presents the baseline results without additional covariates, while the bottom panel includes additional covariates. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.10: The Short-run and Long-run Effect of GP Practice Mergers

Variables	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat $\times$ SR	0.024	0.221***	0.970***	0.104	-0.150	-2.505***	-2.640***
	(0.016)	(0.027)	(0.309)	(0.161)	(0.289)	(0.421)	(0.631)
Treat $\times$ LR	0.033	0.197***	0.538	-0.099	-0.672*	-3.410***	-3.160***
	(0.021)	(0.034)	(0.412)	(0.210)	(0.371)	(0.529)	(0.853)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	15,536	18,144	15,536

Notes: This table presents the short and long-run merger effect from a DiD specification with covariates. The SR represents the short-run period, which covers 1 to 2 years after the merger, while the LR represents the long-run period, covering 3 to 5 years after the merger. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.11: Heterogeneity Effects: Different Pre-merger Sizes of Merging Parties

Size of Merging Parties	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Small Practices Merge	0.023	0.253***	1.163***	0.140	-0.114	-2.551***	-2.058***
	(0.026)	(0.036)	(0.416)	(0.218)	(0.403)	(0.493)	(0.795)
Small and Large Merge	0.033*	0.175***	0.473	-0.094	-0.469	-2.460***	-3.074***
	(0.018)	(0.031)	(0.362)	(0.186)	(0.322)	(0.540)	(0.804)
Large Practices Merge	0.024	0.197***	-0.433	0.174	-0.125	-4.717***	-6.357***
	(0.021)	(0.059)	(0.570)	(0.291)	(0.478)	(1.066)	(1.632)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,393	17,393	17,393	17,393	14,890	17,393	14,890
Test equality of coefficients (p-value)							
Small Practices Merge=Small and Large Merge	0.717	0.037	0.150	0.347	0.430	0.881	0.300
Small Practices Merge=Large Practices Merge	0.955	0.319	0.009	0.915	0.984	0.039	0.010

Notes: This table examines the heterogeneity effects of mergers based on the sizes of the merging parties from a DiD specification with time-varying covariates. I compare three cases: mergers between small practices, mergers between large practices, and mergers involving both small and large practices. The size of a practice is determined by the number of registered patients in the pre-merger year, where a practice is considered small if its patient number is below the national average and large if it is above. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.12: Heterogeneity Effects: Within-market vs. Cross-market mergers

Merger Types	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Within-market Mergers	0.025	0.218***	0.979***	0.154	-0.078	-2.818***	-2.822***
	(0.016)	(0.028)	(0.325)	(0.168)	(0.300)	(0.443)	(0.679)
Cross-market Mergers	0.029	0.207***	0.555	-0.215	-0.695	-2.473***	-2.472**
	(0.028)	(0.041)	(0.468)	(0.243)	(0.435)	(0.633)	(0.980)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	15,536	18,144	15,536
Test equality of coefficients (p-value)							
Within-market=Cross-market	0.890	0.753	0.347	0.117	0.141	0.587	0.728

Notes: This table compares the effects of within-market mergers and cross-market mergers using a DiD specification with time-varying covariates. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.13: Heterogeneity Effects: Different Claimed Merger Motivations

Merger Motivations	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
For survival	0.019	0.202***	1.719***	0.475	0.395	-1.558**	-1.839
	(0.014)	(0.030)	(0.506)	(0.294)	(0.575)	(0.623)	(1.296)
For efficiency	0.040	0.228***	0.234	-0.245	-0.689*	-3.826***	-3.942***
	(0.025)	(0.038)	(0.387)	(0.203)	(0.364)	(0.553)	(0.742)
Neutral mergers	0.014	0.207***	1.068**	0.139	-0.035	-2.149***	-1.740*
	(0.023)	(0.036)	(0.457)	(0.212)	(0.354)	(0.595)	(0.957)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	15,536	18,144	15,536
Test equality of coefficients (p-value)							
For survival=For efficiency	0.368	0.454	0.006	0.018	0.062	0.002	0.120

Notes: This table shows the heterogeneous effect based on claimed merger motivations from a DiD specification with time-varying covariates. Mergers motivated by failure and owner retirement are grouped as "for survival" mergers. Mergers aimed at achieving efficiency are labelled as "for efficiency" mergers. Neutral mergers refer to mergers where the motivations are not identifiable. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.14: Test Market Concentration Mechanism

Pre-merger Market Competition	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
Low comp	0.905**	0.049	-0.218	-2.671***	-3.553***
	(0.418)	(0.215)	(0.432)	(0.639)	(1.108)
High comp	1.526***	0.429*	0.299	-2.232***	-2.024**
	(0.469)	(0.236)	(0.450)	(0.602)	(0.899)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	11,685	11,685	9,994	11,685	9,994
Test equality of coefficients (p-value)					
Low comp=High comp	0.206	0.141	0.270	0.536	0.187

Notes: This table examines whether changes in market concentration are responsible for the decline in quality after mergers. Focusing on the subsample of within-market mergers, I compare merged entities in highly competitive markets with those in low competitive markets. The level of competition is determined by the number of competitors the acquirer has one year before the merger. A market with an above-average number of competitors is considered highly competitive, while a market with a below-average number of competitors is considered low competition. The market is defined as a 2km radius around each practice. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.15: Test Incompetency Mechanism

Size of Merging Parties	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
Small Practices Merge	0.433	-0.461	-1.149	-2.920***	-2.366
	(0.890)	(0.492)	(0.952)	(0.971)	(1.536)
Small and Large Merge	0.110	-0.395	-0.686	-3.286***	-2.978**
	(0.609)	(0.274)	(0.430)	(0.878)	(1.309)
Large Practices Merge	-1.625*	-0.725	-1.828**	-5.034**	-5.471**
	(0.957)	(0.509)	(0.744)	(1.977)	(2.683)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	6,221	6,221	5,333	6,221	5,333
Test equality of coefficients (p-value)					
Small Practices Merge=Small and Large Merge	0.739	0.897	0.633	0.741	0.735
Small Practices Merge=Large Practices Merge	0.060	0.673	0.517	0.277	0.267

Notes: This table examines whether incompetency is responsible for the decline in quality after mergers. Focusing on the subsample of cross-market mergers, I compare merged entities of different sizes. I compare three cases: mergers between small general practices, mergers between large general practices, and mergers involving both small and large general practices. The size of a practice is determined by the number of registered patients in the pre-merger year, where a practice is considered small if its patient number is below the national average and large if it is above. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates are not shown to save space. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figures

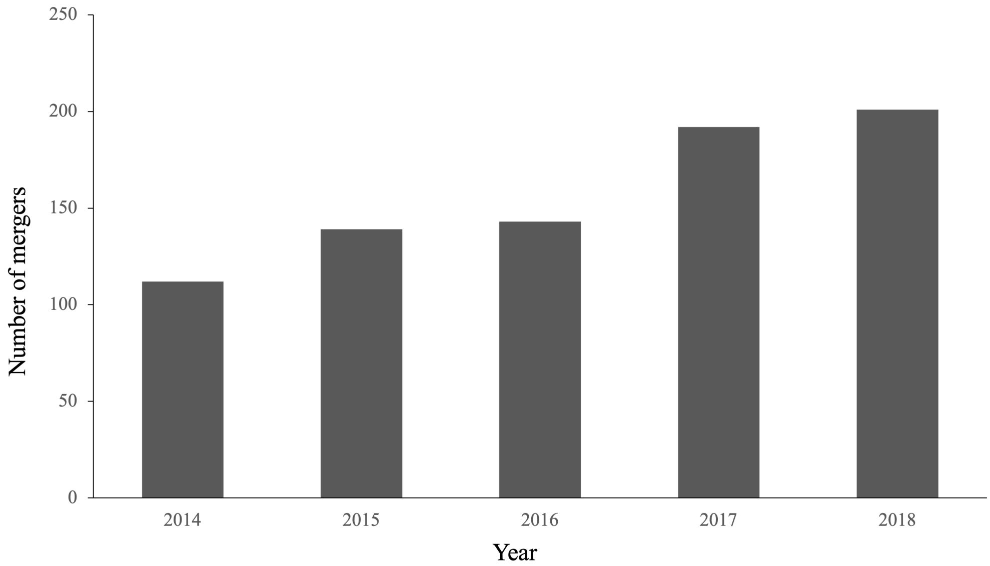
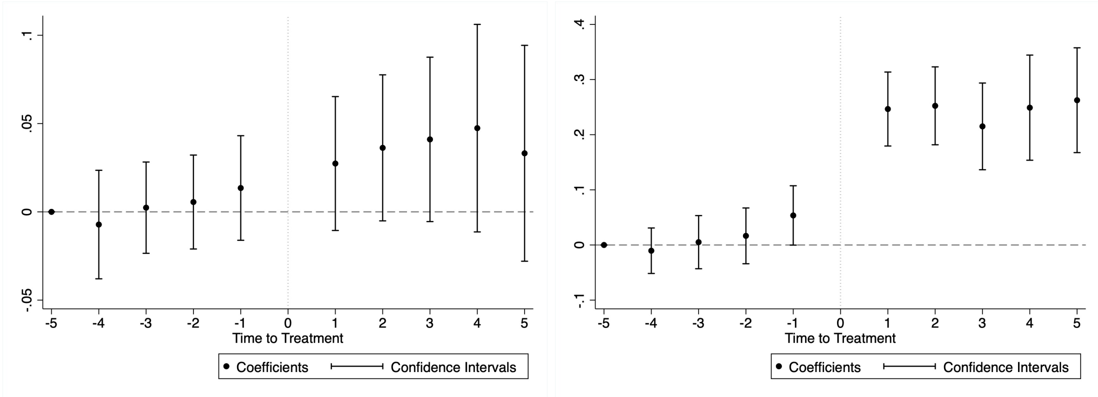


Figure 1.1: Number of GP Practice Mergers in England

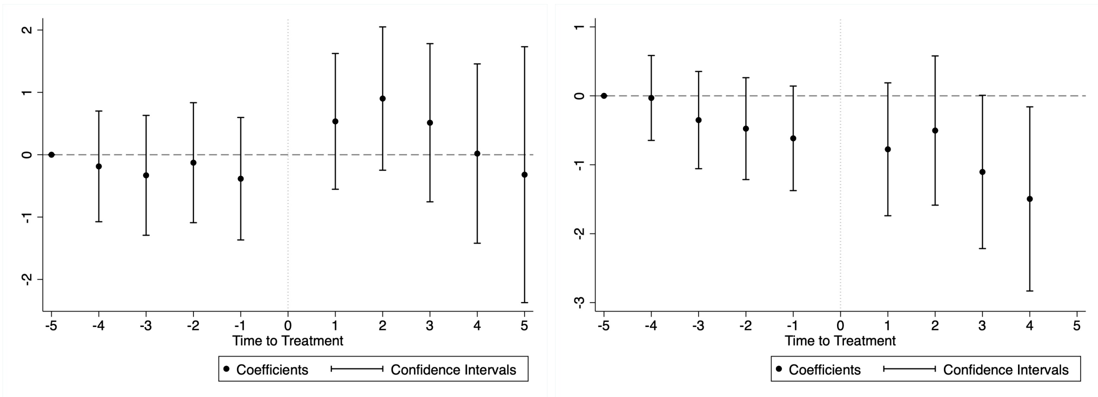
Notes: This figure displays the yearly count of GP practice mergers in England. Each bar represents the number of merger events that take place in a particular year. A merger event refers to the amalgamation of an acquiring practice and a target practice. The data used in this figure is collected and calculated by the author.



(a) Revenue per patient **(b)** Revenue per FTE GP

Figure 1.2: Estimation of Dynamic Treatment Effect on Financial Performance

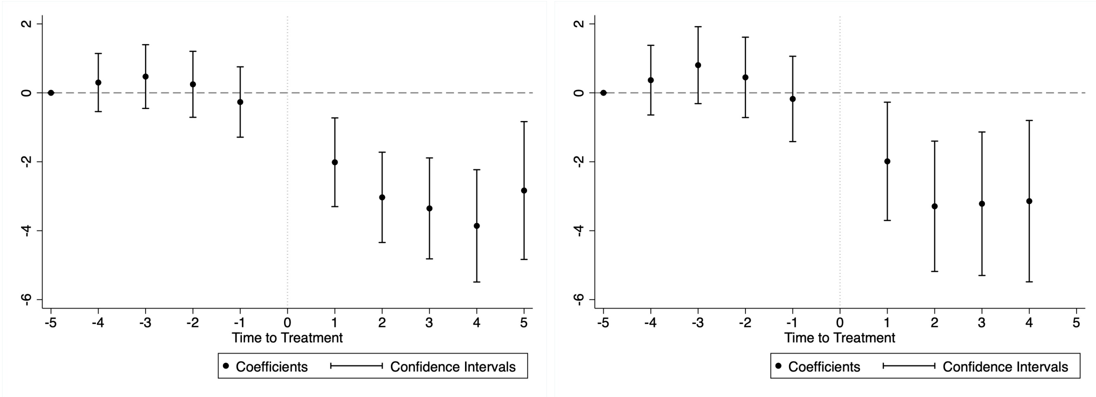
Notes: This figure plots the estimated coefficients from the dynamic DID with covariates. The reference period is represented by -5. The vertical axis represents the magnitude of the effect relative to the effect in period -5, which is normalized to 0. The 95% confidence intervals are included. The confidence intervals of estimates for pre-merger years contain 0, indicating that the parallel trend assumption is satisfied. Standard errors are clustered at the cohort-practice level.



(a) QOF Performance **(b)** True Clinical Quality

Figure 1.3: Estimation of Dynamic Treatment Effect on clinical Quality

Notes: This figure plots the estimated coefficients from the dynamic DID with covariates. In Panel (a), the dependent variable is qofOutcome, while in Panel (b), the dependent variable is ClinPA2. The reference period is represented by -5. The vertical axis represents the magnitude of the effect relative to the effect in period -5, which is normalized to 0. The 95% confidence intervals are included. The confidence intervals of estimates for pre-merger years contain 0, indicating that the parallel trend assumption is satisfied. Standard errors are clustered at the cohort-practice level.



(a) Overall Satisfaction Level **(b)** Recommendation Rate

Figure 1.4: Estimation of Dynamic Treatment Effect on Patient Experience

Notes: This figure plots the estimated coefficients from the dynamic DiD with covariates. The reference period is represented by -5. The vertical axis represents the magnitude of the effect relative to the effect in period -5, which is normalized to 0. The 95% confidence intervals are included. The confidence intervals of estimates for pre-merger years contain 0, indicating that the parallel trend assumption is satisfied. Standard errors are clustered at the cohort-practice level.

Appendix

Appendix A: Collect GP Practice Merger Data

I manually assemble the GP practice merger information through two steps. The first step involves compiling a comprehensive list of general practices that might have engaged in mergers during the sample period. Then, I perform a manual check to determine if each practice on the list participated in a merger and, if so, its merging partner(s).

In order to generate a comprehensive list of potential general practice mergers, three different data sources are used. The first data source is the QOF data published annually by the NHS. When the NHS published the QOF data, they also made available the "Practice Validation Comments" documents, which contain important notes made by the area team staff during their data validation process. This document includes the status of a practice, such as if the practice is merged or closed. My investigation reveals that some practices classified as "closed" in this document are actually merged ones. The status "closed" could either mean the practice ceased operation permanently or it had merged with other practice(s) and is now operating under a new contract. As such, I extract the list of both "merged practice" and "closed practice" from this document.

The second data source is the Ebranchs data published monthly by NHS Digital. This data includes a list of GP branch surgeries in England. These branch surgeries could be merged as, according to NHS England, when GP practices merge, one site becomes the main site, while the other site(s) will operate as branches.[35] I, therefore, extract the list of branch surgeries and verify in the second step if they are merged ones.

While both of these aforementioned datasets are necessary, they have limitations. The QOF data excludes practices that do not participate in this scheme, while the Ebranchs data only includes the main branch surgeries and excludes peripheral ones. To ensure a comprehensive list is composed, a third data source is used. I specifically utilize a dataset that incorporates the full list of active practices in England. One such dataset is the "Practice Level List Size" dataset published monthly by NHS England. By comparing the list of active practices from year to year, I identify all practices that became inactive over the period. These inactive practices could be either merged ones that stopped reporting data to the NHS under their old contract or practices that were closed permanently.

By using these three data sources, I am confident that a complete list of potential merged practices is documented in my first step of the data collection process.

The second step is to go through the list I generate and determine if each of the practices on the list was merged, and if so, with whom. This information is obtained manually by searching local news, reports, and individual practice website announcements. In cases where these sources do not provide relevant information, I turn to some consumer review websites, such as Google Reviews, BestCareCompare, and iWantGreat-

[35] See the information published by NHS England

Care websites, where some registered patients may have left comments suggesting the status change of their practice, including the mergers.[36] In particular, I try to identify the motivation behind each practice merger when I collect the data. For most merged practices, their motivation is explicitly outlined.[37]

[36]One example of such review is "This really was a decent surgery when it was Greenway. worst thing they ever done was merge the 2 together".

[37]For example, the merger between Bank Street Surgery and Castlehead Medical Centre is due to the owner's retirement at Bank Street Surgery. This merger motivation is clearly outlined in the practice website announcement to patients.

Appendix B: Additional Tables

Table B1: Factors that Predict the GP Practice Mergers

Variables	Cohort2014	Cohort2015	Cohort2016	Cohort2017	Cohort2018
NumComp	0.003	0.005	-0.026*	0.003	-0.014
	(0.019)	(0.015)	(0.015)	(0.012)	(0.013)
NumPatient	-9.29e-05	-0.0001**	-0.0001***	-4.95e-05	-0.0002***
	(6.02e-05)	(5.62e-05)	(3.94e-05)	(3.85e-05)	(3.82e-05)
IMD	1.99e-05	1.06e-06	-1.92e-05*	-2.71e-06	-5.62e-06
	(1.41e-05)	(1.24e-05)	(1.10e-05)	(9.45e-06)	(9.60e-06)
PreHYP	0.037	0.057	0.073**	0.046	0.061*
	(0.047)	(0.040)	(0.031)	(0.033)	(0.033)
PreSTIA	-0.138	0.045	0.096	0.056	-0.118
	(0.274)	(0.225)	(0.201)	(0.167)	(0.184)
PreCHD	-0.134	-0.251	-0.110	-0.031	-0.132
	(0.182)	(0.154)	(0.105)	(0.117)	(0.127)
PreAST	0.137	0.011	-0.044	0.062	-0.025
	(0.087)	(0.074)	(0.058)	(0.058)	(0.058)
PreCOPD	-0.034	0.032	0.282***	0.033	-0.144
	(0.160)	(0.128)	(0.098)	(0.102)	(0.103)
PreCAN	0.214	0.277	-0.224	-0.108	0.042
	(0.214)	(0.171)	(0.138)	(0.121)	(0.125)
PreMH	0.026	0.059	0.166	-0.077	0.030
	(0.203)	(0.133)	(0.145)	(0.138)	(0.090)
PreDM	0.064	0.012	-0.187***	-0.004	0.005
	(0.070)	(0.060)	(0.055)	(0.044)	(0.044)
PreEP	0.232	0.053	0.016	-0.283	0.707***
	(0.417)	(0.357)	(0.064)	(0.307)	(0.265)
GpFTE	0.086	0.049	0.034	-0.058	-0.056
	(0.057)	(0.053)	(0.046)	(0.047)	(0.046)
NurseFTE	0.023	0.055	0.112*	0.095	0.204***
	(0.100)	(0.086)	(0.064)	(0.064)	(0.055)
AdminFTE	0.038	0.054	0.037	-6.67e-05	0.037
	(0.043)	(0.039)	(0.026)	(0.027)	(0.027)
Urban	0.992**	0.382	0.125	-0.0006	0.016
	(0.420)	(0.324)	(0.279)	(0.250)	(0.239)
Dispensing	-0.219	-0.473	-0.636*	-0.882***	-0.904***
	(0.394)	(0.354)	(0.331)	(0.311)	(0.291)
Observations	6,770	6,830	6,772	6,739	6,423

Notes: I run separate logit regressions for each year of mergers, i.e. each cohort. In each regression, I include all practices that have not yet merged in the given year. It's important to note that for merged practices, I use the observations for both the acquirer and target practices, rather than using the constructed pseudo-merged values. For the merged practices, only values from the year before the merger are used in the regression. The dependent variable is whether the practice merged. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B2: Factors that Predict the Timing of GP Practice Mergers

Variables	Cohort14	Cohort15	Cohort16	Cohort17
NumComp	-0.003	0.0003	-0.003	-0.0002
	(0.012)	(0.010)	(0.008)	(0.006)
NumPatient	4.01e-05	6.84e-05**	1.25e-05	-7.26e-06
	(3.33e-05)	(2.99e-05)	(1.33e-05)	(6.56e-06)
IMD	-3.98e-06	-1.85e-07	4.06e-06	4.45e-06
	(8.74e-06)	(7.85e-06)	(5.83e-06)	(4.58e-06)
PreHYP	-0.016	0.009	-0.0004	0.001
	(0.034)	(0.030)	(0.024)	(0.017)
PreSTIA	0.063	0.125	0.027	0.129
	(0.204)	(0.182)	(0.091)	(0.110)
PreCHD	0.077	0.025	-0.092	-0.099
	(0.120)	(0.102)	(0.090)	(0.068)
PreAST	-0.023	0.032	-0.009	-0.043
	(0.063)	(0.052)	(0.040)	(0.030)
PreCOPD	-0.113	-0.120	-0.094	-0.048
	(0.107)	(0.089)	(0.072)	(0.054)
PreCAN	-0.146	-0.162	0.122	0.002
	(0.157)	(0.131)	(0.095)	(0.073)
PreMH	0.022	-0.136	0.050	-0.034
	(0.188)	(0.152)	(0.084)	(0.090)
PreDM	0.067	0.054	0.070**	0.016
	(0.045)	(0.036)	(0.029)	(0.021)
PreEP	0.445	0.269	0.164	0.461***
	(0.336)	(0.279)	(0.215)	(0.176)
GpFTE	-0.076**	-0.049	-0.057***	-0.033**
	(0.035)	(0.031)	(0.021)	(0.016)
NurseFTE	-0.007	-0.012	-0.018	0.007
	(0.051)	(0.041)	(0.030)	(0.021)
AdminFTE	0.003	-0.032*	0.010	0.017*
	(0.020)	(0.017)	(0.010)	(0.009)
Urban	-0.206	0.073	-0.065	-0.042
	(0.221)	(0.190)	(0.152)	(0.110)
Dispensing	0.932***	0.715***	0.228	0.154
	(0.263)	(0.228)	(0.177)	(0.126)
Observations	740	641	528	392

Notes: This table analyzes how observable practice and local characteristics predict the timing of the merger conditional on eventually merging. In each regression, I consider only practices that have not yet merged in a given year but will merge in the future. For these practices, I use the constructed pseudo-merged values before the merger as predictors in the regression. The dependent variable is the year in which the practice merges. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B3: Changes in Patient Mix After the Merger

Variable	(1) PreHYP	(2) PreSTIA	(3) PreCHD	(4) PreAST	(5) PreCOPD	(6) PreCAN	(7) PreMH	(8) PreDM	(9) PreEP
Treat × Post	-0.117*	0.007	-0.004	-0.066**	0.014	0.020	0.0001	-0.036	0.012
	(0.061)	(0.014)	(0.018)	(0.026)	(0.015)	(0.021)	(0.008)	(0.030)	(0.010)
Additional Controls	No	No	No	No	No	No	No	No	No
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	18,144	18,144	18,144	18,144	18,144

Notes: The dependent variables represent the practice prevalence rates of long-term illnesses. From Column (1) to Column (9), the dependent variables respectively represent the prevalence rates of hypertension, stroke, coronary heart disease, asthma, chronic obstructive pulmonary disease, cancer, serious mental illness, diabetes, and epilepsy in the regression. All specifications include cohort-specific practice and year fixed effects. No additional time-varying covariates are included here. For merged practices, the year of the merger is entirely excluded. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B4: Effect of GP Practice Mergers on Outcomes – Event Study

Variables	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × (T-4)	-0.007	-0.010	-0.187	0.024	-0.031	0.299	0.368
	(0.016)	(0.021)	(0.453)	(0.220)	(0.314)	(0.430)	(0.516)
Treat × (T-3)	0.002	0.005	-0.331	-0.146	-0.352	0.472	0.803
	(0.013)	(0.025)	(0.490)	(0.244)	(0.360)	(0.472)	(0.569)
Treat × (T-2)	0.006	0.017	-0.128	-0.181	-0.476	0.247	0.449
	(0.014)	(0.026)	(0.491)	(0.250)	(0.377)	(0.488)	(0.595)
Treat × (T-1)	0.014	0.054*	-0.385	-0.226	-0.617	-0.267	-0.177
	(0.015)	(0.028)	(0.501)	(0.254)	(0.387)	(0.521)	(0.631)
Treat × (T+1)	0.027	0.247***	0.535	-0.184	-0.776	-2.015***	-1.988**
	(0.019)	(0.034)	(0.555)	(0.282)	(0.492)	(0.657)	(0.875)
Treat × (T+2)	0.036*	0.252***	0.901	0.086	-0.504	-3.032***	-3.292***
	(0.021)	(0.036)	(0.586)	(0.307)	(0.552)	(0.668)	(0.965)
Treat × (T+3)	0.041*	0.215***	0.513	-0.127	-1.104*	-3.353***	-3.218***
	(0.024)	(0.040)	(0.647)	(0.332)	(0.567)	(0.747)	(1.062)
Treat × (T+4)	0.047	0.249***	0.018	-0.439	-1.495**	-3.860***	-3.143***
	(0.030)	(0.049)	(0.734)	(0.365)	(0.681)	(0.831)	(1.194)
Treat × (T+5)	0.033	0.263***	-0.320	-0.601		-2.835***	
	(0.031)	(0.049)	(1.047)	(0.465)		(1.020)	
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,144	18,144	18,144	18,144	15,536	18,144	15,536

Notes: This table shows the estimated coefficients from the dynamic DiD specification with covariates. The additional covariates consist of the following: at the practice level, the number of competitors within a 2km radius market, the total number of registered patients, a set of prevalence rates, staff FTEs (including GPs, nurses, and admins), and dispensing status; at the local level, the IMD rank and a dummy variable for urban status. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are available for the subset of years from 2013 to 2018. T represents the year of the merger and is specified separately for each cohort. $(T-5)$ is used as the reference year. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B5: Heterogeneity Effects: Different Pre-merger Sizes

Size of Merging Parties	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Small Practices Merge	0.053	0.298***	3.324***	1.085***	1.228*	-2.410***	-2.421**
	(0.057)	(0.068)	(0.712)	(0.377)	(0.700)	(0.755)	(1.078)
Small and Large Merge	0.017	0.091**	1.833***	0.253	0.132	-2.574***	-3.514***
	(0.021)	(0.038)	(0.640)	(0.324)	(0.575)	(0.902)	(1.280)
Large Practices Merge	0.008	0.100	0.246	0.102	-0.161	-6.355***	-6.904**
	(0.035)	(0.082)	(1.116)	(0.663)	(1.110)	(2.192)	(2.963)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,634	6,634	6,634	6,634	5,679	6,634	5,679
Test equality of coefficients (p-value)							
Small Practices Merge=Small and Large Merge	0.585	0.007	0.066	0.046	0.151	0.865	0.437
Small Practices Merge=Large Practices Merge	0.560	0.062	0.006	0.126	0.208	0.060	0.114

Notes: I distinguish between three cases: mergers involving only small practices, mergers involving only large practices, and mergers involving both small and large practices. The size of a practice is measured by the number of registered patients in the year before the merger. Merging practices are considered small if their patient numbers are below the 25th percentile and large if they are above the 75th percentile. The omitted group consists of never-merged practices selected by PSM. All specifications include cohort-specific practice and year fixed effect. Coefficients on the additional time-varying covariates are not shown to save space. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Potential Reasons for Heterogeneous Outcomes: Workload Effect

Size of Merging Parties	(1) PatientsPerGP	(2) PatientsPerNurse	(3) PatientsPerAdmin
Small Practices Merge	684.0***	1,596*	-9.891
	(160.3)	(859.0)	(76.54)
Small and Large Merge	695.0	-693.6	-41.38
	(706.3)	(539.7)	(105.1)
Large Practices Merge	198.4	-816.6	-8.280
	(385.3)	(1,346)	(212.8)
Additional Controls	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	6,634	6,634	6,634
Test equality of coefficients (p-value)			
Small Practices Merge=Small and Large Merge	0.988	0.040	0.721
Small Practices Merge=Large Practices Merge	0.202	0.202	0.993

Notes: I distinguish between three cases: mergers involving only small practices, mergers involving only large practices, and mergers involving both small and large practices. The size of a practice is measured by the number of registered patients in the year before the merger. Merging practices are considered small if their patient numbers are below the 25th percentile and large if they are above the 75th percentile. The omitted group consists of never-merged practices selected by PSM. All specifications include cohort-specific practice and year fixed effect. Coefficients on the additional time-varying covariates are not shown to save space. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B7: Test Market Concentration Mechanism

Pre-merger Market Competition	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
Low comp	1.003	0.241	-0.173	-3.677***	-4.942***
	(0.759)	(0.388)	(0.886)	(1.034)	(1.834)
High comp	1.179*	0.423	-0.063	-2.333***	-2.819**
	(0.605)	(0.295)	(0.583)	(0.761)	(1.205)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	6,048	6,048	5,169	6,048	5,169
Test equality of coefficients (p-value)					
Low comp=High comp	0.210	0.222	0.996	0.464	0.881

Notes: The level of competition is determined by the number of competitors the acquirer practice has in the year before the merger. I categorize the pre-merger competition level into two groups: low competitive market (below the 25th percentile) and high competitive market (above the 75th percentile). The market is defined as a 2km radius around each practice. The reference group consists of never-merged practices selected by PSM. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. Coefficients on the additional time-varying covariates are not shown to save space. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B8: Test Market Concentration Mechanism: Calculate Market Concentration Changes Based on the HHI Index

Merger Types	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
No Concerns	1.723***	0.589**	0.677	-1.639**	-0.852
	(0.617)	(0.294)	(0.555)	(0.708)	(1.041)
Low Risk	1.208***	0.071	-0.352	-2.978***	-3.523***
	(0.450)	(0.251)	(0.450)	(0.651)	(1.009)
High Risk	0.739	0.103	-0.116	-2.625***	-3.892***
	(0.474)	(0.238)	(0.511)	(0.788)	(1.404)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	11,685	11,685	9,994	11,685	9,994
Test equality of coefficients (p-value)					
No Concerns=Low Risk	0.441	0.128	0.093	0.115	0.032
No Concerns=High Risk	0.150	0.144	0.212	0.295	0.041

Notes: This table calculates changes in market concentration based on the HHI index, using the number of registered patients. I focus on the subsample of within-market mergers and calculate the HHI index for one year before and after the merger, along with the change in the HHI index. Following the guidelines by the Federal Trade Commission, merger types are classified into three categories: those that should raise no competitive concerns (“No concerns”), mergers that are unlikely to be challenged (“low risk”), and mergers that should raise serious competitive concerns (“high risk”). The reference group consists of never-merged practices selected by PSM. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. Coefficients on the additional time-varying covariates are not shown to save space. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B9: Test Market Concentration Mechanism: Excluding the Potential Effect of Incompetency

Merger Types	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
Within-market Mergers	0.181	0.308	0.382	-4.041***	-5.261**
	(0.692)	(0.350)	(0.539)	(1.427)	(2.146)
Cross-market Mergers	-0.245	-0.031	-0.863	-2.754	-3.247
	(0.825)	(0.484)	(0.669)	(2.158)	(2.861)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	3,336	3,336	2,861	3,336	2,861
Test equality of coefficients (p-value)					
Within-market=Cross-market	0.646	0.470	0.077	0.498	0.466

Notes: This table examines the subsample of mergers between large general practices, defined as those above the mean size. The reference group consists of never-merged practices selected by PSM. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. Coefficients on the additional time-varying covariates are not shown to save space. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B10: Test Market Concentration Mechanism: Excluding the Potential Effect of Incompetency

Merger Types	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
Within-market Mergers	1.508***	0.493*	0.371	-2.865***	-2.558**
	(0.575)	(0.284)	(0.504)	(0.681)	(1.090)
Cross-market Mergers	0.850	-0.106	-0.430	-2.738**	-2.741*
	(0.975)	(0.564)	(0.987)	(1.064)	(1.611)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	6,779	6,779	5,794	6,779	5,794
Test equality of coefficients (p-value)					
Within-market=Cross-market	0.483	0.266	0.400	0.903	0.912

Notes: This table examines the subsample of mergers between small general practices, defined as those below the mean size. The reference group consists of never-merged practices selected by PSM. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. Coefficients on the additional time-varying covariates are not shown to save space. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B11: Test Incompetency Mechanism

Size of Merging Parties	(1) qofOutcome	(2) ClinPA	(3) ClinPA2	(4) OverallSat	(5) Recommend
Small Practices Merge	3.904***	1.108	1.696	-1.928	0.042
	(1.407)	(0.694)	(1.215)	(1.688)	(2.618)
Small and Large Merge	2.660**	0.083	-0.364	-2.850*	-1.390
	(1.241)	(0.548)	(0.861)	(1.557)	(2.266)
Large Practices Merge	0.007	-1.057	-2.484	-1.920	-1.196
	(2.276)	(1.231)	(1.759)	(3.270)	(4.538)
Additional Controls	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	2,090	2,090	1,787	2,090	1,787
Test equality of coefficients (p-value)					
Small Practices Merge=Small and Large Merge	0.427	0.164	0.111	0.640	0.617
Small Practices Merge=Large Practices Merge	0.083	0.074	0.029	0.998	0.779

Notes: Focusing on the subsample of cross-market mergers, I compare three cases: mergers between small general practices, mergers between large general practices, and mergers involving both small and large general practices. The size of a practice is determined by the number of registered patients in the pre-merger year, where a practice is considered small if its patient number is below the 25th percentile, and large if its patient number is above the 75th percentile. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates are not shown to save space. Columns (3) and (5) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. All specifications include cohort-specific practice and year fixed effect. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B12: Robustness check: Staggered Roll-out Design

	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
ATT	0.013	0.157***	0.942***	-0.128	-0.454**	-2.839***	-2.394***
	(0.010)	(0.019)	(0.234)	(0.114)	(0.185)	(0.298)	(0.445)
Observations	42,878	42,878	42,878	42,878	36,526	42,878	36,526

Notes: This table uses the estimator proposed by Callaway and Sant'Anna (2021), which is specifically designed for the staggered roll-out design. Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. Bootstrapped Standard errors are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B13: Robust check: Varying the Number of Neighbors in PSM

Panel A. #Neighbors=1

Variable	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × Post	0.030*	0.216***	0.774**	-0.009	-0.279	-2.523***	-2.305***
	(0.017)	(0.029)	(0.348)	(0.168)	(0.296)	(0.454)	(0.668)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,689	8,689	8,689	8,689	7,442	8,689	7,442

Panel B. #Neighbors=5

Variable	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × Post	0.028*	0.231***	0.829***	0.062	-0.276	-2.708***	-2.838***
	(0.016)	(0.029)	(0.303)	(0.158)	(0.293)	(0.398)	(0.623)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,517	27,517	27,517	27,517	23,565	27,517	23,565

Panel C. #Neighbors=7

Variable	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × Post	0.034**	0.238***	1.035***	0.140	-0.216	-2.978***	-3.361***
	(0.016)	(0.029)	(0.305)	(0.159)	(0.311)	(0.401)	(0.637)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	36,835	36,835	36,835	36,835	31,541	36,835	31,541

Notes: This table tests the sensitivity of the findings to the number of matches selected by the PSM. Panel A, B, and C display the findings when using 1, 5, and 7 closest matches from the PSM as control practices, respectively. The specification includes cohort-specific practice and year fixed effect, as well as additional time-varying covariates (not shown to save space). Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B14: Robustness Check: Using all Never-merged Practices as the Control Group

Variable	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × Post	0.072***	0.294***	1.102***	0.215	0.178	-4.114***	-6.274***
	(0.022)	(0.031)	(0.294)	(0.150)	(0.269)	(0.417)	(0.666)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	197,948	197,948	197,948	197,948	168,900	197,948	168,900

Notes: This table examines the sensitivity of the results when using all never-merged practices as the control group. The specification includes cohort-specific practice and year fixed effect, as well as additional time-varying covariates (not shown to save space). Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B15: Robustness Check: Using Merges that Occur At Least Two Years Later as the Control Group

Variable	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × Post	0.017	0.174***	0.551	-0.167	-0.306	-0.989*	-1.452**
	(0.023)	(0.044)	(0.478)	(0.225)	(0.342)	(0.575)	(0.722)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,983	4,983	4,983	4,983	4,851	4,983	4,851

Notes: This table uses mergers that occur at least two years later as the comparison group. The specification includes cohort-specific practice and year fixed effects, as well as time-varying covariates (not shown to save space). Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B16: Robustness Check: Drop Mergers Involving Poor-quality Merging Parties

Variable	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Treat × Post	0.022	0.212***	0.681**	-0.075	-0.344	-3.426***	-3.243***
	(0.018)	(0.029)	(0.326)	(0.161)	(0.283)	(0.419)	(0.640)
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Practice FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,330	16,330	16,330	16,330	13,976	16,330	13,976

Notes: Mergers involving merging parties rated as inadequate by the Care Quality Commission are excluded. The control group for each merged unit consists of the three closest never-merged practices selected by PSM. The specification includes cohort-specific practice and year fixed effects, as well as time-varying covariates (not shown to save space). Columns (5) and (7) have fewer observations because measures of ClinPA2 and Recommend are only available for the subset of years from 2013 to 2018. The year of the merger is dropped altogether for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

CHAPTER 2

THE EFFECT OF COMPETITION AND PRICE DISPERSION ON SEARCH BEHAVIOR

2.1 Introduction

Economic models for homogeneous products suggest that under frictionless trade conditions, perfect competition among firms should result in the "law of one price". Yet, in many homogeneous product markets, price dispersion still exists. The seminal paper by Stigler (1961) provides the first consumer search-based rationale for price dispersion, which has since gained considerable attention.

A large body of theoretical work has analyzed how, in the presence of varying levels of consumer information, consumer search and price dispersion occur in equilibrium (see, for instance, Varian, 1980; Stahl, 1989; Chandra and Tappata, 2011). Specifically, search and price dispersion are jointly determined: increased consumer search reduces price dispersion, while lower price dispersion leads to decreased consumer search, all else being equal. While empirical research has primarily focused on exploring the impact of consumer search on price dispersion, limited attention has been given to the converse relationship—how price dispersion affects consumer search. This gap in the literature is important, as failing to account for the reverse causality of price dispersion could potentially introduce biases (Noel, 2018). Our paper provides empirical evidence on how changes in price dispersion affect consumer search behavior, while credibly controlling for market structure.

Identifying the effect of price dispersion on consumer search is challenging. First, establishing a causal claim can be difficult without an exogenous shock to price dispersion.[1] Second, consumer search behavior is often unobserved, and previous literature usually captures consumer search indirectly by using variables related to search costs or benefits as proxies. Third, given that market structure affects both search and price dispersion, it is essential to credibly control for market concentration. Defining market boundaries and assessing effective competition is another important empirical difficulty.

We leverage an excise duty tax increase policy as a plausible exogenous shock to price dispersion and study how consumer search behavior responds. The validity of this method relies on the assumption that consumer search reacts to the tax shock only indirectly through the effect of price dispersion. We support this assumption by showing that consumer search remains unresponsive to the public announcement of the tax policy; however, it increases substantially when actual price changes following the tax shock.

Our data come from the retail petroleum market (unleaded 95 and heating oil) on various small isolated Greek islands. We directly measure consumer search using the number of user visits to fuelGR, a Greek gasoline price information platform. Similar applications using search queries or web traffic from price information websites to measure consumer search can be seen in, for instance, Noel (2018); Lewis and Marvel (2011); Byrne et al. (2015); and Byrne and de Roos (2017). However, as we will explain later, a key advantage of our data is that we are able to uniquely identify each user through a random ID. This allows us to explore whether the changes in search activity come from existing users of the platform or from new users.

The main analysis proceeds in two parts. First, we use a difference-in-difference method to estimate the tax pass-through. As excise duty increased only for unleaded 95 but remained unchanged for heating oil products, we use heating oil as the control group in our analysis. As stations respond to the tax shock at different speeds, this generates cross-sectional price dispersion on the market. Next, we use an event study approach to directly examine the impact of the tax shock on both price dispersion and consumer search. The event study method allows us to investigate the dynamics of the effect over time.

We also perform a heterogeneous analysis, exploring whether the market competition level influences the relationship between search behavior and price dispersion. Theoretical work suggests that market competition impacts both price dispersion and the share of informed consumers (see, for instance, Chandra and Tappata, 2011; Pennerstorfer et al., 2020). As a result, the interplay between competition, search behavior, and price dispersion is complicated. However, existing empirical studies fail to disentangle the effect of competition and price dispersion on consumer search. Our paper addresses this

[1] For instance, Byrne et al. (2015) and Byrne and de Roos (2017) are two papers that study the impact of price dispersion on consumer search in retail gasoline markets. They use price dispersion resulting from price cycles and establish causality by arguing the plausible exogeneity of price cycles to their consumer search measure—website traffic. However, given that price cycles usually span about a week, without a shock to the price cycles, it can be hard to disentangle the effect of price dispersion and regular search behavior due to commuter traffic patterns (Noel, 2018).

challenge and contributes new evidence regarding the influence of competition on the relationship between search behavior and price dispersion. The unique context of the isolated Greek islands clearly defines local markets and provides an exogenous variation in market size, which facilitates our analysis of the competition effect.

We find four main results. First, the tax gradually transmits to the retail price, with an average pass-through rate of 0.625 within 10 days of the excise duty change. Due to varying speeds of station responses to the tax shock, this leads to price dispersion across stations in the market. Second, we find direct evidence that price dispersion rises in response to the tax shock, with the most significant impact observed in the initial seven days following the tax shock. Third, consumer search increases in response to increased price dispersion. However, this effect is of relatively short duration, occurring only during the initial four days after the tax shock before diminishing. Finally, the effect of price dispersion on consumer search remains regardless of market competition level.

The rest of this paper is structured as follows. Section 2.2 reviews related literature. Section 2.3 introduces the Greek petroleum market. Data sources are described in Section 2.4. Section 2.5 and Section 2.6 present our empirical models and findings. Finally, Section 2.7 concludes the paper.

2.2 Literature Review

Our paper adds to the consumer search literature. There is a rich theoretical literature on equilibrium consumer search and price dispersion. The classic Diamond paradox (Diamond 1971) states that if all consumers face positive search costs and products are homogeneous, equilibrium in oligopolistic markets leads to firms setting a monopoly price and consumers not searching. However, the model is based on the assumption of identical buyers. Later, Varian (1980) and Stahl (1989) introduce heterogeneity in consumer search costs: some consumers bear zero search costs and always buy from the lowest-priced seller, while others incur costs to become informed. In such setup, price dispersion and consumer search occur in equilibrium. Chandra and Tappata (2011) further demonstrates that the relationship between price dispersion and search is non-monotonic. Depending on the initial search intensity, more search can either increase or decrease price dispersion.

Several empirical studies have explored the effect of consumer search behavior on price dispersion, often finding a negative association. Due to the challenge of directly observing search behavior, many of these studies use variables related to search costs or benefits as proxies. For instance, Dahlby and West (1986) demonstrates that car insurance premiums exhibit lower dispersion among drivers with lower search costs. Brown and Goolsbee (2002) investigates the impact of Internet comparison shopping sites on life insurance markets, finding that increased usage of these platforms leads to reduced price dispersion. An example using search benefits as a proxy for search is Sorensen (2000).

Using medication purchase frequency as a measure of search benefits, he finds that medications requiring more frequent purchases tend to have less price dispersion.[2]

Within the context of the retail gasoline market, Chandra and Tappata (2011) compare price dispersion among different fuel types and find that those with higher associated search costs tend to exhibit greater price dispersion. Noel and Qiang (2019) examines products with varying search costs based on whether their prices are prominently displayed on large signboards in front of gasoline stations or not. The study concludes that products with displayed prices on large signboards experience lower price dispersion. Pennerstorfer et al. (2020) construct a novel measure of the share of informed consumers using commuting behavior data and establish a non-monotonic relationship between the share of informed consumers and price dispersion.[3]

While much research investigates how consumer search affects price dispersion, less attention has been given to the reverse: the impact of price dispersion on consumer search. Byrne et al. (2015) and Byrne and de Roos (2017) investigate consumer search responses to price dispersion stemming from Edgeworth price cycles and identify a positive correlation. Different from their approach, we examine the effects of an unanticipated common cost shock on price dispersion and subsequent consumer behavior. Notably, Noel (2018) is the only study we know of that uses an exogenous shock to investigate how price dispersion affects consumer search. He uses the event of a refinery fire that shuts down price cycles and the high-frequency intraday price dispersion generated by the price cycles to examine search responses. He finds a decrease in consumer search coinciding with reduced price dispersion following the fire. This suggests a positive relationship between price dispersion and consumer search. Our study differs from Noel (2018) in two key aspects. First, while Noel (2018) analyzes a market scenario of a negative shock, we examine a positive tax shock. Considering the potential non-linear relationship between consumer search and price dispersion (Pennerstorfer et al., 2020) and the parallel literature on the ”rockets and feathers” phenomenon, whether the findings in Noel (2018) extend to the case when price dispersion increases is unclear. Our analysis shows that consumers react to price dispersion increases by searching more during such events. Thus, our work complements Noel (2018) and expands the understanding of consumer behavior under different conditions of price dispersion changes. Second, our unique setting of isolated small islands allows us to examine if different market structure affects the relationship between price dispersion and consumer search, for which there is no evidence in the literature.

[2]Other studies explore similar topic in various markets, such as Brynjolfsson and Smith (2000); Tang et al. (2010) for online markets; Hortaçsu and Syverson (2004) for mutual fund market; Sherman and Weiss (2017) for grocery market; Orlov (2011) for airline markets.

[3]Another related paper is Lewis and Marvel (2011). They analyze asymmetric consumer search patterns in response to price changes. While they do not directly test the relationship between price dispersion and consumer search due to data limitations, they infer that increasing prices are associated with more search activity and less price dispersion.

Our study also contributes to the body of literature examining the transmission of cost shocks to prices, particularly in terms of the interplay of competition and station responses to these shocks. Prior research has shown that firms operating in more competitive markets tend to adjust prices more frequently and exhibit greater pass-through, as demonstrated by Genakos and Pagliero (2022) and Gopinath and Itskhoki (2010). We expand on this literature by providing new evidence regarding the impact of cost shocks on cross-sectional price dispersion. Our findings suggest that price dispersion increases as stations adjust prices at different speeds. Moreover, the impact on cross-sectional price dispersion remains regardless of the level of market competition.

2.3 Institutional Background

There are four different segments in the petroleum industry: oil extraction and refining, transportation and storage, wholesale distribution, and retail distribution. Oil is the primary energy source in Greece, contributing to over half of the total final energy consumption.[4] Domestic crude oil production is insignificant and thus the country is mainly dependent on imports. Imported crude oil is then refined in four domestic refineries, of which three are owned by Hellenic Petroleum and one by Motor Oil Hellas. In 2017, there were two oil pipelines in Greece. Most crude oil and products are transported by trucks and ships within Greece. There are ten oil terminals, seven of which are located in the Attica area (Athens) and three in the Thessaloniki area. Greece has sufficient storage capacity for industry operations and mandatory industry stocks. Hellenic Petroleum controls 65% of the wholesale market. As for the retail sector, the market is in general well supplied. Hellenic Petroleum and Motor Oil dominate the market, operating various brands. The marginal cost of petroleum products relies on long-term contracts between retailers and trade companies. Over the short period under consideration, it is reasonable to assume a constant marginal cost. Two main taxes are applied to energy products: an excise tax (a per unit tax) and an ad valorem tax, such as VAT (value added tax), which is a percentage-based tax. The retail gasoline price is determined as $P_{retail} = (P_{refinery} + exciseduty\&fees + margins) \times (1 + VAT)$. In this paper, we focus on the event of an excise duty change, thus calculating the prices net of VAT.

In 2010 the inability of the Greek government to borrow funds from the international markets led to a €110 billion bailout loan from the European Commission, the European Central Bank, and the International Monetary Fund. As part of the loan agreement, the Greek government agreed and implemented a series of austerity measures. The third and last economic adjustment programme was signed by the Greek government in July 2015. One of the austerity measures taken by the Greek government to increase tax revenues

[4]See International Energy Agency (2017).

was to increase excise duties on fuel. On January 2017 excise duties on unleaded 95 (by 24 percent, from 330 to 410 per 1000 litres) and diesel fuel were raised. Importantly, the excise duty for heating oil products remained unchanged during this period.[5] Therefore, we use heating oil products as the control group in our empirical analysis.

2.4 Data

In our analysis, we combine datasets on daily gas station prices and consumer visits to a price information platform.

2.4.1 Market Definition

Before describing our data, we first clarify the market in our analysis. We focus on 26 small isolated islands that have at least two but fewer than ten gas stations (Table A1). Islands with only one gas station are excluded because our study examines cross-sectional price dispersion, which requires at least two stations for comparison. These small isolated islands clearly define the local market in our analysis. Arbitrage across islands is almost impossible. Refueling a car by travelling to a different island is prohibitively expensive, and privately importing fuel in tanks or similar containers is dangerous and illegal. By focusing on these small isolated islands, we not only ensure reliable measurement of local price dispersion but also establish clear levels of market competition. This enables us to explore the interplay between market competition, price dispersion, and consumer search.

2.4.2 Retail Price Data

We obtain daily station-level retail prices for unleaded 95 and heating oil products during 2016 and 2017. These data were officially collected by the Greek Ministry of Development and Competitiveness, through a reporting system where managers of each station are mandated to report retail prices on a daily basis. The aim of this system is to stimulate comparison and reduce search costs for consumers.

We use the daily price data to construct measures of price dispersion on each island. Following the literature, we proceed in two steps (see for example, Lach 2002; Lewis 2008; Luco 2019; Pennerstorfer et al. 2020). First, we construct 'clean' or 'residual' prices. These are the price level net of any persistent seller heterogeneity and are obtained

[5] Heating oil is chemically identical to diesel (although coloured differently to avoid replacement) and is available at the same petrol stations across the country. It is subject to a lesser excise duty since it is deemed a necessity, as the vast majority of families use heating oil rather than gas in the winter. Obviously, it is illegal to sell and use heating oil for transportation, and the law is strictly enforced.

from a regression of raw prices on station fixed effects. Then, we construct measures of price dispersion using these 'clean' prices.

The idea is that although petroleum products are considered as homogeneous, there is station heterogeneity, such as their locations, the service provided, etc., that might explain part of the price differences. It is therefore important to control for these sources of heterogeneity when measuring price dispersion. Our first step is to ensure that we remove these influences and get the 'cleaned' prices. We use a regression of the form:

$$P_{kist} = \alpha_0 + \alpha_{ks} + u_{kist} \tag{2.1}$$

where P_{kist} is the raw price of product k at station s on island i on date t. Station-product fixed effects α_{ks} capture both observed and unobserved time-invariant differences in seller and product characteristics. Residuals of this regression are obtained as the cleaned prices net of dispersion caused by station and product heterogeneity and are used to calculate measures of price dispersion. We denote the clean prices as $\widehat{RP_{kist}}$ calculated as $\widehat{RP_{kist}} = \widehat{u_{kist}}$.

Following Pennerstorfer et al. (2020), we construct three measures of price dispersion for each gasoline product at the island level using these 'clean' prices. The first measure is the sample range SR_{kit}, calculated as the difference between the maximum and the minimum price in the market, that is, $SR_{kit} = RP_{ikt}^{\max} - RP_{ikt}^{\min}$. This measure captures, on average, the most a consumer can save by searching every gas station in the market. However, this measure might be strongly influenced by outliers, so we construct the second measure using the sample standard deviation at each island SD_{kit}, calculated as $SD_{kit} = \sqrt{\sum_{s \in i} \frac{(RP_{\text{kist}} - \overline{RP}_{\text{kit}})^2}{N_i}}$. This measure does not rely on extreme values and is commonly used as a measure of price dispersion in the literature (see for instance, Noel 2018). For our last measure, we construct gains from search (GS_{kit}) for consumers $\mathrm{GS}_{kit} = E_i(RP_{kist} - RP_{ikst}^{\min})$, which is the difference between the expected price and the expected minimum price in each market i (Chandra and Tappata 2011). We calculate the expected price for each gasoline product in the market using the average market price for that product.

In our analysis, we focus on a period of 20 days around the tax shock. Approximately 80% of the stations changed their prices at least once within 20 days after the tax shock. Panel A in Table 2.1 shows the summary statistics of different measures of price dispersion for the 20 days around the tax shock. The price dispersion for heating oil is slightly higher than that of the unleaded 95, though the difference seems small.

To investigate the possibility of price dispersion due to permanent price differences, we conduct a rank reversal test following the approach of Chandra and Tappata (2011). Focusing on unleaded 95 products, for each pair of stations m and n within each island, we measure rank reversals rr_{mn} using $rr_{mn} = \frac{1}{T_{mn}} \sum_{t=1}^{T_{mn}} \mathbf{I}_{RP_{mt} > RP_{nt}}$. T_{mn} denotes the total number of days where price data is available for both stations. Over these T_{mn} days, station m is defined to be the usually cheaper station such that $RP_{mt} \leq RP_{nt}$ is observed for most of the days. rr_{mn} thus measures the frequency of which the usually cheaper station charges a higher price. Using the whole year of 2016 data, we find an average rank reversal of 0.14 for unleaded 95 products. This means that a station that typically offers lower prices exhibits higher prices on 14% of the days, suggesting that stations indeed mix strategies.

2.4.3 Consumer Search Data

We obtain consumer search data through a collaboration with fuelGR, a mobile application providing fuel price information in Greece since December 2015. This user-friendly app allows individuals to conveniently explore fuel prices without requiring registration. Upon launching the app, users can select their preferred gasoline products. Then the app generates a map displaying nearby fuel stations based on the user's current location, along with the corresponding fuel prices for each station. Users can zoom in or out on the map, which triggers updates to display price information for stations within the map's view.

Similar applications by using search queries or web traffic from price information websites to measure consumer search can be seen in studies such as Noel (2018); Lewis and Marvel (2011); Byrne et al. (2015); and Byrne and de Roos (2017). Importantly, our data offers a unique advantage: we can identify each user through a randomized ID assigned to their mobile device.[6] This feature allows us to investigate whether consumer search responses come from a small group of active existing users or a larger number of unique new users.

We obtain information on all search queries for both the unleaded 95 and heating oil products over a 20-day period before and after the tax shock. The data are recorded at the user level, including information on the timestamp and launch location of each query. Fig. 2.1 shows an example of the raw data. The map illustrates the search activity on the island of Agistri on a day in 2017, with each dot representing a user-initiated search query.[7]

We construct two measures of consumer search using this data. Our main measure is the daily total number of search queries for each gasoline product on each island. Given that users may access the platform multiple times within a short period, we treat

[6]Each user is assigned a unique ID for each mobile device. Although individuals may own multiple phones or acquire new mobile devices over time, this concern is minor given the limited 40-day data window we use.

[7]The figure plots the raw data where there are some search queries that appear to originate from water. This occurs due to inaccuracies in GPS locations recorded by users' mobile devices. In such instances, users might adjust the map's zoom level to improve location accuracy, eventually stabilizing the location within the island. We address this issue during data cleaning to make sure that each search query is correctly tagged to the corresponding island.

queries as distinct if they are separated by at least a 5-minute interval. We do this because following discussions with the platform owner, users might zoom in or out the map just to refine location accuracy. Therefore, we consider such actions part of a single search session. Nonetheless, as a robustness check, we also use the raw total number of search queries to assess consumer search intensity. In addition, we construct a second measure of consumer search, which is the total number of unique searchers for each gasoline product each day at each market.

Panel B in Table 2.1 presents the summary statistics of the consumer search measures across islands for the 20 days around the tax shock. On average, the platform records about 22 daily searches for unleaded 95, significantly higher than the searches for heating oil, which is less than 1 per day on average. The number of individuals searching on the platform closely matches the total number of searches conducted. This suggests that most users only search once on the platform. Overall, search activity on the platform is relatively modest. This might not be surprising given our focus on small Greek islands where search costs are low, and consumers can easily visit gasoline stations to compare prices. Our approach, which relies on search queries from a price information platform, effectively assumes that consumers are more actively engaged in price search on days when fuelGR reports higher query volumes. As such, our results should be interpreted as representing the lower bound of the actual search activity changes.

Fig. 2.2 plots measures of consumer search with varying numbers of gasoline stations for unleaded 95. There seems to be a positive correlation between market competition and consumer search. Islands with more gasoline stations tend to have more searchers and higher levels of consumer search activity.

2.5 Empirical Methodology

Our goal is to explore the effect of price dispersion on consumer search intensity. To establish causality, we use a tax increase event as a plausible exogenous shock to price dispersion and analyze its impact on consumer search behavior. The idea is that retail gasoline stations may react to the tax shock at varying speeds, thus generating cross-sectional price dispersion across the market. Consumers may then respond following these price dispersion changes.

The validity of this method relies on the assumption that the tax shock only affects consumer search indirectly through price dispersion. Assessing this is challenging, as the tax shock and station price changes may occur simultaneously, making it difficult to determine whether search behavior changes are a direct response to the tax shock or an indirect consequence of subsequent price adjustments. To address this issue, we examine a subset of islands where the initial price adjustments occurred at least three days after the tax policy implementation. In this subset, a minimum 3-day interval exists between the tax policy implementation and any subsequent price changes. If we observe a significant increase in search activity during this interval, it implies consumer responsiveness to the public announcement of the tax shock. On the other hand, if search patterns remain unchanged during this period and only increase after price adjustments, it suggests that the tax shock affects consumer search indirectly through its effect on price dispersion. We will examine this assumption formally in Section 2.6.1.

Our main analysis proceeds as follows. We first study the effect of the tax on prices by using a standard difference-in-difference method to estimate the tax pass-through. Then, we use an event study approach to directly estimate the effect of the tax shock on price dispersion. Finally, we explore the causal impact of the tax-induced price dispersion on consumer search behavior using an event study method. As the tax change was only implemented on unleaded 95, whereas heating oil remained unaffected, we use the heating oil as the control group throughout our analysis.

Our estimation of the tax pass-through follows Genakos and Pagliero (2022). We use a difference-in-difference approach, comparing the prices of the treatment (unleaded 95) and the control group product (heating oil) on two different dates (before and after the tax change). Our main estimation equation is as follows:

$$P_{kist} = \beta_0 + \rho Tax_{kt} + \alpha_{ks} + \alpha_t + \epsilon_{kist} \tag{2.2}$$

where P_{kist} represents the retail price of product k at gas station s on island i on day t $\in \{\tau - 1, \tau + \delta\}$, where τ is the date of the tax change and $\delta = 1, ..., 20$ is the length of the adjustment period considered. Tax_{kt} denotes the excise duty, and its coefficient ρ is our estimand of interest indicating the tax pass-through. We add product-gas station and day fixed effects. Product-station fixed effects α_{ks} capture both observed and unobserved time-invariant differences in product and seller characteristics, while day fixed effects α_t control pricing trends common to all stations. Standard errors are clustered at the island level.

Next, we use a dynamic difference-in-difference model to directly examine the impact of the unexpected tax shock on the cross-sectional price dispersion in the market. We estimate the following equation:

$$PD_{kit} = \alpha_t + \alpha_i + \sum_{q=-2}^{T_1=-10} \beta_t \times Treat_k \times Day_{t=q} + \sum_{q=0}^{T_2=20} \beta_t \times Treat_k \times Day_{t=q} + \gamma Treat_k + \epsilon_{kit} \tag{2.3}$$

where PD_{kit} represents various measures of price dispersion for product k (unleaded 95 or heating oil products) on island i on day t. $Treat_k$ is a dummy variable for treatment products, which equals 1 for unleaded 95, and 0 for heating oil. The set of $Day_{t=q}$ denotes indicator variables for each day within our study period. The parameters T_1 and T_2 indicate the number of days considered before and after the tax shock, respectively. We perform the regression to a window from 10 days before to 20 days after the tax change. The day prior to the tax shock ($t = -1$) is used as the reference. Our primary interest is the coefficients of β_t, which reveal the dynamic treatment effects relative to the day prior to the tax shock ($t = -1$). To account for common trends across days and persistent differences between islands, we incorporate day and island fixed effects in the regression. Standard errors are clustered at the island level.

Lastly, we analyze the impact of the tax shock-induced price dispersion on consumer search behavior. Here, we use the same dynamic event study framework as Eq. (2.3). The dependent variable is the measure of consumer search for each product on each island each day. The estimation equation is as follows:

$$Search_{kit} = \alpha_t + \alpha_i + \sum_{q=-2}^{T_1=-10} \beta_t \times Treat_k \times Day_{t=q} + \sum_{q=0}^{T_2=20} \beta_t \times Treat_k \times Day_{t=q} + \gamma Treat_k + \epsilon_{kit} \tag{2.4}$$

Our primary focus remains on the β_t, which gives the dynamic treatment effect.

2.6 Empirical results

2.6.1 Consumers Respond Indirectly to the Tax shock

The validity of our method relies on the assumption that consumer search responds to the tax policy indirectly through its influence on price changes. To investigate this, we examine a subset of islands where the first price adjustment was observed at least three days after the tax shock. This subset comprises approximately 36% of the markets in our sample. In this subset, a minimum 3-day interval exists between the policy announcement and any subsequent price changes. We estimate Eq. (2.4) on this subsample of islands. Should our assumption hold, we expect that estimates of β_t will show significance only from day three onwards.

Our findings, presented in Table 2.2, show that the estimated coefficients for the first two day indicators after the tax shock, when the tax policy is already in effect but price changes have not yet occurred, are statistically insignificant at 5% level. A joint significance test fails to reject the null hypothesis. However, these coefficients become significant on the third day when actual price changes take place. This result suggests that consumer search activity is primarily influenced by actual price fluctuations rather than the public announcement of the tax shock. We plot the estimated coefficients in Fig. 2.3. The figure clearly illustrates that the search effect is not statistically different from zero during the first two days, but there is a substantial increase in consumer search on the third day when actual price adjustments occur.[8] Overall, these results imply that consumer search behavior mainly responds to actual price changes after the tax shock.

[8] For additional robustness check, we also examine a subsample of islands where the first price increase

2.6.2 Effect of the Tax Shock on Prices

Table 2.3 reports the average pass-through for a 20-day adjustment period by estimating Eq. (2.2). The average pass-through over a 20-day period is approximately 0.755, which is not significantly different from unit pass-through. To further explore the speed of adjustment, we estimate Eq. (2.2) for different adjustment periods. Results, as reported in Table 2.4 and graphically depicted in Fig. 2.4, highlight a gradual adjustment of prices. A shorter adjustment period corresponds to a lower average pass-through. Unit tax pass-through is achieved around the 17th day after the tax shock. The speed of adjustment in our study is slower than that observed in Genakos and Pagliero (2022). There are two possible reasons for this difference. First, Genakos and Pagliero (2022) examine three tax policy events, while we focus only on one tax change, thereby resulting in limited variability in our sample. Second, the tax event we examine is the fourth tax increase since the beginning of the debt crisis in Greece in 2010, leaving very little room for stations to exercise market power.

2.6.3 Effect of the Tax Shock on Price Dispersion

Given the potential variations in the rate of adjustment among stations, this may generate cross-sectional price dispersion across the market. We study the impact of the tax shock on price dispersion by estimating Eq. (2.3). The results are reported in Table 2.5. Regardless of the measures of price dispersion used, the estimated coefficients on day dummies prior to the tax shock are statistically insignificant, suggesting that the parallel trend assumption is satisfied. This implies that prices for treated products did not change prior to the tax reform in anticipation of the policy change. Following the tax shock, there is a significant increase in price dispersion across stations. This effect appears to be long-lasting, extending up to 20 days after the tax shock. [9] This trend is more clearly shown in Fig. 2.5, where we plot the estimated coefficients. The figure illustrates a sharp increase in price dispersion after the tax shock. In the following analysis, we will show how this unforeseen increase in price dispersion may affect consumer search.

2.6.4 Effect of the Tax Shock on Consumer Search

First, as descriptive evidence, we plot the number of searches for unleaded 95 and heating oil by day for the 20 days around the tax shock in Fig. 2.6. Since heating oil is infrequently purchased and searched, we use 4-day moving averages to smooth out the trend. The figure shows a similar trend in searches for both products prior to the tax

---

occurred on the second day or later. The results, as shown in Table A2 and illustrated in Fig. B1, show a search spike after stations start to adjust prices.

[9] We also perform a robustness check using raw prices to construct measures of price dispersion. Our main findings are robust, see Table A3.

shock at day τ. However, after the tax shock, there is a significant increase in search for unleaded 95 only. This visualization provides direct evidence that consumer search increases following the price dispersion shock.

Next, we examine the impact of price dispersion on consumer search by estimating Eq. (2.4). The results are presented in Table 2.6. Column (1) uses the number of total searches as the dependent variable. The estimates indicate a significant increase in consumer search activities following the price dispersion shock. To determine whether this increase in search comes from more intensive searches by existing users or from new users on the platform, we estimate Eq. (2.4) using the number of unique searchers as the dependent variable. Column (2) in Table 2.6 gives the estimated coefficients. The results are quantitatively similar to those of Column (1), implying that the increase in search activity comes mostly from new users on the platform. Most users do not search intensively, searching only once on the platform.

We plot the estimated coefficients in Fig. 2.7. Fig. 2.7a shows the effect on the number of search queries around the tax shock, while Fig. 2.7b illustrates the effect on the number of people searching. These two figures closely resemble each other, as most users only search once on the platform. The confidence intervals of estimates for days prior to the tax shock contain 0, suggesting that the parallel trend assumption is satisfied. There is a noticeable increasing trend of consumer search after the tax shock, which peaks around the third day before declining. These results suggest that consumers respond by searching more following price dispersion increases. However, the effect is temporary, lasting for around four days.

In our main analysis, we group search queries conducted by the same user within a five-minute interval to account for the possibility that users trigger multiple searches by zooming in or out the map to improve location pinpointing. As a robustness check, we use the number of raw search queries as the dependent variable. Our findings remain robust (see Table A4).

Additionally, we perform a robustness analysis by estimating an event study regression with day dummies on the treatment group only. This analysis will demonstrate how search changes over time for the treatment group. Results as given in Table A5 remain similar with those of the main analysis. The coefficients on day dummies for the period before the tax shock are statistically insignificant, indicating no significant changes in consumer search behavior leading up to the treatment. Following the tax shock, there is an immediate and significant increase in consumer search. This robustness test confirms the positive association between price dispersion and consumer search.

Byrne and De Roos (2022) documents how a large exogenous shock to price dispersion generates a substantial and permanent increase in search. Our finding of a small and temporary increase in search following a price dispersion shock can be rationalised in various ways. First, there is a difference in the magnitude of the price dispersion shock we examine. They use a price war that causes a prolonged disruption in a stable pricing equilibrium, while we focus on the impact of a common tax shock to price dispersion. Second, the tax increase in our case is an exogenous and unanticipated one-off event. In their case, there is an evolution of the whole industry equilibrium and consumers do not know in advance (or it is not clear) what would be the overall effect or even its direction. Hence, no wonder that consumers keep on searching for a longer time period. We view the results in the two papers as complementary, where the common thread is a better understanding of how a price dispersion shock generates a (temporary or not) increase in search activity.

In conclusion, our analysis suggests that an unexpected increase in price dispersion leads to a rise in consumer search activity. Coupled with the evidence from Noel (2018), our results imply that the effect of price dispersion on consumer search is positive irrespective of the direction of price dispersion changes.

2.6.5 Heterogeneous Effect of Market Competition

Having established that an unexpected increase in price dispersion leads to an increase in consumer search, we investigate whether this effect varies based on market competition levels.

We begin by examining how tax pass-through varies with market competition levels by estimating the following equation:

$$P_{kist} = \beta_0 + \rho Tax_{kt} + \lambda Tax_{kt} \times Comp_i + \alpha_{ks} + \alpha_t + \epsilon_{kist} \tag{2.5}$$

where $Comp_i$ is the number of gas stations on island i. We present the results in Table 2.7. Column (1) estimates the equation on day t $\in \{\tau - 1, \tau + 20\}$, while Column (2) uses t $\in \{\tau - 1, \tau + 10\}$. τ is the date of the tax change. The results suggest that competition does not seem to have significant impact on tax pass-through. Furthermore, we categorize islands into low-competition (2 to 4 gas stations) and high-competition (5 to 9 gas stations) markets, and report the average pass-through for these two groups during different adjustment periods in Table 2.8 and Fig. 2.8. The speed of adjustment for the two groups closely resembles each other, confirming the findings.[10] This seems to contrast with Genakos and Pagliero (2022), who report a positive and decreasing impact of competition on pass-through. One potential explanation for this difference is that Genakos and Pagliero (2022) examine three tax policy events, whereas we focus only on one tax change, thereby resulting in limited variability in our sample.

We then study the impact of competition on price dispersion after the tax shock. First, we estimate a simple two-period difference-in-difference regression. The regression equation is as follows:

[10] For robustness, we also estimate a non-parametric specification using $P_{kist} = \beta_0 + \sum_j \rho_j I(n_i = j) \times Tax_{kt} + \alpha_{ks} + \alpha_t + \epsilon_{kist}$, where I is an indicator variable for each observed number of gas stations. Our main findings remain robust (see Fig. B2 and Table A6).

$$
\begin{aligned}
PD_{kit} = \beta_0 + \beta_1 \times Treat_k + \beta_2 \times Treat_k \times Post_t + \beta_3 \times Lowcompetition_i \times Treat_k + \\
\beta_4 \times Lowcompetition_i \times Post_t + \beta_5 \times Lowcompetition_i \times Treat_k \times Post_t + \\
\alpha_i + \alpha_t + \epsilon_{kit} \qquad (2.6)
\end{aligned}
$$

where PD_{kit} represents different measures of price dispersion for product k on island i on day t. $Lowcompetition_i$ is a dummy variable for islands with 2 to 4 gas stations. $Treat_k$ equals 1 for unleaded 95 and 0 for heating oil. $Post_t$ is a dummy for post-tax shock period. Standard errors are clustered at the island level. The coefficient on the triple interaction term $Lowcompetition_i \times Treat_k \times Post_t$ captures the potential heterogeneous effect.

Results are presented in Table 2.9, where Column (1) to (3) use a time window from 10 days before the shock to 20 days after. Column (4) to (6) narrow this window to a 10-day before and after the shock. The coefficients on the triple interaction term are insignificant across all columns, implying that changes in price dispersion are similar across markets with varying levels of competition after the tax shock.

As a robustness check, we also estimate a non-parametric specification using the following equation:

$$
PD_{kit} = \beta_0 + \sum_j \rho_j I(n_i = j) \times Treated_{kt} + \alpha_i + \alpha_t + \epsilon_{kit}. \qquad (2.7)
$$

$Treated$ is a dummy variable that equals 1 for unleaded 95 in post-tax periods. I is an indicator variable for each observed number of gas stations.

Results presented in Table A7 and illustrated in Fig. B3, show that the impact on price dispersion does not significantly differ based on market competition. Given that the magnitude of the price dispersion shock is same across islands with different market competition levels, we then study if consumer search behavior differs with competition. Should price dispersion be the primary driver for consumer search, we expect to see a similar consumer response across islands with different competition levels as well.

We first estimate Eq. (2.6) using measures of consumer search as the dependent variable. In particular, we adjust the search measure by dividing it by the population size of each island, to ensure comparability across markets.[11] To make the values of the search measure larger after the adjustment, we rescale them by a constant factor of 1000. Estimation results are presented in Table 2.10. Column (1) and (2) use a window from 10 days before to 20 days after the shock. However, the longer the window, the more likely other factors could influence consumer search after the shock. Therefore, we also use a 10-day window before and after the shock for estimation, with results reported in Column (3) and (4). Across all specifications, the coefficients on the triple interaction

[11] Information on population size comes from Hellenic Statistical Authority (2010).

term are insignificant, suggesting that the price dispersion shock affects consumer search similarly across islands of different market structure.

Moreover, we estimate a non-parametric specification Eq. (2.7) with consumer search as the dependent variable. The results presented in Table A8 and plotted in Fig. B4 confirm that consumer response across islands with varying competition levels does not appear to differ significantly.

This result helps us disentangle the effect of competition and price dispersion on consumer search behavior. We find that when there is a similar price dispersion shock across different market structure, search reactions do not differ across different market structure as well. This implies that controlling for price dispersion changes, market competition does not directly and differentially affect consumer search. This finding is new to the literature, and contributes to our understanding of the interplay between consumer search, market competition and price dispersion.

2.7 Conclusion

This paper provides new empirical evidence on the effects of price dispersion on consumer search. Using the retail gasoline market on oligopolistic markets of various sizes, as defined by small Greek islands, we exploit an excise duty tax increase policy as a plausibly exogenous shock to price dispersion. We directly measure consumer search using the number of user visits to a price information platform and mobile application. As stations adjust to the tax shock at different speed, this generates price dispersion on the market, which in turn leads to a temporary increase in consumer search. Combined with the finding of Noel (2018), this suggests that the positive impact of price dispersion on consumer search persists regardless of the direction of the price dispersion changes.

Moreover, we are also able to disentangle the effect of competition and price dispersion on consumer search behavior. Our clean environment of Greek small islands provides exogenous variation in market size, which enables us to explore the interplay between search and price dispersion across different competition level. We find that the effect of price dispersion on consumer search remains unaffected by market structure.

We acknowledge that Greek islands are not necessarily representative of oligopolistic markets for other products. However, we select this environment precisely because it provides clean variation in the competitive environment and allows us to compare the impact of a common shock across different markets within the same country. We believe that the results contribute to our understanding of the complex interplay among competition, price dispersion and search behavior.

Tables

Table 2.1: Summary Statistics

Panel A

	Mean	SD	Median	10th percentile	90th percentile	Observations
Unleaded95						
PD1	10.555	9.978	8.388	2.157	16.853	984
PD2	4.526	4.238	3.796	1.308	6.342	984
PD3	5.800	6.235	4.081	1.055	10.506	984
Heating oil						
PD1	11.444	9.356	9.797	0.821	27.535	927
PD2	6.101	5.516	5.196	0.416	10.688	927
PD3	6.119	5.240	4.779	0.491	14.421	927

Panel B

	Mean	SD	Median	10th percentile	90th percentile	Observations
Search Queries						
Unleaded95	22.220	12.199	18	12	39	1025
Heating oil	0.220	0.469	0	0	1	1025
Num of People Searching						
Unleaded95	19.976	11.258	16	11	32	1025
Heating oil	0.317	0.561	0	0	1	1025

Notes: This table presents summary statistics of our measures of price dispersion and consumer search. $PD1$ is the the sample range. $PD2$ is sample standard deviation, and $PD3$ is the gains from search.

Table 2.2: Consumer Search Responds to Actual Price Changes

Variable	Search
Day($\tau = -5$)$\times$ Treat	0.333
	(0.203)
Day($\tau = -4$)$\times$ Treat	-0.333
	(0.321)
Day($\tau = -3$)$\times$ Treat	0.333
	(0.593)
Day($\tau = -2$)$\times$ Treat	0.125
	(0.597)
Day($\tau = 0$)$\times$ Treat	-0.500
	(0.413)
Day($\tau = 1$)$\times$ Treat	0.792*
	(0.442)
Day($\tau = 2$)$\times$ Treat	0.625
	(0.594)
Day($\tau = 3$)$\times$ Treat	1.625***
	(0.540)
Day($\tau = 4$)$\times$ Treat	0.833
	(0.764)
Day($\tau = 5$)$\times$ Treat	0.167
	(0.296)
Day($\tau = 6$)$\times$ Treat	0
	(0.249)
Day($\tau = 7$)$\times$ Treat	0
	(0.352)
Day($\tau = 8$)$\times$ Treat	0.333
	(0.733)
Day($\tau = 9$)$\times$ Treat	-0.333
	(0.321)
Day($\tau = 10$)$\times$ Treat	0.667
	(0.813)
F-test: p-value	0.364
Island FE	Yes
Day FE	Yes
Observations	630

Notes: The dependent variables is the daily number of search queries conducted through fuelgr at the island level. The sample comprises islands where the initial price adjustment took place at least three days following the tax shock. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.3: Estimated Tax Pass-through

Variables	Price
Tax	0.755***
	(0.142)
Test of unit pass-through (p-value)	0.099
Day FE	Yes
Product-station FE	Yes
Observations	406

Notes: The dependent variable is the price of each gasoline product for each gas station on each island, and day t $\in \{\tau$-1, τ+20$\}$, where τ is the date of the tax shock. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.4: Pass-through and Speed of Adjustment

Variables	Price
Tax	0.080
(τ-1,τ+1)	(0.051)
Tax	0.277**
(τ-1,τ+2)	(0.128)
Tax	0.256*
(τ-1,τ+3)	(0.133)
Tax	0.382***
(τ-1,τ+4)	(0.122)
Tax	0.415***
(τ-1,τ+5)	(0.121)
Tax	0.415***
(τ-1,τ+6)	(0.128)
Tax	0.426***
(τ-1,τ+7)	(0.126)
Tax	0.485***
(τ-1,τ+8)	(0.150)
Tax	0.440***
(τ-1,τ+9)	(0.149)
Tax	0.625***
(τ-1,τ+10)	(0.139)
Tax	0.620***
(τ-1,τ+11)	(0.131)
Tax	0.670***
(τ-1,τ+12)	(0.135)
Tax	0.688***
(τ-1,τ+13)	(0.135)
Tax	0.702***
(τ-1,τ+14)	(0.137)
Tax	0.687***
(τ-1,τ+15)	(0.149)
Tax	0.596***
(τ-1,τ+16)	(0.140)
Tax	0.766***
(τ-1,τ+17)	(0.130)
Tax	0.727***
(τ-1,τ+18)	(0.139)
Tax	0.740***
(τ-1,τ+19)	(0.139)
Tax	0.755***
(τ-1,τ+20)	(0.142)
Day FE	Yes
Product-station FE	Yes

Notes: The dependent variable is the price of each gasoline product for each gas station on each island, and day t $\in \{\tau\text{-1}, \tau+\delta\}$, where τ is the date of the tax shock and δ is the length of the adjustment period. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.5: Effect of the Tax Shock on Price Dispersion

Variables	PD1	PD2	PD3
Day(τ =-10)$\times$ Treat	0.409	0.467*	0.429
	(0.438)	(0.262)	(0.297)
Day(τ =-9)$\times$ Treat	0.298	0.449	0.123
	(0.642)	(0.356)	(0.414)
Day(τ =-8)$\times$ Treat	0.502	0.438	0.292
	(0.415)	(0.259)	(0.262)
Day(τ =-7)$\times$ Treat	0.488	0.438	0.310
	(0.416)	(0.259)	(0.258)
Day(τ =-6)$\times$ Treat	0.518	0.359	0.254
	(0.412)	(0.249)	(0.241)
Day(τ =-5)$\times$ Treat	-0.163	-0.094	-0.128
	(0.283)	(0.149)	(0.207)
Day(τ =-4)$\times$ Treat	0.072	0.024	0.035
	(0.093)	(0.046)	(0.108)
Day(τ =-3)$\times$ Treat	-0.052	-0.032	-0.072
	(0.045)	(0.027)	(0.065)
Day(τ =-2)$\times$ Treat	-0.044	-0.0153	-0.019
	(0.044)	(0.024)	(0.032)
Day(τ =0)$\times$ Treat	0.096	0.132	-0.003
	(0.321)	(0.198)	(0.194)
Day(τ =1)$\times$ Treat	1.422*	0.668*	0.267
	(0.705)	(0.375)	(0.308)
Day(τ =2)$\times$ Treat	3.029**	1.179**	1.251
	(1.083)	(0.485)	(0.766)
Day(τ =3)$\times$ Treat	2.943**	1.163**	1.322
	(1.180)	(0.527)	(0.859)
Day(τ =4)$\times$ Treat	3.311***	1.332**	1.649*
	(1.172)	(0.518)	(0.875)
Day(τ =5)$\times$ Treat	3.186**	1.300**	1.393
	(1.195)	(0.523)	(0.933)
Day(τ =6)$\times$ Treat	3.815***	1.533***	1.929**
	(1.217)	(0.539)	(0.914)
Day(τ =7)$\times$ Treat	3.977***	1.577***	1.959**
	(1.205)	(0.534)	(0.913)
Day(τ =8)$\times$ Treat	4.140***	1.706***	2.190**
	(1.150)	(0.541)	(0.892)
Day(τ =9)$\times$ Treat	3.654***	1.500**	1.895*
	(1.159)	(0.546)	(0.937)
Day(τ =10)$\times$ Treat	4.054***	1.694***	2.461**
	(1.200)	(0.542)	(0.944)
Day(τ =11)$\times$ Treat	4.427***	1.960***	2.620**
	(1.344)	(0.582)	(1.050)
Day(τ =12)$\times$ Treat	4.098***	1.861***	2.500**
	(1.427)	(0.614)	(1.112)
Day(τ =13)$\times$ Treat	4.775***	2.164***	2.984***
	(1.340)	(0.587)	(1.046)
Day(τ =14)$\times$ Treat	4.661***	2.133***	2.989***
	(1.363)	(0.593)	(1.045)
Day(τ =15)$\times$ Treat	4.405***	2.032***	2.835**
	(1.303)	(0.586)	(1.025)
Day(τ =16)$\times$ Treat	4.592***	2.080***	2.754**
	(1.367)	(0.562)	(1.030)
Day(τ =17)$\times$ Treat	4.810***	2.230***	2.995***
	(1.426)	(0.619)	(1.063)
Day(τ =18)$\times$ Treat	5.244***	2.371***	3.259***
	(1.476)	(0.654)	(1.093)
Day(τ =19)$\times$ Treat	5.250***	2.383***	3.304***
	(1.476)	(0.653)	(1.090)
Day(τ =20)$\times$ Treat	5.494***	2.462***	3.394***
	(1.545)	(0.673)	(1.111)
Island FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Observations	1,441	1,441	1,441

Notes: The dependent variables are different measures of price dispersion at the island level. $PD1$ is the the sample range. $PD2$ is sample standard deviation, and $PD3$ is the gains from search. τ is the day of the tax shock. $\tau - 1$ is used as the reference. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Effect of Price Dispersion on Consumer Search

Variable	(1) Search	(2) Searcher
Day(τ=-10)$\times$ Treat	-0.240	-0.200
	(0.198)	(0.177)
Day(τ=-9)$\times$ Treat	-0.400**	-0.400**
	(0.186)	(0.156)
Day(τ=-8)$\times$ Treat	0.040	-0.040
	(0.273)	(0.239)
Day(τ=-7)$\times$ Treat	0.040	-0.160
	(0.273)	(0.201)
Day(τ=-6)$\times$ Treat	-0.280	-0.240
	(0.182)	(0.170)
Day(τ=-5)$\times$ Treat	0.160	0.120
	(0.210)	(0.189)
Day(τ=-4)$\times$ Treat	-0.080	-0.040
	(0.155)	(0.150)
Day(τ=-3)$\times$ Treat	0	0.040
	(0.236)	(0.216)
Day(τ=-2)$\times$ Treat	0	-0.080
	(0.270)	(0.212)
Day(τ=0)$\times$ Treat	0.120	0.120
	(0.223)	(0.198)
Day(τ=1)$\times$ Treat	0.960***	0.840***
	(0.208)	(0.193)
Day(τ=2)$\times$ Treat	1.760***	1.640***
	(0.534)	(0.428)
Day(τ=3)$\times$ Treat	1.760***	1.600***
	(0.446)	(0.391)
Day(τ=4)$\times$ Treat	0.960**	0.880**
	(0.361)	(0.319)
Day(τ=5)$\times$ Treat	0.480*	0.480*
	(0.237)	(0.237)
Day(τ=6)$\times$ Treat	0.600**	0.520**
	(0.263)	(0.229)
Day(τ=7)$\times$ Treat	0.160	0.200
	(0.183)	(0.177)
Day(τ=8)$\times$ Treat	0.280	0.240
	(0.247)	(0.238)
Day(τ=9)$\times$ Treat	0.880**	0.600**
	(0.355)	(0.236)
Day(τ=10)$\times$ Treat	0.640*	0.480
	(0.322)	(0.289)
Day(τ=11)$\times$ Treat	0.040	0
	(0.191)	(0.177)
Day(τ=12)$\times$ Treat	0.400*	0.400*
	(0.220)	(0.204)
Day(τ=13)$\times$ Treat	0	-0.080
	(0.312)	(0.242)
Day(τ=14)$\times$ Treat	0.040	-0.040
	(0.191)	(0.125)
Day(τ=15)$\times$ Treat	0.080	0.080
	(0.195)	(0.166)
Day(τ=16)$\times$ Treat	0.400	0.360
	(0.250)	(0.211)
Day(τ=17)$\times$ Treat	-0.120	-0.120
	(0.170)	(0.170)
Day(τ=18)$\times$ Treat	0	0
	(0.220)	(0.167)
Day(τ=19)$\times$ Treat	0.320	0.200
	(0.385)	(0.289)
Day(τ=20)$\times$ Treat	0	-0.080
	(0.228)	(0.203)
Island FE	Yes	Yes
Day FE	Yes	Yes
Observations	1,550	1,550

Notes: Column (1) uses the number of total search queries as the dependent variables. Column (2) uses the number of people searching as the dependent variable. τ is the day of the tax shock. $\tau - 1$ is used as the reference. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.7: Competition and Tax Pass-through

Variables	(1) Price	(2) Price
Tax	1.324***	0.946**
	(0.356)	(0.391)
Tax$\times$ Comp	-0.105*	-0.059
	(0.060)	(0.064)
Day FE	Yes	Yes
Product-station FE	Yes	Yes
Observations	406	413

Notes: The dependent variable is the price of each gasoline product for each gas station on each island. Column (1) estimates on and day t $\in \{\tau$-1, τ+20$\}$, while Column (2) estimates on and day t $\in \{\tau$-1, τ+10$\}$. τ is the date of the tax shock. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.8: Competition and Speed of Adjustment

Variables	(1) Price	(2) Price	(3) Price	(4) Price	(5) Price
	(τ-1,τ+1)	(τ-1,τ+2)	(τ-1,τ+3)	(τ-1,τ+4)	(τ-1,τ+5)
Tax × Low competition	0.248	0.510	0.513	0.508*	0.650**
	(0.156)	(0.312)	(0.328)	(0.284)	(0.265)
Tax × High competition	0.003	0.170	0.138	0.324**	0.308**
	(0.092)	(0.143)	(0.146)	(0.139)	(0.138)
Test equality of coefficients (p-value)	0.265	0.346	0.318	0.580	0.265
Day FE	Yes	Yes	Yes	Yes	Yes
Product-station FE	Yes	Yes	Yes	Yes	Yes
Observations	412	412	412	412	412
Variables	(6) Price	(7) Price	(8) Price	(9) Price	(10) Price
	(τ-1,τ+6)	(τ-1,τ+7)	(τ-1,τ+8)	(τ-1,τ+9)	(τ-1,τ+10)
Tax × Low competition	0.518*	0.518*	0.494*	0.542*	0.707**
	(0.281)	(0.281)	(0.280)	(0.272)	(0.289)
Tax × High competition	0.368**	0.383**	0.480**	0.393*	0.588***
	(0.171)	(0.168)	(0.193)	(0.197)	(0.174)
Test equality of coefficients (p-value)	0.681	0.711	0.969	0.680	0.741
Day FE	Yes	Yes	Yes	Yes	Yes
Product-station FE	Yes	Yes	Yes	Yes	Yes
Observations	412	412	411	411	413
Variables	(11) Price	(12) Price	(13) Price	(14) Price	(15) Price
	(τ-1,τ+11)	(τ-1,τ+12)	(τ-1,τ+13)	(τ-1,τ+14)	(τ-1,τ+15)
Tax × Low competition	0.633**	0.679**	0.783**	0.862**	0.851**
	(0.287)	(0.284)	(0.293)	(0.311)	(0.311)
Tax × High competition	0.614***	0.665***	0.644***	0.627***	0.611***
	(0.156)	(0.157)	(0.156)	(0.156)	(0.172)
Test equality of coefficients (p-value)	0.956	0.966	0.692	0.524	0.516
Day FE	Yes	Yes	Yes	Yes	Yes
Product-station FE	Yes	Yes	Yes	Yes	Yes
Observations	415	413	413	413	413
Variables	(16) Price	(17) Price	(18) Price	(19) Price	(20) Price
	(τ-1,τ+16)	(τ-1,τ+17)	(τ-1,τ+18)	(τ-1,τ+19)	(τ-1,τ+20)
Tax × Low competition	0.779**	0.973***	1.023***	1.041***	1.056***
	(0.288)	(0.297)	(0.322)	(0.315)	(0.311)
Tax × High competition	0.511***	0.670***	0.587***	0.597***	0.612***
	(0.178)	(0.160)	(0.194)	(0.195)	(0.203)
Test equality of coefficients (p-value)	0.464	0.411	0.304	0.291	0.291
Day FE	Yes	Yes	Yes	Yes	Yes
Product-station FE	Yes	Yes	Yes	Yes	Yes
Observations	413	413	407	407	406

Notes: *Low Competition* is a dummy equals 1 for islands with 2 to 4 gas stations. *High Competition* is a dummy equals 1 for islands with at least 5 stations. The dependent variable is the price of each gasoline product for each gas station on each island, and day t $\in \{\tau\text{-}1, \tau+\delta\}$, where τ is the date of the tax shock and δ is the adjustment period. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.9: Heterogeneous Effect on Price Dispersion: Market Competition

Variables	(1) PD1	(2) PD2	(3) PD3	(4) PD1	(5) PD2	(6) PD3
Treat	-6.349*	-3.865***	-3.999*	-6.349*	-3.865***	-3.999*
	(3.176)	(1.306)	(2.174)	(3.178)	(1.307)	(2.175)
Treat × Post	5.545**	1.886**	3.545**	4.498**	1.417*	2.427
	(2.150)	(0.745)	(1.654)	(1.864)	(0.699)	(1.453)
Treat × Low competition	7.089	2.741	5.018	7.177	2.819	5.055
	(5.532)	(2.709)	(3.439)	(5.531)	(2.706)	(3.437)
Post × Low competition	2.283	0.960	1.610	2.001	0.949	1.200
	(2.210)	(0.908)	(1.709)	(1.973)	(0.835)	(1.582)
Treat × Post × Low competition	-3.034	-0.657	-2.400	-2.698	-0.622	-1.746
	(2.262)	(0.857)	(1.726)	(1.978)	(0.810)	(1.505)
Island FE	Yes	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Day Window	$[\tau-10,\tau+20]$	$[\tau-10,\tau+20]$	$[\tau-10,\tau+20]$	$[\tau-10,\tau+10]$	$[\tau-10,\tau+10]$	$[\tau-10,\tau+10]$
Observations	1,441	1,441	1,441	984	984	984

Notes: The dependent variables are different measures of price dispersion at the island level. $PD1$ is the the sample range. $PD2$ is sample standard deviation, and $PD3$ is the gains from search. Low Competition is a dummy equals 1 for islands with 2 to 4 gas stations. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.10: Heterogeneous Effect on Consumer Search: Market Competition

Variables	(1) Search	(2) Searcher	(3) Search	(4) Searcher
Treat	0.0864***	0.0752***	0.0864***	0.0752***
	(0.0171)	(0.0143)	(0.0171)	(0.0144)
Treat × Post	0.0879***	0.0807***	0.141***	0.130***
	(0.0204)	(0.0177)	(0.0357)	(0.0318)
Treat × Low competition	0.0809	0.0773*	0.0809	0.0773*
	(0.0497)	(0.0450)	(0.0497)	(0.0450)
Post × Low competition	-0.000867	-0.000963	0.000132	-0.000521
	(0.00303)	(0.00327)	(0.00362)	(0.00422)
Treat × Post × Low competition	0.0738	0.0752	0.144	0.135
	(0.0686)	(0.0595)	(0.105)	(0.0900)
Island FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Observations	1,550	1,550	1,050	1,050

Notes: Column (1) and (2) use a window from 10 days before to 20 days after the shock. Column (3) and (4) use a window of 10 days before and after the shock. Low Competition is a dummy equals 1 for islands with 2 to 4 gas stations. Standard errors, clustered at the island level, are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figures

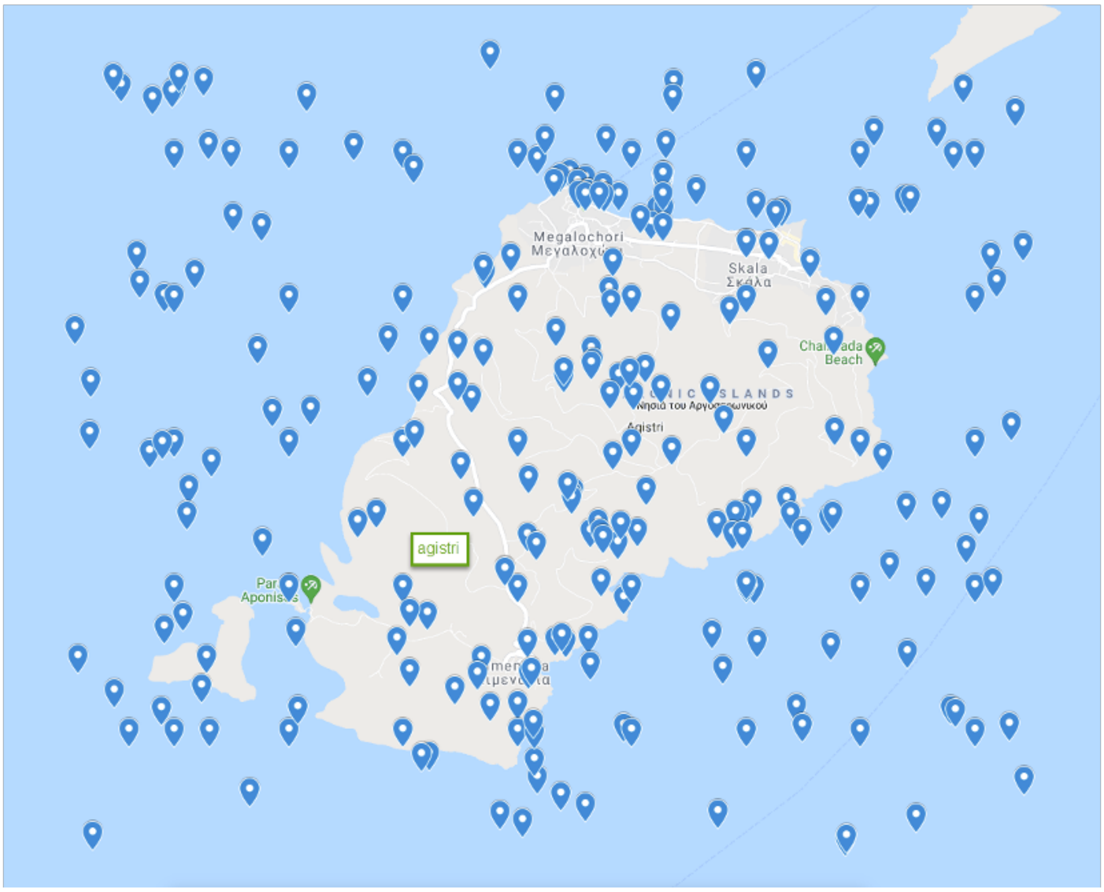


Figure 2.1: An Illustration of the Consumer Search Data

Notes: This figure shows a snapshot of our search data for Agistri Island on a day in 2017. Each dot corresponds to a distinct search query carried out by a unique user. Our dataset includes information on the timestamp and location of each search query.

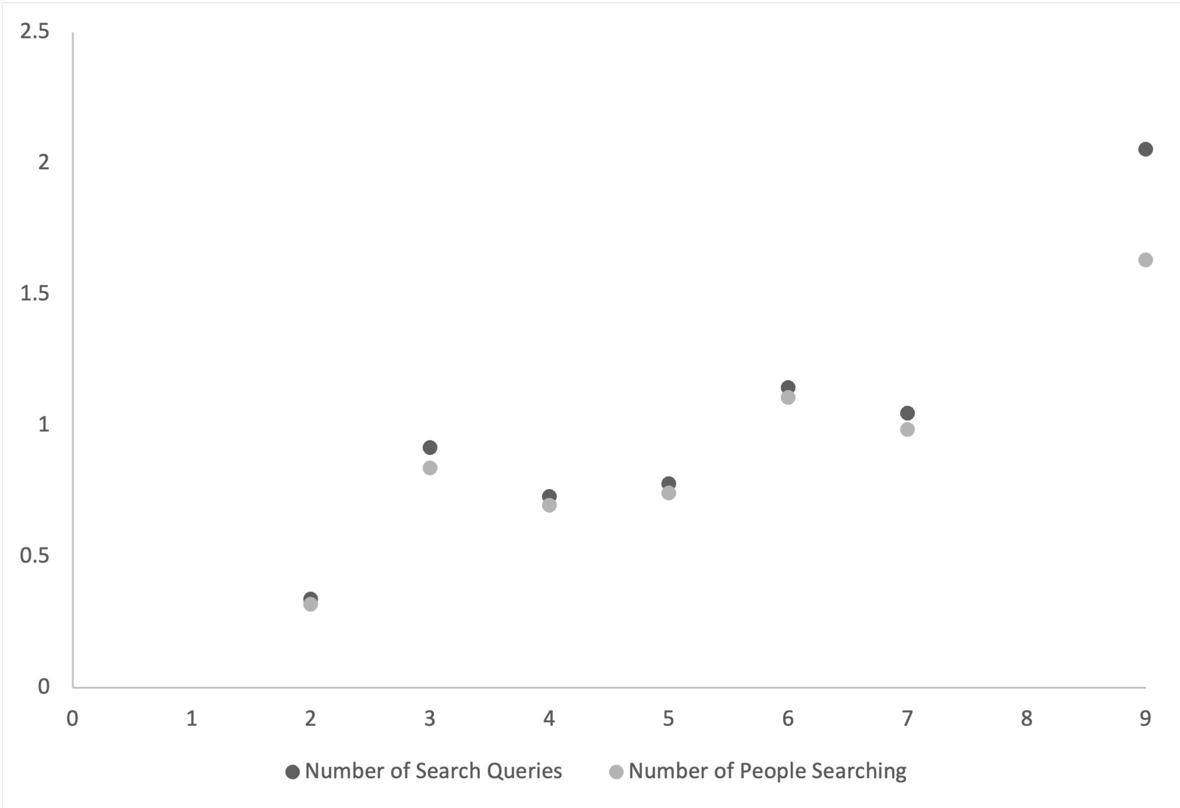


Figure 2.2: Consumer Search and Number of Gas Stations

Notes: Islands with one gas station are excluded.

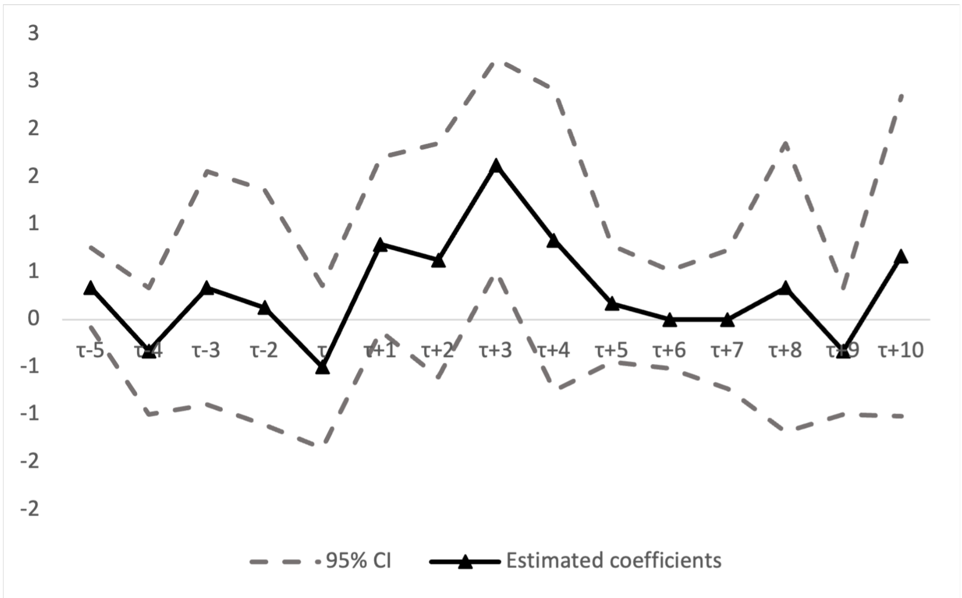


Figure 2.3: Consumer Search Responds to Actual Price Changes

Notes: This figures plots the estimated coefficients from Table 2.2. The samples are islands where the initial price change of unleaded 95 took place at least three days after the tax shock. τ is the day of the tax shock. $\tau - 1$ is used as the reference.

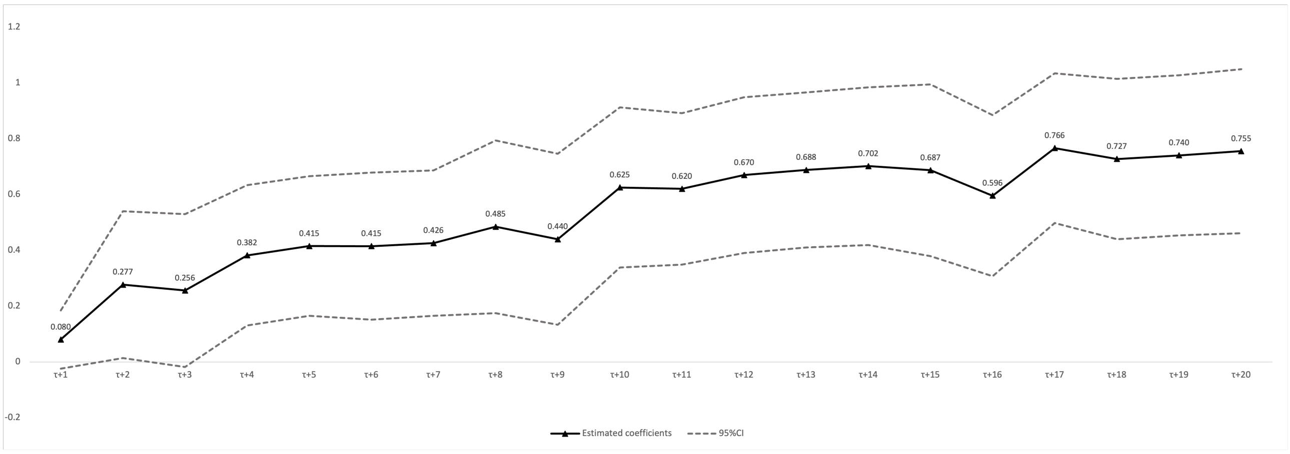


Figure 2.4: Speed of Adjustment

Notes: This figure plots coefficients from Table 2.4. τ is the day of the tax shock.

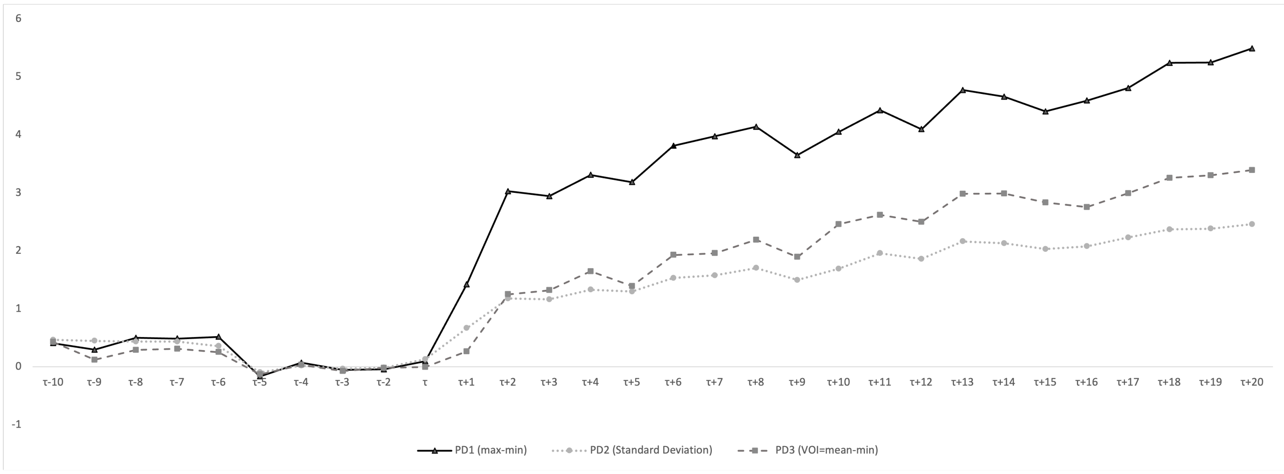


Figure 2.5: Effect of the Tax Shock on Price Dispersion

Notes: This figures plots the estimated coefficients from Table 2.5. τ is the day of the tax shock. $\tau - 1$ is used as the reference. $PD1$ is the the sample range. $PD2$ is sample standard deviation, and $PD3$ is the gains from search.

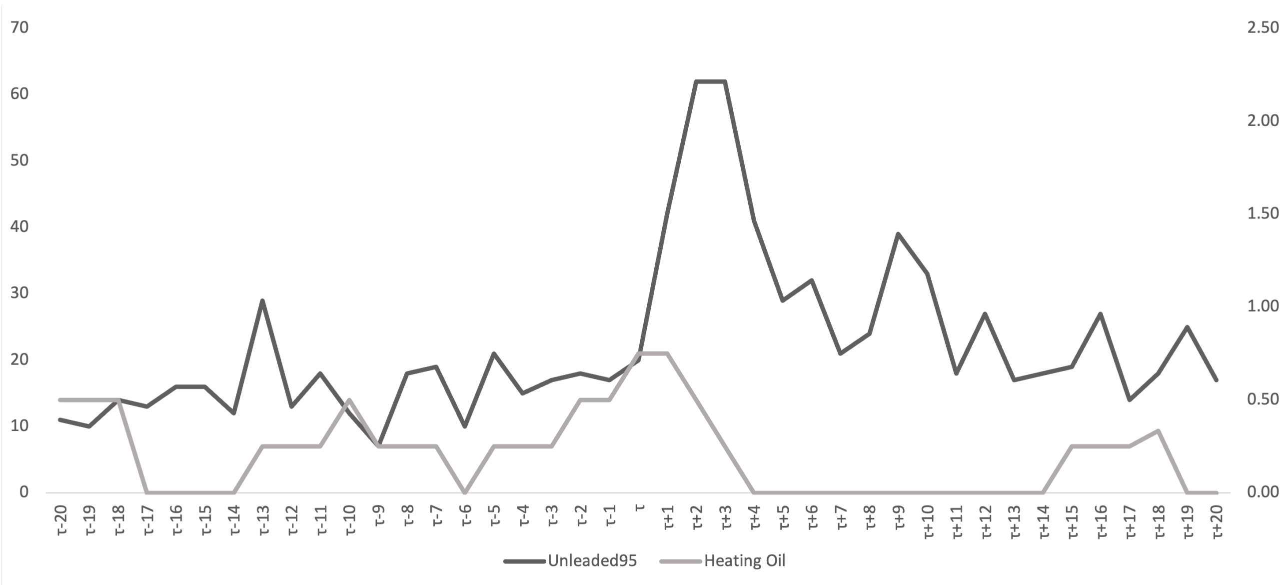
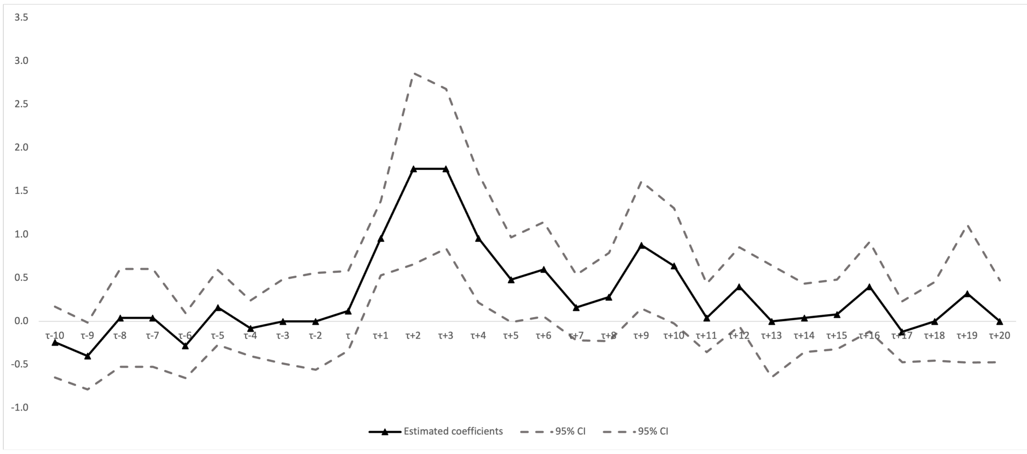
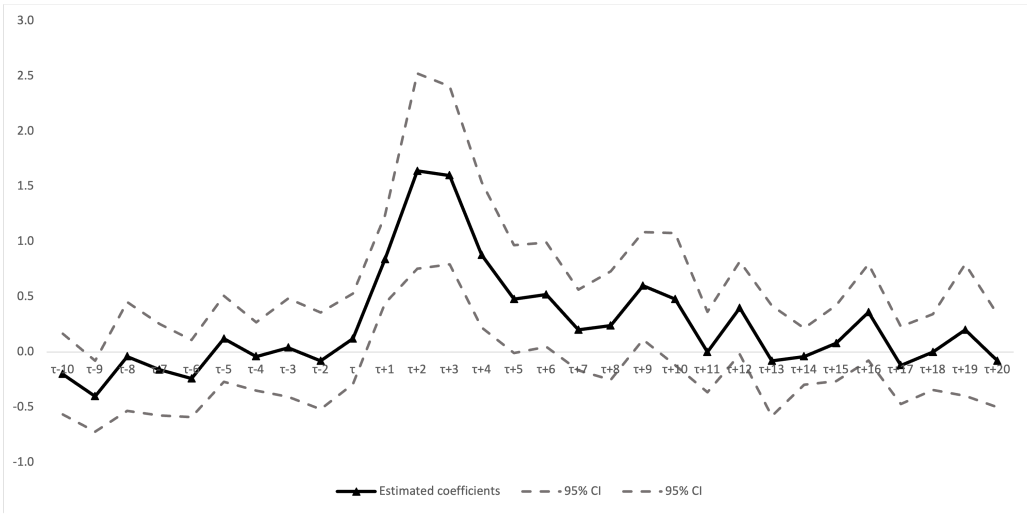


Figure 2.6: Consumer Search around the Tax Shock

Notes: This figure plots the number of searches for unleaded 95 and heating oil by day for the 20 days around the tax shock at day τ. We calculate and plot the 4-day moving averages for heating oil.



(a) Effect on Total Number of Search Queries



(b) Effect on the Number of Unique Searchers

Figure 2.7: Effect on Consumer Search

Notes: This figures plots the estimated coefficients from Table 2.6. τ is the day of the tax shock. $\tau - 1$ is used as the reference.

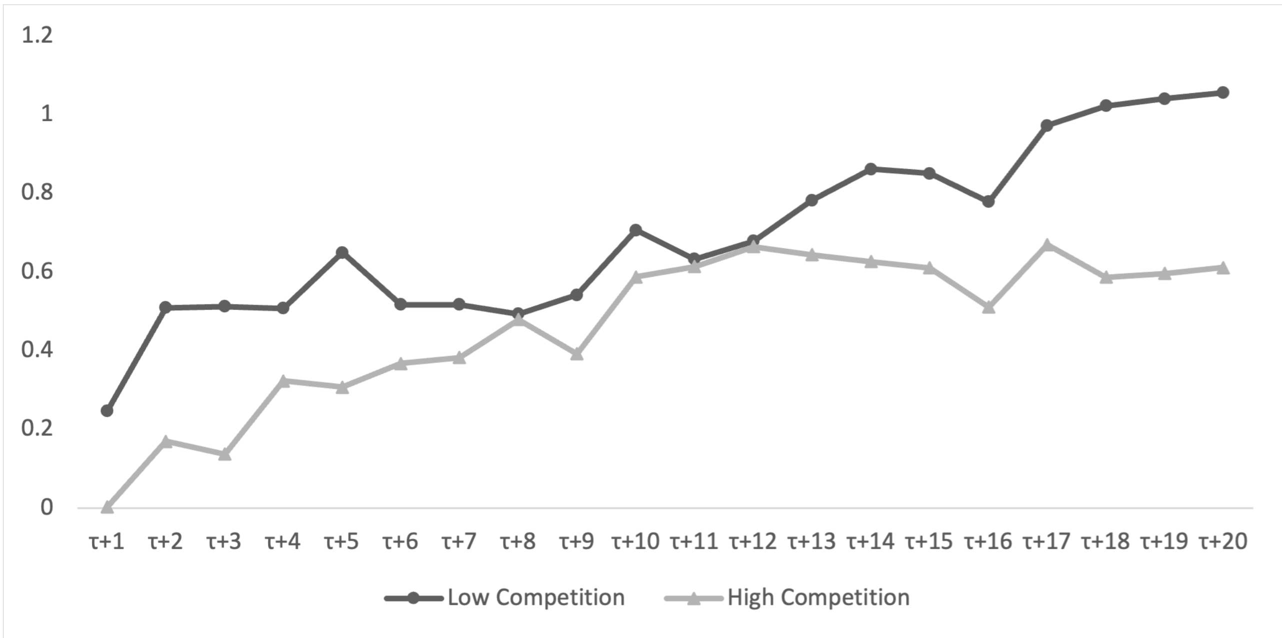


Figure 2.8: Competition and Speed of Adjustment

Notes: This figure plots estimated coefficients from Table 2.8. *Low Competition* is a dummy equals 1 for islands with 2 to 4 gas stations. *High Competition* is a dummy equals 1 for islands with at least 5 stations. τ is the date of the tax shock.

Appendix

Appendix A: Additional Tables

Table A1: Greek Islands

Number of gas stations	2	3	4	5	6	7	9
	Poros	Kea	Kythira	Milos	Andros	Mykonos	Ikaria
	Antiparos	Kythnos	Karpathos	Skiathos	Kalymnos	Tinos	Aigina
	Amorgos	Leros	Skopelos				Syros
	Astypalaia	Alonnisos					
	Folegandros	Ios					
	Spetses	Skyros					
	Ithaki	Sifnos					

Notes: This table shows the small isolated Greek islands included in our study, along with the number of gas stations on each island. We focus on islands with at least two gas stations.

Table A2: Robustness: Consumer Search Responds to Actual Price Changes

Variable	Search
Day($\tau = -5$)$\times$ Treat	0.154
	(0.155)
Day($\tau = -4$)$\times$ Treat	0.0769
	(0.212)
Day($\tau = -3$)$\times$ Treat	0.385
	(0.333)
Day($\tau = -2$)$\times$ Treat	-0.0417
	(0.295)
Day($\tau = 0$)$\times$ Treat	-0.0769
	(0.290)
Day($\tau = 1$)$\times$ Treat	0.958***
	(0.293)
Day($\tau = 2$)$\times$ Treat	1.804*
	(0.915)
Day($\tau = 3$)$\times$ Treat	2.266***
	(0.766)
Day($\tau = 4$)$\times$ Treat	1.077**
	(0.463)
Day($\tau = 5$)$\times$ Treat	0.846**
	(0.358)
Day($\tau = 6$)$\times$ Treat	0.385
	(0.268)
Day($\tau = 7$)$\times$ Treat	0.308
	(0.265)
Day($\tau = 8$)$\times$ Treat	0.385
	(0.387)
Day($\tau = 9$)$\times$ Treat	0.385
	(0.333)
Day($\tau = 10$)$\times$ Treat	0.923*
	(0.463)
F-test: p-value	0.0891
Island FE	Yes
Day FE	Yes
Observations	592

Notes: The sample comprises islands where the initial price adjustment occurs at least two days following the tax shock. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Robustness Check: Effect of the Tax Shock on Price Dispersion

	PD1	PD2	PD3
Day(τ=-10)× Treat	-1.201	-0.836	-0.517
	(0.910)	(0.618)	(0.496)
Day(τ=-9)× Treat	-1.476*	-0.913*	-0.772*
	(0.797)	(0.527)	(0.435)
Day(τ=-8)× Treat	-0.521**	-0.305**	-0.265*
	(0.246)	(0.141)	(0.133)
Day(τ=-7)× Treat	-0.538**	-0.298**	-0.240*
	(0.247)	(0.141)	(0.127)
Day(τ=-6)× Treat	-0.519**	-0.253*	-0.241*
	(0.218)	(0.130)	(0.125)
Day(τ=-5)× Treat	-0.206	-0.121	-0.218
	(0.161)	(0.082)	(0.145)
Day(τ=-4)× Treat	-0.054	-0.031	-0.055
	(0.112)	(0.059)	(0.066)
Day(τ=-3)× Treat	-0.103	-0.046	-0.054
	(0.103)	(0.047)	(0.055)
Day(τ=-2)× Treat	-0.035	-0.017	-0.015
	(0.036)	(0.021)	(0.026)
Day(τ=0)× Treat	0.142	0.094	0.060
	(0.147)	(0.104)	(0.075)
Day(τ=1)× Treat	1.029*	0.633*	0.487*
	(0.508)	(0.340)	(0.258)
Day(τ=2)× Treat	2.347**	1.060**	1.229*
	(0.952)	(0.452)	(0.691)
Day(τ=3)× Treat	2.500**	1.119**	1.315
	(1.028)	(0.474)	(0.770)
Day(τ=4)× Treat	3.144***	1.366***	1.651**
	(1.004)	(0.460)	(0.771)
Day(τ=5)× Treat	2.920**	1.290**	1.465*
	(1.060)	(0.476)	(0.811)
Day(τ=6)× Treat	3.310***	1.435***	1.814**
	(1.020)	(0.470)	(0.782)
Day(τ=7)× Treat	3.310***	1.450***	1.838**
	(1.020)	(0.467)	(0.781)
Day(τ=8)× Treat	3.524***	1.497***	2.139***
	(0.987)	(0.472)	(0.748)
Day(τ=9)× Treat	3.357***	1.432***	1.956**
	(1.037)	(0.493)	(0.788)
Day(τ=10)× Treat	3.596***	1.562***	2.325***
	(1.093)	(0.491)	(0.789)
Day(τ=11)× Treat	3.696***	1.527***	2.367***
	(1.152)	(0.515)	(0.814)
Day(τ=12)× Treat	3.420**	1.411**	2.261**
	(1.261)	(0.560)	(0.888)
Day(τ=13)× Treat	3.416**	1.366**	2.271**
	(1.246)	(0.545)	(0.879)
Day(τ=14)× Treat	3.471***	1.382**	2.249**
	(1.225)	(0.539)	(0.885)
Day(τ=15)× Treat	3.376**	1.323**	2.150**
	(1.227)	(0.530)	(0.859)
Day(τ=16)× Treat	3.261**	1.359**	2.085**
	(1.240)	(0.525)	(0.879)
Day(τ=17)× Treat	3.556**	1.503**	2.176**
	(1.280)	(0.562)	(0.900)
Day(τ=18)× Treat	3.839***	1.506**	2.339**
	(1.312)	(0.662)	(0.921)
Day(τ=19)× Treat	3.824***	1.513**	2.376**
	(1.312)	(0.661)	(0.919)
Day(τ=20)× Treat	3.951***	1.561**	2.455**
	(1.342)	(0.672)	(0.933)
Island FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Observations	1,441	1,441	1,441

Notes: The dependent variables are different measures of price dispersion constructed using the raw prices of the product. $PD1$ is the the sample range. $PD2$ is sample standard deviation, and $PD3$ is the gains from search. τ is the day of the tax shock. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Robustness: Effect on the Raw Number of Search Queries

Variable	Search
Day(τ=-10)× Treat	-2.680*
	(1.308)
Day(τ=-9)× Treat	-2.840*
	(1.651)
Day(τ=-8)× Treat	-0.480
	(1.419)
Day(τ=-7)× Treat	4.880
	(6.774)
Day(τ=-6)× Treat	-3.280**
	(1.515)
Day(τ=-5)× Treat	-0.880
	(1.651)
Day(τ=-4)× Treat	-0.600
	(1.247)
Day(τ=-3)× Treat	-0.360
	(1.352)
Day(τ=-2)× Treat	0.040
	(2.083)
Day(τ=0)× Treat	3.760
	(2.701)
Day(τ=1)× Treat	7.520**
	(3.433)
Day(τ=2)× Treat	14.80**
	(6.565)
Day(τ=3)× Treat	11.16**
	(4.873)
Day(τ=4)× Treat	8.680***
	(2.967)
Day(τ=5)× Treat	4.760
	(5.067)
Day(τ=6)× Treat	6.160
	(4.852)
Day(τ=7)× Treat	2.120
	(2.602)
Day(τ=8)× Treat	2.600
	(2.506)
Day(τ=9)× Treat	7.200**
	(3.022)
Day(τ=10)× Treat	2.560
	(2.728)
Day(τ=11)× Treat	-0.680
	(1.424)
Day(τ=12)× Treat	2.920
	(1.803)
Day(τ=13)× Treat	0.080
	(1.897)
Day(τ=14)× Treat	-1.560
	(1.264)
Day(τ=15)× Treat	-0.280
	(1.990)
Day(τ=16)× Treat	1.320
	(1.962)
Day(τ=17)× Treat	2.320
	(3.987)
Day(τ=18)× Treat	-1.120
	(1.817)
Day(τ=19)× Treat	3.360
	(4.468)
Day(τ=20)× Treat	-0.120
	(1.872)
Island FE	Yes
Day FE	Yes
Observatinos	1,550

Notes: The dependent variable is the raw number of search queries conducted on the platform. τ is the day of the tax shock. $\tau - 1$ is used as the reference. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Robustness: Event Study Regression on Treatment Group Only

Variables	(1) Search	(2) Searchers	(3) Raw Search
Day(τ=-10)	-0.200	-0.160	-2.560*
	(0.186)	(0.163)	(1.296)
Day(τ=-9)	-0.400**	-0.400**	-2.840*
	(0.186)	(0.156)	(1.651)
Day(τ=-8)	0.040	-0.040	-0.480
	(0.273)	(0.239)	(1.418)
Day(τ=-7)	0.080	-0.120	4.960
	(0.263)	(0.189)	(6.761)
Day(τ=-6)	-0.280	-0.240	-3.280**
	(0.182)	(0.169)	(1.515)
Day(τ=-5)	0.160	0.120	-0.880
	(0.210)	(0.189)	(1.650)
Day(τ=-4)	-0.080	-0.040	-0.600
	(0.155)	(0.150)	(1.246)
Day(τ=-3)	0	0.040	-0.360
	(0.236)	(0.216)	(1.352)
Day(τ=-2)	0.040	0	0.280
	(0.260)	(0.212)	(2.095)
Day(τ=0)	0.120	0.120	3.760
	(0.223)	(0.198)	(2.700)
Day(τ=1)	1***	0.880***	7.560**
	(0.212)	(0.198)	(3.430)
Day(τ=2)	1.800***	1.680***	14.84**
	(0.536)	(0.436)	(6.563)
Day(τ=3)	1.800***	1.640***	11.48**
	(0.456)	(0.407)	(4.905)
Day(τ=4)	0.960**	0.880**	8.680***
	(0.361)	(0.319)	(2.967)
Day(τ=5)	0.480*	0.480*	4.760
	(0.236)	(0.236)	(5.065)
Day(τ=6)	0.600**	0.520**	6.160
	(0.263)	(0.229)	(4.851)
Day(τ=7)	0.160	0.200	2.120
	(0.183)	(0.177)	(2.601)
Day(τ=8)	0.280	0.240	2.600
	(0.246)	(0.238)	(2.505)
Day(τ=9)	0.880**	0.640**	7.440**
	(0.355)	(0.242)	(3.016)
Day(τ=10)	0.640*	0.480	2.560
	(0.322)	(0.289)	(2.727)
Day(τ=11)	0.040	0	-0.680
	(0.191)	(0.177)	(1.424)
Day(τ=12)	0.400*	0.400*	2.920
	(0.220)	(0.204)	(1.802)
Day(τ=13)	0	-0.080	0.080
	(0.312)	(0.242)	(1.896)
Day(τ=14)	0.040	-0.040	-1.560
	(0.191)	(0.125)	(1.263)
Day(τ=15)	0.080	0.080	-0.280
	(0.195)	(0.166)	(1.990)
Day(τ=16)	0.400	0.360	1.320
	(0.250)	(0.211)	(1.962)
Day(τ=17)	-0.120	-0.120	2.320
	(0.170)	(0.170)	(3.986)
Day(τ=18)	0.040	0.040	-1.080
	(0.216)	(0.171)	(1.818)
Day(τ=19)	0.320	0.200	3.360
	(0.385)	(0.288)	(4.467)
Day(τ=20)	0	-0.080	-0.120
	(0.228)	(0.203)	(1.872)
Island FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Observations	775	775	775

Notes: This table shows estimated coefficients from an event study estimation on the subsample of treatment group only (unleaded95). Column (1) uses the number of searches as the dependent variable. Column (2) uses the number of unique searchers as the dependent variable. Column (3) uses the raw number of search queries as the dependent variable. τ is the day of the tax shock. $\tau - 1$ is used as the reference. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Robustness: Competition and Pass-through

Variables	Price	Price
Tax × Two competitors	1.068**	0.779
	(0.406)	(0.456)
Tax × Three competitors	1.048**	0.651*
	(0.498)	(0.346)
Tax × Four competitors	1.532**	1.567*
	(0.666)	(0.793)
Tax × Five competitors	0.540*	0.403
	(0.280)	(0.292)
Tax × ≥Six competitors	0.420*	0.413*
	(0.224)	(0.218)
Day FE	Yes	Yes
Product-station FE	Yes	Yes
Observations	406	413

Notes: Column (1) reports estimates on $\{\tau - 1, \tau + 20\}$; while Column (2) uses $\{\tau - 1, \tau + 10\}$. τ is the day of the tax shock. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Robustness: Competition and Price Dispersion

	(1)	(2)	(3)
Variables	PD1	PD2	PD3
Treated × Two competitors	-0.313	-1.221	-0.530
	(1.964)	(1.124)	(0.914)
Treated × Three competitors	5.475	0.731	3.755
	(7.498)	(3.995)	(4.656)
Treated × Four competitors	3.701**	1.137	2.703**
	(1.429)	(0.715)	(1.087)
Treated × Five competitors	-5.105	-3.794	-3.319
	(3.640)	(2.539)	(2.775)
Treated × ≥Six competitors	0.855	-1.315	0.671
	(1.753)	(0.804)	(1.144)
Day FE	Yes	Yes	Yes
Island FE	Yes	Yes	Yes
Observations	1,441	1,441	1,441

Notes: The dependent variable are different measures of price dispersion of each gasoline product on each island for $[\tau - 1, \tau + 20]$, where τ is the date of the tax shock. Treated is a dummy equals 1 for unleaded95 for post-shock days. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Robustness: Competition and Consumer Search

Variables	(1) Search	(2) Searcher	(3) Search	(4) Searcher
Treated × Two competitors	0.350***	0.329***	0.514**	0.477**
	(0.120)	(0.105)	(0.201)	(0.172)
Treated × Three competitors	0.355***	0.327***	0.449***	0.403***
	(0.0759)	(0.0710)	(0.0713)	(0.0662)
Treated × Four competitors	0.189***	0.185***	0.266***	0.258***
	(0.0249)	(0.0238)	(0.0296)	(0.0268)
Treated × Five competitors	0.215***	0.211***	0.344***	0.328***
	(0.0531)	(0.0469)	(0.0609)	(0.0561)
Treated × >=Six competitors	0.179***	0.156***	0.215***	0.191***
	(0.0231)	(0.0164)	(0.0405)	(0.0299)
Island FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Observations	1,550	1,550	1,050	1,050

Notes: Column (1) and Column (2) use a window $[\tau - 1, \tau + 20]$, while Column (3) and (4) use $[\tau - 1, \tau + 10]$. τ is the date of the tax shock. Standard errors, clustered at the station level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B: Additional Figures

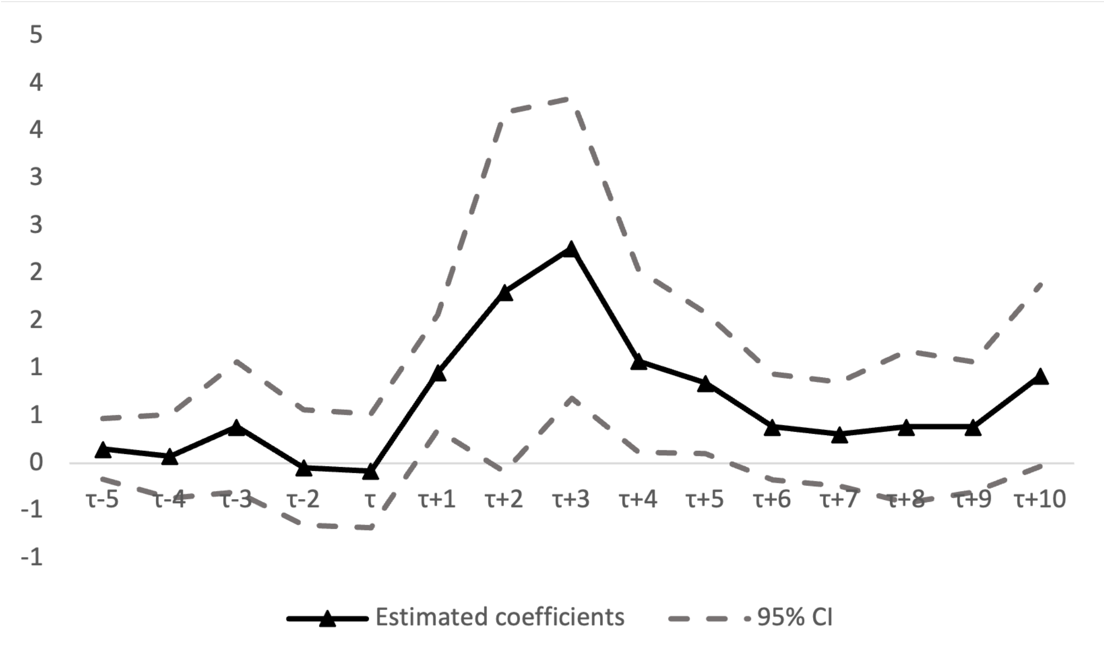


Figure B1: Robustness: Consumer Search Responds to Actual Price Changes

Notes: The dependent variables is the daily number of search queries conducted through fuelgr at the island level. The sample comprises islands where the initial price adjustment took place at least two days following the tax shock. Standard errors, clustered at the island level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

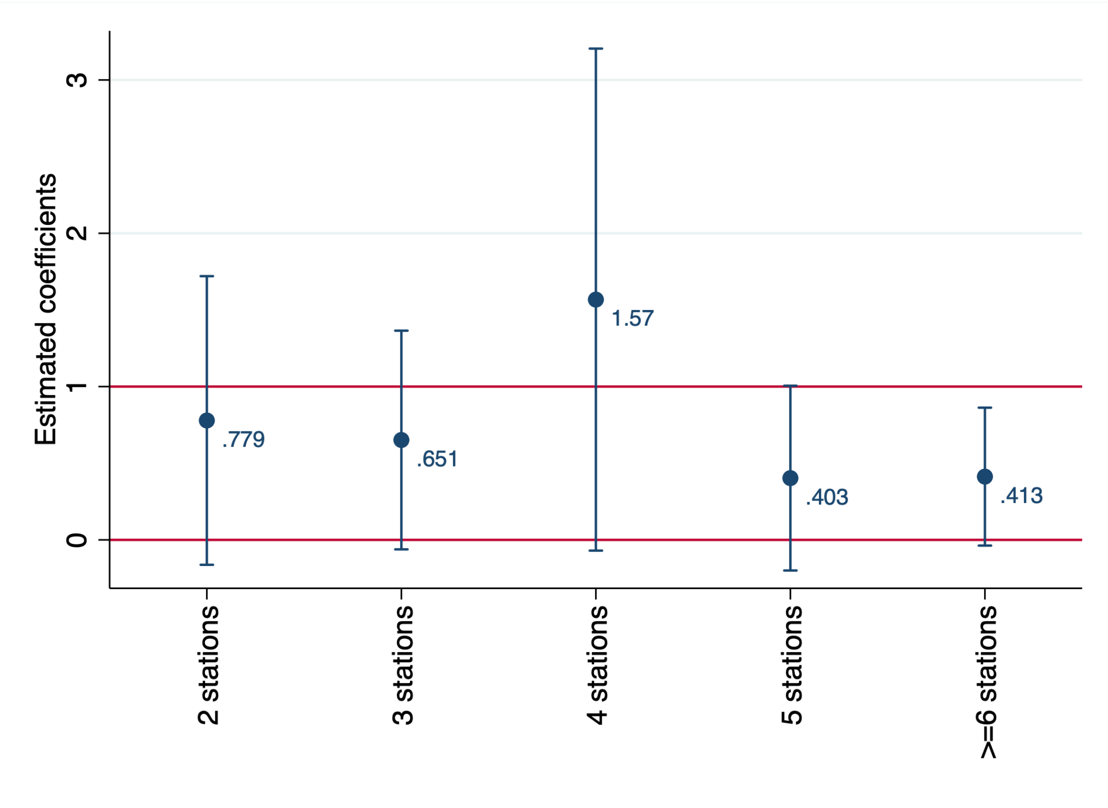
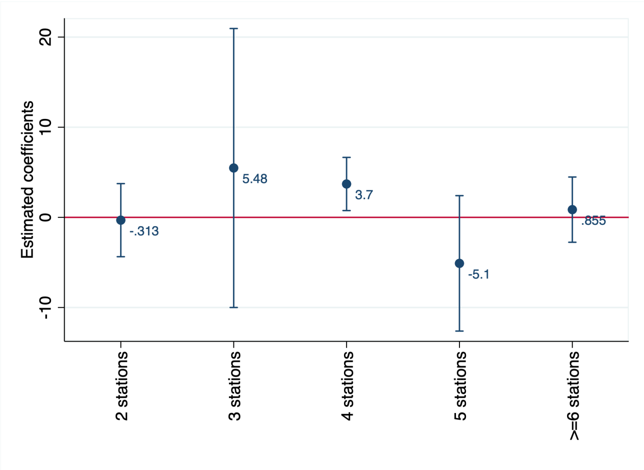
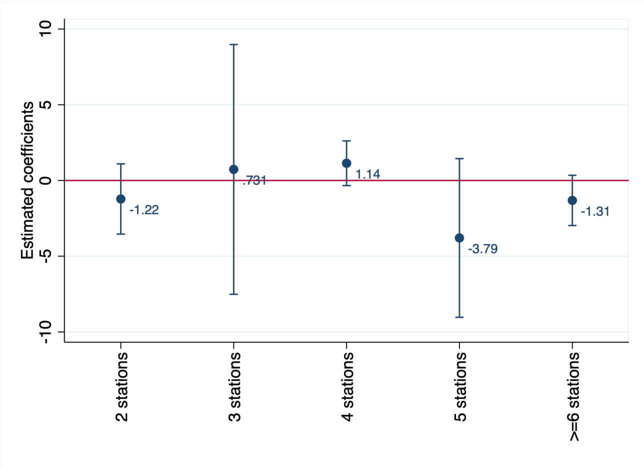


Figure B2: Robustness: Competition and Tax Pass-through

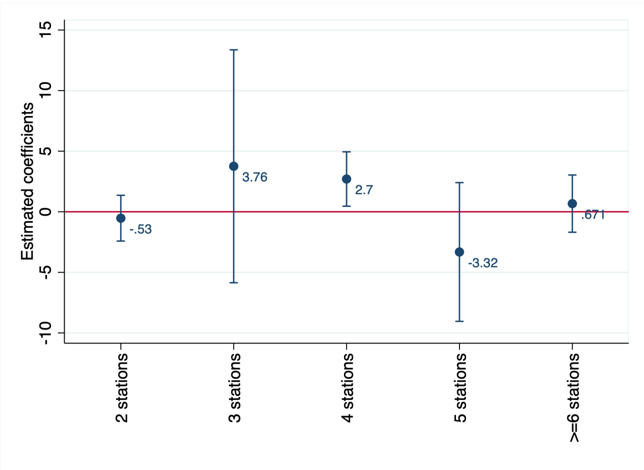
Notes: This figure plots estimated coefficients from Column (2) of the Table A6.



(a) Price Dispersion 1: Sample Range



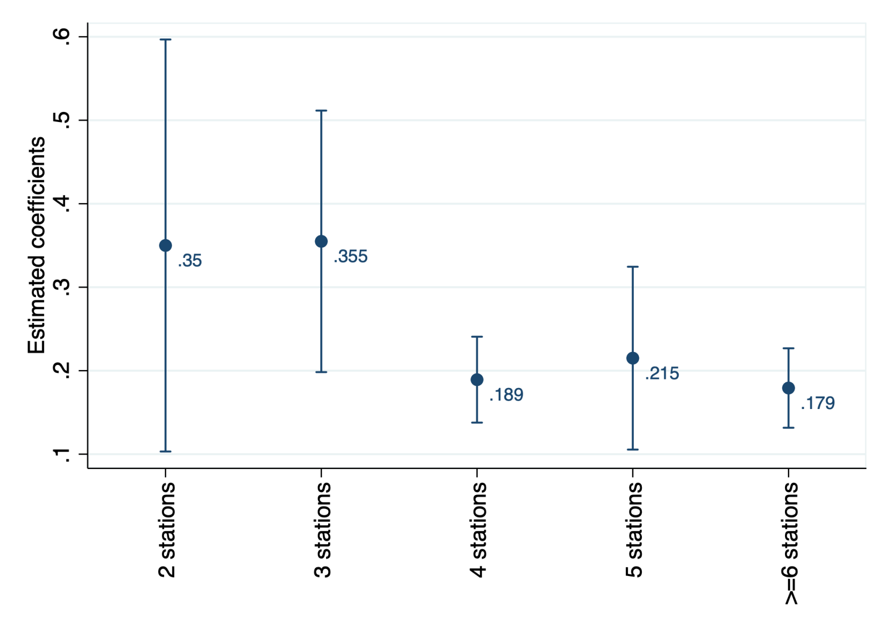
(b) Price Dispersion 2: Sample Standard Deviation



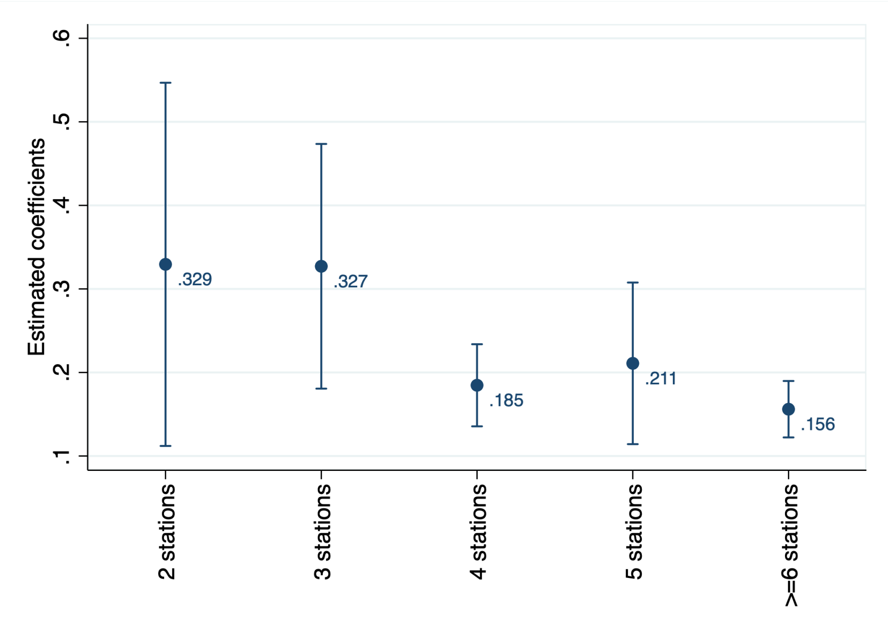
(c) Price Dispersion 3: Gains from Search

Figure B3: Competition and Price Dispersion

Notes: These figures plot estimated coefficients from Table A7.



(a) Competition and the Number of Searches



(b) Competition and the Number of Unique Searchers

Figure B4: Competition and Consumer Search

Notes: Fig. B4a and Fig. B4b plot estimated coefficients of Column (1) and Column (2) respectively from Table A8.

CHAPTER 3

Asymmetric Pass-Through and Competition

3.1 Introduction

Over the past thirty years, a large body of literature showed that retail prices tend to respond faster to marginal cost increases than to decreases. This asymmetric pass-through, or asymmetric price adjustment, also known as the "rockets and feathers" phenomenon, was first studied by Bacon (1991) in relation to several enquires by the Monopoly and Merger Commission in the UK gasoline market. Empirical evidence has then accumulated not only on the gasoline market (Borenstein et al., 1997), but on a large set of different markets, including food markets and financial markets. For example, Peltzman (2000) documents that asymmetric pass-through is common across a variety of industries in the US economy. Recent surveys of the large and growing literature on the topic (Frey and Manera, 2007; Bakucs et al., 2014) confirm that, although not ubiquitous, asymmetric pass-through is a general phenomenon across industries, countries, and periods.

Empirical research on the asymmetric pass-through has been largely devoted to establishing whether pass-through is symmetric or not. However, a more recent literature is developing combining theoretical and empirical analysis in an attempt to understand the causes of asymmetric price adjustments. Although several competing explanations exists, including market power and collusion (Bacon, 1991, Borenstein et al., 1997), inventory management (Borenstein et al., 1997), menu costs (Meyer and von Cramon-Taubadel, 2004), and search costs (Yang and Ye, 2008; Tappata, 2009; Cabral and Gilbukh, 2020), there is no consensus on the relative merits of the different explanations.

Although market power was the first conjectured explanation for asymmetric response, surprisingly few empirical papers provide specific evidence on the relation between competition and asymmetric pass-through. The main reason is that it is hard to simultaneously identify the asymmetry of price responses and the relation between asymmetry and competition. In fact, most studies rely on industry-specific data on exogenous input cost shocks (positive and negative) to estimate the asymmetric response of prices. However, since market structure is fixed at the industry level, there is typically little to no variability that can be used to identify the interaction between competition and asymmetric pass-through. A second problem is that market structure is likely to be endogenous, hence the intensity of competition may well be determined in equilibrium together with firms' pricing strategies.

This paper directly tackles the issue of estimating the relation between competition and asymmetric pass-through. This is a critical piece of evidence to assess the relevance of standard models of competition and their ability to explain observed pricing patterns in the data. We fill this gap in the literature by using the approach of Genakos and Pagliero (2022), who exploit repeated, large, and unexpected changes in excise duties for petroleum products and exogenous variability in market size across Greek islands. Small islands precisely define oligopolistic retail markets for petroleum products, as some of these are so small to accommodate only one or few gas stations, providing variability in number of competitors that is driven by the specific geography of the region.

We find that, on average, the tax pass-through is 0.7 for tax increases and 0.14 for tax decreases (measured after 10 days). For monopoly markets, the pass-through is about 0.4 for tax increases and 0.2 for tax decreases. The pass-through of a tax increase then grows with competition and converges towards 1, but that of a tax decrease remains constant. The asymmetry in price response grows from 0.2 to 0.8 as the number of competitors increases from one to four/five, but it does not further grow in larger and more competitive markets.

We also find a significant asymmetry in the speed of price adjustments to positive and negative shocks. The average pass-through for tax increases grows significantly with the adjustment period considered, going from 23 percent after one day to 71 percent after ten days. On the contrary, the pass-through for the tax decrease grows very slowly, going from 1 percent after 1 day to only 14 percent after ten days.

In monopoly markets, where asymmetry is smallest, we cannot reject the null hypothesis of a symmetric response. Although the precision of our estimates is limited, the failure to reject symmetry in monopoly markets is consistent with standard monopoly pricing. However, as competition increases, pass-through converges towards one, as predicted by oligopoly models, only for positive cost shocks. Although this asymmetry is not consistent with oligopoly models (Weyl and Fabinger, 2013; Miklós-Thal and Shaffer, 2021; Adachi and Fabinger, 2022), it is consistent with different models of asymmetric pass-through, for example with a dynamic equilibrium in which firms collude by not responding to negative cost shocks, as conjectured by the early literature on asymmetric pass-through, or with search cost models, where positive cost shocks trigger more active search by consumers than negative shocks and firms adjust their prices accordingly.

Our results on the relation between competition and asymmetric pass-through have also two additional implications for the growing empirical research on asymmetric pass-through. First, asymmetry is not necessarily apparent in very concentrated markets (monopolies in our case) and certain degree of competition might be necessary to observe a statistically significant asymmetry. Second, the relevant range for variability in competition is between one and four/five competitors. Beyond this point, the relation between competition and asymmetric pass-through flattens substantially. Hence, if one wants to test the impact of competition on asymmetric pass-through, it is critical to correctly identify oligopolistic markets with the appropriate number of competitors.

3.2 The Market for Petroleum Products and the Data

We focus on the retail market for petroleum products on a sample of small Greek islands. For gas stations, the marginal cost of petroleum products depends on long-term contracts with trade companies and is reasonably constant. Petroleum products are subject to excise duties, which is a unit tax rate (€-cents per liter), and the Value Added Tax (VAT), which is a percentage tax. The retail price is determined as $P_{retail} = (P_{refinery} + exciseduty\&fees + margins) \times (1 + VAT)$. We focus on the impact of changes in excise duties on prices, which are reported net of VAT.

We augment the data of Genakos and Pagliero (2022) to cover the 2010-2013 period. Our sample covers three substantial increases in excise duties occurred in 2010 and a subsequent drop in 2012. Each of these tax changes was announced and implemented the day after the decision was made, as typically happens in order to reduce opportunities for arbitrage. Table 3.1 shows that the tax changes were significant (between -20% and +29%) and different across products. Remarkably, in each of these four events the excise duties of one of the products shown on the table remained unchanged. Hence, for the three increases we use heating oil, whereas for the decrease we use unleaded 95 as our control group.

The data set includes daily station-level retail prices for five different gasoline products: unleaded 95, unleaded 100, super (or leaded gasoline), diesel, and heating oil. Our sample includes 37 islands with at most 131 gas stations and about 14,435 daily price observations.[1] Gas stations in our sample are independently operated and pricing decisions are taken locally. The data also includes socioeconomic (e.g., education, income, number of tourist arrivals) and geographic (size, distance from Piraeus[2], distance from mainland, number of ports and airports etc.) characteristics of each island from the Hellenic Statistical Authority (2010). We measure the number of gas stations operating in each island using independent information from Yellow Pages (2018) and company reports. We verify and update information on the number of gas stations as in some islands the number of gas stations slightly changed between 2010 and 2013. The key feature of our data is that arbitrage across islands is impossible, since transportation of petroleum products is too costly (and dangerous) to be economically viable, on top of being illegal. Hence, substitution effects across islands are reasonably absent and each island can be considered as an independent market.

3.3 Identification and Empirical Methodology

We use the econometric approach of Genakos and Pagliero (2022), but we allow for a different pass-through for the tax increases and decreases. Hence, results exactly replicate those for the tax increases, but differ for the tax decrease, allowing to test the asymmetry in estimated pass-through. Our baseline estimation framework is as follows:

[1]See Genakos and Pagliero (2022) for a discussion of the representativeness of the data and summary statistics.

[2]The primary distribution center for gasoline products in Greece.

$$P_{kist} = \beta_0 + \rho_U(U_t \times Tax_{kt}) + \rho_D(D_t \times Tax_{kt}) + \beta_{1ks}U_t + \beta_{2ks}D_t + \beta_t + e_{kist} \quad (3.1)$$

where P_{kist} denotes the retail price of product k, on island i, in gas station s, on day $t \in \{\tau(n) - 1,\ \tau(n) + \delta\}$, where $\tau(n)$ is the date of each of the four excise duty changes $(n = 1, \ldots 4)$ and $\delta = 1, \ldots, \delta_n$ is the length of the adjustment period considered. U_t is an indicator variable equal to one for the observations around the tax increases, and D_t is the corresponding indicator for periods around the tax decrease. Tax_{kt} is the excise duty, and the coefficients ρ_U and ρ_D capture the tax pass-through for positive and negative tax changes. Finally, the model includes product-gas station and calendar day fixed effects that capture any unobserved permanent differences across stations (geographical location, brand name, reputation etc) as well as macroeconomic common time shocks.

We then allow for a more flexible specification by making the parameters ρ_U and ρ_D depend on island characteristics:

$$P_{kist} = \beta_0 + \rho_U(n_i, Z_i)(U_t \times Tax_{kt}) + \rho_D(n_i, Z_i)(D_t \times Tax_{kt})\beta_{1ks}U_t + \beta_{2ks}D_t + \beta_t + e_{kist} \quad (3.2)$$

where the pass-through $\rho_U(n_i, Z_i)$ may be a linear function $\rho_U(n_i, Z_i) = \vartheta_{U0} + \vartheta_{U1}n_i + \vartheta_{U2}Z_i$ of the number of competitors n_i and other island specific characteristics Z_i and, similarly, $\rho_D(n_i, Z_i) = \vartheta_{D0} + \vartheta_{D1}n_i + \vartheta_{D2}Z_i$. The number of competitors and island characteristics do not vary around each excise duty change as we consider a relatively small time window (10 days) in each case. For simplicity, we omit the subscript t for variables n_i and Z_i in equation Eq. (3.2).

Alternatively, the relation between pass-through ρ_U and ρ_D and number of stations j can be non-parametrically estimated replacing $\rho_U(n_i) = \sum_j \rho_{Uj} I(n_i = j)$ and $\rho_D(n_i) = \sum_j \rho_{Dj} I(n_i = j)$, where I is an indicator variable for each observed number of gas stations on island i. The identifying assumption is $E(e_{kist}|X) = 0$, where X is the matrix of all covariates.

Although variables in Z_i capture the potential effect of other observed island characteristics on pass-through, in Section 3.4.1 we will also report IV estimates of Eq. (3.2), where exogenous variability in market size is used to estimate the impact of the number of competitors on pass-through. The rationale for this approach is based on the observation that market size is a crucial determinant of entry and competition, while it is arguably uncorrelated with unobservable determinants of the pass-through (such as demand convexity). Hence, the IV approach assumes that market size can be excluded from Z_i, while being correlated with measures of competition.

3.3.1 Parallel Trends

Following Ashenfelter et al. (2013) and Genakos and Pagliero (2022), we conduct two tests of the parallel trend assumption. First, we estimate the following equation:

$$P_{kist} = \beta_0 + \gamma Trend_t + \gamma_T Trend_t \times Treat + \beta_k + \beta_s + e_{kist} \tag{3.3}$$

where P_{kist} denotes the retail price of product k, on island i, in gas station s, on day t and $Treat$ is an indicator variable for products in the treatment group. We separately estimate Eq. (3.3) using data for the 10 days before each excise duty change. We then test and cannot reject the null hypothesis that the coefficient γ_T is equal to zero at the 5 percent confidence level (Table A1).

Second, we replace the trend variable in Eq. (3.3) with more flexible period-specific dummies β_t. We also replace the interaction of trend and the treatment group indicator with $\beta_t \times Treat$ and then test the null hypothesis that the coefficients of the period-specific interactions are all equal to zero (individually and jointly). Even with this more flexible specification, we cannot reject the null hypothesis of parallel trends at the 5 percent confidence level (Table A2).

3.4 Empirical Results

3.4.1 Asymmetric Pass-through and Competition

We use Eq. (3.1) and Eq. (3.2) to estimate the "average" pass-through and the "conditional" pass-through ("conditional on starting to adjust"), using respectively all the data or only the data for firms that have changed their prices at least once by a given date. We separately report results on average and conditional pass-through as they measure the "extensive" and the "intensive" margins of adjustment.

Table 3.2, Columns 1 and 2 report the estimated conditional and average pass-through for a 10-day adjustment period. The 10-day adjustment period is chosen so that it is close enough to the change in excise duty, but is also long enough for almost all of the gas stations (94%) to have changed their prices for the tax increases.[3] The average pass-through is 0.7 for a tax increase and 0.14 for a tax decrease. The conditional pass-through is only slightly higher. The differences between the pass-through for tax increases and decreases are significant at 1 percent confidence level.

Table 3.2, Columns 3-8 report the results of Eq. (3.2). Column 3 shows that average pass-through increases with competition for positive tax changes but does not depend on competition when it comes to a tax decrease. Column 4 adds additional interactions with variables Z_i. Finally, Column 5 reports IV estimates of the impact of competition on pass-through. The results are not substantially affected. One additional competitor implies an increase in the asymmetry of the average pass-through between 7 and 9 percent.

Fig. 3.1a reports the average pass-through obtained using the semiparametric version of Eq. (3.2), where pass-through is estimated using interactions with dummies for the number of competitors. Table A3 in the Appendix reports the corresponding regression coefficients and standard errors. For monopoly markets, the level of pass-through is about 0.427 for tax increases and 0.178 for tax decreases. In spite of the large difference in point estimates, we cannot reject the symmetry of pass-through for monopoly markets (p-value = 0.270).[4]

For duopoly markets the estimated pass-through is 0.54 for tax increases and 0.147 for the tax decrease and their difference is statistically significant (p-value = 0.036). Then the pass-through of a tax increase sharply grows with competition and converges to 1 in markets with 4 competitors, but that of a tax decrease does not show any systematic correlation with competition. The asymmetry in price response grows systematically as the number of competitors increases. In markets with 6 or more firms, the estimated pass-through for tax increases is still not significantly different from one and that of tax decreases is still not significantly different from zero. The difference in pass-through is about 0.8.

Fig. 3.1b shows the same pattern for the conditional pass-through. Since this is computed using only data for firms that have adjusted their price at least once, the conditional

[3] For the tax decrease, we also estimated the specifications using longer time windows (40 days) after the policy change and the results remain unchanged (see Table A5).

[4] Refer to Table A4 for the results of the coefficient equality tests for Table A3.

pass-through is systematically higher than the average pass-through. However, differences are small and do not impact our general results on the asymmetric pass-through. Even conditional on adjusting their prices at least once, firms change their prices very little in response to a drop in taxes, no matter what the level of competition is. On the other hand, firms fully adjust their prices to tax hikes when competition is sufficiently intense.

3.4.2 Asymmetric Pass-through and Speed of Adjustment

We re-estimate Eq. (3.1) changing the adjustment window around each tax change from 1 to 10 days. The results are reported in Fig. 3.2, in which each pair of points corresponds to the pass-through for positive and negative tax changes estimated for a different adjustment period. Each pair of points comes from a separate regression using Eq. (3.1). Table A5 in the Appendix reports the corresponding regression results.

Fig. 3.2a shows that the average pass-through for a tax increase grows sharply with the adjustment period, from about 0.23 one day after the policy change to 0.71 after 10 days. On the contrary, the pass-through for a tax decrease grows very slowly, from about 0.01 one day after the change to 0.14 after 10 days. The wedge between the two lines becomes statistically significant (at 5% confidence level) when the adjustment period is two days and grows thereafter.

Fig. 3.2b reports the corresponding values for the conditional pass-through. The conditional pass-through does not substantially change with the length of the adjustment period for tax increases. This reflects the fact that when firms adjust their prices, they tend to do that fully to the new level. However, there is a very slow increase of the conditional pass-through for the tax decrease, as firms partially adjust prices even after ten days after the tax change. The estimated pass-through (average and conditional) for longer adjustment periods are not significantly different. The results are reported in Table A5, Columns 3 and 4 in the Appendix.

We also compute the results of Fig. 3.2 and Table A5 splitting islands into two groups. The "low competition" group includes those with 1 to 3 competitors and the "high competition" group those with 4 or more competitors. Although the speed of adjustment is higher for more competitive markets when taxes increase, we do not detect any significant difference between more and less competitive markets when taxes decrease. The results are reported in Table A6 in the Appendix.

Finally, Fig. 3.3 reports the cumulative frequency of station-product combinations that changed their prices between τ and $\tau + \delta$ for the tax decrease, on islands with 1-3 (low competition) and 4-7 (high competition) gas stations. The Kolmogorov-Smirnov test does not reject the equality of the CDFs for the two groups of islands at the 1 percent confidence level. This implies that competition does not significantly affect the speed of price adjustment when taxes decrease. This stands in contrast with the corresponding results for the tax increases (Figure 6 in Genakos and Pagliero (2022)), which show significant differences between the two CDFs for islands with low and high competition for the tax increases.

3.4.3 Consistency of Empirical Results on Pass-through Asymmetry

Borenstein et al. (1997) find a pass-through of 0.55 of positive shocks after two weeks but no significant response to negative shocks (see, their Figure 3) using US data. Pass-through of negative shocks grows slowly over time generating asymmetry that becomes insignificant after six weeks. In more recent studies, Johnson (2002) finds a cumulative pass-through of 0.50 vs. 0.16 for gasoline after two weeks, and 0.70 vs. 0.40 for diesel (see Table V), while in Verlinda (2008) the pass-through after three weeks is 1.10 for wholesale cost increases, but only 0.83 for cost decreases. Finally, Montag et al. (2021) also confirm this asymmetry in Germany, estimating a pass-through of 0.34 to 0.79 for the VAT tax cut and 0.69 to 0.92 for the tax rise.

Beyond gasoline, Peltzman (2000) reports a pass-through of 0.235 for a one percent input price increase vs. 0.127 for an equivalent decrease in input cost for a large and diverse sample of consumer and producer goods. Similarly, Benzarti et al. (2020) report a pass-through of 35% to prices of VAT increases in Europe for a wide variety of goods, while VAT decreases are pass-through only 6% one month after the reforms. Therefore, although the pass-through magnitudes differ across products, time and countries, the ratio between cost or tax increases and decreases varies between two and five times, similar to what we observe in the Greek islands environment.

3.4.4 Consistency of Results with Alternative Theoretical Models

In this subsection, we discuss the most relevant possible explanations that the literature has provided to explain the asymmetric price response.

Menu costs: Menu costs may generate asymmetric responses if drops in marginal costs are short lived, leading to temporary adjustments only, which do not allow to recover the fixed costs involved in changing prices (Blinder, 1982; Ball and Mankiw, 1994; Kovenock and Widdows, 1998). This type of explanation is unlikely to explain our results for two reasons. First, menu costs are negligible for gas stations, which typically adjust prices at a very high frequency and do not face any physical cost or information cost involved in adjusting prices. Second, tax changes are typically long lived, and the tax drop of January 2013 was no exception. Differently from most of the literature, in our analysis we do not use high frequency changes in crude oil prices as a source of identifying variation, but long lasting changes in taxes.

Inventory management: The quantity adjustment caused by a drop in price is constrained by existing inventories. Hence, in principle, this may limit the ability to decrease prices. However, this explanation is unlikely to hold for retail markets (Borenstein and Shepard, 1996; Borenstein et al., 1997), as gas stations generally hold sufficient inventories in underground tanks to accommodate the increased demand and may receive new deliveries at short notice.

Market power and collusion: The oldest explanation for asymmetric pass-through posits that firms collude on prices when costs fall using dynamic strategies based on the threat of a punishment phase (e.g. a price war) in case of deviations. Among the infinite number of strategies that can support collusion in repeated games, those using the old retail price as focal point for collusion after a cost drop seem natural candidates for collusion (Bacon, 1991; Borenstein et al., 1997).

On monopoly islands there is no competition and no role for collusion, hence we should see no asymmetry. This is consistent with our results. In all other market configurations, we cannot reject that the pass-through for negative tax changes is equal to that of a monopolist. Taken together, our results show no relation between competition and pass-through for negative tax changes.

In principle, these two results are consistent with gas stations competing when costs increase but colluding and pricing “like a monopolist” when costs fall. However, the collusion explanation requires firms not only being able to monitor (at least imperfectly) each other and having a focal point ("keep price constant when cost falls"), but also having a sufficiently high discount factor. In most collusive equilibria in repeated games, there is a threshold value of the discount factor that guarantees the stability of collusion. This threshold generally increases with the number of firms, making collusion more difficult as competition increases. Hence, in more competitive markets, deviations from the "constant price strategy when costs fall" should be more likely, and, on average, we should observe an increasing pass-through of negative cost shocks as the number of competitors increases.

We do not find evidence of this general comparative static result. First, we do not observe higher pass-through in markets with more firms when tax decreases. This is surprising, since we do see a large increase in pass-through for positive cost changes, which is consistent with a significant increase in competition. (Hence, it seems that the range in which the number of firms is varying in our sample is the relevant range for measuring competition). Second, we do not find any direct evidence of price wars in which collusion breaks down in any period in our sample. We do not observe sudden drops in prices on any island in any period, not even on islands with six or more firms.[5] Instead we observe a very slow and gradual adjustment of prices even after forty days since the tax decrease (Fig. 3.2, Table A5). Although observing price wars is not theoretically a necessary condition for the collusion explanation (in equilibrium we might not observe any price war; Green and Porter, 1984), the empirical literature has emphasized the empirical relevance of price wars in collusive markets (Byrne and De Roos, 2019). Hence, it seems unlikely that collusion is prevalent in practice, but we do not find any direct evidence of sudden island-specific price drops that are specific to a given island.

Deltas (2008) uses monthly state level data and shows that markets with high average retail-wholesale margins (high market power) experience a slower adjustment and a more asymmetric response. Verlinda (2008) studies how local market power affects asymmetric pass-through. Exploiting variation local supply and demand conditions to proxy for market power, he finds that increasing the number of rivals within one mile

[5]The evidence on this point is obtained by plotting the difference between the price for each station-product and the average price for each island-product. Sudden increases of this difference for some stations and decreases for others should signal a price war (results not reported here, available on request).

and decreasing the distance to the nearest rival decreases asymmetry. Assuming these variables are negatively correlated with market power, these results are consistent with more market power increasing the likelihood of collusion and asymmetric pass-through.[6] However, both papers, as well as the literature that follows, typically defined markets based on the distance between gas stations (Shepard, 1991; Barron et al., 2004; Eckert and West, 2005; Hosken et al., 2008). While realistic, this approach cannot guarantee the absence of substitution effects with firms outside the geographical area considered. In contrast, Greek islands clearly define local markets and allows us to measure market power in a clean way.

Consumers' search: Our results are consistent with asymmetric pass-through being caused by consumers' search behavior rather than collusion. On monopoly islands, there is no reason for consumers to search for the lowest price. Hence, if asymmetric pass-through is caused by search, we should see no asymmetry on monopoly islands also in this case. This simple prediction is consistent with our results.

In search models, more price dispersion generally leads to more search, hence higher pass-through. Asymmetric pass-through occurs because incentives to search are higher when costs increase than decrease. Although the specific mechanism that leads to consumers' search depends on the modelling assumptions, the intuition is that, faced with more consumers searching, firms compete more intensely.

Although we do not directly observe search, we use the daily price data to construct different measures of price dispersion for each gasoline product on each island. First, we construct 'clean' or 'residual' prices. These are the price level net of any persistent seller heterogeneity and are obtained from a regression of raw prices on station fixed effects. Then, we construct three measures of price dispersion using these 'clean' prices. The first measure is the sample range SR_{kit}, calculated as the difference between the maximum and the minimum price in the market, that is, $SR_{kit} = RP^{\max}_{ikt} - RP^{\min}_{ikt}$. This measure captures, on average, the most a consumer can save by searching every gas station in the market. However, this measure might be strongly influenced by outliers, so we construct the second measure using the sample standard deviation at each island SD_{kit}, calculated as $SD_{kit} = \sqrt{\sum_{s\in i} \frac{(RP_{\text{kist}} - \overline{RP}_{\text{kit}})^2}{N_i}}$. This measure does not rely on extreme values and is commonly used as a measure of price dispersion in the literature (see for instance,

[6]However, they are also consistent with search theories, since the number of rivals in close proximity and the distance from rivals should decrease search costs and therefore reduce asymmetry.

Noel 2018). For our last measure, we construct gains from search (GS_{kit}) for consumers $\mathrm{GS}_{kit} = E_i(RP_{kist} - RP_{ikst}^{\min})$, which is the difference between the expected price and the expected minimum price in each market i (Chandra and Tappata 2011). We calculate the expected price for each gasoline product in the market using the average market price for that product.

We then estimate a diff-in-diff model similar to Eq. (3.1),

$$PD(P_{\text{kit}}) = \beta_0 + \rho_U(U_t \times \text{Tax}_{kt}) + \rho_D(D_t \times \text{Tax}_{kt}) + \beta_{1k}U_t + \beta_{2k}D_t + \beta_t + e_{kit} \quad (3.4)$$

where $PD(P_{kit})$ is any of the three measures of price dispersion we introduced above for product k on island i in period t. This regression provides evidence on whether tax increases and decreases affect differentially price variability across firms. We then allow parameters ρ_U and ρ_D to depend on island characteristics, $\rho_U(n_i, Z_i)$ and $\rho_D(n_i, Z_i)$.

Table 3.3 summarizes the results. As we can see in the first three columns, increases in taxes led to increases in the variability of prices, no matter which measure of variability we use. On the contrary, decreases in taxes had no effect on price variability. The last three columns confirm that competition has a positive effect on price variability, but only when taxes increase. Overall, our results are consistent with consumers searching more when costs increase than decrease and searching more in more competitive markets.

Tappata (2009) explains asymmetric pass-through in a model of all-or-nothing simultaneous endogenous search with cost persistence (modeled as a Markov process). The intuition is that when marginal costs are expected to remain relatively high, consumers expect prices to remain high and have little dispersion. Hence, consumers have little incentive to search. In this case, if costs unexpectedly fall, sellers will not adjust prices downwards, as consumers will tend not to react to the lower prices. On the other hand, when marginal costs are expected to remain relatively low, consumers expect prices to remain low and have large dispersion. Hence, incentives to search are high. In this case, if costs unexpectedly increase, sellers will adjust prices upwards as consumers will tend to react to the higher prices. In practice, firms face more inelastic demands when the cost drops than when it goes up and this generates the asymmetric pass-through.

Yang and Ye (2008) also propose a model of search. The model shares a number of features with Tappata (2009), such as non-sequential search and Markov dynamics with persistent cost shocks. However, consumers never observe past cost realizations and gradually learn the true state. In equilibrium, consumers quickly learn about cost increases and slowly learn about cost decreases, leading to faster upward adjustment of prices. As the cost shocks become more persistent, the downward price adjustment on average spreads over longer periods of time. These characteristics of the equilibrium are in line with the differences in the speed of price adjustments that we estimate for tax increases and decreases, which are relatively infrequent and persistent policy changes.

Lewis (2011) develops a search model in which consumers' expectations of prices are based on mean prices observed during previous periods, so that expectations (or reference prices) are adaptive. The model generates asymmetric pass-through and predicts that consumers search less when prices are falling, which results in higher profit margins and a slower price response to cost changes. The model has only two firms and there are no comparative statics with respect to competition. Implications regarding margins cannot be tested as we do not have margins. Finally, Remer (2015) shows that prices for premium gasoline fall more slowly than regular gasoline, which is reasonably purchased by consumers with lower search costs. We cannot test this as our treatment group includes only one type of fuel.

3.4.5 Implications for Research on Asymmetric Pass-through

Our results on the relation between competition and asymmetric pass-through have also two additional implications for the growing empirical research on asymmetric pass-through. First, asymmetry is not necessarily apparent in very concentrated markets (monopolies, for example) and certain degree of competition might be necessary to observe a statistically significant asymmetry. Second, the relevant range for variability in competition is between one and four/five competitors. Beyond this point, the relation between competition and asymmetric pass-through flattens substantially. Hence, if one wants to test the impact of competition on asymmetric pass-through, it is critical to correctly identify oligopolistic markets with the appropriate number of competitors.

3.5 Conclusions

This paper contributes to the empirical literature by providing new evidence on the relation between competition and asymmetric pass-through that could be useful in the search for the causes of the "rockets and feathers" phenomenon. We document a strong asymmetry in retail prices with the average tax pass-through for tax increases to be five times larger than that of tax decreases (0.7 vs. 0.14). Most importantly, we show that the pass-through of a tax increase grows with competition and converges to 1 after four/five competitors, whereas that of the tax decrease does not vary with competition (asymmetric competition effect). We also find a significant asymmetry in the speed of price adjustments, with tax increases been transmitted much faster than tax decreases. These finding have important policy implications as often times governments around the world have been modifying tax rates trying to raise more revenue or provide a fiscal stimulus (for example, during the recent Covid19 pandemic).

We provide evidence that potential explanations, such as menu costs, inventory management and, particularly, market power and collusion are unlikely to be the sources for the "rockets and feathers" phenomenon. Our results are consistent with consumers searching more when there is a positive than a negative shock and searching more in more competitive markets. More research is needed in this direction to more precisely understand the exact mechanism that search is affecting price dispersion and pass-through.

We acknowledge that Greek islands are not necessarily representative of oligopolistic markets for other products. However, we selected this environment precisely because it provides clean variation in the competitive environment and allows us to compare different tax shocks across different markets within the same country. We believe that the results contribute to our understanding of asymmetric price adjustment, by showing new evidence on relationships that may be present in other settings and in larger markets.

Tables

Table 3.1: Excise Duty Tax Changes (€cents per litre and $\Delta\%$)

Type of energy product	(1) Unleaded 95	(2) Unleaded 100	(3) Diesel	(4) Super (leaded)	(5) Heating oil
before	41	41	30.2	42.1	2.1
10-Feb-10	53	53	35.2	54.1	2.1
	(29%)	(29%)	(17%)	(29%)	(0%)
04-Mar-10	61	61	38.2	62.1	2.1
	(15%)	(15%)	(9%)	(15%)	(0%)
03-May-10	67	67	41.2	68.1	2.1
	(10%)	(10%)	(8%)	(10%)	(0%)
15-Oct-12	67	67	33	68.1	33
	(0%)	(0%)	(-20%)	(0%)	(1471%)

Notes: The table reports the level and percentage changes in excise duties by product.

Table 3.2: Excise Duty Pass-through and Competition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation Method	FE	FE	FE	FE	IV	FE	FE	IV
Pass-through Definition	Average	Conditional	Average	Average	Average	Conditional	Conditional	Conditional
Dependent Variable	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$
Sample	$\tau-1,\tau+10$	$\tau-1,\tau+10$	$\tau-1,\tau+10$	$\tau-1,\tau+10$	$\tau-1,\tau+10$	$\tau-1,\tau+10$	$\tau-1,\tau+10$	$\tau-1,\tau+10$
Tax_{kt} × Increase	0.713***	0.767***	0.409***	-0.465	0.403***	0.449***	-0.668	0.464***
	(0.101)	(0.097)	(0.106)	(0.858)	(0.129)	(0.101)	(0.851)	(0.109)
Tax_{kt} × Increase × Number of competitors			0.082***	0.066*	0.083***	0.086***	0.079**	0.082***
			(0.023)	(0.034)	(0.024)	(0.024)	(0.032)	(0.023)
Tax_{kt} × Decrease	0.142**	0.175***	0.217*	-0.694	0.316**	0.330**	-0.785	0.458***
	(0.053)	(0.059)	(0.127)	(0.757)	(0.135)	(0.156)	(0.763)	(0.165)
Tax_{kt} × Decrease × Number of competitors			-0.013	-0.014	-0.029*	-0.024	-0.024	-0.045**
			(0.018)	(0.019)	(0.017)	(0.022)	(0.020)	(0.020)
First stage F-test (Increase × Number of competitors)					26.87			30.64
First stage F-test (Decrease × Number of competitors)					21.31			21.95
Observations	1341	1253	1341	1341	1341	1253	1253	1253
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Product × Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Product type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls				Yes			Yes	
Test equality of coefficients (p-value)								
Tax_{kt} × Increase=Tax_{kt} × Decrease	0.000	0.000						
Tax_{kt} × Increase × Number of competitors= Tax_{kt} × Decrease × Number of competitors			0.002	0.018	0.000	0.001	0.004	0.000

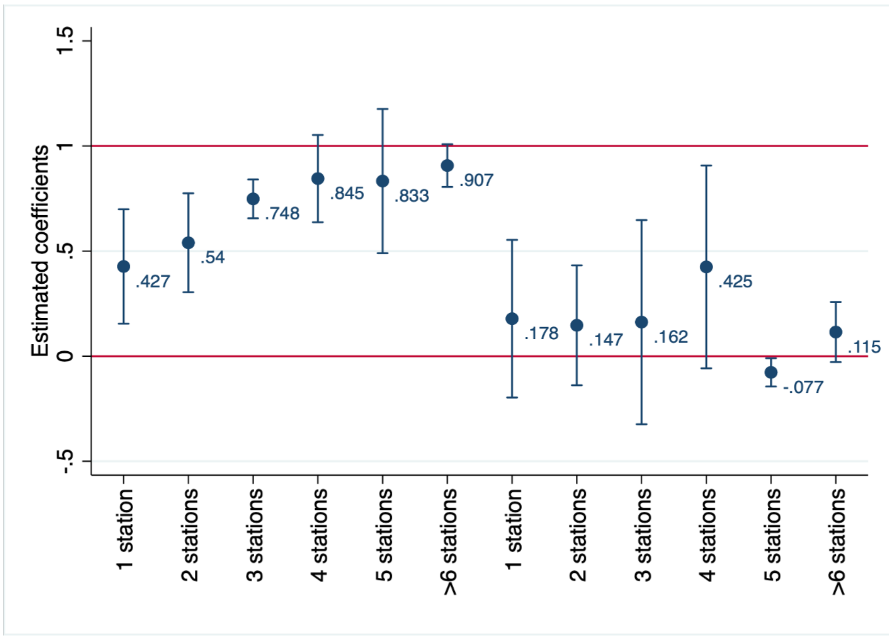
Notes: The dependent variable is the retail price of product k, on island i , in gas station s, and day $t \in \{\tau-1, \tau+10\}$, where τ is the date of each of the four excise duty changes. The pass-through is estimated using observations for station-product combinations that have changed the price at least once between τ and τ+10 (conditional pass-through), or all the available data (average pass-through). Standard errors clustered at the island level are reported in parentheses below coefficients. Additional controls include interactions with income, education, number of ports, and airports, distance from Piraeus and tourist arrivals. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.3: Test Consumer Search Explanation

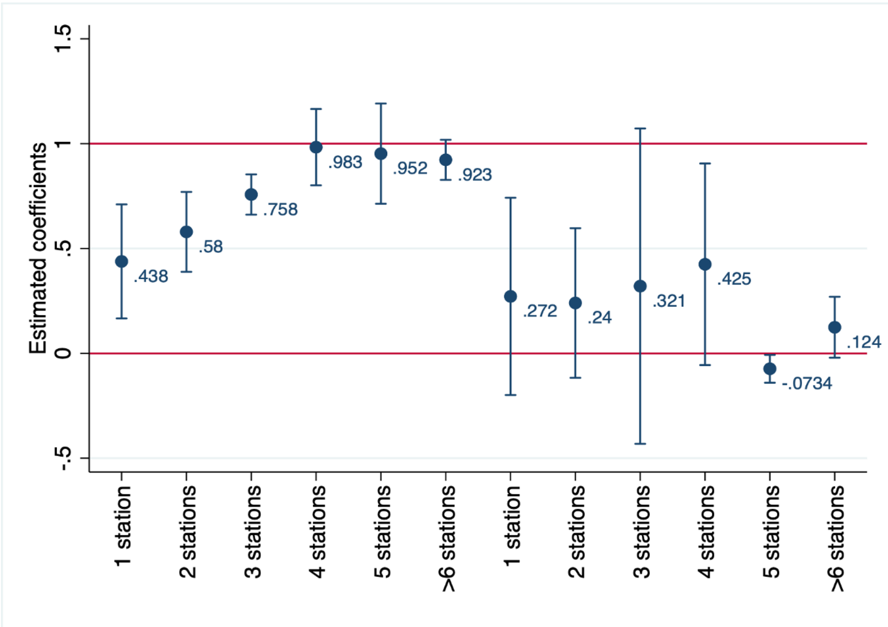
	(1)	(2)	(3)	(4)	(5)	(6)
Estimation Method	FE	FE	FE	FE	FE	FE
Dependent variable	PD1	PD2	PD3	PD1	PD2	PD3
Sample	τ-1, τ+10	τ-1, τ+10	τ-1, τ+10	τ-1, τ+10	τ-1, τ+10	τ-1, τ+10
Tax_{kt} × Increase	0.330***	0.191***	0.166***	0.113	0.072	0.057
	(0.101)	(0.063)	(0.052)	(0.185)	(0.121)	(0.086)
Tax_{kt} × Increase × Number of competitors				0.017***	0.007***	0.008***
				(0.003)	(0.002)	(0.001)
Tax_{kt} × Decrease	-0.004	0.025	0.002	-0.026	0.067	-0.008
	(0.080)	(0.039)	(0.062)	(0.322)	(0.187)	(0.165)
Tax_{kt} × Decrease × Number of competitors				-0.001	-0.003	-0.0001
				(0.007)	(0.004)	(0.004)
Observations	399	399	399	399	399	399
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Product type FE	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls				Yes	Yes	Yes

Notes: The dependent variable is the different measure of price dispersion of product k, on island i and day $t \in \{\tau - 1, \tau + 10\}$, where τ is the date of each of the four excise duty changes. PD1 is the the sample range. PD2 is sample standard deviation, and PD3 is the gains from search. Additional controls include interactions with income, education, number of ports, and airports, distance from Piraeus and tourist arrivals. *** p<0.01, ** p<0.05, * p<0.1.

Figures



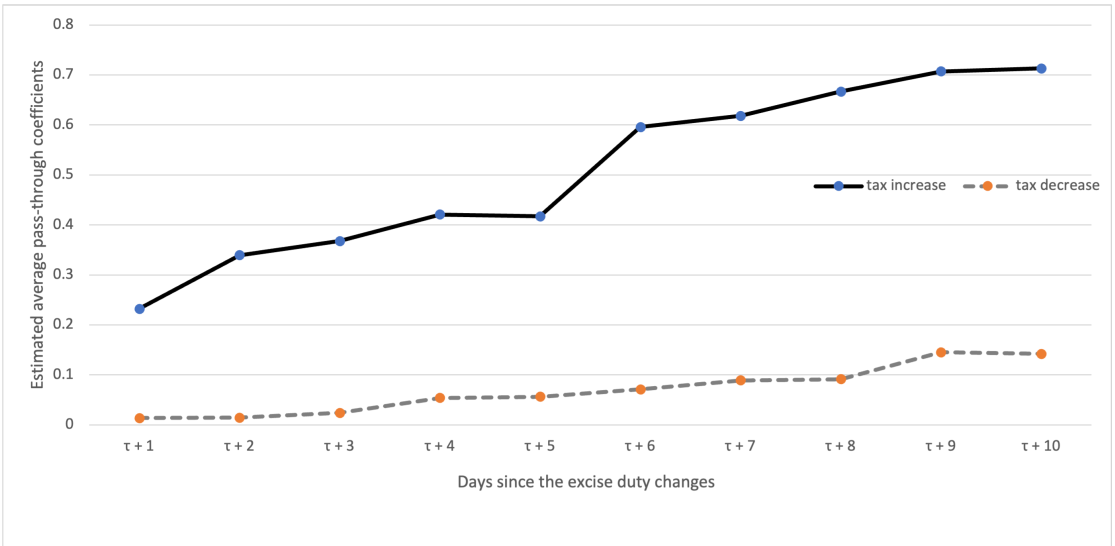
(a) Average Pass-through and Competition



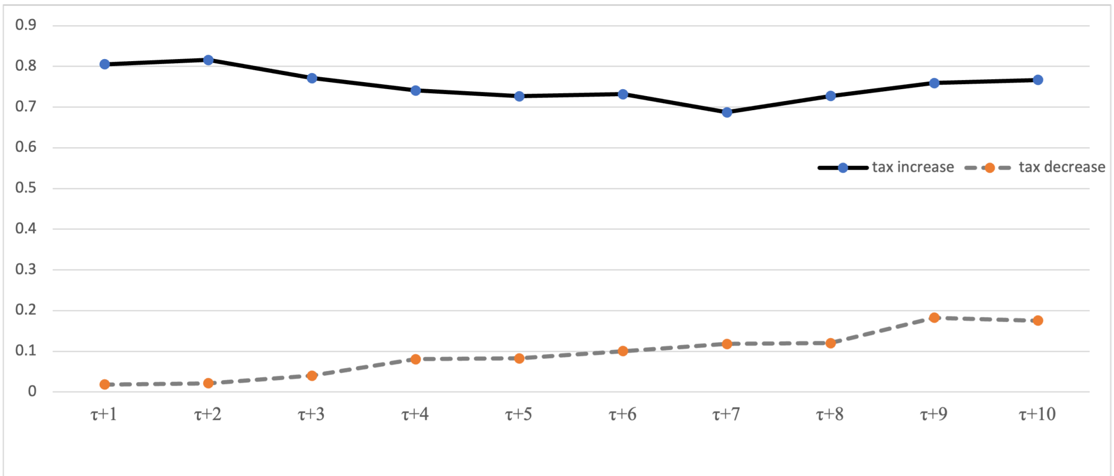
(b) Conditional Pass-through and Competition

Figure 3.1: Pass-through and Competition

Notes: The figure plots the estimated coefficients from Table A3, together with the 95% confidence interval.



(a) Average Pass-through and Speed of Adjustment



(b) Conditional Pass-through and Speed of Adjustment

Figure 3.2: Pass-through and Speed of Adjustment

Notes: The figure plots the estimated coefficients from Table A5. τ is the date of each of the four excise duty changes.

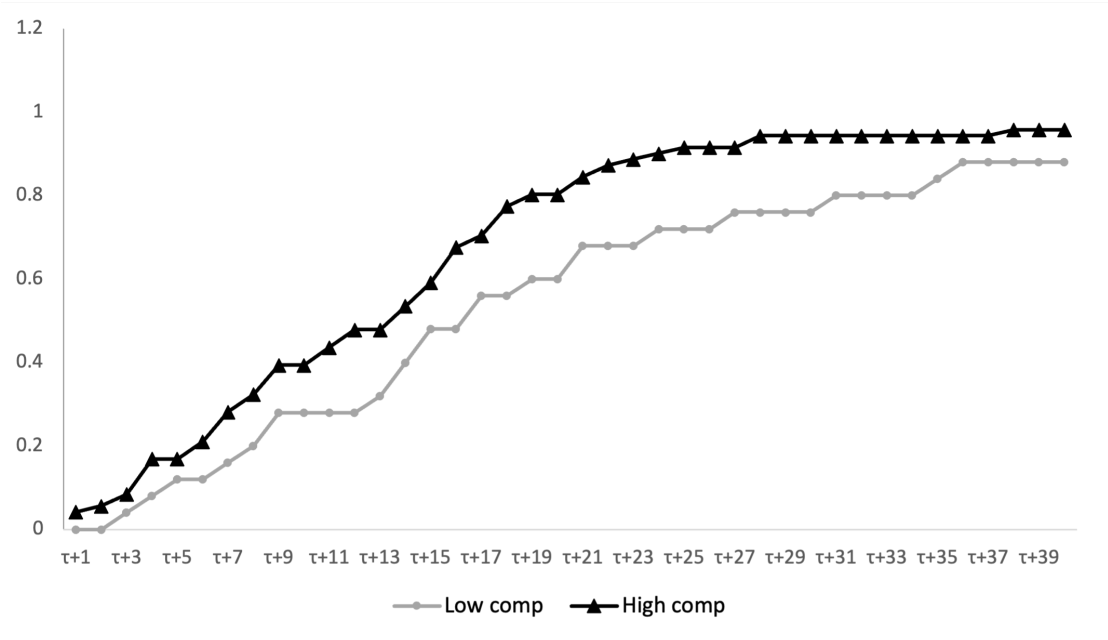


Figure 3.3: Cumulative Frequency of Price Changes

Notes: The figure plots the cumulative frequency of station–product combinations that changed their prices between τ and $\tau + \delta$, where τ is the date of the tax decrease event and δ=1,...,40, on islands with one to three (“low competition”) and four to seven (“high competition”) gas stations.

Appendix

Appendix A: Additional Tables

Table A1: Parallel Trend Tests

	(1)	(2)	(3)	(4)
Dependent variable	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$	$Price_{kist}$
Sample	Tax Increase 1	Tax Increase 2	Tax Increase 3	Tax Decrease
$Trend_t$	-0.017	0.052	0.041	-0.038
	(0.029)	(0.074)	(0.039)	(0.023)
$Trend_t$ × Treat	0.039	0.166	0.023	0.025
	(0.037)	(0.097)	(0.051)	(0.023)
Window before the event	$[\tau - 10, \tau - 1]$	$[\tau - 10, \tau - 1]$	$[\tau - 10, \tau - 1]$	$[\tau - 10, \tau - 1]$
Observations	1,196	1,552	1,750	2,011
Product type FE	Yes	Yes	Yes	Yes
Station FE	Yes	Yes	Yes	Yes

Notes: The table reports results for the parallel trend assumption test based on Eq. (3.3) in the main text. Standard errors clustered at the island are reported in parentheses below coefficients.

Table A2: Parallel Trend Tests (Non-parametric)

Dependent Variable Sample	(1) $Price_{kist}$ Tax Increase 1	(2) $Price_{kist}$ Tax Increase 2	(3) $Price_{kist}$ Tax Increase 3	(4) $Price_{kist}$ Tax Decrease
Day (T-10) $\times$ Treat	-0.395	-1.666	-0.329	-0.170
	(0.305)	(0.733)	(0.464)	(0.164)
Day (T-9) $\times$ Treat	-0.473	-1.078	-0.175	-0.166
	(0.373)	(0.756)	(0.372)	(0.164)
Day (T-8) $\times$ Treat	-0.341	-0.849	-0.017	-0.178
	(0.321)	(0.725)	(0.360)	(0.168)
Day (T-7) $\times$ Treat	-0.337	-0.909	0.107	-0.061
	(0.320)	(0.718)	(0.261)	(0.140)
Day (T-6) $\times$ Treat	-0.283	-0.909	0.073	0.021
	(0.314)	(0.717)	(0.237)	(0.116)
Day (T-5) $\times$ Treat	-0.263	-0.257	0.121	0.098
	(0.177)	(0.570)	(0.237)	(0.088)
Day (T-4) $\times$ Treat	-0.274	-0.553	-0.020	0.050
	(0.163)	(0.540)	(0.086)	(0.070)
Day (T-3) $\times$ Treat	-0.274	-0.282	-0.038	0.010
	(0.163)	(0.223)	(0.070)	(0.060)
Day (T-2) $\times$ Treat	-0.146	0.000	-0.022	0.007
	(0.130)	(0.000)	(0.017)	(0.011)
Joint test of significance (F-test)	1.579	2.106	0.0298	0.216
(p-value)	(0.231)	(0.167)	(0.864)	(0.645)
Window before the event	$[\tau-10,\tau-1]$	$[\tau-10,\tau-1]$	$[\tau-10,\tau-1]$	$[\tau-10,\tau-1]$
Observations	1,196	1,552	1,750	2,011
Time FE	Yes	Yes	Yes	Yes
Product type FE	Yes	Yes	Yes	Yes
Station FE	Yes	Yes	Yes	Yes

Notes: The table reports results for the parallel trend assumption test based on Eq. (3.3) in the main text, where the trend is replaced by day binary indicators. Only the interaction effects of day fixed effects with the treat variable are reported here. Standard errors clustered at the island level are reported in parentheses below coefficients.

Table A3: Pass-through and Competition (Non-linear)

	(1)	(2)
Estimation Method	FE	FE
Pass-through Definition	Average	Conditional
Dependent Variable	$Price_{kist}$	$Price_{kist}$
Sample	τ-1, τ+10	τ-1, τ+10
Tax_{it} × One competitor × U	0.427***	0.438***
	(0.133)	(0.134)
Tax_{it} × Two competitors x U	0.540***	0.580***
	(0.115)	(0.094)
Tax_{it} × Three competitors x U	0.748***	0.758***
	(0.045)	(0.047)
Tax_{it} × Four competitors x U	0.845***	0.983***
	(0.102)	(0.090)
Tax_{it} × Five competitors x U	0.833***	0.952***
	(0.168)	(0.118)
Tax_{it} × Six+ competitors x U	0.907***	0.923***
	(0.050)	(0.047)
Tax_{it} × One competitor × D	0.178	0.272
	(0.184)	(0.231)
Tax_{it} × Two competitors × D	0.147	0.240
	(0.140)	(0.176)
Tax_{it} × Three competitors × D	0.162	0.321
	(0.238)	(0.370)
Tax_{it} × Four competitors × D	0.425*	0.425*
	(0.236)	(0.237)
Tax_{it} × Five competitors × D	-0.077**	-0.073**
	(0.033)	(0.033)
Tax_{it} × Six+ competitors × D	0.115	0.124*
	(0.070)	(0.071)
Observations	1,341	1,286
Time FE	Yes	Yes
Product × Station FE	Yes	Yes
Excise incident × Product type FE	Yes	Yes
Excise incident × Station FE	Yes	Yes

Notes: The dependent variable is the retail price of product k, on island i, in gas station s, and day $t \in \{\tau - 1, \tau + 10\}$, where τ is the date of each of the four excise duty changes. In Column (1) we use all available data (average pass-through), whereas in Column (2) we use observations for station-product combinations that have changed the price at least once between τ and τ +10 (conditional pass-through). Standard errors clustered at the island level are reported in parentheses below coefficients. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Test Equality of Coefficients for Pass-through and Competition (Non-linear)

	Tax_{it} × 1 comp × D	Tax_{it} × 2 comp × D	Tax_{it} × 3 comp × D	Tax_{it} × 4 comp × D	Tax_{it} × 5 comp × D	Tax_{it} × 6 comp × D
Tax_{it} × 1 comp × U	0.270	0.153	0.341	0.994	0.001	0.043
Tax_{it} × 2 comp x U	0.111	0.036	0.164	0.670	0.000	0.003
Tax_{it} × 3 comp x U	0.007	0.000	0.025	0.181	0.000	0.000
Tax_{it} × 4 comp x U	0.003	0.000	0.005	0.103	0.000	0.000
Tax_{it} × 5 comp x U	0.015	0.004	0.030	0.177	0.000	0.001
Tax_{it} × 6 comp x U	0.001	0.000	0.004	0.060	0.000	0.000

Notes: This table reports the p-values from the test of equality of coefficients presented in Table A3, Column (1).

Table A5: Speed of Adjustment

	(1)	(2)	(3)	(4)
Dependent Variable	Pricekist	Pricekist	Pricekist	Pricekist
Pass-through definition	Average	Conditional	Average	Conditional
Sample	Symmetric (10 days)	Symmetric (10 days)	Asymmetric (10 days increase, 40 days decrease)	Asymmetric (10 days increase, 40 days decrease)
$Tax_{it} \times$ U	0.232**	0.805***	0.232**	0.805***
(τ-1, τ+1)	(0.109)	(0.128)	(0.109)	(0.128)
$Tax_{it} \times$ U	0.339***	0.816***	0.339***	0.816***
τ-1, τ+2	(0.099)	(0.125)	(0.099)	(0.125)
$Tax_{it} \times$ U	0.368***	0.771***	0.368***	0.771***
τ-1, τ+3	(0.097)	(0.116)	(0.097)	(0.116)
$Tax_{it} \times$ U	0.421***	0.741***	0.421***	0.741***
τ-1, τ+4	(0.089)	(0.107)	(0.089)	(0.107)
$Tax_{it} \times$ U	0.417***	0.727***	0.417***	0.727***
τ-1, τ+5	(0.089)	(0.106)	(0.089)	(0.106)
$Tax_{it} \times$ U	0.596***	0.732***	0.596***	0.732***
τ-1, τ+6	(0.111)	(0.109)	(0.111)	(0.109)
$Tax_{it} \times$ U	0.618***	0.687***	0.618***	0.687***
τ-1, τ+7	(0.116)	(0.108)	(0.116)	(0.108)
$Tax_{it} \times$ U	0.667***	0.727***	0.667***	0.727***
τ-1, τ+8	(0.115)	(0.111)	(0.115)	(0.111)
$Tax_{it} \times$ U	0.707***	0.759***	0.707***	0.759***
τ-1, τ+9	(0.103)	(0.100)	(0.103)	(0.100)
$Tax_{it} \times$ U	0.713***	0.767***	0.713***	0.767***
τ-1, τ+10	(0.101)	(0.096)	(0.101)	(0.096)
$Tax_{it} \times$ D	0.013	0.018	0.013	0.018
τ-1, τ+1	(0.018)	(0.031)	(0.018)	(0.031)
$Tax_{it} \times$ D	0.014	0.021	0.014	0.021
τ-1, τ+2	(0.018)	(0.030)	(0.018)	(0.030)
$Tax_{it} \times$ D	0.024	0.040	0.024	0.040
τ-1, τ+3	(0.020)	(0.031)	(0.020)	(0.031)
$Tax_{it} \times$ D	0.054**	0.081**	0.054**	0.081**
τ-1, τ+4	(0.023)	(0.034)	(0.023)	(0.034)
$Tax_{it} \times$ D	0.056**	0.083**	0.056**	0.083**
τ-1, τ+5	(0.023)	(0.033)	(0.023)	(0.033)
$Tax_{it} \times$ D	0.071**	0.100**	0.071**	0.100**

Continued on next page

τ-1, τ+6	(0.027)	(0.038)	(0.027)	(0.038)
Tax_{it} × D	0.089**	0.118**	0.089**	0.118**
τ-1, τ+7	(0.036)	(0.047)	(0.036)	(0.047)
Tax_{it} × D	0.091**	0.120**	0.091**	0.120**
τ-1, τ+8	(0.038)	(0.048)	(0.038)	(0.048)
Tax_{it} × D	0.145***	0.182***	0.145***	0.182***
τ-1, τ+9	(0.052)	(0.058)	(0.052)	(0.058)
Tax_{it} × D	0.142**	0.175***	0.142**	0.175***
τ-1, τ+10	(0.052)	(0.058)	(0.052)	(0.058)
Tax_{it} × D			0.165**	0.197***
τ-1, τ+11			(0.062)	(0.068)
Tax_{it} × D			0.165**	0.199***
τ-1, τ+12			(0.063)	(0.070)
Tax_{it} × D			0.174**	0.207***
τ-1, τ+13			(0.065)	(0.071)
Tax_{it} × D			0.191***	0.230***
τ-1, τ+14			(0.068)	(0.075)
Tax_{it} × D			0.205***	0.231***
τ-1, τ+15			(0.067)	(0.071)
Tax_{it} × D			0.216***	0.245***
τ-1, τ+16			(0.067)	(0.070)
Tax_{it} × D			0.222***	0.246***
τ-1, τ+17			(0.068)	(0.070)
Tax_{it} × D			0.230***	0.256***
τ-1, τ+18			(0.068)	(0.070)
Tax_{it} × D			0.232***	0.253***
τ-1, τ+19			(0.071)	(0.072)
Tax_{it} × D			0.221***	0.242***
τ-1, τ+20			(0.072)	(0.073)
Tax_{it} × D			0.249***	0.265***
τ-1, τ+21			(0.065)	(0.066)
Tax_{it} × D			0.225***	0.242***
τ-1, τ+22			(0.071)	(0.072)
Tax_{it} × D			0.237***	0.255***
τ-1, τ+23			(0.073)	(0.074)
Tax_{it} × D			0.231***	0.231***
τ-1, τ+24			(0.070)	(0.070)
Tax_{it} × D			0.253***	0.253***
τ-1, τ+25			(0.063)	(0.064)
Tax_{it} × D			0.263***	0.263***

Continued on next pag

τ-1, τ+26			(0.068)	(0.068)
Tax_{it} × D			0.259***	0.259***
τ-1, τ+27			(0.068)	(0.068)
Tax_{it} × D			0.241***	0.241***
τ-1, τ+28			(0.076)	(0.076)
Tax_{it} × D			0.246***	0.246***
τ-1, τ+29			(0.074)	(0.074)
Tax_{it} × D			0.225***	0.225***
τ-1, τ+30			(0.058)	(0.058)
Tax_{it} × D			0.221***	0.221***
τ-1, τ+31			(0.049)	(0.049)
Tax_{it} × D			0.231***	0.231***
τ-1, τ+32			(0.044)	(0.044)
Tax_{it} × D			0.237***	0.237***
τ-1, τ+33			(0.043)	(0.043)
Tax_{it} × D			0.239***	0.239***
τ-1, τ+34			(0.043)	(0.043)
Tax_{it} × D			0.244***	0.244***
τ-1, τ+35			(0.042)	(0.042)
Tax_{it} × D			0.252***	0.252***
τ-1, τ+36			(0.042)	(0.042)
Tax_{it} × D			0.271***	0.271***
τ-1, τ+37			(0.039)	(0.039)
Tax_{it} × D			0.267***	0.267***
τ-1, τ+38			(0.043)	(0.043)
Tax_{it} × D			0.258***	0.258***
τ-1, τ+39			(0.045)	(0.045)
Tax_{it} × D			0.264***	0.264***
τ-1, τ+40			(0.046)	(0.046)
Day FE	Yes	Yes	Yes	Yes
Product × Station FE	Yes	Yes	Yes	Yes
Excise change × Product type FE	Yes	Yes	Yes	Yes
Excise change × Station FE	Yes	Yes	Yes	Yes

Notes: Each coefficient comes from a separate regression. The dependent variable is the retail price of product k, on island i, in gas station s, and day $t \in \{\tau - 1, \tau + \delta\}$, where τ is the date of each of the three excise duty changes and δ=1,...,40 is the adjustment period. The average pass-through is estimated using all the data. The conditional pass-through is estimated using observations for station-product combinations that have changed the price at least once between τ and $\tau + \delta$. Standard errors clustered at the island level are reported in parentheses below coefficients. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Speed of Adjustment and Competition

Panel A: Average Pass-through

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Estimation Method	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Dependent Variable	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$
Sample	All changes	All changes	All changes	All changes	All changes	All changes	All changes	All changes	All changes	All changes
	(τ-1, τ+1)	(τ-1, τ+2)	(τ-1, τ+3)	(τ-1, τ+4)	(τ-1, τ+5)	(τ-1, τ+6)	(τ-1, τ+7)	(τ-1, τ+8)	(τ-1, τ+9)	(τ-1, τ+10)
Tax_{it} × Low competition × U	0.136*	0.200**	0.198**	0.273***	0.268***	0.410***	0.443***	0.456***	0.519***	0.531***
(1-3 competitors)	(0.072)	(0.080)	(0.076)	(0.063)	(0.063)	(0.103)	(0.102)	(0.101)	(0.092)	(0.090)
Tax_{it} × High competition × U	0.301*	0.433***	0.500***	0.534***	0.534***	0.747***	0.766***	0.831***	0.855***	0.856***
(4-7 competitors)	(0.149)	(0.126)	(0.116)	(0.109)	(0.109)	(0.123)	(0.132)	(0.115)	(0.105)	(0.105)
Test equality of coefficients (p-value)	0.216	0.051	0.007	0.008	0.007	0.010	0.017	0.004	0.006	0.008
Tax_{it} × Low competition × D	0.056	0.055	0.092	0.090	0.088	0.085	0.067	0.064	0.052	0.052
(1-3 competitors)	(0.053)	(0.053)	(0.061)	(0.061)	(0.061)	(0.061)	(0.062)	(0.062)	(0.062)	(0.062)
Tax_{it} × High competition × D	-0.005	-0.002	0.003	0.039	0.033	0.060	0.036	0.046	0.115	0.111
(4-7 competitors)	(0.010)	(0.011)	(0.015)	(0.024)	(0.025)	(0.037)	(0.052)	(0.056)	(0.076)	(0.076)
Test equality of coefficients (p-value)	0.263	0.298	0.172	0.453	0.420	0.740	0.708	0.829	0.507	0.541
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Product × Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Product type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,320	1,316	1,351	1,352	1,355	1,363	1,372	1,390	1,396	1,395

Panel B: Conditional Pass-through

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Estimation Method	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Dependent Variable	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$	$Price_{ist}$
Sample	All changes	All changes	All changes	All changes	All changes	All changes	All changes	All changes	All changes	All changes
	(τ-1, τ+1)	(τ-1, τ+2)	(τ-1, τ+3)	(τ-1, τ+4)	(τ-1, τ+5)	(τ-1, τ+6)	(τ-1, τ+7)	(τ-1, τ+8)	(τ-1, τ+9)	(τ-1, τ+10)
Tax_{it} × Low competition × U	0.639***	0.614***	0.528***	0.528***	0.523***	0.509***	0.486***	0.502***	0.552***	0.565***
(1-3 competitors)	(0.106)	(0.089)	(0.076)	(0.070)	(0.070)	(0.103)	(0.094)	(0.091)	(0.083)	(0.080)
Tax_{it} × High competition × U	0.888***	0.952***	0.966***	0.953***	0.932***	0.939***	0.886***	0.926***	0.948***	0.951***
(4-7 competitors)	(0.116)	(0.079)	(0.056)	(0.059)	(0.065)	(0.088)	(0.093)	(0.087)	(0.074)	(0.074)
Test equality of coefficients (p-value)	0.048	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000
Tax_{it} × Low competition × D	0.146	0.145	0.148	0.146	0.119	0.115	0.096	0.093	0.080	0.081
(1-3 competitors)	(0.125)	(0.124)	(0.088)	(0.088)	(0.077)	(0.077)	(0.078)	(0.077)	(0.078)	(0.078)
Tax_{it} × High competition × D	-0.012	-0.005	0.005	0.052	0.045	0.076*	0.044	0.052	0.120	0.116
(4-7 competitors)	(0.017)	(0.017)	(0.021)	(0.031)	(0.032)	(0.045)	(0.058)	(0.061)	(0.078)	(0.078)
Test equality of coefficients (p-value)	0.220	0.242	0.132	0.330	0.394	0.669	0.590	0.677	0.717	0.746
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Product × Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Product type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excise change × Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	811	904	1,002	1,079	1,100	1,188	1,248	1,287	1,306	1,307

Notes: The dependent variable is the retail price of product k, on island i, in gas station s, and day $t \in \{\tau - 1, \tau + \delta\}$, where τ is the date of each of the four excise duty changes and δ=1,...,10. Standard errors clustered at the gas station level are reported in parentheses below coefficients. *** p<0.01, ** p<0.05, * p<0.1.

REFERENCES

Adachi, T. and M. Fabinger (2022). Pass-through, welfare, and incidence under imperfect competition. *Journal of Public Economics 211*, 104589.

Ahmed, K., S. Hashim, M. Khankhara, I. Said, A. T. Shandakumar, S. Zaman, and A. Veiga (2021). What drives general practitioners in the uk to improve the quality of care? a systematic literature review. *BMJ Open Quality 10*(1), e001127.

American Medical Association (2022). Policy research perspectives.

Ashenfelter, O. C., D. S. Hosken, and M. C. Weinberg (2013). The price effects of a large merger of manufacturers: A case study of maytag-whirlpool. *American Economic Journal: Economic Policy 5*(1), 239–261.

Asker, J. and V. Nocke (2021). Collusion, mergers, and related antitrust issues. In *Handbook of industrial organization*, Volume 5, pp. 177–279. Elsevier.

Bacon, R. W. (1991). Rockets and feathers: the asymmetric speed of adjustment of uk retail gasoline prices to cost changes. *Energy economics 13*(3), 211–218.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Bakucs, Z., J. Fałkowski, and I. Fertő (2014). Does market structure influence price transmission in the agro-food sector? a meta-analysis perspective. *Journal of Agricultural Economics 65*(1), 1–25.

Ball, L. and N. G. Mankiw (1994). Asymmetric price adjustment and economic fluctuations. *The Economic Journal 104*(423), 247–261.

Barron, J. M., B. A. Taylor, and J. R. Umbeck (2004). Number of sellers, average prices, and price dispersion. *International Journal of Industrial Organization 22*(8-9), 1041–1066.

Beckert, W., M. Christensen, and K. Collyer (2012). Choice of nhs-funded hospital services in england. *The Economic Journal 122*(560), 400–417.

Benzarti, Y., D. Carloni, J. Harju, and T. Kosonen (2020). What goes up may not come down: asymmetric incidence of value-added taxes. *Journal of Political Economy 128*(12), 4438–4474.

Berry, S. T. and J. Waldfogel (2001). Do mergers increase product variety? evidence from radio broadcasting. *The Quarterly Journal of Economics 116*(3), 1009–1025.

Blinder, A. S. (1982). Inventories and sticky prices: More on the microfoundations of macroeconomics. *American Economic Review 72*(3), 334–348.

Borenstein, S., A. C. Cameron, and R. Gilbert (1997). Do gasoline prices respond asymmetrically to crude oil price changes? *The Quarterly journal of economics 112*(1), 305–339.

Borenstein, S. and A. Shepard (1996). Sticky prices, inventories, and market power in wholesale gasoline markets.

Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

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

Brynjolfsson, E. and M. D. Smith (2000). Frictionless commerce? a comparison of internet and conventional retailers. *Management science 46*(4), 563–585.

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

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

Byrne, D. P. and N. De Roos (2022). Start-up search costs. *American Economic Journal: Microeconomics 14*(2), 81–112.

Byrne, D. P., G. W. Leslie, and R. Ware (2015). How do consumers respond to gasoline price cycles? *The Energy Journal 36*(1).

Cabral, L. and S. Gilbukh (2020). Rational buyers search when prices increase. *Journal of Economic Theory 187*, 104998.

Caliendo, M. and S. Kopeinig (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys 22*(1), 31–72.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230.

Capps, C. S. (2005). *The quality effects of hospital mergers*. Number 6. US Department of Justice, Antitrust Division.

Casalino, L. P. (2006). Which type of medical group provides higher-quality care? *Annals of Internal Medicine 145*(11), 860–861.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134*(3), 1405–1454.

Chandra, A. and M. Tappata (2011). Consumer search and dynamic price dispersion: an application to gasoline markets. *The RAND Journal of Economics 42*(4), 681–704.

Chauhan, B. F., M. Jeyaraman, A. S. Mann, J. Lys, B. Skidmore, K. M. Sibley, A. Abou-Setta, and R. Zarychanksi (2017). Behavior change interventions and policies influencing primary healthcare professionals' practice—an overview of reviews. *Implementation Science 12*(1), 1–16.

Cutler, D., L. Dafny, and C. Ody (2017). How does competition impact the quality of price of outpatient service facilities? a case study of the us dialysis industry. Technical report, Working Paper.

Dafny, L. (2021). How health care consolidation is contributing to higher prices and spending, and reforms that could bolster antitrust enforcement and preserve and promote competition in health care markets. *testimony before the US Committee on the Judiciary, Subcommittee on Antitrust, Commercial and Administrative Law*.

Dafny, L., K. Ho, and R. S. Lee (2019). The price effects of cross-market mergers: theory and evidence from the hospital industry. *The RAND Journal of Economics 50*(2), 286–325.

Dahlby, B. and D. S. West (1986). Price dispersion in an automobile insurance market. *Journal of Political Economy 94*(2), 418–438.

Deltas, G. (2008). Retail gasoline price dynamics and local market power. *The Journal of Industrial Economics 56*(3), 613–628.

Deshpande, M. and Y. Li (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy 11*(4), 213–48.

Diamond, P. A. (1971). A model of price adjustment. *Journal of economic theory 3*(2), 156–168.

Dietrichson, J., L. M. Ellegård, and G. Kjellsson (2020). Patient choice, entry, and the quality of primary care: Evidence from swedish reforms. *Health Economics 29*(6), 716–730.

Dossa, A. R., J. Moisan, L. Guénette, S. Lauzier, and J.-P. Grégoire (2017). Association between interpersonal continuity of care and medication adherence in type 2 diabetes: an observational cohort study. *Canadian Medical Association Open Access Journal 5*(2), E359–E364.

Eckert, A. and D. S. West (2005). Price uniformity and competition in a retail gasoline market. *Journal of Economic Behavior & Organization 56*(2), 219–237.

Eliason, P. J., B. Heebsh, R. C. McDevitt, and J. W. Roberts (2020). How acquisitions affect firm behavior and performance: Evidence from the dialysis industry. *The Quarterly Journal of Economics 135*(1), 221–267.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review 103*(5), 1598–1628.

Fan, Y. and C. Yang (2022). Estimating discrete games with many firms and many decisions: An application to merger and product variety. Technical report, National Bureau of Economic Research.

Freeman, G., J. Hughes, et al. (2010). Continuity of care and the patient experience.

Frey, G. and M. Manera (2007). Econometric models of asymmetric price transmission. *Journal of Economic surveys 21*(2), 349–415.

Fulton, B. D. (2017). Health care market concentration trends in the united states: evidence and policy responses. *Health Affairs 36*(9), 1530–1538.

Gaynor, M. (2004). Quality and competition in health care markets: What do we know? what don't we know. *Economie Publique 15*(2), 87–124.

Gaynor, M. (2006). What do we know about competition and quality in health care markets? unpublished paper.

Gaynor, M., K. Ho, and R. J. Town (2015). The industrial organization of health-care markets. *Journal of Economic Literature 53*(2), 235–284.

Gaynor, M., M. Laudicella, and C. Propper (2012). Can governments do it better? merger mania and hospital outcomes in the english nhs. *Journal of health economics 31*(3), 528–543.

Gaynor, M. and R. J. Town (2011). Competition in health care markets. *Handbook of health economics 2*, 499–637.

Genakos, C. and M. Pagliero (2022). Competition and pass-through: evidence from isolated markets. *American Economic Journal: Applied Economics 14*(4), 35–57.

George, L. M. (2002). Ownership concentration and product variety in daily newspaper markets. *Communications policy and information technology: Promises, problems, prospects*, 235–253.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277.

Gopinath, G. and O. Itskhoki (2010). Frequency of price adjustment and pass-through. *The Quarterly Journal of Economics 125*(2), 675–727.

Gravelle, H., D. Liu, C. Propper, and R. Santos (2019). Spatial competition and quality: Evidence from the english family doctor market. *Journal of Health Economics 68*, 102249.

Gravelle, H., D. Liu, and R. Santos (2022). How do clinical quality and patient satisfaction vary with provider size in primary care? evidence from english general practice panel data. *Social Science & Medicine 301*, 114936.

Green, E. J. and R. H. Porter (1984). Noncooperative collusion under imperfect price information. *Econometrica: Journal of the Econometric Society*, 87–100.

Hellenic Statistical Authority (2010). Greece in figures. https://www.statistics.gr/en/home. Accessed July 2018.

Ho, V. and B. H. Hamilton (2000). Hospital mergers and acquisitions: does market consolidation harm patients? *Journal of Health Economics 19*(5), 767–791.

Hortaçsu, A. and C. Syverson (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly journal of economics 119*(2), 403–456.

Hosken, D. S., R. S. McMillan, and C. T. Taylor (2008). Retail gasoline pricing: What do we know? *International Journal of Industrial Organization 26*(6), 1425–1436.

International Energy Agency (2017). *Energy Policies of IEA Countries: Greece 2017 Review*. Paris: IEA. Licence: CC BY 4.0.

Jeziorski, P. (2014). Effects of mergers in two-sided markets: The us radio industry. *American Economic Journal: Microeconomics 6*(4), 35–73.

Johnson, R. N. (2002). Search costs, lags and prices at the pump. *Review of Industrial Organization 20*, 33–50.

Kajaria-Montag, H. and M. Freeman (2020). Explaining the erosion of relational care continuity: an empirical analysis of primary care in england.

Kelly, E. and G. Stoye (2014). *Does GP practice size matter? GP practice size and the quality of primary care*. Number R101. IFS Report.

Kontopantelis, E., D. Springate, D. Reeves, D. M. Ashcroft, J. M. Valderas, and T. Doran (2014). Withdrawing performance indicators: retrospective analysis of general practice performance under uk quality and outcomes framework. *Bmj 348*.

Kovenock, D. and K. Widdows (1998). Price leadership and asymmetric price rigidity. *European Journal of Political Economy 14*(1), 167–187.

Lach, S. (2002). Existence and persistence of price dispersion: an empirical analysis. *Review of economics and statistics 84*(3), 433–444.

Lewis, M. (2008). Price dispersion and competition with differentiated sellers. *The Journal of Industrial Economics 56*(3), 654–678.

Lewis, M. S. (2011). Asymmetric price adjustment and consumer search: An examination of the retail gasoline market. *Journal of Economics & Management Strategy 20*(2), 409–449.

Lewis, M. S. and H. P. Marvel (2011). When do consumers search? *The Journal of Industrial Economics 59*(3), 457–483.

Lewis, M. S. and K. E. Pflum (2017). Hospital systems and bargaining power: evidence from out-of-market acquisitions. *The RAND Journal of Economics 48*(3), 579–610.

Luco, F. (2019). Who benefits from information disclosure? the case of retail gasoline. *American Economic Journal: Microeconomics 11*(2), 277–305.

McCauley, L., R. L. Robert L. Phillips, Jr., M. Meisnere, and S. K. Robinson (2021). *Implementing high-quality primary care: rebuilding the foundation of health care*. National Academies Press (US).

Mehrotra, A., A. M. Epstein, and M. B. Rosenthal (2006). Do integrated medical groups provide higher-quality medical care than individual practice associations? *Annals of Internal Medicine 145*(11), 826–833.

Meyer, J. and S. von Cramon-Taubadel (2004). Asymmetric price transmission: a survey. *Journal of agricultural economics 55*(3), 581–611.

Miklós-Thal, J. and G. Shaffer (2021). Pass-through as an economic tool: on exogenous competition, social incidence, and price discrimination. *Journal of Political Economy 129*(1), 323–335.

Minchin, M., M. Roland, J. Richardson, S. Rowark, and B. Guthrie (2018). Quality of care in the united kingdom after removal of financial incentives. *New England Journal of Medicine 379*(10), 948–957.

Montag, F., A. Sagimuldina, and M. Schnitzer (2021). Does tax policy work when consumers have imperfect price information? theory and evidence.

Noel, M. D. (2018). Gasoline price dispersion and consumer search: evidence from a natural experiment. *The Journal of Industrial Economics 66*(3), 701–738.

Noel, M. D. and H. Qiang (2019). The role of information in retail gasoline price dispersion. *Energy Economics 80*, 173–187.

Orlov, E. (2011). How does the internet influence price dispersion? evidence from the airline industry. *The Journal of Industrial Economics 59*(1), 21–37.

Pál, V., G. Lados, Z. Ilcsikné Makra, L. Boros, A. Uzzoli, and S. Fabula (2021). Concentration and inequality in the geographic distribution of physicians in the european union, 2006–2018. *Regional Statistics 11*(3), 3–28.

Peltzman, S. (2000). Prices rise faster than they fall. *Journal of political economy 108*(3), 466–502.

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

Remer, M. (2015). An empirical investigation of the determinants of asymmetric pricing. *International Journal of Industrial Organization 42*, 46–56.

Renneboog, L. and C. Vansteenkiste (2019). Failure and success in mergers and acquisitions. *Journal of Corporate Finance 58*, 650–699.

Robinson, J. C. (1999). *The corporate practice of medicine: competition and innovation in health care*, Volume 1. Univ of California Press.

Roland, M. and B. Guthrie (2016). Quality and outcomes framework: what have we learnt? *Bmj 354*.

Roland, M. and F. Olesen (2016). Can pay for performance improve the quality of primary care? *Bmj 354*.

Rosenbaum, P. R. and D. B. Rubin (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician 39*(1), 33–38.

Santos, R., H. Gravelle, and C. Propper (2017). Does quality affect patients' choice of doctor? evidence from england. *The Economic Journal 127*(600), 445–494.

Schmitt, M. (2017). Do hospital mergers reduce costs? *Journal of health economics 52*, 74–94.

Scott, A., J. Li, H. Gravelle, and M. McGrail (2022). Physician competition and low-value health care. *American Journal of Health Economics 8*(2), 252–274.

Shanafelt, T. D., C. M. Balch, G. Bechamps, T. Russell, L. Dyrbye, D. Satele, P. Collicott, P. J. Novotny, J. Sloan, and J. Freischlag (2010). Burnout and medical errors among american surgeons. *Annals of surgery 251*(6), 995–1000.

Shepard, A. (1991). Price discrimination and retail configuration. *Journal of Political Economy 99*(1), 30–53.

Sherman, J. and A. Weiss (2017). On fruitful and futile tests of the relationship between search and price dispersion. *Economic Inquiry 55*(4), 1898–1918.

Sorensen, A. T. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy 108*(4), 833–850.

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

Stiebale, J. and F. Szücs (2022). Mergers and market power: evidence from rivals' responses in european markets. *The RAND Journal of Economics 53*(4), 678–702.

Stigler, G. J. (1961). The economics of information. *Journal of political economy 69*(3), 213–225.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.

Sweeting, A. (2010). The effects of mergers on product positioning: evidence from the music radio industry. *The RAND Journal of Economics 41*(2), 372–397.

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

Tappata, M. (2009). Rockets and feathers: Understanding asymmetric pricing. *The RAND Journal of Economics 40*(4), 673–687.

Van den Hombergh, P., B. Künzi, G. Elwyn, J. van Doremalen, R. Akkermans, R. Grol, and M. Wensing (2009). High workload and job stress are associated with lower practice performance in general practice: an observational study in 239 general practices in the netherlands. *BMC Health Services Research 9*(1), 1–8.

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

Verlinda, J. A. (2008). Do rockets rise faster and feathers fall slower in an atmosphere of local market power? evidence from the retail gasoline market. *The Journal of Industrial Economics 56*(3), 581–612.

Weyl, E. G. and M. Fabinger (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of political economy 121*(3), 528–583.

Weyrauch, K. F., L. Rhodes, B. M. Psaty, and D. Grubb (1995). The role of physicians' personal knowledge of the patient in clinical practice. *Journal of Family Practice 40*(3), 249–256.

Yang, H. and L. Ye (2008). Search with learning: understanding asymmetric price adjustments. *The RAND Journal of Economics 39*(2), 547–564.

Yellow Pages (2018). Yellow pages (greece). https://www.xo.gr. Accessed July 2018.